Posters

August 28, 2023 (Mon.) [Poster Session I]

Efficient Learning of Mixed-State Tomography for Photonic Quantum Walk

Qin-Qin Wang,[1,2] Shaojun Dong,[3] Xiao-Wei Li,[4] Xiao-Ye Xu,[1,2,5,*] Chao Wang,[3] Shuai Han,[6] Man-Hong Yung,[7,†] Yong-Jian Han,[1,2,3,5,‡] Chuan-Feng Li,[1,2,5,§] and Guang-Can Guo[1,2,5]

[1] *CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China*
[2] *CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*
[3] *Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230031, China*
[4] *Department of Physics, Southern University of Science and Technology, Shenzhen 518055, China*
[5] *Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*
[6] *Yangtze Delta Region Industrial Innovation Center of Quantum and Information Technology, Suzhou 215100, China*
[7] *Institute for Quantum Science and Engineering, Southern University of Science and Technology, Shenzhen 518055, China*

The use of noise-enhanced applications in the field of open quantum walk (QW) has recently seen a surge due to their ability to improve performance by introducing artificial noise to quantum evolution. However, verifying the success of open QW remains a significant challenge, as state tomography for density matrix is a resource-intensive process, and implementing all required measurements is almost impossible due to various physical constraints. To address this challenge, we present a neural-network-based method for reconstructing mixed states with a high fidelity ($\sim 98\%$), while costing only 50% of the number of measurements typically required for open QW. Our method employs a neural density operator that models both the system and the environment, followed by a generalized natural-gradient-descent procedure that significantly speeds up the training process. In addition, we introduce a compact realization of the interferometric measurements, which improves the scalability of our photonic QW setup that enables experimental learning of the mixed state. Our results demonstrate that highly expressive neural networks can serve as powerful and practical alternatives to traditional state tomography, enabling researchers to study a broader range of problems in both Hermitian and non-Hermitian physics using photonic platforms.

Introduction

Quantum Walk (QW) provides a basic framework for developing effective quantum algorithms and simulating complex phenomena [1–3]. Through interaction with a certain amount of noise, the open (i.e. noisy) QW can produce a remarkable improvement in quantum transport over the noise-free QW [4, 5]. And such a noise-enhanced feature aids in problem-solving efficiency in tasks such as graph isomorphism testing [6], a maze escape [7, 8], and ranking elements [9] in large networks. While as a prototypical dynamical process, the open QW can model the dissipative evolution of quantum neural networks [10, 11], and such simulation enables better performance to process various issues like pattern recognition [12]. Adding controlled noise into the quantum evolution, QW can be dynamically initialized in any high-dimensional form [13–15] and generate the Haar random unitary operators [16, 17] required for quantum computation. Implementing the open QW and validating these noise-assisted computational and simulated performances demand a complete density-matrix characterization. However, identifying such a system with inherent high-dimensional structures is indeed a fundamental task to be solved in quantum information science [18].

For a discrete-time QW [19], the usual local tomographic technique was experimentally demonstrated that enables one to access the two correlated subsystems of coin [20–22] and position [23, 24], respectively. The tomography was subsequently extended to the walker's complete wave-function with a pure-state hypothesis [25], or a low-rank state when decoherence should be estimated [26]. An alternative tomographic method was recently developed with machine-learning techniques [27–29], where a trained neural network can be used to recognize and classify QW states [30, 31]. The complicated mapping from the measured data to the probability amplitudes can also be learned supervisedly [32], and a "physical" wave-function of QW is obtained by projecting the un-physical neural-network outputs [33]. Nevertheless, these methods mainly focus on the tomography of a closed QW, and none of them could be applicable for the mixed density-matrix reconstruction of an open QW, whose number of parameters scales quadratically with the pure-state one.

Here, we experimentally realize the full density-matrix tomography of an open QW with arbitrary mixing. We achieve this by parameterizing the complex-valued matrix elements using a neural density operator (NDO) [34], because such an effective ansatz has been shown to have better performance than standard maximum likelihood estimation in terms of the reconstruction fidelity and the number of measurements required. For training NDO to maximize the measurement data likelihood more efficiently, the generalized natural gradient descent (GNGD) procedure, recently developed in our group [35], is adopted to achieve about an order of magnitude of convergence rate than traditional gradient descent procedure. The effective mixed-state tomography of open

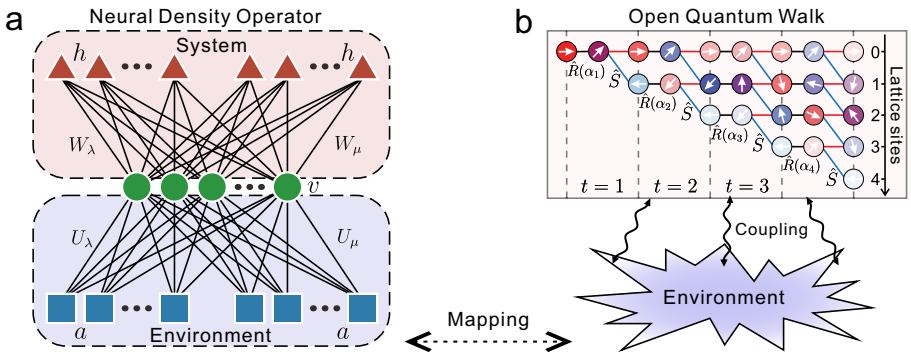


FIG. 1. **A schematic diagram of NDO and open QW.** (a) $|v\rangle$ is encoded in the visible layer (green circles), a hidden layer h (red triangles) captures the correlation within the physical system, and an ancillary layer a (blue squares) encodes the coupling between the system and the environment. Matrices $W_{\lambda,\mu}$ and $U_{\lambda,\mu}$ are the connecting weights. (b) Open QW consists of two degrees of freedom: the coin (white arrows) and lattice (colored circles), and the coupling environment introduces the mixing of this system.

QW, with a lower requirement of the number of measurements and training iterations for high-fidelity reconstructions, is benchmarked on synthetic datasets. Moreover, we further experimentally demonstrate that the trained NDO can learn noisy quantum states from partial measurements on a photonic QW. And the amount of mixing of the QW caused by the interaction with the environment can be well captured.

Results

Neural-network ansatz

The quantum state of open systems is generally described by the density matrix $\rho = \sum_{v,v'} \rho(v,v') |v\rangle \langle v'|$ in a basis $|v\rangle$. To obtain the complex-valued matrix elements, a tomography with a series of measurements on a collection of bases $\{v^n\}$ is necessary. The substantial demand for the number of complete tomographic measurements and a complicated data post-processing process make it impractical for a complex system [18], whereas neural-network-assisted tomography enables high-fidelity reconstruction of quantum states with fewer measurement resources [36]. The core step is an effective variational parametrization of the measurement distributions on a quantum state in terms of a neural-network model to infer the complex-valued probability amplitudes [37], and the density matrix elements that rely on a linear reconstruction from the trained model distributions [38]. The linear reconstruction of ρ lacks the positivity constraint such that it cannot give a physical quantum state. To give a physical output, a NDO ansatz can be constructed by purifying the mixed state with an auxiliary Hilbert space $\mathcal{H}_a$ [34]. Thus, ρ is written as the reduced state of a pure one $|\Psi\rangle = \sum_{va} \Psi(v,a) |v\rangle |a\rangle$, that is, $\rho = \mathrm{Tr}_a[|\Psi\rangle\langle\Psi|]$. The probability amplitude $\Psi(v,a)$ can be parameterized by a three-layer network shown by Fig. 1(a):

$$\Psi_{\lambda\mu}(v,a) = Z_\lambda^{-\frac{1}{2}} \sqrt{p_\lambda(v,a)} e^{i \log p_\mu(v,a)/2}. \quad (1)$$

The joint distribution of the visible and the ancillary layers $p_\theta(v,a) = e^{\sum_i \log(1+e^{W_\theta^{[i]} v + c_\theta^{[i]}}) + a^{\mathrm{T}} U_\theta v + b_\theta^{\mathrm{T}} v + d_\theta^{\mathrm{T}} a}$ ($\theta = \{\lambda,\mu\}$), where $A^{[i]}$ denotes ith row of A and A^{T} is the transpose of A. Vectors b_θ, c_θ, and d_θ are the biases coupled to the visible, hidden, and ancillary neurons, respectively. Z_λ is a normalization constant. The NDO ansatz of the density matrix can be obtained by tracing out the auxiliary system [39–41]:

$$\rho_{\lambda\mu} = \sum_{v,v'} [\sum_a \Psi_{\lambda\mu}(v,a) \Psi^*_{\lambda\mu}(v,a)] |v\rangle \langle v'| . \quad (2)$$

Thus, the state tomography is mapped into an unsupervised learning task for training the network parameters such that the trained NDO gives an approximate physical state, i.e., $\rho_{\lambda\mu} \simeq \rho$. The expressive ability of the NDO ansatz depends on the number of the hidden and ancillary neurons. We now show how to accomplish the NDO tomography for an open QW.

Open quantum walk model

As shown in Fig. 1(b), QW is composed of two interacting subsystems of the coin and the lattice, in which the probability amplitudes of a quantum walker spreads between lattice sites depending on the internal coin states [19]. The quantum state of QW is fully described by a density matrix acting on the product Hilbert space $\mathcal{H} \equiv \mathcal{H}_C \otimes \mathcal{H}_l$, where $\mathcal{H}_C \equiv \mathrm{span}\{|s\rangle : \mu = \uparrow, \downarrow\}$ and $\mathcal{H}_l \equiv \mathrm{span}\{|l\rangle : l \in \mathbb{Z}\}$ represent the coin and the lattice subspace, respectively. Then the reference base is a tensor product of the basis for each subsystem $|v\rangle = |s\rangle |l\rangle$. And the density matrix on this base takes the form:

$$\rho = \sum_{s,l;s',l'} \rho_{sl,s'l'} |s\rangle |l\rangle \langle s'| \langle l'| . \quad (3)$$

The open QW at each step t follows the non-unitary dynamics [42]: $\rho_{t+1} = (1-w)\hat{U}\rho_t\hat{U}^\dagger + w\sum_i \hat{K}_i\hat{U}\rho_t\hat{U}^\dagger\hat{K}_i^\dagger$, which describes a coherent QW evolution mixed with the

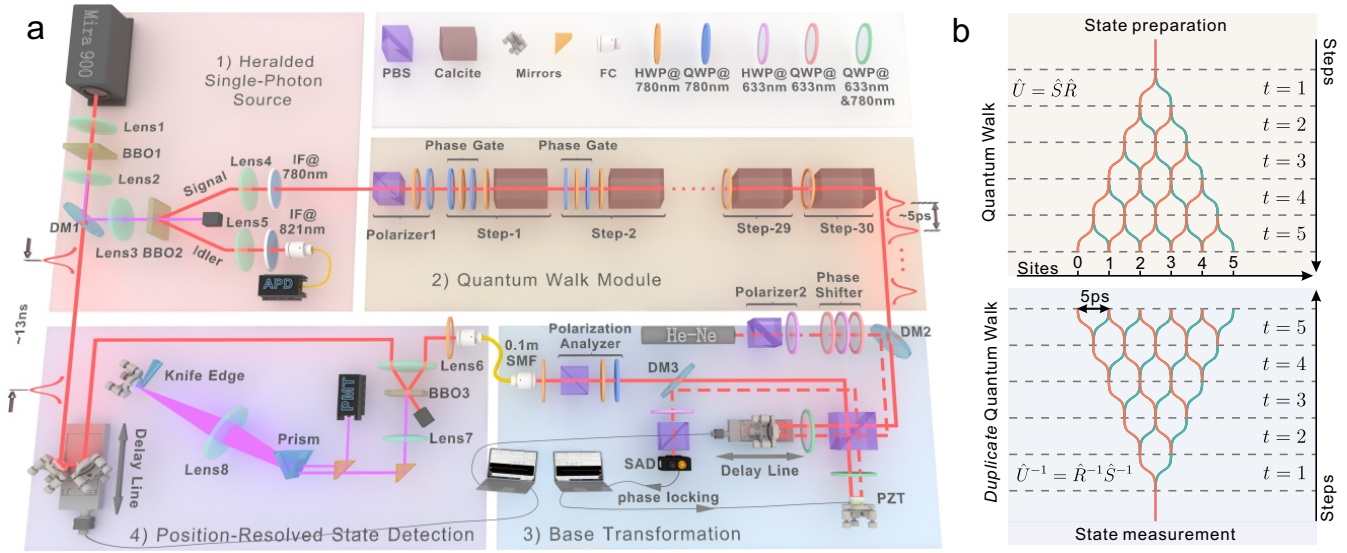


FIG. 2. **Photonic open QW.** (a) The experimental setup mainly has four central parts: 1) Spontaneous parameter down-conversion generates the time-correlated photon pairs, where signal photons as the walker and the idler photons serve to herald; 2) The open QW, as reported in the upper panel of (b), involves the cascade of HWPs (for operator $\hat{R}(\alpha)$) and calcites (for operator $\hat{S}$) that realizes the unitary dynamics $\hat{U}^t$. An additional varying phase gate $\hat{R}(\beta)$ with the QWP-HWP-QWP setting is added to each time step to introduce incoherent contributions $\hat{K}_i$; 3) A Michelson interferometer together with a polarization analyzer carries out the base transformation $\hat{\mathcal{U}}(v^n, v)$ on the reference base $|v\rangle$, which mimics the effect of a *duplicate* QW in the lower panel of (b). The PZT-driven mirror is applied for phase-locking by using a reference He-Ne laser, and another moveable mirror is used for adjustable arm length difference; 4) A single-photon frequency up-conversion implements the position-resolved detection of the walker in each base. (b) The schematic diagrams show the QW dynamics of a localized initial state (top) and the duplicate QW with time inversion constructed for the state measurements in different bases (bottom), respectively. A list of abbreviations: $\beta - BaB_2O_4$ (BBO); dichroic mirror (DM); interference filter (IF); polarization-dependent beam splitter (PBS); half-wave plate (HWP); quarter-wave plate (QWP); piezoelectric ceramic (PZT); fiber collimator (FC); single-mode fiber (SMF); Si amplified detector (SAD); photomultiplier tube (PMT); avalanche photodiode detector (APD).

incoherent contributions given by the Kraus operators $\hat{K}_i = |i\rangle\langle i|$. The mixing parameter $w \in [0, 1]$ quantifies the transition from the closed QW to the classical random walk (CRW). The unitary operator $\hat{U}$ for a time step is given by $\hat{U} = \hat{S}\hat{R}$, where coin-flip operator $\hat{R}(\alpha) = e^{-i\alpha\hat{\sigma}_y}\hat{\sigma}_z$ and conditional shift operator $\hat{S} = \sum_l (|\uparrow\rangle\langle\uparrow| \otimes |l+1\rangle\langle l| + |\downarrow\rangle\langle\downarrow| \otimes |l\rangle\langle l|)$. $\hat{\sigma}_y$ and $\hat{\sigma}_z$ are the Pauli matrices. After a N-step walk, the size of the high-dimensional lattice state $|l\rangle$ is $N + 1$. Considering the coin subspace is modeled as a two-level system, the total dimension of ρ is thus $4(N + 1)^2$. Here, we utilize the NDO ansatz $\rho_{\lambda\mu}$ in Eq. (2) as the variational representation of open QW state ρ.

Experimental implementation

QW can be realized on multiple physical platforms [43], and our experimental setup for photonic QW is shown in Fig. 2. More details are given in the Methods. In the experiment, we adopted the heralded single photons as the walker. The coin $\{|\uparrow\rangle, |\downarrow\rangle\}$ and the lattice $\{|l\rangle\}$ are encoded in the two polarizations $\{|H\rangle, |V\rangle\}$ and the different arrival times of photons $\{|t_l\rangle\}$, respectively. Polarization flip $\hat{R}(\alpha)$ is realized by tuning the angle of half-wave plates. A calcite crystal delays $|V\rangle$ by one time-bin length ($\Delta t \sim 5$ ps) relative to $|H\rangle$, which realizes the shift operator $\hat{S}$ on the time-bin modes. Considering each state $|t_l\rangle$ in the time basis is represented by a photonic wave packet with a width of ~140 fs that is far less than the time interval Δt, thus the overlap between different time bins is negligible. Through cascading N times the two operators of $\hat{R}$ and $\hat{S}$, one can implement an N-step walk dynamics. The outgoing $N + 1$ time bins are then injected into a Michelson interferometer for measurement base transformation, followed by an upconverted single-photon detector.

The Michelson interferometer in Fig. 2(a3) is a core ingredient for full state tomography and eases the usual demand of constructing a *duplicate* QW system (see lower panel in Fig. 2(b)) to obtain the phase relation between lattice sites [20]. A moveable mirror in the interferometer introduces a controllable time delay between the two arms, which makes the photons with horizontal polarization $|H\rangle$ travel the integer multiples of 5 ps faster (or later) than those with vertical polarization $|V\rangle$. After the interferometer, a polarization analyzer with a QWP-HWP-PBS setting performs the single-qubit Pauli measurements on $\hat{\sigma}_x$-, $\hat{\sigma}_y$- and $\hat{\sigma}_z$-basis. Therefore, the interferometric measurements between all time bins here

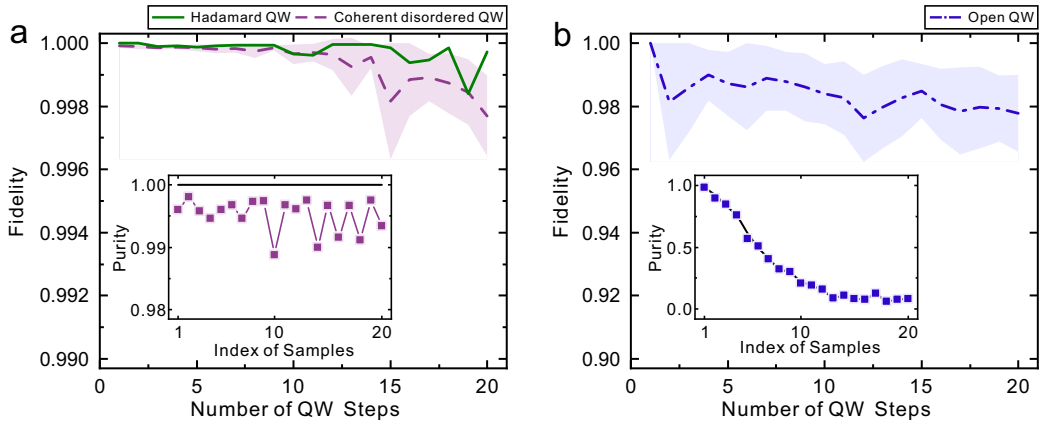


FIG. 3. **Benchmarking NDO tomography using partial measurements.** NDO reconstruction fidelity $F(\rho,\rho_{\lambda\mu}) = \mathrm{Tr}(\sqrt{\sqrt{\rho}\rho_{\lambda\mu}\sqrt{\rho}})^2$ as a function of the number of steps for (a) Hadamard QW with $\alpha \equiv \pi/4$ (green solid line), coherent disordered QW with α_t being time-dependent (red dashed line), and (b) open QW with arbitrary mixing (blue dash-dotted line). The shaded regions for coherent disordered and open QW are the standard errors of NDO reconstruction with 20 random samples for each step, and the lines are the averaging results. The insets in (a) and (b) show the purity of reconstructed states $\rho_{\lambda\mu}$ of the 20 samples for a 20-step coherent disordered (red squares) and open QW (blue squares), respectively. The black solid lines in the insets are the theoretical values of the purity for target density matrix ρ.

can be viewed as the system's projections on a series of basis $\{|\uparrow\rangle|l\rangle\}$, $\{|\downarrow\rangle|l\rangle\}$, $\{\frac{1}{\sqrt{2}}(|\uparrow\rangle|l\rangle \pm i|\downarrow\rangle|l'\rangle)\}$, and $\{\frac{1}{\sqrt{2}}(|\uparrow\rangle|l\rangle \pm |\downarrow\rangle|l'\rangle)\}$ $(l,l' = 0,1,\cdots,N)$. Moreover, even with the number of steps N increasing, the required amounts of the optical elements for these interferometric measurements remain constant thanks to the compact structure. Consequently, the compact interferometer can save 50% of optical resources as there is no need to build another duplicate QW to implement these measurements.

One can classify these projection measurements into a set of bases $\{v^n\}$. The first base $\{v^0\}$ is the reference base $\{v\} = \{|\uparrow\rangle|l\rangle, |\downarrow\rangle|l\rangle\}_{l=0}^{N}$, while the bases $\{v^{2k-1}\}$ and $\{v^{2k}\}$ $(k \in \{1,2,\cdots,N+1\})$ can be written as $\{\frac{1}{\sqrt{2}}(|\uparrow\rangle|l\rangle \pm i|\downarrow\rangle|[l-(k-1)]\mathrm{mod}(N+1)\rangle)\}_{l=0}^{N}$ and $\{\frac{1}{\sqrt{2}}(|\uparrow\rangle|l\rangle \pm |\downarrow\rangle|[l-(k-1)]\mathrm{mod}(N+1)\rangle)\}_{l=0}^{N}$, respectively. Throughout the experiment, the probability distribution $P(v^n)$ in each base is analyzed by utilizing the up-converted detector in Fig. 2(a4). With a spatial delay line, one can scan a strong pump light of 300 mW to upconvert the signal photons in all time bins with an interval of 5 ps. The up-converted photons are filtered by a spectrum filter that consists of a dispersion prism in the 4-f system to reduce extra scattering noise. And the filtered photon counts are then measured by a PMT, with which we can obtain experimental distributions $P_{\mathrm{exp}}(v^n)$. Note that, using the linear reconstruction directly [38, 44], one can only access $2(N+1)^2+2(N+1)$ matrix elements according to the measurements on these $2(N+1)+1$ sets of bases. Herein, we resort to the effective NDO method to learn the full $4(N+1)^2$ matrix elements from the incomplete measurements, which is an extension of learning some particular properties of quantum states using partial measurements [45, 46].

Performance of the efficient tomography for QW

Before training the neural network with noisy experimental results, we start by benchmarking the NOD tomography of QW using synthetic datasets. We generate the synthetic distribution $P(v^n)$ imposed by the target QW state ρ in a collection of bases $\{v^n\}$. The training process minimizes the cost function, defined as the total statistical distance $\mathcal{D}_\theta = \sum_n \sum_{v^n} P(v^n)\log[P(v^n)/P_\theta(v^n)]$, between the target and the reconstructed distributions. The process starts with the network parameters being initialized to random values, and in each optimization iteration i, the parameters are updated according to the GNGD procedure [35]: $\theta_{i+1} = \theta_i - \eta_i G^{-1}\nabla\mathcal{D}_\theta$, where the searching step η_i is determined by the line search process. To determine a well-performed metric G that can speed up the gradient-based optimization, the core idea of GNGD is to introduce a proper reference space, which is regarded as a flat space here. By choosing the identity matrix as the metric of reference space $G_{ij}^{\mathrm{ref}} = \delta_{i,j}$, the metric for the cost function G can be obtained through the conversion of coordinates (See Methods for details). All the NDO reconstructions given below were obtained in this way.

All the initial state is set to be $\rho_0 = |\psi_0\rangle\langle\psi_0|$ with $|\psi_0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\rangle + i|\downarrow\rangle)\otimes|0\rangle$. We first focus on the simplest model where the flip angle α is time-independent and the mixing parameter $w = 0$. Consider that the coin operator is a Hadamard gate ($\alpha = \pi/4$), and the corresponding walk is the typical Hadamard QW. The fidelity of the NDO tomography performed on the synthetic datasets is reported by the green solid line in Fig. 3(a), and a uni-

formly high fidelity is achieved even with an increasing number of time steps. In a bit complicated case, the angle $\alpha \in [0, \pi)$ is set to be time-dependent and becomes completely random for each step. For a given step of this coherent disordered QW [47, 48], twenty disordered configurations are generated, corresponding to twenty different QW dynamics of the quantum states. The red dashed line in Fig.3(a) shows the mean performance of the reconstruction fidelity obtained by averaging over the results of the twenty samples. A reconstruction fidelity error below 4×10^{-3} was found for the coherent walk with 400 samples in total. We then turn to a more challenging scenario of open QW interacting with a simulated environment [17, 49]. Decoherence can be introduced by inserting an extra phase gate $\hat{R}(\beta) = e^{\frac{i}{2}\beta\hat{\sigma}_z}$ with a fast fluctuating phase $\beta \in [-\delta\beta, \delta\beta]$ into each step of coherent QW. The fluctuating degree $\delta\beta \in [0, \pi]$ determines the system's mixing parameter $w \in [0, 1]$. We also generated 20 samples of the open QW with different mixing for each time step, and the NDO reconstruction fidelity shown in Fig.3(b) can still reach a high value of $\sim 98\%$ with the partial measurements. Beyond fidelity, an important question is whether the mixing of QW is reproduced accurately in the NDO reconstruction. The insets in Fig.3 display the reconstructed purity $\mathrm{Tr}[\rho^2_{\lambda\mu}]$ for the coherent and open QW, finding a close agreement with the expected results of the target state ρ. And the mean value of the reconstruction errors for purity is 5×10^{-3} and 3×10^{-3}, respectively.

Experimental mixed-state reconstructions

After benchmarking it on synthetic data, we now demonstrate the efficient learning of open QW with noisy experimental results. Using Polarizer1 in Fig.2(a2), the walker's initial state is fixed to be a product state $\frac{1}{\sqrt{2}}(|H\rangle + i\,|V\rangle) \otimes |t_{l=0}\rangle$. The first experiment we performed was a five-step Hadamard walk interacting with a controllable simulated environment. Besides a Hadamard gate realized by an HWP with its optical axis oriented at 22.5°, the extra phase gate $\hat{R}(\beta) = |H\rangle\langle H| + e^{i\beta}\,|V\rangle\langle V|$ is introduced by a configuration of QWP-HWP-QWP for each time step. The configuration features two QWPs rotated to an angle of 45° and a sandwiched HWP whose rotation angle controls the relative phase β between horizontal and vertical polarization [50]. To mimic the decoherence effect, the five sandwiched HWPs are installed on the motorized stages to constantly change their rotation angles with a controlled fluctuating degree $\delta\beta = 0, \pi/8, \pi/4, \pi/2, 3\pi/4$ and π. Then an ensemble measurement of $P_{\mathrm{exp}}(v^n)$ is performed on 13 sets of bases by unitizing the Michelson interferometer (see Fig.2(a3)) and the up-converted detector (see Fig.2(a4)). As reported in Fig.4(a), NDO can learn to reproduce the purity of open QW's mixed states present in the experimental data well. And the fidelity of the reconstructed full

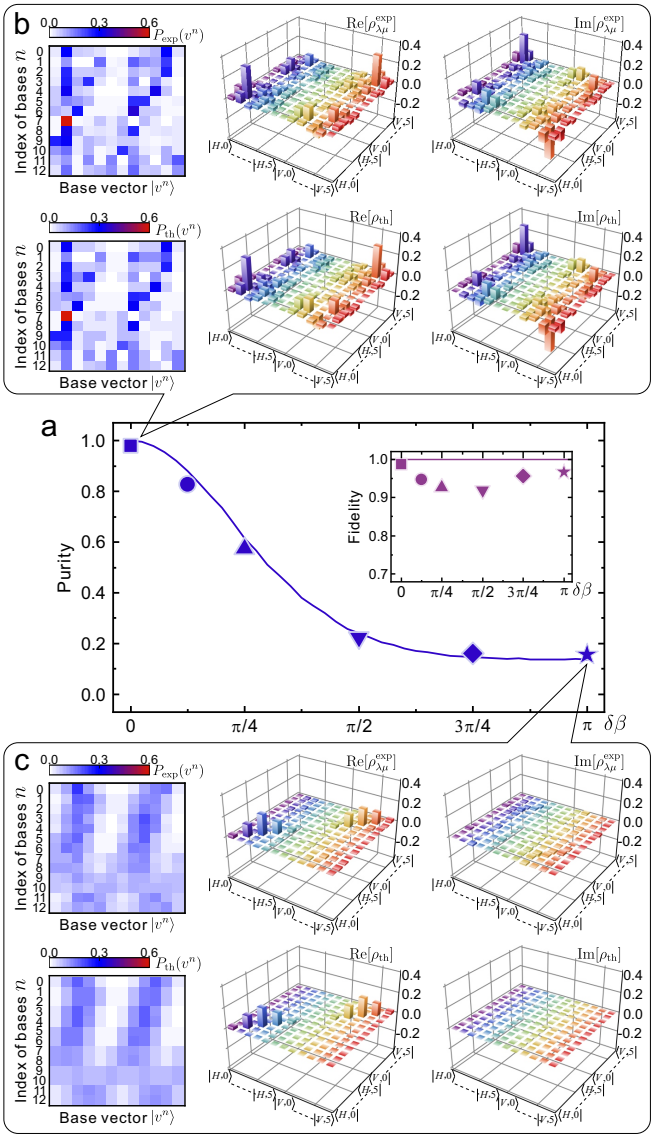


FIG. 4. **Experimental NDO tomography of open QW in a simulated environment.** (a) Purity of the NDO reconstructed states $\rho^{\mathrm{exp}}_{\lambda\mu}$ (blue symbols), trained on the experimental measurement data $P_{\mathrm{exp}}(v^n)$ for a five-step QW with six distinct mixings. The red symbols in the inset report the reconstruction fidelity. The blue and red solid lines give the theoretical results of target states ρ_{th}. Error bars considering the statistical noise are smaller than the symbol size. (b) and (c) show the measured probability distributions $P_{\mathrm{exp}}(v^n)$ on 13 sets of bases, real and imaginary parts of $\rho^{\mathrm{exp}}_{\lambda\mu}$ (from left to right) for the Hadamard QW with $\delta\beta = 0$ and the fully decoherent QW with $\delta\beta = \pi$, respectively. The lower panels in (b) and (c) represent the corresponding theoretical expectations.

density matrix is greater than 0.9, as the inset shows. These results allow us to investigate the QW-to-CRW transition in greater depth than by only considering their transport behaviors [51] exhibited in the diagonal terms of the density matrix. For the Hadamard walk ($w = 0$) shown in Fig.4(b), beyond the typical characterization

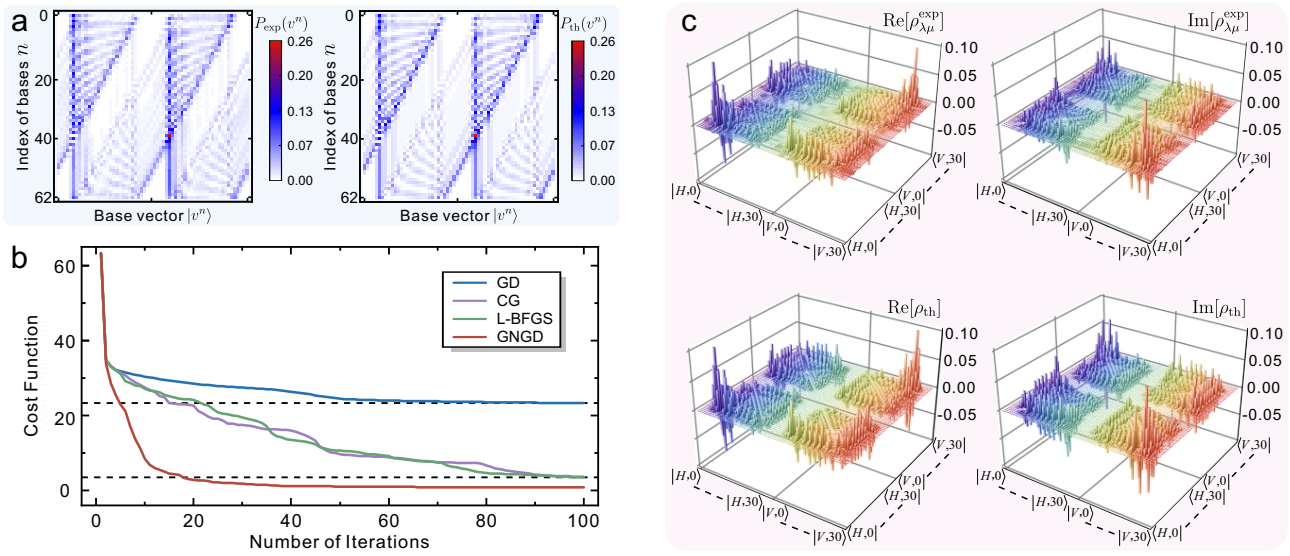


FIG. 5. **Experimental GNGD-enhanced NDO tomography of QW in a real environment.** (a) Measured probability distributions $P_{\rm exp}(v^n)$ on 63 sets of bases for a 30-step Hadamard walk. The right panel is the theoretical expectation. (b) Cost function $\mathcal{D}_\theta$ versus the number of training iterations for NDO tomography using GD, CG, L-BFGS, and GNGD optimizer. (c) Real and imaginary parts of the NDO reconstructed state $\rho^{\rm exp}_{\lambda\mu}$, and the theoretical expectations are shown in the lower panels.

with pronounced side peaks, the underlying coherence is clearly presented in the complex-valued off-diagonal matrix elements of $\rho^{\rm exp}_{\lambda\mu}$. The decoherence introduced by the simulated environment can completely destroy the phase relation between lattice sites such that CRW ($w = 1$) dominates. Consequently, it can bring the off-diagonal matrix elements to zero and cause the diagonal elements to have a classical binomial distribution, as shown in Fig. 4(c).

As the total QW steps N increase, the learning of mixed states with the dimension of $4(N+1)^2$ demands the complicated training of a larger neural network to keep its high expressive power. Thus, we introduce the GNGD procedure [35] to enhance the training efficiency of NDO. To better demonstrate the performance of GNGD for addressing complex networks of NDO, we extend the evolutionary time $N = 30$, and a thirty-step Hadamard walk is experimentally realized. Interferometric measurements on $\mathcal{N}_b = 63$ sets of bases are performed, and the obtained distributions $P_{\rm exp}(v^n)$ are displayed in Fig. 5(a). The mean value of classical indicator similarity of the measured distributions, defined as $\mathcal{S} = \sum_n \sum_{v^n} \sqrt{P_{\rm exp}(v^n) P_{\rm th}(v^n)}/\mathcal{N}_b$, reads 0.965 ± 0.008, giving good agreement with the theoretical predictions. In Fig. 5(b), we compare the cost function $\mathcal{D}_\theta$ as a function of the number of iterations using gradient descent (GD), conjugate gradient (CG) [52], L-BFGS [53], and GNGD methods for training NDO, respectively. It can be found that the GNGD optimizer enables a lower value of $\mathcal{D}_\theta$ in one order of magnitude fewer iterations than the traditional GD-based optimization, and achieves a 5-fold speedup in the training process over CG and L-BFGS. As shown in Fig. 5(c), the GNGD-enhanced NDO can efficiently output the experimental mixed state $\rho^{exp}_{\lambda\mu}$ with 62^2 complex-valued matrix elements, using fewer training iterations and partial measurements. The emblematic side peaks and the off-diagonal elements that represent coherence are still clear after the 30-step Hadamard walk. The reconstructed fidelity and purity are 0.789 ± 0.011 and 0.675 ± 0.013, respectively. A test on synthetic data yields the values of fidelity and purity both greater than 0.995, indicating that the uncontrolled decoherence introduced by the real environment is the major deviation for the experimental reconstruction.

Discussion

In this work, we have experimentally demonstrated an effective learning of mixed states for open photonic QW in terms of the number of training iterations, the amount of measurement resources, and the reconstruction fidelity. The learning method is achieved by training a neural-network-parameterized density operator on measurement datasets and then outputting complex-valued matrix elements accelerated by the GNGD procedure. The GNGD optimizer can also be directly applied to various neural-network architectures for quantum state tomography with gradient-based optimization [54–57]. To capture the correlations within the physical system and the environment more accurately, we usually increase the number of hidden and ancillary neurons, respectively. One can introduce the deep restricted Boltzmann machine with improved representational power [58], at the cost of growing training complexity, to further develop the NDO method to learn the open quantum systems.

Overall, our approach provides a promising avenue for addressing the challenges associated with verifying open QW. We expect that these results will lead to new insights and discoveries in this exciting area of research on numerous physical systems such as fiber loop [59], spatial path [60], orbital angular momentum [61], transverse momentum [62, 63], hybrid architecture [64], etc. Moreover, the full quantum state tomography technique can be combined with the known abilities of arbitrary initialization [13] and flexible manipulation, which can inspire a prospective QW platform for developing novel applications in a range of fields.

Methods

Training the neural network by GNGD

The training of the neural network is to learn the optimal parameters $\theta = \{\lambda, \mu\}$ that minimize the cost function D_θ. The derivative of the cost function with respect to the network parameters reads:

$$\begin{aligned}\nabla\mathcal{D}_\theta = &-\sum\nolimits_n \sum\nolimits_{v^n} P(v^n)/\rho_\theta^n(v^n, v^n) \\ &\sum\nolimits_{\alpha,\beta} \mathcal{U}^n(v^n, \alpha)\rho_\theta(\alpha, \beta)\nabla A_\theta(\alpha, \beta)\mathcal{U}^{n,\dagger}(\beta, v^n) \\ &+\mathcal{N}_b \sum\nolimits_v \rho_\theta(v, v)\nabla A_\theta(v, v),\end{aligned} \tag{4}$$

where $\rho_\theta^n(v, v') = \sum_{\alpha,\beta}\mathcal{U}^n(v, \alpha)\rho_\theta(\alpha, \beta)\mathcal{U}^{n,\dagger}(\beta, v')$, $\mathcal{U}^n = \langle v^n|v\rangle$ is the base transformation matrix from the reference base $|v\rangle$ to $|v^n\rangle$, $\mathcal{N}_b$ is the total number of measurement bases. $\rho_{\lambda\mu}(\alpha, \beta) = Z_\lambda^{-1}e^{A_{\lambda\mu}}$ is the rebuild density matrix of the system, where

$$\begin{aligned}A_{\lambda\mu}(v, v') =&\Gamma_\lambda^+(v, v') + i\Gamma_\mu^-(v, v') + \Pi_{\lambda\mu}(v, v'), \\ \Gamma_\theta^\pm(v, v') =&\frac{1}{2}[\sum\nolimits_i \log(1 + e^{W_\theta^{[i]}v + c_\theta^{[i]}}) \\ &\pm \log(1 + e^{W_\theta^{[i]}v' + c_\theta^{[i]}}) + b_\theta^T(v \pm v')], \\ \Pi_{\lambda\mu}(v, v') =&\sum\nolimits_i \log\{1 + \exp[\frac{1}{2}U_\lambda^{[i]}(v + v') \\ &+ \frac{i}{2}U_\mu^{[i]}(v - v') + d_\lambda^{[i]}]\}.\end{aligned} \tag{5}$$

The superscript $[i]$ means the ith row elements. The network parameters to be optimal here are $\{W_\theta, U_\theta, b_\theta, c_\theta, d_\theta\}$. W_θ and U_θ are the weight matrices, whose dimension is $m \times d$ with d being the degree of freedom of the physical system and m being the number of hidden or ancillary neurons. b_θ, c_θ and d_θ are the vectors of dimension d representing the biases coupled to the visible, hidden, and ancillary neurons, respectively.

To accelerate the convergence efficiency of training, the GNGD method emphasizes introducing a proper reference space [35]. Here, the space span by the density matrix $\rho_\theta(\alpha, \beta)$ is selected as the reference manifold. We choose a simple identity matrix $G^{\text{ref}}_{\alpha,\beta;\alpha',\beta'} = \delta_{\alpha,\alpha'}\delta_{\beta,\beta'}$ as the metric of reference space. Through the conversion of coordinates, the metric G for the cost function can be determined as:

$$G_{i,j} = \sum\nolimits_{\alpha,\beta;\alpha',\beta'} \frac{\partial\rho_\theta(\alpha, \beta)}{\partial\theta_i} G^{\text{ref}}_{\alpha,\beta;\alpha',\beta'} \frac{\partial\rho_\theta(\alpha', \beta')}{\partial\theta_j} \tag{6}$$

With this metric, the network parameters are updated according to the gradient-based optimization as:

$$\theta_{i+1} = \theta_i - \eta_i G^{-1}\nabla\mathcal{D}_\theta \tag{7}$$

where η_i, which is determined by the line search process, is the searching step at iteration i. We directly sum over all the terms in $\mathcal{D}_\theta$ and $G_{i,j}$ during training for being free of the sampling error of Monte Carlo sampling based on the probability $P(v^n)$.

Heralded single-photon source

A Ti: sapphire laser source (Mira 900, Coherent) launches a series of optical pulses with a temporal pulse width of 140 fs, a repetition rate of 76 MHz, the wavelength λ=800 nm, and an average power of 1 W. The ultrafast optical pulses focused by Lens1 (focal length f=100 mm) pump the first $\beta - BaB_2O_4$ (BBO1), causing the second harmonic generation. The frequency-doubled pulses with horizontal polarization and the residual fundamental pulses with vertical polarization are collimated by Lens2 and then spatially separated by the first dichroic mirror (DM1). The fundamental pulses go through a pair of cylinder lenses (not shown) to reform the beam's profile into a Gaussian shape and are then reused for position-resolved state detection. The frequency-doubled pulses with λ=400 nm and an average power of 200 mW are used as a pump source and focused on the BBO2 by lens3 (focal length f=100 mm). Herein, BBO2 is designed for type-II, non-degenerate, and "beam-like" spontaneous parametric down-conversion (SPDC). The generated time-correlated photon pairs (namely, signal photons with λ=780 nm and horizontal polarization, and idler photons with λ=821 nm and vertical polarization) are spatially separated by the non-collinear SPDC process with a half opening angle of 3°. The idler photons, collimated by the Lens5 and cleaned by a spectrum filter, are coupled to a single-mode fiber and then detected by the avalanche photodiode detector (SPCM-AQRH-14-FC, Excelitas) to herald the appearance of the signal photons. The coincidence count rate between idler and signal photons is 4.5×10^2 pairs/(s mW). The collimated and filtered signal photons are adopted as the walker and sent to the quantum walk (QW) module.

Time-multiplexing photonic quantum walk

The time-multiplexing QW is achieved by encoding the walker's lattice and coin space in the arrival time and polarization of signal photons, respectively. The signal photons at the input of the QW module are set to be in the first time bin, i.e., the origin in lattice space with

$l = 0$. Then, the initial polarization state of signal photons $|\uparrow_y\rangle = \frac{1}{\sqrt{2}}(|H\rangle + i\,|V\rangle)$ is prepared through a polarization-dependent beam splitter (PBS), a half-wave plate (HWP), followed by a quarter-wave plate (QWP). $|H\rangle$ and $|V\rangle$ represent horizontal and vertical polarization, respectively. Thus, the initial state of the walker is $|\psi_0\rangle = |l = 0\rangle \otimes |\uparrow_y\rangle$. In the protocol of Hadamard QW, the unitary operator of a single step $\hat{U} = \hat{S}\hat{R}$ is composed of an HWP rotated to a specific angle at 22.5° (for operator $\hat{R}$) and a calcite crystal (for operator $\hat{S}$). Each piece of calcite is cut parallel to its optical axes to a length of 8.98 mm, whose birefringence $\Delta = 0.167$ at λ=800 nm will induce a 5 ps time shift between the horizontal $|H\rangle$ and vertical polarization $|V\rangle$ of the signal photons. In the experiment, N sets of HWPs and calcite crystals are positioned to realize an N-step Hadamard walk. Thus, the lattice space of the walker at the end of the QW module consists of the superposition of $N + 1$ time bins with a time interval of 5 ps. And for each lattice site, the walker has an internal coin state, which is usually distinct at a different site. A QWP-HWP-QWP setting for realizing the phase gate operator $\hat{R}(\beta) = e^{\frac{i}{2}\beta\hat{\sigma}_z}$ can be introduced at each QW step to mimic a controlled decoherence effect. Two QWPs are rotated to an angle at 45°. HWP in the setting is installed on the motorized stage (PR50PP, Newport) to change its rotation angle constantly. The fluctuating range $[-\delta\beta, \delta\beta]$ determines the degree of decoherence and has a similar effect as the the mixing parameter w: $\delta\beta = 0$ is the coherent QW with $w = 0$, and $\delta\beta = \pi$ results in a fully decoherent classical walk with $w = 1$.

Base transformation and phase-locking technique

The base transformation is achieved by a Michelson interferometer consisting of a PBS, two QWPs rotated to an angle at 45°, and two mirrors. One mirror is piezoelectric ceramic (PZT)-driven to compensate for the phase fluctuation of the interferometer, and the other mirror is installed on a motorized positioning system (KXL06100, Suruga). The moveable mirror is controlled to induce a time shift between the two arms of the interferometer, which is accurately the integer multiples of 5 ps for ensuring maximal visibility between different time bins of the signal photons and avoiding temporal mode mismatching. The phase difference between the two arms can be controlled by the phase shifter (PS) and locked in a fixed value by an ancillary Helium-Neon (He-Ne) laser using a proportional-integral-differential (PID) feedback unit [65]. The PS is the configuration of QWP-HWP-QWP, where two QWPs are rotated to an angle at 45° and the rotation angle of HWP controls the relative phase between $|H\rangle$ and $|V\rangle$ states. Initially, the He-Ne laser source (HNL050LB, Thorlabs) with λ=632.8 nm is prepared in the state $|\downarrow_x\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle)$ through the Polarizer2 and then launched to the same propagation path with the signal photons through DM2. The signal photons and the He-Ne laser pass through the interferometer and are spatially separated by DM3. The reflected He-Ne laser is projected in the $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$ state and detected by a amplified Si photodetector (PDA8A/M, Thorlabs). The detected He-Ne laser intensity is initially set near a reference value through PS. Within the acquisition time of time-correlated photon pairs, the PID unit compares the detected intensity using the amplified Si photodetector to the reference value and then uses the difference to calculate a new input value through the PID algorithm that is designed to keep the detected intensity at the reference value. Thus, the unstable phase difference of the interferometer for signal photons can then be locked with the He-Ne laser. PS can be used to compensate the phase difference of the signal photons introduced by the two arms of the interferometer to zero. After passing through the Michelson interferometer and the polarization analyzer, the signal photons are collected into a 0.1 m-long single-mode fiber (SMF) for spatial filtering. The two HWPs at the input and output of the SMF are used for polarization maintenance.

Frequency up-conversion single-photon detector

To measure the probability distribution $P(v^n)$ in each base v^n, a position-resolved detection of the signal photons is constructed. The lattice space of the signal photons consists of a pulse train with a time interval of 5 ps. Herein, we utilize a single-photon frequency upconversion system converting the high time resolution to the spatial resolution. Specifically, the residual fundamental pulses with λ=800 nm and average power 300 mW during the second harmonic generation serve as the strong pump pulse. The up-converted photons are generated in BBO3, where the signal photons and the strong pump pulse meet, and then detected by a photomultiplier tube (H10682, Hamamatsu). In the up-conversion process, the propagation path of the strong pump pulse is delayed and aligned using two moveable mirrors installed on a motorized positioning system (KXL06100, Suruga). And the alignment of the signal photons is achieved by replacing them with a reference coherent laser to optimize the spatial mode matching. In addition, the delay of the strong pump pulse is fine-tuned to match each pulse in the signal-photon pulse train in the time domain. A spectrum filter composed of a dispersion prism is used to reduce the background.

Acknowledgements

We thank Zhih-Ahn Jia, Hai-Lan Ma and Zi-Wei Cui for useful discussions. This work was supported by Innovation Program for Quantum Science and Technology (No. 2021ZD0301200), National Natural Science Foundation of China (Nos. 12022401, 62075207, 11874343, 12104433, 11821404), the Fundamental Research Funds for the Central Universities

(No. WK2470000030), the CAS Youth Innovation Promotion Association (No. 2020447). Q.-Q. Wang acknowledges support from China Postdoctoral Science Foundation (No. 2021M703108) and National Natural Science Foundation of China (No. 12204468).

Author contributions

Q.-Q. W. and S. D. contributed equally to this work. Q.-Q. W. performed the experiments under the guidance of X.-Y. X. and C.-F. L. S. D. performed the theoretical works under the guidance of Y.-J. H. and C. Wang. X.-W. L., S. H. and M.-H. Y. provided theoretical supports. Q.-Q. W. and S. D. drafted the manuscript, with revisions from M.-H. Y., X.-Y. X., Y.-J. H. and C.-F. Li. C.-F. L. and G.-C. G. supervised the project.

Data Availability.

The data for all the pictures in this paper are available from the corresponding authors upon justified request.

Competing Interests

The authors declare no competing interests.

* xuxiaoye@ustc.edu.cn
† yung@sustech.edu.cn
‡ smhan@ustc.edu.cn
§ cfli@ustc.edu.cn

[1] R. Portugal, *Quantum Walks and Search Algorithms* (Springer, New York, NY, 2013).
[2] F. Flamini, N. Spagnolo, and F. Sciarrino, *Photonic quantum information processing: a review,* Rep. Prog. Phys. **82**, 016001 (2018).
[3] A. W. Young, W. J. Eckner, N. Schine, A. M. Childs, and A. M. Kaufman, *Tweezer-programmable 2d quantum walks in a hubbard-regime lattice,* Science **377**, 885–889 (2022).
[4] F. Caruso, *Universally optimal noisy quantum walks on complex networks,* New J. Phys. **16**, 055015 (2014).
[5] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. Wong, T. Baehr-Jones, M. Hochberg, *et al.*, *Quantum transport simulations in a programmable nanophotonic processor,* Nat. Photonics **11**, 447–452 (2017).
[6] M. Bruderer and M. B. Plenio, *Decoherence-enhanced performance of quantum walks applied to graph isomorphism testing,* Phys. Rev. A **94**, 062317 (2016).
[7] F. Caruso, A. Crespi, A. Ciriolo, F. Sciarrino, and R. Osellame, *Fast escape of a quantum walker from an integrated photonic maze,* Nat. Commun. **7**, 11682 (2016).
[8] N. Dalla Pozza, L. Buffoni, S. Martina, and F. Caruso, *Quantum reinforcement learning: the maze problem,* Quantum Mach. Intell. **4**, 11 (2022).
[9] H. Tang, R. Shi, T.-S. He, Y.-Y. Zhu, T.-Y. Wang, M. Lee, and X.-M. Jin, *Tensorflow solver for quantum pagerank in large-scale networks,* Sci. Bull. **66**, 120–126 (2021).
[10] M. Schuld, I. Sinayskiy, and F. Petruccione, *Quantum walks on graphs representing the firing patterns of a quantum neural network,* Phys. Rev. A **89**, 032333 (2014).
[11] H. Tang, Z. Feng, Y.-H. Wang, P.-C. Lai, C.-Y. Wang, Z.-Y. Ye, C.-K. Wang, Z.-Y. Shi, T.-Y. Wang, Y. Chen, J. Gao, and X.-M. Jin, *Experimental quantum stochastic walks simulating associative memory of hopfield neural networks,* Phys. Rev. Appl. **11**, 024020 (2019).
[12] L.-J. Wang, J.-Y. Lin, and S. Wu, *Implementation of quantum stochastic walks for function approximation, two-dimensional data classification, and sequence classification,* Phys. Rev. Res. **4**, 023058 (2022).
[13] T. Giordani, E. Polino, S. Emiliani, A. Suprano, L. Innocenti, H. Majury, L. Marrucci, M. Paternostro, A. Ferraro, N. Spagnolo, and F. Sciarrino, *Experimental engineering of arbitrary qudit states with discrete-time quantum walks,* Phys. Rev. Lett. **122**, 020503 (2019).
[14] L. Innocenti, H. Majury, T. Giordani, N. Spagnolo, F. Sciarrino, M. Paternostro, and A. Ferraro, *Quantum state engineering using one-dimensional discrete-time quantum walks,* Phys. Rev. A **96**, 062326 (2017).
[15] T. Giordani, L. Innocenti, A. Suprano, E. Polino, M. Paternostro, N. Spagnolo, F. Sciarrino, and A. Ferraro, *Entanglement transfer, accumulation and retrieval via quantum-walk-based qubit–qudit dynamics,* New J. Phys. **23**, 023012 (2021).
[16] L. Banchi, D. Burgarth, and M. J. Kastoryano, *Driven quantum dynamics: Will it blend?* Phys. Rev. X **7**, 041015 (2017).
[17] H. Tang, L. Banchi, T.-Y. Wang, X.-W. Shang, X. Tan, W.-H. Zhou, Z. Feng, A. Pal, H. Li, C.-Q. Hu, M. S. Kim, and X.-M. Jin, *Generating haar-uniform randomness using stochastic quantum walks on a photonic chip,* Phys. Rev. Lett. **128**, 050503 (2022).
[18] G. M. D'Ariano, M. G. Paris, and M. F. Sacchi, *Quantum tomography,* Adv. Imaging Electron Phys. **128**, 206–309 (2003).
[19] Y. Aharonov, L. Davidovich, and N. Zagury, *Quantum random walks,* Phys. Rev. A **48**, 1687–1690 (1993).
[20] M. Karski, L. Förster, J.-M. Choi, A. Steffen, W. Alt, D. Meschede, and A. Widera, *Quantum walk in position space with single optically trapped atoms,* Science **325**, 174–177 (2009).
[21] S. Barkhofen, L. Lorz, T. Nitsche, C. Silberhorn, and H. Schomerus, *Supersymmetric polarization anomaly in photonic discrete-time quantum walks,* Phys. Rev. Lett. **121**, 260501 (2018).
[22] Q.-Q. Wang, S.-J. Tao, W.-W. Pan, Z. Chen, G. Chen, K. Sun, J.-S. Xu, X.-Y. Xu, Y.-J. Han, C.-F. Li, *et al.*, *Experimental verification of generalized eigenstate thermalization hypothesis in an integrable system,* Light: Sci. Appl. **11**, 194 (2022).
[23] S. Barkhofen, T. Nitsche, F. Elster, L. Lorz, A. Gábris, I. Jex, and C. Silberhorn, *Measuring topological invariants in disordered discrete-time quantum walks,* Phys. Rev. A **96**, 033846 (2017).
[24] Q. Lin, T. Li, L. Xiao, K. Wang, W. Yi, and P. Xue, *Observation of non-hermitian topological anderson insulator in quantum dynamics,* Nat. Commun. **13**, 3229 (2022).
[25] X.-Y. Xu, Q.-Q. Wang, W.-W. Pan, K. Sun, J.-S. Xu, G. Chen, J.-S. Tang, M. Gong, Y.-J. Han, C.-F. Li, and G.-C. Guo, *Measuring the winding number in a large-scale chiral quantum walk,* Phys. Rev. Lett. **120**, 260501 (2018).
[26] X.-Y. Xu, Q.-Q. Wang, M. Heyl, J. C. Budich, W.-W.

Pan, Z. Chen, M. Jan, K. Sun, J.-S. Xu, Y.-J. Han, C.-F. Li, and G.-C. Guo, *Measuring a dynamical topological order parameter in quantum walks,* Light: Sci. Appl. **9**, 7 (2020).
[27] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Machine learning and the physical sciences,* Rev. Mod. Phys. **91**, 045002 (2019).
[28] Z.-A. Jia, B. Yi, R. Zhai, Y.-C. Wu, G.-C. Guo, and G.-P. Guo, *Quantum neural network states: A brief review of methods and applications,* Adv. Quantum Technol. **2**, 1800077 (2019).
[29] V. Gebhart, R. Santagati, A. A. Gentile, E. M. Gauger, D. Craig, N. Ares, L. Banchi, F. Marquardt, L. Pezzè, and C. Bonato, *Learning quantum systems,* Nat. Rev. Phys. **5**, 141–156 (2023).
[30] T. Giordani, A. Suprano, E. Polino, F. Acanfora, L. Innocenti, A. Ferraro, M. Paternostro, N. Spagnolo, and F. Sciarrino, *Machine learning-based classification of vector vortex beams,* Phys. Rev. Lett. **124**, 160401 (2020).
[31] A. Suprano, D. Zia, E. Polino, T. Giordani, L. Innocenti, M. Paternostro, A. Ferraro, N. Spagnolo, and F. Sciarrino, *Enhanced detection techniques of orbital angular momentum states in the classical and quantum regimes,* New J. Phys. **23**, 073014 (2021).
[32] H. Ma, D. Dong, I. R. Petersen, C.-J. Huang, and G.-Y. Xiang, *A comparative study on how neural networks enhance quantum state tomography,* arXiv:2111.09504 [quant-ph] (2021).
[33] D. Zia, R. Checchinato, A. Suprano, T. Giordani, E. Polino, L. Innocenti, A. Ferraro, M. Paternostro, N. Spagnolo, and F. Sciarrino, *Regression of high-dimensional angular momentum states of light,* Phys. Rev. Res. **5**, 013142 (2023).
[34] G. Torlai and R. G. Melko, *Latent space purification via neural density operators,* Phys. Rev. Lett. **120**, 240503 (2018).
[35] S. Dong, F. Le, M. Zhang, S.-J. Tao, C. Wang, Y.-J. Han, and G.-P. Guo, *Generalization to the natural gradient descent,* arXiv:2210.02764 [math.OC] (2022).
[36] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, *Neural-network quantum state tomography,* Nat. Phys. **14**, 447–450 (2018).
[37] G. Torlai, B. Timar, E. P. L. van Nieuwenburg, H. Levine, A. Omran, A. Keesling, H. Bernien, M. Greiner, V. Vuletić, M. D. Lukin, R. G. Melko, and M. Endres, *Integrating neural networks with a quantum simulator for state reconstruction,* Phys. Rev. Lett. **123**, 230504 (2019).
[38] J. Carrasquilla, G. Torlai, R. G. Melko, and L. Aolita, *Reconstructing quantum states with generative models,* Nat. Mach. Intell. **1**, 155–161 (2019).
[39] M. J. Hartmann and G. Carleo, *Neural-network approach to dissipative quantum many-body dynamics,* Phys. Rev. Lett. **122**, 250502 (2019).
[40] A. Nagy and V. Savona, *Variational quantum monte carlo method with a neural-network ansatz for open quantum systems,* Phys. Rev. Lett. **122**, 250501 (2019).
[41] F. Vicentini, A. Biella, N. Regnault, and C. Ciuti, *Variational neural-network ansatz for steady states in open quantum systems,* Phys. Rev. Lett. **122**, 250503 (2019).
[42] M. A. Broome, A. Fedrizzi, B. P. Lanyon, I. Kassal, A. Aspuru-Guzik, and A. G. White, *Discrete single-photon quantum walks with tunable decoherence,* Phys. Rev. Lett. **104**, 153602 (2010).
[43] K. Manouchehri and J. Wang, *Physical implementation of quantum walks* (Springer, 2014).
[44] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Measurement of qubits,* Phys. Rev. A **64**, 052312 (2001).
[45] J. Gao, L.-F. Qiao, Z.-Q. Jiao, Y.-C. Ma, C.-Q. Hu, R.-J. Ren, A.-L. Yang, H. Tang, M.-H. Yung, and X.-M. Jin, *Experimental machine learning of quantum states,* Phys. Rev. Lett. **120**, 240501 (2018).
[46] M. Yang, C.-l. Ren, Y.-c. Ma, Y. Xiao, X.-J. Ye, L.-L. Song, J.-S. Xu, M.-H. Yung, C.-F. Li, and G.-C. Guo, *Experimental simultaneous learning of multiple nonclassical correlations,* Phys. Rev. Lett. **123**, 190401 (2019).
[47] A. Geraldi, A. Laneve, L. D. Bonavena, L. Sansoni, J. Ferraz, A. Fratalocchi, F. Sciarrino, A. Cuevas, and P. Mataloni, *Experimental investigation of superdiffusion via coherent disordered quantum walks,* Phys. Rev. Lett. **123**, 140501 (2019).
[48] Q.-Q. Wang, X.-Y. Xu, W.-W. Pan, K. Sun, J.-S. Xu, G. Chen, Y.-J. Han, C.-F. Li, and G.-C. Guo, *Dynamic-disorder-induced enhancement of entanglement in photonic quantum walks,* Optica **5**, 1136–1140 (2018).
[49] A. Schreiber, K. N. Cassemiro, V. Potoček, A. Gábris, I. Jex, and C. Silberhorn, *Decoherence and disorder in quantum walks: From ballistic spread to localization,* Phys. Rev. Lett. **106**, 180403 (2011).
[50] B.-G. Englert, C. Kurtsiefer, and H. Weinfurter, *Universal unitary gate for single-photon two-qubit states,* Phys. Rev. A **63**, 032303 (2001).
[51] T. A. Brun, H. A. Carteret, and A. Ambainis, *Quantum to classical transition for random walks,* Phys. Rev. Lett. **91**, 130602 (2003).
[52] R. Fletcher and C. M. Reeves, *Function minimization by conjugate gradients,* The Computer Journal **7**, 149–154 (1964).
[53] D. C. Liu and J. Nocedal, *On the limited memory bfgs method for large scale optimization,* Mathematical Programming **45**, 503–528 (1989).
[54] S. Ahmed, C. Sánchez Muñoz, F. Nori, and A. F. Kockum, *Quantum state tomography with conditional generative adversarial networks,* Phys. Rev. Lett. **127**, 140502 (2021).
[55] T. Schmale, M. Reh, and M. Gärttner, *Efficient quantum state tomography with convolutional neural networks,* npj Quantum Inf. **8**, 115 (2022).
[56] Y. Zuo, C. Cao, N. Cao, X. Lai, B. Zeng, and S. Du, *Optical neural network quantum state tomography,* Advanced Photonics **4**, 026004 (2022).
[57] M. Neugebauer, L. Fischer, A. Jäger, S. Czischek, S. Jochim, M. Weidemüller, and M. Gärttner, *Neural-network quantum state tomography in a two-qubit experiment,* Phys. Rev. A **102**, 042604 (2020).
[58] X. Gao and L.-M. Duan, *Efficient representation of quantum many-body states with deep neural networks,* Nat. Commun. **8**, 662 (2017).
[59] T. Nitsche, S. Barkhofen, R. Kruse, L. Sansoni, M. Stefanak, A. Gábris, V. Potoček, T. Kiss, I. Jex, and C. Silberhorn, *Probing measurement induced effects in quantum walks via recurrence,* Sci. Adv. **4**, eaar6444 (2018).
[60] T. Kitagawa, M. A. Broome, A. Fedrizzi, M. S. Rudner, E. Berg, I. Kassal, A. Aspuru-Guzik, E. Demler, and A. G. White, *Observation of topologically protected bound states in photonic quantum walks,* Nat. Commun. **3**, 882

(2012).
[61] F. Cardano, F. Massa, H. Qassim, E. Karimi, S. Slussarenko, D. Paparo, C. de Lisio, F. Sciarrino, E. Santamato, R. W. Boyd, and L. Marrucci, *Quantum walks and wavepacket dynamics on a lattice with twisted photons,* Sci. Adv. **1**, e1500087 (2015).
[62] C. Esposito, M. R. Barros, A. Durán Hernández, G. Carvacho, F. Di Colandrea, R. Barboza, F. Cardano, N. Spagnolo, L. Marrucci, and F. Sciarrino, *Quantum walks of two correlated photons in a 2d synthetic lattice,* npj Quantum Inf. **8**, 1–7 (2022).
[63] F. D. Colandrea, A. Babazadeh, A. Dauphin, P. Massignan, L. Marrucci, and F. Cardano, *Ultra-long quantum walks via spin–orbit photonics,* Optica **10**, 324–331 (2023).
[64] B. Wang, T. Chen, and X. Zhang, *Experimental observation of topologically protected bound states with vanishing chern numbers in a two-dimensional quantum walk,* Phys. Rev. Lett. **121**, 100501 (2018).
[65] C. Corsi, I. Liontos, S. Cavalieri, M. Bellini, G. Venturi, and R. Eramo, *An ultrastable michelson interferometer for high-resolution spectroscopy in the xuv,* Opt. Express **23**, 4106–4113 (2015).

Efficient information recovery from Pauli noise via classical shadow

Yifei Chen[1] Zhan Yu[1] Chenghong Zhu[1] Xin Wang[1] *

[1] *Institute for Quantum Computing, Baidu Research, Beijing 100193, China*

Abstract. Quantum systems are inherently susceptible to noises, which adversely corrupt the information encoded in quantum systems. In this work, we introduce an efficient algorithm that can recover information from quantum states under Pauli noise. The core idea is to learn the necessary information of the unknown Pauli channel by post-processing the classical shadows of the channel. For a local and bounded-degree observable, only partial knowledge of the channel is required rather than its complete classical description to recover the ideal information. This leads to an algorithm that runs in polynomial-time rather than exponential-time. Furthermore, the channel sample complexity scales logarithmically in the number of qubits. We also prove that the sample complexity of this algorithm is optimal and generalise the algorithm to broader class of channels. As a notable application, our method can be severed as a sample-efficient error mitigation scheme for Clifford circuits.

Note: A technical version of this work is available as arXiv:2305.04148.

Keywords: information recovery, quantum noise, Pauli noise, classical shadow, error mitigation

Introduction. One of the most important ingredients in quantum computing is to extract information from a quantum system by measuring the quantum state, which is described as the expectation value of some observable O of interest. The expectation value of some chosen observable unravels many properties of the quantum system, which is extensively used in many quantum algorithms, including variational quantum eigensolver [1], quantum approximate optimization algoirhtm [2], and quantum machine learning [3]. For an ideal quantum state σ, the information that we seek to obtain is $\mathrm{tr}(O\sigma)$. However, due to the noise present in the quantum computer, the actual state in practice is some *noisy state* $\tilde{\sigma}$ instead.

One of the most standard theoretical models for quantum noise in the study of quantum error correction and mitigation is *Pauli noise*. On one hand, Pauli noise provides a simple model that describes common incoherent noise such as bit-flip, depolarizing, and dephasing. On the other hand, general quantum noise can be mapped to Pauli noise without incurring a loss of fidelity by the technique of randomised compiling [4, 5]. In the theory of quantum information [6–8], noise of quantum systems are modelled by *quantum channels*, which are completely positive and trace-preserving (CPTP) maps between spaces of operators. An n-qubit *Pauli channel* is defined as

$$\mathcal{P}(\sigma) \coloneqq \sum_{P\in\{I,X,Y,Z\}^{\otimes n}} p(P)\cdot P\sigma P^{\dagger}, \tag{1}$$

where P is an n-fold tensor product of Pauli operators in $\{I,X,Y,Z\}$, and p is a probability distribution on $\{I,X,Y,Z\}^{\otimes n}$.

The problem of recovering from Pauli noise is that, given access to an unknown Pauli noise $\mathcal{P}$ and copies of the noisy state $\tilde{\sigma}=\mathcal{P}(\sigma)$, retrieve the information $\mathrm{tr}(O\sigma)$ for some observable O. To recover from a noise $\mathcal{P}$, a natural way is to construct a map $\mathcal{Q}$ such that the composed map $\mathcal{Q}\circ\mathcal{P}$ is an identity map [9], which could covert the noisy state $\mathcal{P}(\sigma)$ to the ideal state σ. Such a map is actually not necessary if we are only concerned with the target expectation value $\mathrm{tr}(O\sigma)$ instead of the ideal state σ. Zhao et al. [10] proved the necessary and sufficient condition for retrieving the target information from noisy quantum states, and utilised semidefinite programming to determine an optimal protocol for constructing the map $\mathcal{D}$ that satisfies $\mathrm{tr}(\mathcal{D}\circ\mathcal{P}(\sigma)O)=\mathrm{tr}(O\sigma)$. While this method is not restricted to the class of Pauli channels, it requires complete information of the quantum noise. Obtaining the full classical description of an unknown Pauli channel often uses techniques like quantum process tomography [11–13], typically requiring a number of copies of the channel that scales exponentially in the number of qubits, which is resource-consuming and inefficient. Furthermore, the map $\mathcal{D}$ proposed in Ref. [10] needs to be simulated via probabilistic sampling, which requires additional resources. Then, how to efficiently recover information from a Pauli channel with no prior information still remains an open and challenging problem.

Overview of results. Motivated by efficient methods such as shadow tomography [14] and classical shadow [15], we make progress towards solving this open problem of information recovery from Pauli noise by obtaining partial information of the unknown Pauli channel and using it to efficiently retrieve the expectation value of given observables. To be specific, we establish the following:

1. The main result is an efficient algorithm that retrieves the information $\mathrm{tr}(O\sigma)$ from unknown Pauli noise $\mathcal{P}$ for arbitrary n-qubit noisy state $\mathcal{P}(\sigma)$ and *bounded-degree k-local observables* O, using $\mathcal{O}(\log(n^k))$ applications of Pauli channel $\mathcal{P}$. We also prove that the channel sample complexity is optimal.

2. We then extend the algorithm to a broader class of quantum channels by establishing a sufficient cri-

*wangxin73@baidu.com

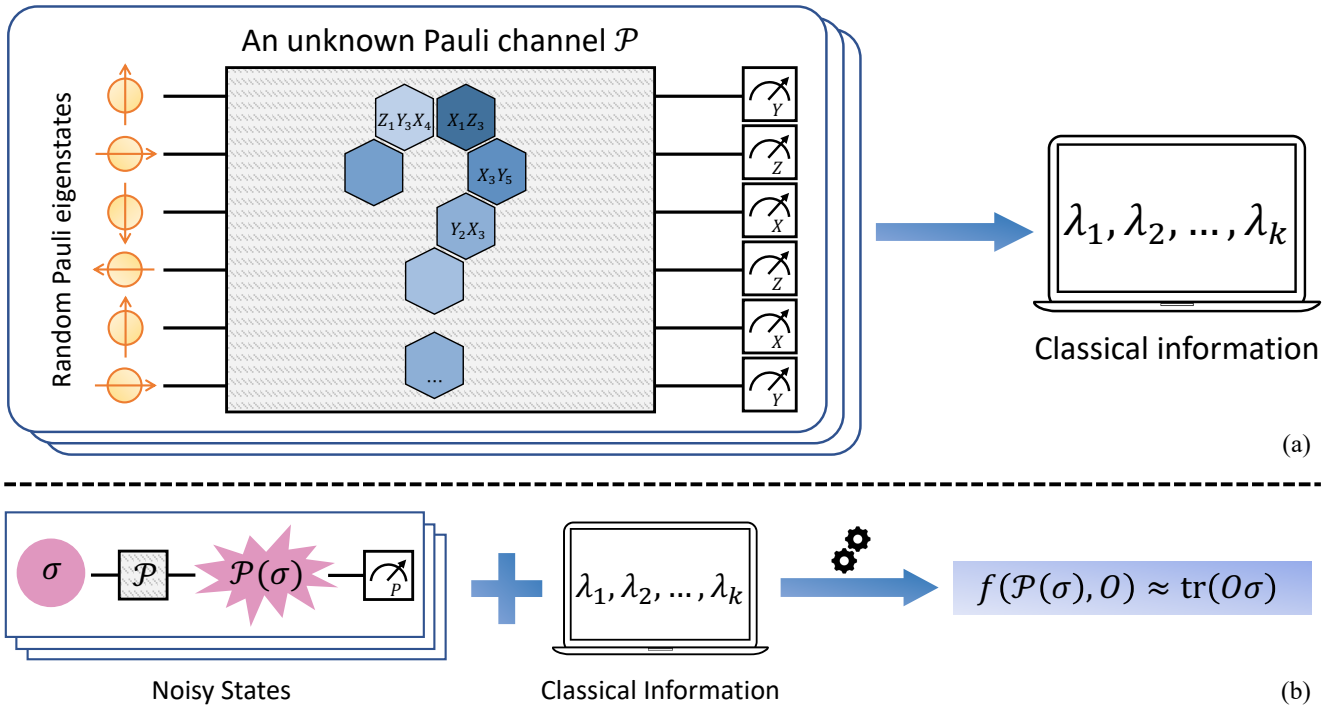


Figure 1: **Illustration of the algorithm for recovering information from Pauli noise.** (a) The classical information contains eigenvalues of the Pauli channel. It is first estimated using the classical shadows of the channel, which is obtained by preparing random Pauli eigenstates as input and measuring the output states in random Pauli basis. (b) Then given any quantum states $\mathcal{P}(\sigma)$ that is subjected to this Pauli noise, estimation of $\mathrm{tr}(O\sigma)$ can be obtained by post-processing measurement results of the noisy state and the classical information we learnt. We note that the same classical information can be reused to recover information for different noisy states.

teria relating to the Pauli transfer matrix of the channel. We also evaluate our algorithm's performance on both Pauli and Clifford shadows.

3. As a notable application, we apply our method to mitigate Pauli errors in Clifford circuits, which leads to a more sample-efficient Pauli error mitigation scheme than previous methods such as probabilistic error cancellation [16].

Algorithm for Information Recovery from Pauli noise. Our first contribution is a polynomial-time algorithm on efficiently recovering information from Pauli noise with logarithmic channel sample complexity. We then prove that the channel sample complexity of the learning algorithm is optimal. The starting point of our method is that under Pauli noise, the Pauli expectation value of a noisy state is a multiple of the noiseless expectation that we seek, scaled by an eigenvalue of the channel. An observable O can be written as a linear combination of Pauli operators, so the value of $\mathrm{tr}(O\sigma)$ is a linear combination of Pauli expectation value of σ. A Pauli noise would then scale each term by a different eigenvalue, so if we can learn the eigenvalues, we would be able to recover the noiseless expectation value.

In the realm of quantum state and quantum process learning, an approach is called classical shadow tomography, which uses randomised measurements to construct a classical estimate of the unknown state [15]. Then the information of the state can be calculated classically. In particular, the sample complexity scales logarithmically with the number of expectation values to predict, hence it is very effective and efficient at estimating multiple information at the same time. As an extension, given an unknown quantum process, one can prepare random input state and perform random measurements on the output state. This is referred to as classical shadow of a quantum process, which can be utilised to construct estimate of the Choi matrix of the process [17, 18], or estimate directly information of a state under this unknown process [19]. This provides us the intuition that the technique of classical shadow tomography has the potential to lead to an efficient method of retrieving information from Pauli noise.

Now we present our algorithm. The main scheme is illustrated in Fig. 1 and the main steps of the algorithm are given in Algorithm 1, where $|P|$ denotes the weight of Pauli operator P, meaning number of tensor factors that are not identity. Our algorithm consists of two parts.

In the first part, we collect classical shadows of the channel, which is the measurement results obtained when one prepares random Pauli eigenstates as input, and measures the output states in random Pauli basis. Using

them, we can calculate the eigenvalues of the noisy Pauli channel. We consider only obtaining eigenvalues for Pauli operators whose weight is less than or equal to k instead of the full 4^n eigenvalues, which would be sufficient for retrieving the expectation value of k-local observables. We show that for local observable, only $\mathcal{O}(\text{poly}(n))$ eigenvalues need to be learnt and only $\mathcal{O}(\log(n))$ applications of the channel suffice, where n is the number of qubits. We then make use of information-theoretic techniques in Refs. [20–22] to prove a lower bound of the sample complexity on this learning task. This matches our upper bound hence shows that this part of the algorithm is optimal.

Algorithm 1 Information recovery from Pauli noise

Require: Access to an unknown n-qubit Pauli channel $\mathcal{P}$, a bounded-degree k-local observable $O = \sum_P \alpha_P P$, copies of unknown noisy state $\mathcal{P}(\sigma)$.

Ensure: An estimation of $\text{tr}(O\sigma)$.

1: Collect classical shadows of the channel.

2: For each n-qubit Pauli operator P with $|P| \leq k$, compute $\widehat{\lambda}_P$, the estimate for eigenvalue of Pauli operator P under Pauli channel $\mathcal{P}$, using the collected shadows.

3: For $O = \sum_{P:|P|\leq k} \alpha_P P$, let $\overleftarrow{\alpha}_P = \alpha_P / \widehat{\lambda}_P$ for $|P| \leq k$.

4: Perform Pauli measurement on the noisy state $\mathcal{P}(\sigma)$ for each Pauli operator P with $|P| \leq k$ and construct the estimation as

$$\text{tr}(O\sigma) \approx \sum_{P:|P|\leq k} \overleftarrow{\alpha}_P \, \text{tr}(P\mathcal{P}(\sigma)).$$

In the second part, using $\mathcal{O}(\text{poly}(n))$ copies of the noisy state, we can obtain desired Pauli expectation values of the noisy state by classical shadow tomography. Combining these with the eigenvalues we learnt, we are able to construct an estimate for the expectation value of the noiseless state. Together, we have the following guarantee:

Theorem 1 (informal version) *To obtain an estimation of* $\text{tr}(O\sigma)$ *for bounded-degree* k*-local observable* O*,* $\mathcal{O}(\log(n^k))$ *applications of the channel* $\mathcal{P}$ *and* $\mathcal{O}(\text{poly}(n))$ *copies of the noisy state* $\mathcal{P}(\sigma)$ *suffice. The total classical computation time is* $\mathcal{O}(\text{poly}(n))$.

Extensions of algorithm. Our second contribution is extending our algorithms on a broader class of channels and evaluating the performance of other classical shadows i.e. the Clifford shadow. We establish a sufficient criterion related to the Pauli transfer matrix of the channel in order to extend the set of channels from which information can be efficiently recovered. One example is the product channel and we present the corresponding algorithm in Section VI of the full version.

As another extension, we explore sample complexity using Clifford shadows, which involves preparing input states from distribution that is invariant under the Clifford group and measuring the output states in random Clifford basis. We find that Clifford shadows would break the efficiency obtained for k-local observables but is potentially better for general observable. A detailed analysis can be found in Appendix D of the full version.

Application. Our third contribution is a Pauli error mitigation scheme that uses our algorithm for mitigating Pauli errors in Clifford circuits. Under the assumption that the circuit $\mathcal{C}$ is consisted of H, S and CNOT gates, and the Pauli noise channel affect each type of the gate is the same. Once again, each Pauli expectation value is scaled by a multiplier that is related to the eigenvalues of the noise channels. We first learn the eigenvalues of the three different noise channels using the same method as before, then using the fact that Pauli operator conjugated by a Clifford gate is another Pauli and this can be computed classically efficiently, we can obtain the noiseless expectation value $\text{tr}(O\mathcal{C}(\sigma))$.

Comparison with related works. Existing information recovery method [10] and quantum error mitigation techniques [16] necessitates the full information of the channel is known, which requires an extensive amount of quantum resources. It also requires the implementation of arbitrary CPTP maps. However, our proposed algorithm of information recovery requires no prior knowledge of a Pauli channel, and the channel sample complexity is logarithmic in the number of qubits. And our method only requires preparing Pauli eigenstates and performing Pauli measurements.

We also note that our proposed algorithm is inspired from [19]. The algorithm in [19] aims to predict the value of $\text{tr}(O\mathcal{P}(\rho))$ from access to $\mathcal{P}$ and ρ, whereas we want to recover the original information from noisy $\mathcal{P}(\rho)$.

Furthermore, there are algorithms that can estimate the error rates for Pauli channels [21–25]. However, learning error rates to high accuracy cannot guarantee to estimate eigenvalues of the channel to the same accuracy and vice versa. In Ref. [23], the authors presents a method for estimating single eigenvalue for a given Pauli operator, then treat it as a query access and use it to estimate the error rates, which is similar to our method in spirit but their estimate is restricted to the specific Pauli that they query.

Concluding remarks. In this work, we introduced an efficient quantum algorithm that could retrieve information from an unknown Pauli noise, using resources that scale polynomially in the number of qubits. The efficiency of the algorithm comes from the fact that only partial knowledge of the channel is required to recover the ideal information for a local and bounded-degree observable. For learning partial eigenvalues of the Pauli channel, we proved it is optimal. We have also shown that the method can be directly applied to recover information from noisy Clifford circuits.

References

[1] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor.

Nature Communications, 5(1):4213, July 2014. ISSN 2041-1723. doi: 10.1038/ncomms5213.

[2] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm, November 2014.

[3] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671): 195–202, September 2017. ISSN 1476-4687. doi: 10.1038/nature23474.

[4] Joel J. Wallman and Joseph Emerson. Noise tailoring for scalable quantum computation via randomized compiling. *Physical Review A*, 94(5): 052325, November 2016. doi: 10.1103/PhysRevA.94.052325. URL https://link.aps.org/doi/10.1103/PhysRevA.94.052325.

[5] Akel Hashim, Ravi K. Naik, Alexis Morvan, Jean-Loup Ville, Bradley Mitchell, John Mark Kreikebaum, Marc Davis, Ethan Smith, Costin Iancu, Kevin P. O'Brien, Ian Hincks, Joel J. Wallman, Joseph Emerson, and Irfan Siddiqi. Randomized Compiling for Scalable Quantum Computing on a Noisy Superconducting Quantum Processor. *Physical Review X*, 11(4):041039, November 2021. doi: 10.1103/PhysRevX.11.041039. URL https://link.aps.org/doi/10.1103/PhysRevX.11.041039.

[6] Mark M Wilde. *Quantum Information Theory*. Cambridge University Press, Cambridge, 2017. ISBN 9781316809976. doi: 10.1017/9781316809976. URL http://ebooks.cambridge.org/ref/id/CBO9781316809976.

[7] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018. ISBN 9781107180567.

[8] Masahito Hayashi. *Quantum Information Theory*. Graduate Texts in Physics. Springer Berlin Heidelberg, Berlin, Heidelberg, 2017. ISBN 978-3-662-49723-4. doi: 10.1007/978-3-662-49725-8. URL http://link.springer.com/10.1007/978-3-662-49725-8.

[9] Jiaqing Jiang, Kun Wang, and Xin Wang. Physical Implementability of Linear Maps and Its Application in Error Mitigation. *Quantum*, 5:600, December 2021. doi: 10.22331/q-2021-12-07-600.

[10] Xuanqiang Zhao, Benchi Zhao, Zihan Xia, and Xin Wang. Information recoverability of noisy quantum states. *Quantum*, 7:978, April 2023. ISSN 2521-327X. doi: 10.22331/q-2023-04-13-978.

[11] Isaac L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, November 1997. ISSN 0950-0340. doi: 10.1080/09500349708231894.

[12] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White. Ancilla-Assisted Quantum Process Tomography. *Physical Review Letters*, 90(19):193601, May 2003. doi: 10.1103/PhysRevLett.90.193601.

[13] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Physical Review A*, 77(3): 032322, March 2008. doi: 10.1103/PhysRevA.77.032322.

[14] Scott Aaronson. Shadow tomography of quantum states. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2018, pages 325–338, New York, NY, USA, June 2018. Association for Computing Machinery. ISBN 978-1-4503-5559-9. doi: 10.1145/3188745.3188802. URL https://doi.org/10.1145/3188745.3188802.

[15] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, October 2020. ISSN 1745-2481. doi: 10.1038/s41567-020-0932-7.

[16] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error Mitigation for Short-Depth Quantum Circuits. *Physical Review Letters*, 119(18):180509, November 2017. doi: 10.1103/PhysRevLett.119.180509.

[17] Ryan Levy, Di Luo, and Bryan K Clark. Classical shadows for quantum process tomography on near-term quantum computers. *arXiv preprint arXiv:2110.02965*, 2021.

[18] Jonathan Kunjummen, Minh C Tran, Daniel Carney, and Jacob M Taylor. Shadow process tomography of quantum channels. *arXiv preprint arXiv:2110.03629*, 2021.

[19] Hsin-Yuan Huang, Sitan Chen, and John Preskill. Learning to predict arbitrary quantum processes, October 2022. URL http://arxiv.org/abs/2210.14894.

[20] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Information-Theoretic Bounds on Quantum Advantage in Machine Learning. *Physical Review Letters*, 126(19):190505, May 2021. doi: 10.1103/PhysRevLett.126.190505. URL https://link.aps.org/doi/10.1103/PhysRevLett.126.190505.

[21] Omar Fawzi, Aadil Oufkir, and Daniel Stilck França. Lower Bounds on Learning Pauli Channels, January 2023.

[22] Senrui Chen, Sisi Zhou, Alireza Seif, and Liang Jiang. Quantum advantages for Pauli channel estimation. *Physical Review A*, 105(3):032435, March

2022. ISSN 2469-9926, 2469-9934. doi: 10.1103/PhysRevA.105.032435. URL http://arxiv.org/abs/2108.08488.

[23] Steven T. Flammia and Ryan O'Donnell. Pauli error estimation via population recovery. *Quantum*, 5:549, sep 2021. doi: 10.22331/q-2021-09-23-549. URL https://doi.org/10.22331%2Fq-2021-09-23-549.

[24] Steven T. Flammia and Joel J. Wallman. Efficient estimation of pauli channels. *ACM Transactions on Quantum Computing*, 1(1):1–32, dec 2020. doi: 10.1145/3408039. URL https://doi.org/10.1145%2F3408039.

[25] Robin Harper, Steven T Flammia, and Joel J Wallman. Efficient learning of quantum noise. *Nature Physics*, 16(12):1184–1188, 2020. doi: 10.1038/s41567-020-0992-8.

A Comprehensive Strategy for Improving Steiner-Gauss Elimination: Qubit Layout Optimization and Circuit Division

Huan Yu[1] * Zanhe Qi[1] † Shigeru Yamashita[1] ‡

[1] *Ritsumeikan University*

Abstract. This paper proposes a comprehensive method to improve Steiner-Gauss elimination for synthesizing an NNA-compliant circuit. This method contains two components. The first component provides an optimized initial qubit layout which can minimize the sum of distances between the control bits and target bits of all CNOT gates in the circuit to improve Steiner-Gauss elimination. The second component divides the circuit into subcircuits based on the "orientation" of CNOT gates to eliminate unnecessary NNA CNOT gates in NNA-compliant circuits to improve Steiner-Gauss elimination. Our experimental results show that our proposed method can reduce 33.18% NNA CNOT gates on average compared to the basic Steiner-Gauss elimination, and reduce 26.20% NNA CNOT gates on average compared to PAQCS and Steiner-Gauss elimination.

Keywords: Nearest Neighboring Architecture constraint, CNOT gate, Steiner-Gauss elimination, dividing CNOT circuit, optimized initial qubit layout

1 INTRODUCTION

Quantum circuits are composed of quantum logic operations. The majority of quantum logic operations can be effectuated using certain universal quantum gate sets, i.e., the Clifford and T gate set. In other words, quantum circuits can be constructed as a sequence of fundamental quantum gates, specifically, one-qubit quantum gates and CNOT gates. To physically realize quantum circuits, an essential step is transforming quantum circuits to meet so-called *Nearest Neighboring Architecture (NNA) constraint* [1]; NNA constraint dictates that all CNOT gates must operate on two physically adjacent qubits, and we refer to CNOT gates and circuits that satisfy NNA constraint as NNA CNOT gates and NNA-compliant circuits.

Many studies have focused on devising efficient methods to transform quantum circuits to meet NNA constraint for one-dimensional [2] [3], two-dimensional [4] [5] and three-dimensional [6] [7] qubit arthitecture. The general strategy is to insert *SWAP* gates to interchange qubits, making the control and target bits of CNOT gates adjacent to satisfy NNA constraint. However, given the high cost of *SWAP* gates, Kissinger et. al proposed a more efficient approach, known as *Steiner-Gauss elimination* [8]. This approach utilizes *Gaussian elimination* [9] to design NNA-compliant circuits with typically fewer NNA CNOT gates than approaches based on inserting SWAP gates.

Our contribution: To improve Steiner-Gauss elimination, we provide a comprehensive proposal comprised of two components:

- A pioneering algorithm enables the efficient determination of an optimal initial qubit layout for a CNOT circuit, a strategy rooted in minimizing the sum of distances between control and target bits of all CNOT gates. This component can reduce the number of NNA CNOT gates in the NNA-compliant circuit to improve Steiner-Gauss elimination.
- An innovative algorithm divides a CNOT circuit into sub-circuits, based on a novel notion termed "orientation" of CNOT gates. This component can eliminate the unnecessary NNA CNOT gates in the NNA-compliant circuit to improve Steiner-Gauss elimination.

We can integrate these two algorithms as a unified, comprehensive proposal. Additionally, it remains a great option to employ either of these algorithms individually in contexts where they demonstrate optimal applicability.

2 PRELIMINARY

2.1 Representing a CNOT circuit by 0/1 matrix

As mentioned previously, NNA constraint mainly imposes the restriction on CNOT gates. Hence, to design NNA-compliant circuits, we focus on CNOT circuits composed exclusively of CNOT gates.

For CNOT gates, we can consider that a CNOT gate computes the XOR operation of the two inputs on the target qubit when the input quantum states are classical (i.e., $|0\rangle$ or $|1\rangle$). Let the inputs of an n-qubit quantum circuit be $q_0, q_1, \cdots, q_{n-1}$ in the following. As shown in Fig. 1 (a), the output state of each qubit of a CNOT circuit can always be expressed by an exclusive sum (i.e., linear combination) of q_i.

As shown in Fig. 1 (b), we obtain the same output state of the CNOT circuit, by performing matrix multiplication between a 0/1 matrix and a vector that represents the inputs of the CNOT circuit. Thus, it is reasonable to regard this matrix as a representation of the functionality of the CNOT circuit. In other words, the 0/1 matrix in Fig. 1 (b) can represent the CNOT circuit in Fig. 1 (a).

*y@ngc.is.ritsumei.ac.jp
†goose@ngc.is.ritsumei.ac.jp
‡ger@cs.ritsumei.ac.jp

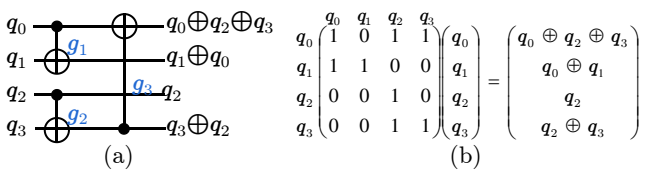


Figure 1: The example of using a 0/1 matrix to represent a CNOT circuit.

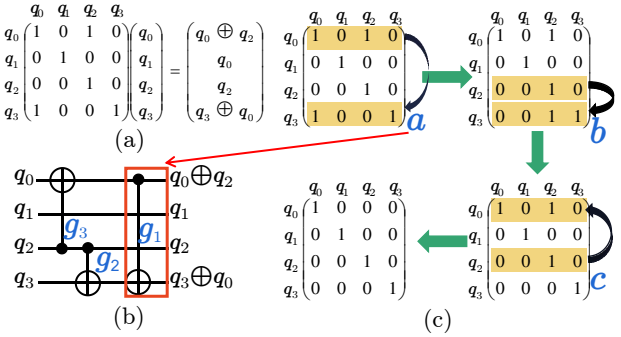


Figure 2: The example of the CNOT circuit synthesis by Gaussian elimination.

In this way, we can use 0/1 matrices to represent CNOT circuits.

2.2 CNOT circuit synthesis by Gaussian elimination

After obtaining a 0/1 matrix that represents a CNOT circuit, the next step is to explore the utilization of this matrix for constructing a new circuit. Gaussian elimination, a renowned mathematical algorithm, serves as a means to solve linear equations by transforming the associated matrix into an identity matrix. We employ Gaussian elimination to transform the 0/1 matrix into an identity matrix.

To achieve this, we implement the following operation: substituting the i-th row with the result of the XOR operation between the i-th and j-th rows. It is worth noting that this operation corresponds to the effect of a CNOT gate whose target and control bits are the i-th and the j-th qubits, respectively. We repeat the row operations above to transform the matrix into an identity matrix, resulting in a sequence of CNOT gates that correspond to the row operations implemented during Gaussian elimination. Because the identity matrix corresponds to the input state of the circuit, the above sequence of gates transforms the output state of the circuit to the initial state. This means that we can get the desired circuit with the same functionality by reversing the above sequence of CNOT gates.

For example, the leftmost 0/1 matrix in Fig. 2 (a) represents the functionality of the CNOT circuit. In Fig. 2 (c), we show a procedure of Gaussian elimination to transform the matrix in Fig. 2 (a) into an identity matrix. The circuit in Fig. 2 (b) is the new synthesized circuit. The row operations a, b, and c in Fig. 2 (c) correspond to the CNOT gates g_1, g_2, g_3 in Fig. 2 (b), respectively.

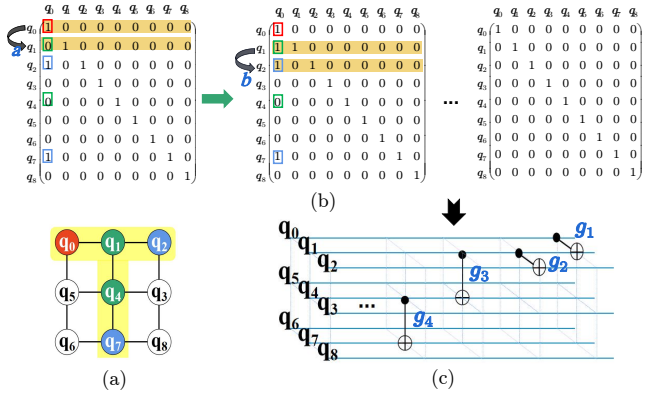


Figure 3: The example of Steiner-Gauss elimination

2.3 NNA-compliant circuit synthesis by Steiner-Gauss elimination

Building upon the preceding introduction, we enforce that solely the row operations associated with the NNA CNOT gates are permitted during the process of Gaussian elimination. Consequently, we derive a sequence of NNA CNOT gates, effectively synthesizing the desired NNA-compliant circuit.

Now let us consider NNA constraint in the two-dimensional 3×3 qubit architecture as shown in Fig. 3 (a). The most-left matrix in Fig. 3 (b) represents a circuit containing two non-NNA CNOT gates, whose control bits are q_0 and target bits are q_2 and q_7, respectively. We utilize the Breadth-First Search algorithm to generate a Steiner tree [10] that connects q_0 and q_2, as well as q_0 and q_7 by inserting the least Steiner points as shown in Fig. 3 (a). Employing the Steiner tree as a guide, we are enabled to utilize only row operations that correspond to NNA CNOT gates to transform the matrix representing the CNOT circuit into an identity matrix. Subsequently, compile the NNA CNOT gates corresponding to row operations utilized during the transformation of the matrix into an identity matrix to synthesize the desired NNA-compliant circuit. For example, the NNA-compliant circuit is shown as Fig. 3 (c), and the row operations a, and b in Fig. 3 (b) correspond to the NNA CNOT gates g_1, and g_2 in Fig. 3 (c).

3 Proposed method

3.1 Component. 1: Optimizing initial qubit layout

When employing Steiner-Gauss elimination to convert a non-NNA CNOT gate into a sequence of NNA CNOT gates, it is evident that the distance between the control bit and the target bit of the CNOT gate directly influences the number of required NNA CNOT gates. As the distance decreases, the control bit and the target bit are closer, and fewer NNA CNOT gates are required.

Hence, we present an initial qubit layout where the distance between the control bit and the target bit is smaller. In this initial layout, the circuit containing this CNOT gate can be converted into an NNA-compliant circuit with fewer or even without additional NNA CNOT

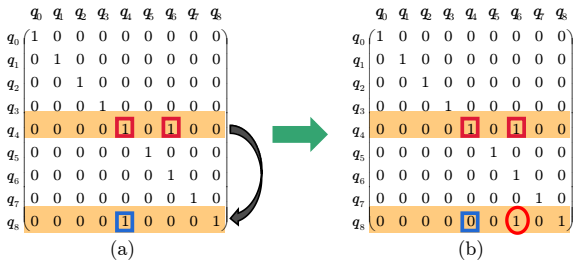


Figure 4: The example of the disadvantage of Steiner-Gauss elimination

gates.

Moreover, for a CNOT circuit comprising some CNOT gates, we can offer an initial qubit layout that reduces the sum of distances of all CNOT gates in this circuit as well. In other words, within this initial qubit layout, the CNOT gates' control bits and target bits become closer generally. In this initial qubit layout, this CNOT circuit containing these CNOT gates can be converted into an NNA-compliant circuit with fewer NNA CNOT gates. Thus, we can improve Steiner-Gauss elimination by utilizing the optimized initial qubit layout which minimizes the sum of distances of all CNOT gates.

To achieve this proposal, we employ Simulated Annealing to generate a list of initial qubit layouts which minimize the sum of distances of all CNOT gates in the CNOT circuit. Then, we select the initial qubit layout in the list which has the minimal NNA CNOT gates in the NNA-compliant circuit generated by Steiner-Gauss elimination as the optimized qubit layout.

3.2 Component. 2: Dividing CNOT circuit

Predicated on our experimental experience, we recognize that there is a flaw in Steiner-Gauss elimination that leads to the generation of unnecessary NNA CNOT gates in the NNA-compliant circuit synthesized by Steiner-Gauss elimination. For example, when eliminating the "1" in the blue box in Fig. 4 (a) using Steiner-Gauss elimination, the undesired "1" in the red circle is added in Fig. 4 (b), which requires additional row operations for elimination. The emergence of this undesired "1" is attributed to the presence of the "1" in the upper triangular matrix.

To avoid the emergence of the undesired "1", we first introduce a new notion termed "*orientation*" of a CNOT gate as follows.

Definition 1 *For a qubit order:* $q_0, q_1, \cdots, q_{n-1}$*, if a CNOT gate whose control bit is in front of the target bit, the "orientation" of the CNOT gate is* **downward**, *or else the "orientation" is* **upward**.

The matrix representing a circuit containing CNOT gates with the same orientation is the lower triangular matrix or upper triangular matrix. In the case of lower triangular matrices and upper triangular matrices, the undesired "1s" do not emerge during Steiner-Gauss elimination. Hence, we propose that while preserving the functionality of the circuit, divide a CNOT circuit into multiple sub-circuits comprised of CNOT gates with the same "*orientation*". Then perform Steiner-Gauss elimination on each subcircuit. In this way, we can improve Steiner-Gauss elimination by eliminating the unnecessary CNOT gates.

In summary, our proposed approach is to provide the optimized initial qubit layout first. Subsequently, divide the circuit in the optimized initial qubit layout into subcircuits comprised of CNOT gates with the same "*orientation*". Perform Steiner-Gauss elimination on each subcircuit to obtain NNA-compliant sub-circuits. Finally, integrate NNA-compliant sub-circuits into the desired NNA-compliant circuit.

Table 1: Experimental Results

circuits	#	Steiner-Gauss	PAQCS	Proposed method
R_cnot_1	10	30	29	14
R_cnot_2	10	24	23	14
R_cnot_3	15	41	47	30
R_cnot_4	15	48	28	23
R_cnot_5	15	42	42	21
R_cnot_6	15	40	37	26
R_cnot_7	15	49	36	25
R_cnot_8	20	67	43	31
R_cnot_9	20	64	49	33
R_cnot_10	20	39	42	25
R_cnot_11	20	60	35	36
R_cnot_12	25	64	65	44
R_cnot_13	25	60	58	46
R_cnot_14	25	51	52	43
R_cnot_15	30	52	63	51
R_cnot_16	30	41	57	45
R_cnot_17	30	63	50	51

4 Experimental results

To assess the effectiveness of our proposed method, we compared the performance of the following three methods:

- Basic Steiner-Gauss elimination
- PAQCS [11] (Providing an algorithm to optimize the initial qubit layout as well) and Steiner-Gauss elimination
- Our proposed method

We used randomly generated CNOT circuits to evaluate the total number of NNA CNOT gates in NNA-compliant circuits by each method. The randomly generated CNOT circuits, *R_cnot_1* to *R_cnot_17*, consist of 9 qubits, and all qubits are placed on 3×3 two-dimensional square architecture. "#" represents the number of CNOT gates in test circuits. The proposed method demonstrated superior performance, with an average reduction of 33.18% NNA CNOT gates compared to Basic Steiner-Gauss elimination, and 26.20% compared to PAQCS and Steiner-Gauss elimination.

For computing time, we set a relatively small number of iterations and a low initial temperature for Simulated Annealing to ensure the computing time of Component. 1 within 1 second, and the computing time of Component. 2 can be considered negligible. Thus, the computing time of the proposed method is still within 1 second, indicating no practical issues.

References

[1] Y. Hirata, M. Nakanishi, S. Yamashita, and Y. Nakashima, "An efficient method to convert arbitrary quantum circuits to ones on a linear nearest neighbor architecture," in 3rd International Conference on Quantum, Nano and Micro Technologies, Feb. 2009, pp. 26–33.

[2] Mehdi Saeedi, Robert Wille, and Rolf Drechsler, "Synthesis of quantum circuits for linear nearest neighbor architectures," Quantum Information Processing, Vol. 10, No. 3, pp. 355–377, 2011.

[3] Alireza Shafaei, Mehdi Saeedi, and Massoud Pedram, "Optimization of quantum circuits for interaction distance in linear nearest neighbor architectures," in 50th ACM/EDAC/IEEE Design Automation Conference (DAC), Austin, TX, USA, 2013, pp. 1-6.

[4] Ken Matsumoto, Kazuyuki Amano, "Representation of quantum circuits with Clifford and $\pi/8$ gates," arXiv preprint arXiv: 0806.3834, 2008.

[5] Alireza Shafaei, Mehdi Saeedi, and Massoud Pedram, "Qubit placement to minimize communication overhead in 2D quantum architectures," in 19th Asia and South Pacific Design Automation Conference (ASP-DAC), Singapore, 2014, pp. 495-500.

[6] Vasmer. Michael, and Dan E. Browne, "Three-dimensional surface codes: Transversal gates and fault-tolerant architectures," Physical Review A, Vol. 100, Iss. 1, pp. 012312, 2019.

[7] Carlos A. Perez-Delgado, Michele Mosca, Paola Cappellaro, and David G. Cory, "Single spin measurement using cellular automata techniques," Phys. Rev. Lett, Vol. 97, No. 10, pp. 100501, 2006.

[8] Aleks Kissinger, and Arianne Meijer-van de Griend, "CNOT circuit extraction for topologically-constrained quantum memories," arXiv preprint arXiv: 1904.00633, pp. 4– 11, 2019.

[9] T. Beth, and M. Rotteler, Quantum algorithms: Applicable algebra and quantum physics (Quantum Information). Springer, 2001.

[10] the free encyclopedia Wikipedia, Steiner tree problem, 2020. [Online]. Available: https://en.wikipedia.org/wiki/Steiner_tree_problem

[11] C.-C. Lin, S. Sur-Kolay and N. K. Jha, "PAQCS: Physical Design-aware Fault-tolerant Quantum Circuit Synthesis," IEEE Transactions on VLSI Systems.

Optimal Strategies of Quantum Metrology with a Strict Hierarchy

Qiushi Liu,[1,*] Zihao Hu,[2,†] Haidong Yuan,[2,‡] and Yuxiang Yang[1,§]

[1] *QICI Quantum Information and Computation Initiative, Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong, China*

[2] *Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong, China*

(Dated: May 18, 2023)

Keywords: quantum metrology, quantum strategies, quantum control, causality, group symmetry

The full article is available on arXiv:2203.09758.

Quantum metrology [1, 2] features a series of promising applications in the near future [3]. In the prototypical setting of quantum metrology, the goal is to estimate an unknown parameter carried by a quantum channel, given N queries to it. A pivotal task is to design a *strategy* that utilizes these N queries to generate a quantum state with as much information about the unknown parameter as possible. This often involves, for example, preparing a suitable input probe state [4–6] and applying intermediate quantum control [7, 8] as well as quantum error correction [9–12].

In reality, strategies that we could implement are often subject to physical restrictions. For example, for systems with short coherence time it might be favorable to adopt the parallel strategy (Fig. 1(a)), where multiple queries of the unknown channel are applied simultaneously on a multipartite entangled state [4]. When the system has longer coherence time and can be better controlled, one could choose to query the channel sequentially (Fig. 1(b)), which may potentially enhance the precision. However, it remains an intriguing open problem whether a sequential strategy can strictly outperform any parallel one for single-parameter estimation. In addition to parallel and sequential strategies, it was recently discovered that the quantum SWITCH [13], a primitive where the order of making queries to the unknown channel is in a quantum superposition (Fig. 1(c)), can be employed to generate new strategies of quantum metrology [14–16] that may even break the Heisenberg limit [16]. Moreover, indefinite causal structures beyond the quantum SWITCH [13, 17, 18] (Fig. 1(d) and (e)) have recently been shown to further boost the performance of certain information processing tasks [19, 20]. Their performance in quantum metrology, however, remains unknown. Overall, the main difficulty is the lack of a systematic method that deals with the optimization of probe state preparation, control, etc., in a strategy in a unified fashion.

It is therefore of utmost importance to identify the *ultimate* precision of quantum metrology

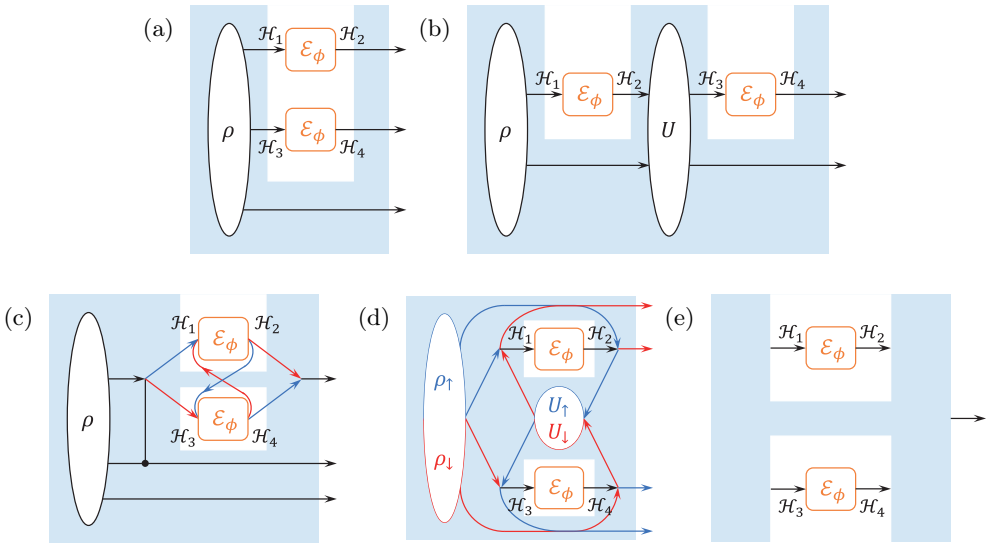


FIG. 1. **Prototypical strategies of quantum metrology** (for the $N = 2$ case). $\mathcal{E}_\phi$ is a quantum channel carrying an unknown parameter ϕ, and the blue shaded area represents a *strategy*. (a) A parallel strategy. (b) A sequential strategy, where U is a control operation. (c) A quantum SWITCH strategy. The blue and red lines respectively correspond to two different execution orders entangled with a control qubit. (d) A causal superposition strategy. Two sequential strategies, plotted in blue and in red respectively, are entangled with a control qubit (not shown in the figure) and the output will be measured with the control qubit collectively. (e) A general indefinite-causal-order strategy.

under each family of strategies and determine whether one family of strategies *strictly* outperforms the other. In this work, our main contributions include:

1. We develop a semidefinite programming method of evaluating the optimal achievable precision characterized by the quantum Fisher information (QFI) and develop an algorithm that yields an optimal strategy (in terms of its process matrix as well as the circuit implementation) attaining this QFI. See Theorem 1 and Algorithm 1 of the full paper.

2. For the strategy set that admits a symmetric structure, we develop a method of reducing the complexity of our algorithms by an exponential factor.

3. We discover a strict hierarchy (see Fig. 2) of the optimal precision for a series of physically relevant families of strategies, including parallel, sequential and indefinite-causal-order [13, 17, 18] ones (see Fig. 1).

4. We discover several intriguing new scenarios where ICO strategies have strictly higher QFI. In particular, there exists a task where a strategy with a simple quantum SWITCH (without any intermediate control operations) beats all causally-ordered strategies.

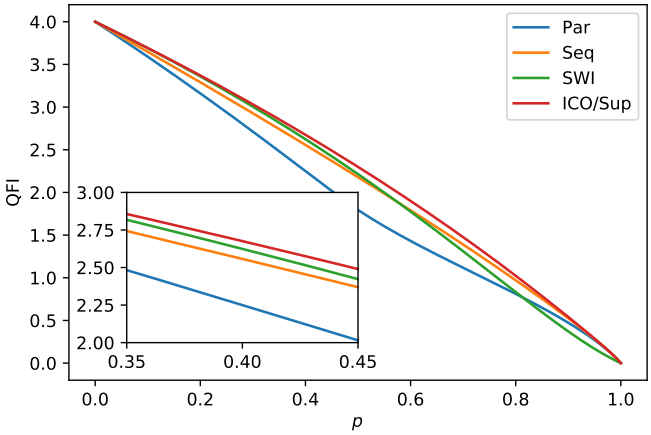


FIG. 2. **Hierarchy of QFI using parallel, sequential, and indefinite-causal-order strategies** ($N = 2$, amplitude damping noise). For each copy of a parameterized quantum channel, the parameter of interest is encoded in a unitary evolution, followed by an amplitude damping channel with the decay parameter p. Some notations: Par for parallel; Seq for sequential; Swi for quantum switch; Sup for causal superposition; ICO for general indefinite-causal-order. A strict hierarchy of Par, Seq and ICO holds, i.e., the QFI $J^{(\mathsf{Par})} < J^{(\mathsf{Seq})} < J^{(\mathsf{ICO})}$. The gaps can be seen more clearly by zooming in on the interval $[0.35, 0.45]$ of the value of p. Moreover, we find it quite surprising that, assisted by the quantum SWITCH (without any additional control operations), we can beat any sequential strategies in certain cases (e.g. $p < 0.5$). The QFI $J^{(\mathsf{Sup})} = J^{(\mathsf{ICO})}$ with an error tolerance of no more than 10^{-8} in this case, but the gap between $J^{(\mathsf{Sup})}$ and $J^{(\mathsf{ICO})}$ could exist with larger N or for other types of noise, which can be observed by randomly sampling noise channels.

Compared to previously known asymptotically saturable bounds for quantum channel estimation, our work fills the gap of fully optimized quantum metrology in the regime when N is not too large. Identifying the strict hierarchy is an important step in the study of quantum metrological strategies (where people mainly focused on the parallel-versus-sequential problem for $N \to \infty$). Moreover, with concrete strategies attaining the optimal precision, our result serves as a versatile tool for the demonstration of optimal quantum metrology and the design of optimal quantum sensors, especially in the context of control optimization [21, 22] and indefinite causal orders [13–20]. The result is appealing not only to the general community of quantum metrology and quantum information, but also to researchers who are seeking advantages that new space-time structures could introduce to quantum information processing and quantum computing [14–16, 20, 23, 24]. Therefore, we believe that our work makes a substantial contribution to quantum metrology and quantum information processing, and will be of interest to the general audience of AQIS.

* qsliu@cs.hku.hk

† zhhu@mae.cuhk.edu.hk

‡ hdyuan@mae.cuhk.edu.hk

§ yuxiang@cs.hku.hk

[1] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photonics **5**, 222 (2011).

[2] C. L. Degen, F. Reinhard, and P. Cappellaro, Quantum sensing, Rev. Mod. Phys. **89**, 035002 (2017).

[3] J. M. Martinis, Qubit metrology for building a fault-tolerant quantum computer, npj Quantum Inf. **1**, 15005 (2015).

[4] H. Lee, P. Kok, and J. P. Dowling, A quantum rosetta stone for interferometry, J. Mod. Opt. **49**, 2325 (2002).

[5] V. Bužek, R. Derka, and S. Massar, Optimal Quantum Clocks, Phys. Rev. Lett. **82**, 2207 (1999).

[6] M. Kitagawa and M. Ueda, Squeezed spin states, Phys. Rev. A **47**, 5138 (1993).

[7] H. Yuan and C.-H. F. Fung, Optimal Feedback Scheme and Universal Time Scaling for Hamiltonian Parameter Estimation, Phys. Rev. Lett. **115**, 110401 (2015).

[8] S. Pang and A. N. Jordan, Optimal adaptive control for quantum metrology with time-dependent hamiltonians, Nat. Commun. **8**, 14695 (2017).

[9] W. Dür, M. Skotiniotis, F. Fröwis, and B. Kraus, Improved Quantum Metrology Using Quantum Error Correction, Phys. Rev. Lett. **112**, 080801 (2014).

[10] E. M. Kessler, I. Lovchinsky, A. O. Sushkov, and M. D. Lukin, Quantum Error Correction for Metrology, Phys. Rev. Lett. **112**, 150802 (2014).

[11] R. Demkowicz-Dobrzański, J. Czajkowski, and P. Sekatski, Adaptive Quantum Metrology under General Markovian Noise, Phys. Rev. X **7**, 041009 (2017).

[12] S. Zhou, M. Zhang, J. Preskill, and L. Jiang, Achieving the heisenberg limit in quantum metrology using quantum error correction, Nat. Commun. **9**, 78 (2018).

[13] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, Quantum computations without definite causal structure, Phys. Rev. A **88**, 022318 (2013).

[14] F. Chapeau-Blondeau, Noisy quantum metrology with the assistance of indefinite causal order, Phys. Rev. A **103**, 032615 (2021).

[15] C. Mukhopadhyay, M. K. Gupta, and A. K. Pati, Superposition of causal order as a metrological resource for quantum thermometry (2018), arXiv:1812.07508 [quant-ph].

[16] X. Zhao, Y. Yang, and G. Chiribella, Quantum Metrology with Indefinite Causal Order, Phys. Rev. Lett. **124**, 190503 (2020).

[17] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nat. Commun. **3**, 1092 (2012).

[18] M. Araújo, C. Branciard, F. Costa, A. Feix, C. Giarmatzi, and Č. Brukner, Witnessing causal nonseparability, New J. Phys. **17**, 102001 (2015).

[19] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Reversing Unknown Quantum Transformations: Universal Quantum Circuit for Inverting General Unitary Operations, Phys. Rev. Lett. **123**, 210502 (2019).

[20] J. Bavaresco, M. Murao, and M. T. Quintino, Strict Hierarchy between Parallel, Sequential, and Indefinite-Causal-Order Strategies for Channel Discrimination, Phys. Rev. Lett. **127**, 200504 (2021).

[21] Z. Hou, R.-J. Wang, J.-F. Tang, H. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Control-Enhanced Sequential Scheme for General Quantum Parameter Estimation at the Heisenberg Limit, Phys. Rev. Lett. **123**, 040501 (2019).

[22] Z. Hou, Y. Jin, H. Chen, J.-F. Tang, C.-J. Huang, H. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, "Super-Heisenberg" and Heisenberg Scalings Achieved Simultaneously in the Estimation of a Rotating Field, Phys. Rev. Lett. **126**, 070503 (2021).

[23] D. Ebler, S. Salek, and G. Chiribella, Enhanced Communication with the Assistance of Indefinite Causal Order, Phys. Rev. Lett. **120**, 120502 (2018).

[24] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. Allard Guérin, A. A. Abbott, C. Branciard, v. Brukner, G. Chiribella, and P. Walther, Experimental quantum communication enhancement by superposing trajectories, Phys. Rev. Research **3**, 013093 (2021).

Observation of a dynamical topological quantization in noisy photonic quantum walks

Qin-Qin Wang,[1,2,*] Si-Jing Tao,[1,2,*] Xiao-Ye Xu,[1,2,3,†] Ying Hu,[4,5] Shao-Jun Dong,[6] Ze-Shan He,[1,2] Yang Xue,[1,2] Yong-Jian Han,[1,2,3,6,‡] Jan Carl Budich,[7,§] Chuan-Feng Li,[1,2,3,¶] and Guang-Can Guo[1,2,3]

[1] *CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China*
[2] *CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China*
[3] *Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*
[4] *State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Laser Spectroscopy, Shanxi University, Taiyuan 030006, China*
[5] *Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China*
[6] *Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230031, China*
[7] *Institute of Theoretical Physics, Technische Universität Dresden and Würzburg-Dresden Cluster of Excellence ct.qmat, 01062 Dresden, Germany*

The discovery of dynamical signatures of topology has recently broadened the notion of topological matter beyond the realm of thermal equilibrium. In coherent quench dynamics, a topological discrepancy between a Hamiltonian changing its topological properties over a parameter quench and a topologically inert quantum state naturally occurs. Identifying and observing topological properties that unambiguously characterize such a complex scenario remains an important challenge. Here, we experimentally demonstrate how the presence of noise in quench dynamics enables the observation of a topologically quantized bulk response known as the mean chiral displacement, in agreement with the dynamical buildup of a topological mixed state. To this end, we realize photonic quantum walks with engineered noise to simulate noisy quench dynamics. The difficulty of complete density-matrix reconstruction in noisy scenarios is efficiently addressed via machine learning of the distributions of multiple interferometric measurements, allowing us to monitor the state topology.

Introduction.—The topological quantization of physical observables is among the most fascinating phenomena in nature [1–5], and has crucial implications for the classification of matter as well as for new metrology and information technology [6–9]. While early work on topological matter has focused on low-temperature physics beyond the celebrated Landau-Ginzburg-Wilson paradigm [10], more recently numerous dynamical topological properties have been explored in dissipative systems [11, 12] and non-equilibrium settings [13–17], respectively. There, however, even though a range of topological phenomena without a direct equilibrium counterpart have been discovered, topological quantization is often challenged and protecting symmetries may be dynamically broken [18–20].

For the generic non-equilibrium scenario of a quantum quench, an intriguing discrepancy may arise between the topological properties of the quantum state that are inert under coherent time evolution and those of the Hamiltonian that may change over the parameter quench [11, 21–26]. Reconciling this topological discrepancy and identifying clear experimental signatures that characterize such complex non-equilibrium situations remains an important challenge, from both a conceptual and experimental perspective, where techniques of real-time full quantum state control and tomography in momentum space need to be combined.

In this work, using the platform of photonic quantum walks (QWs), we theoretically propose (see Fig. 1) and experimentally demonstrate (see Fig. 3) how a topologically quantized observable generically emerges in the noisy quench dynamics of a one-dimensional (1D) chiral-symmetric two-banded system. Specifically, a bulk response known as the mean chiral displacement (MCD) [27–30] is found to exhibit a quantized value that is stabilized by dephasing noise in two important aspects. First, by controlling the interaction of QWs with a specific environment, a pure dephasing of the quantum state in momentum space is introduced. Consequently, we observe that dephasing can lead to a topologically quantized value of the instantaneous MCD by damping its non-equilibrium oscillations, thus alleviating the need for observing long-time averages. Second, while the protecting chiral symmetry is dynamically broken [18] by coherent superpositions of ground- and excited states, it is restored in the ensemble-averaged mixed state of the noisy system. Experimentally, by developing a unique mixed-state learning technique for our photonic QWs, we are able to track the time-dependent topology of the density matrix. Interestingly, we find that the Uhlmann phase [31] of the mixed state exhibits a similar temporal behavior as the observed MCD, thus resolving the topological discrepancy that characterizes the coherent limit.

We note that the dynamical buildup of an asymptotic topological quantization has been theoretically studied in the context of 2D Chern insulators [11], where unitary time evolution keeps the Chern number of states constant [21, 22], while the non-equilibrium Hall response may exhibit a strongly oscillatory time-dependent behavior [22, 23]. There, however, only under specific circumstances, including the limit of a slow quench [11], a nearly quantized Hall response may be recovered in the long-

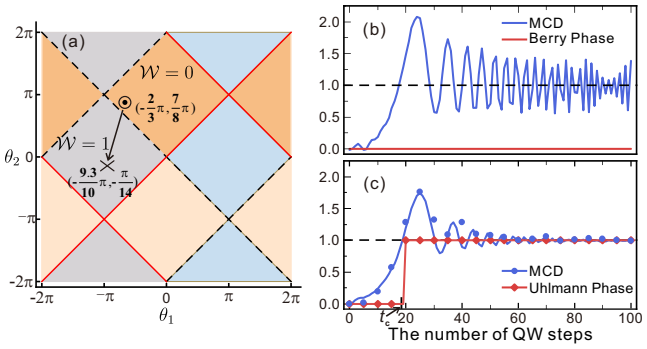


FIG. 1. Theoretical analysis of topological properties in QWs. (a) Phase diagram of the model system and illustration of the parameter quench. (b) Numerical simulation of the MCD and the Berry phase in coherent QWs. (c) Numerical simulation of the Uhlmann phase divided by π and the MCD (solid lines) in noisy QWs with pure dephasing. The blue circles and red diamonds represent the corresponding results in the 100-step noisy QWs interacting with a simulated environment. Decay rate $\gamma_k \equiv 0.05$, fluctuating degree $\delta\theta_1 = 0.2$, and sample size is $M = 200$.

time average, and may be smoothed in the presence of dephasing. By contrast, in our present setting, a topologically quantized instantaneous MCD reflecting uniquely the instantaneous Hamiltonian topology is stabilized by noise, irrespective of details of the quench protocol.

Noisy quantum walk model.—We start by theoretically showing how a dynamical topological quantization can arise in noisy QWs. We consider the split-step QWs [32], whose each time-step dynamic is governed by a unitary operator $U = R_y(\frac{\theta_1}{2})S_\downarrow R_y(\theta_2)S_\uparrow R_y(\frac{\theta_1}{2})$. $R_y(\theta) = e^{-i\theta/2\sigma_y}$ denotes the rotation of the internal spin state of the walker by an angle of θ along the y-axis. The shift operators $S_\uparrow = |x+1\rangle\langle x| \otimes |\uparrow\rangle\langle\uparrow| + |x\rangle\langle x| \otimes |\downarrow\rangle\langle\downarrow|$ and $S_\downarrow = |x\rangle\langle x| \otimes |\uparrow\rangle\langle\uparrow| + |x-1\rangle\langle x| \otimes |\downarrow\rangle\langle\downarrow|$ move the walker to the neighboring sites (labeled by $|x\rangle$) according to its spin states (labeled by $\{|\uparrow\rangle, |\downarrow\rangle\}$). The dynamics U^t can stroboscopically simulate a time evolution described by the effective Hamiltonian $\mathcal{H}_{\rm eff}$, with $U^t = e^{-i\mathcal{H}_{\rm eff}t}$. The effective Hamiltonian in quasi-momentum k space reads (see Supplementary Sect. I): $\mathcal{H}_{\rm eff}(\theta_1,\theta_2) = \int_{-\pi}^{\pi} dk \mathcal{H}_{\rm eff}(k) = \int_{-\pi}^{\pi} dk[E(k)\mathbf{n}_{\mathcal{H}}(k)\cdot\vec{\sigma}] \otimes |k\rangle\langle k|$, where $\vec{\sigma} = (\sigma_x,\sigma_y,\sigma_z)$, $\mathbf{n}_{\mathcal{H}}(k)$ represents the eigenvector of the Bloch Hamiltonian $\mathcal{H}_{\rm eff}(k)$, and $E(k)$ is the quasi-energy dispersion. The topological phase diagram of QWs [33] as classified by the winding number $\mathcal{W}$ is shown in Fig. 1(a). We consider the quench strategy shown by the black arrow in Fig. 1(a): the system initially stays at the ground state of a trivial Hamiltonian $\mathcal{H}^i_{\rm eff}$, and the control parameters $\{\theta_1,\theta_2\}$ are changed with velocity v to quench the trivial Hamiltonian to reach a topologically non-trivial one $\mathcal{H}^f_{\rm eff}$.

In purely coherent quench scenario, while the time-dependent Hamiltonian keeps all symmetries, the unitary evolution breaks the time-reversal symmetry (TRS) and the chiral symmetry of the time-evolving state [18]. However, the particle-hole symmetry of the state is preserved, so the symmetry-protected topological invariant remains quantized. As shown by the red line in Fig. 1(b), the Berry phase of the time-evolving wavefunction is $\Phi_{\rm B}(t) = \int_{-\pi}^{\pi} dk\langle\psi_k(t)|i\partial_k|\psi_k(t)\rangle \equiv 0$. Nonetheless, a bulk response (blue line), namely MCD $\mathcal{C} = \frac{1}{\pi}\int_{-\pi}^{\pi} dk\langle\psi_k|\Gamma(i\partial_k)|\psi_k\rangle$ [28], dynamically builds up. Since the parameter quench between different topological phases inevitably goes through a critical time t_c with gap closing, the dynamics is nonadiabatic, and the coherent superposition of the two energy bands leads to persistent oscillations of MCD. In marked contrast with the zero Berry phase of the instantaneous state at all times, the non-equilibrium MCD can only be quantized in a long-time average of oscillations to an integer value determined by the final Hamiltonian [27], i.e., $\overline{\mathcal{C}} = \lim_{\mathcal{N}\to\infty} 1/\mathcal{N}\int_0^{\mathcal{N}} \mathcal{C}(t)dt \to \mathcal{W}_f = 1$.

To obtain a stationary quantized MCD within a finite time, as will be the case in realistic measurement, and to reconcile its topological discrepancy with the state as shown above, we introduce classical noise into QWs to induce dephasing. The time-dependent density matrix of spin $\rho_k(t)$ is governed by the master equation [11]:

$$\partial_t\rho_k = -i[\mathcal{H}_{\rm eff}(k),\rho_k] + \gamma_k[\tilde{\sigma}^z_k\rho_k\tilde{\sigma}^z_k - \rho_k], \qquad (1)$$

where $\tilde{\sigma}^z_k(t) = \mathbf{n}_{\mathcal{H}}(k,t)\cdot\vec{\sigma}$ is the Pauli operator in the eigenbasis of $\mathcal{H}_{\rm eff}(k,t)$, and γ_k defines the decay rate of the phase coherence for each k. The noisy term dynamically decays the coherent superposition of the ground state and excited state to be classical and induces a purely dephasing effect that preserves the band occupation. For a completely dephased state, it is diagonal in the eigenbasis of the instantaneous Hamiltonian and takes the form $\rho^{\rm de}_k = \frac{1}{2}[\mathbb{1} + \tilde{n}^z_k(t)\tilde{\sigma}^z_k(t)]$, where $\mathbb{1}$ is the identity matrix. For the two-band QWs, the contributions to MCD from the positive-energy and the negative-energy band are equal and additive [28]. Thus, MCD of the dephased state is $\mathcal{C}_{\rm de} = \frac{1}{\pi}\int_{-\pi}^{\pi} dk{\rm Tr}[\Gamma(i\partial_k)\rho^{\rm de}_k] \equiv 1$, irrespective of the quench velocity v. As shown by the blue line of Fig. 1(c), a stationary quantized MCD reflecting the topology of the instantaneous Hamiltonian arises dynamically by adding noise.

It turns out that the anti-unitary symmetries such as TRS and chiral symmetry [34], broken in the coherently evolving state [18], reappear in the dephased mixed state $\rho^{\rm de}_k$. This is because the eigenstates of the instantaneous Hamiltonian respect all the symmetries, and so does the dephased state as a classical mixture of these eigenstates. Such noise-assisted symmetry recovery plays a vital role in reconciling the topological properties of the mixed state with those of the post-quench Hamiltonian, as shown below.

For topological characterization of the mixed state, we note that since the pioneering work by Uhlmann [35], several complementary approaches have been reported [31,

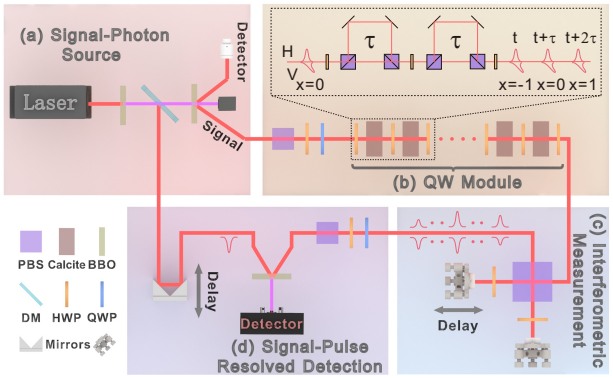


FIG. 2. Photonic experimental setup. The setup mainly includes four parts (see Supplementary Sect. V for detailed designs and descriptions): (a) a spontaneous parametric down-conversion process occurring in BBO generates pairs of photons, in which the idler photons serve as the herald for the signal ones (walker); (b) a QW module that consists of a series of polarization rotation operators and polarization-dependent shift operators; the inset displays the single-step unitary operator; (c) a Michelson interferometer for realizing the interference measurements between different sites; (d) a polarization projection and an up-converted detector for analyzing the details of the signal-photon pulse train. PBS, polarizing beam splitter; BBO: β-BaB_2O_4; DM, dichroic mirror; HWP, half-wave plate; QWP, quarter-wave plate.

[35–41] to generalize the geometric phases and topological invariants to the realm of open systems. For our model system, we find that the temporal evolution of the topologically quantized Uhlmann phase [31] corresponds well to the change of the quantized MCD from a trivial initial to a non-zero post-quench value (cf. Fig. 1). Specifically, with the spectral decomposition of the density matrix $\rho_k = \sum_i p_k^i |\phi_k^i\rangle\langle\phi_k^i|$, the Uhlmann connection is defined as $A_{\rm U} = \sum_{i,j} \frac{|\phi_k^i\rangle\langle\phi_k^i|[\partial_k\sqrt{\rho_k},\sqrt{\rho_k}]|\phi_k^j\rangle\langle\phi_k^j|}{p_k^i+p_k^j} dk$, and the associated Uhlmann phase $\Psi_{\rm U} = \arg{\rm Tr}[\rho_{-\pi} e^{\int_{-\pi}^{\pi} dk A_{\rm U}}]$. In the limit of a rank-1 pure state, $\Psi_{\rm U}$ approaches the corresponding Berry phase $\Phi_{\rm B}$. By contrast, the Uhlmann phase of the dephased state $\rho_k^{\rm de}$ reads as (see Supplementary Sect. IIC): $\Psi_{\rm U}^{\rm de} = \arg(\cos[|\vec{U}|])$, with $|\vec{U}| \approx |\pi\mathcal{W} + \mathcal{O}(v^{\frac{1}{2}})|$ and $\mathcal{W} = 1$ is the winding number of the instantaneous Hamiltonian in the topological phase (cf. Fig. 1(a)). At least for small v, a non-trivial mixed state, with the topologically quantized Uhlmann phase $\Psi_{\rm U}^{\rm de} \equiv \pi$, is generically built up dynamically from a trivial initial state, as shown by the red line in Fig. 1(c). That way, the topological discrepancy between time-dependent Hamiltonian and state, characteristic of the coherent case, is reconciled in noisy QWs.

Experimental implementation.—Experimentally, we realize photonic QWs with a setup shown in Fig. 2, where the heralded signal photons act as the walker. We encode the internal spin states $\{|\uparrow\rangle, |\downarrow\rangle\}$ in the horizontal and vertical polarization of photons $\{|H\rangle, |V\rangle\}$, and the lattice sites $|x\rangle$ are composed of photon arrival times (see the inset of Fig. 2(b)). The single-step unitary operator of QWs U is composed of three HWPs (for spin rotation $R_y(\theta)$) and two calcite crystals (for shift operator $S_{\uparrow,\downarrow}$), with each calcite inducing a time shift $\tau \simeq 5$ ps between the two polarizations. Then, the lattice space of the walker after the QWs module consists of the superposition of $2\mathcal{N}+1$ time bins with a time interval τ.

The key new ingredient in our setup is a Michelson interferometer with controllable arm length difference, used to realize the full quantum state tomography and thus detect the state topology. The interferometer in Fig. 2(c) consists of a PBS, two mirrors, and two QWPs rotated to an angle at 45°. The movable mirror is controlled to introduce a time difference between the two arms, so that the photons with horizontal polarization travel integer multiples of 5 ps faster (or later) than those with vertical polarization. After the interferometer, both the local polarization states and the arrival times of signal photons are analyzed in Fig. 2(d). Finally, interferometric measurements between site x and site $x+i$ ($i = 0, \pm1, \cdots, \pm\mathcal{N}$) are implemented, and a collection of $2\mathcal{N}+1$ sets of photon counts can be obtained. These interferometric measurements can be seen as base transfor-

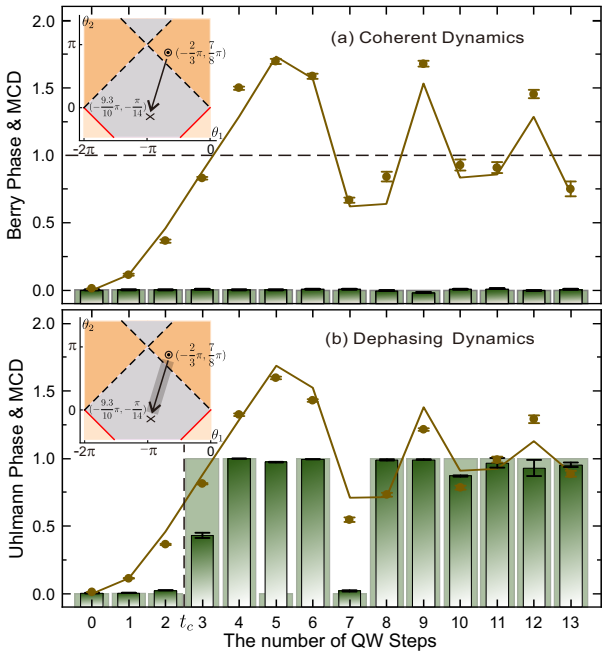


FIG. 3. Experimental bulk response and state topology. (a) Measured MCD (orange dots) and Berry phase (green opaque bars) of the time-evolving wavefunction in coherent quenched QWs. (b) Measured MCD (orange dots) and Uhlmann phase (green opaque bars) of the time-evolving mixed state in noisy QWs. The yellow solid lines and the green transparent bars in (a-b) represent the theoretical predictions of the time-evolving MCD and geometric phase, respectively. The errors are estimated using numerical Monte Carlo simulations considering photon-counting noise.

mations and projecting the system onto the sets of bases $\{|H\rangle\otimes|x\rangle\}$, $\{|V\rangle\otimes|x\rangle\}$, $\{\frac{1}{\sqrt{2}}(|H\rangle\otimes|x\rangle\pm i|V\rangle\otimes|x'\rangle)\}$, and $\{\frac{1}{\sqrt{2}}(|H\rangle\otimes|x\rangle\pm|V\rangle\otimes|x'\rangle)\}$ $(x,x'=0,\pm1,\cdots,\pm\mathcal{N})$.

Observation of topological discrepancy.—We first experimentally validate the topological discrepancy in coherent quenched QWs. The system is initialized in the ground state of the lower energy band of a trivial Hamiltonian $\mathcal{H}^i_{\rm eff}$ with $(\theta^i_1=-2\pi/3,\theta^i_2=7\pi/8)$. This is prepared via the adiabatic evolution from a localized state $|x=0\rangle\otimes\frac{1}{\sqrt{2}}(|H\rangle-i|V\rangle)$, which is the ground state of the trivial flat-band Hamiltonian ($\theta_2=\pi$) with degenerate quasi-energy $E(k)\equiv-\pi$ [42]. Then, the control parameters $\{\theta_1,\theta_2\}$ are varied dynamically to drive the trivial Hamiltonian towards a topologically non-trivial one $\mathcal{H}^f_{\rm eff}$ with $(\theta^f_1=-9.3\pi/10,\theta^f_2=-\pi/14)$.

To monitor the dynamics of MCD, $\mathcal{C}(t)=2\langle\Gamma X\rangle=2\sum_x x[P_{\uparrow_\Gamma}(x,t)-P_{\downarrow_\Gamma}(x,t)]$ [27], we measure the two projection probability distributions $P_{\{\uparrow_\Gamma,\downarrow_\Gamma\}}(x)$ for each time-step t, on the basis of the chiral operator $\{\uparrow_\Gamma,\downarrow_\Gamma\}=\{\frac{1}{\sqrt{2}}(|H\rangle\pm|V\rangle)\}$. As shown by the orange dots in Fig. 3(a), the non-equilibrium MCD displays a marked oscillation around an integer value of 1 determined by the topological Hamiltonian $\mathcal{H}^f_{\rm eff}$. To obtain the Berry phase of the time-evolving state $\Phi_{\rm B}(t)$, we use discrete-time-resolved wavefunction tomography [43] performed on the collection of $2\mathcal{N}+1$ sets of photon counts, and reconstruct the spinor states $|\psi_k(t)\rangle$ for each k in the first Brillouin zone with the Fourier transform. We observe $\Phi_{\rm B}(t)\simeq0$ throughout the coherent evolution, as shown by the green opaque bars in Fig. 3(a), with an obvious discrepancy to the finite MCD.

Observing dynamical topological quantization.—To engineer the type of noise that induces dephasing in QWs, we utilize a slowly fluctuating environment [44, 45], with the time scale of the control parameters changing larger than a single QW realization. Under this environment, the parameter θ_1 is randomly selected on the interval $[\theta_1-\delta\theta_1,\theta_1+\delta\theta_1]$ for a single realization, and the final result is the ensemble average over the multiple realizations. The fluctuating degree $\delta\theta_1$ here plays a similar effect as the decay rate γ_k (see Supplementary Sect. IIIB for details). In the experiment, we have $M=21$ random settings for the parameter θ_1 with $\delta\theta_1=0.2$. Then, the measured probability distributions in noisy QWs are the ensemble average over the 21 samples.

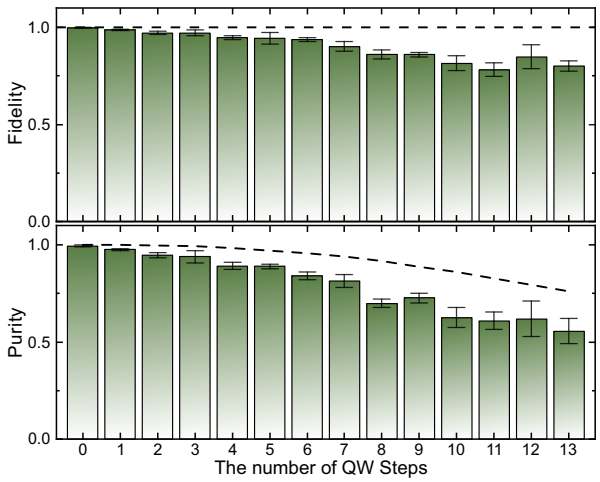


FIG. 4. Experimental density-matrix tomography. (a) Fidelity $(\mathrm{Tr}\sqrt{\sqrt{\varrho}\varrho_{\boldsymbol{\lambda},\boldsymbol{\mu}}\sqrt{\varrho}})^2$ and (b) Purity $\mathrm{Tr}[\varrho^2_{\boldsymbol{\lambda},\boldsymbol{\mu}}]$ (green bars) of the NDO reconstructed mixed states for noisy QWs, given the experimental photon-counting collection. The black dashed lines represent the theoretical predictions for the exact density matrices.

In this noisy scenario, however, the full density-matrix reconstruction was unusual: the real-space system is generally described by a density operator ϱ with the dimension of $4(2\mathcal{N}+1)^2$ after a $\mathcal{N}$-step walk; theoretically, the photon-counting collection of the interferometric measurements, complete for tomography of rank-1 pure state, is incomplete for directly reconstructing ϱ. Here we utilize an effective neural-network quantum state tomography [46], where a neural density operator (NDO) approximates the target density matrix $\varrho_{\boldsymbol{\lambda},\boldsymbol{\mu}}\simeq\varrho$ after training the network parameters $\{\boldsymbol{\lambda},\boldsymbol{\mu}\}$ with given the $2\mathcal{N}+1$ sets of photon counts (see Supplementary Sect. IV for details). Then the reduced density matrix of spin ρ_k in quasi-momentum space can be obtained via Fourier transform on $\varrho_{\boldsymbol{\lambda},\boldsymbol{\mu}}$ and taking the trace over k.

Thanks to the unique ability of our photonic QWs, we can now directly probe the topological properties of the time-dependent mixed states that are characterized by the Uhlmann phase $\Phi_{\rm U}$. The green opaque bars in Fig. 3(b) show that during the noisy quench dynamics, the Uhlmann phase starts with an initial trivial value of 0 and then experiences an unstable time region where the phase coherence gradually disappears. The appearance of the unstable region depends on the quench velocity v determined by the total QWs step $\mathcal{N}$ (see Supplementary Sect. IIIC for further discussion). Finally, a topologically non-trivial mixed state with $\Phi^f_{\rm U}/\pi=0.951\pm0.016$, close to the quantized value of 1, is found to build up, compatible with the underlying topological Hamiltonian $\mathcal{H}^f_{\rm eff}$. Meanwhile, in contrast with the coherent scenario, we observe that the amplitude of the periodic oscillations of MCD is significantly dampened upon adding noise, as shown by the orange dots in Fig. 3(b). The MCD of the final state $\mathcal{C}_f=0.885\pm0.022$ approaches the quantized topological number of $\mathcal{H}^f_{\rm eff}$. The experimentally observed deviation ~0.1 from the quantized $\mathcal{C}_{\rm de}\equiv1$ is mainly because the coherence on the eigenstates of $\mathcal{H}^f_{\rm eff}$ cannot be entirely destroyed by the 13-step noisy dynamics, compared to the numerical results for the 100-step scenario (blue circles in Fig. 1(c)). However, the onset of a damping of the persistent oscillations found in the coherent limit is clearly visible even within the currently available

experimental resources.

As shown in Fig. 4, extra uncontrollable decoherence posed by realistic experimental environments degrades the quality of QWs with the increasing number of steps $\mathcal{N}$. First, the calibration error of optical elements scaling linearly with $\mathcal{N}$ causes an unexpected phase shift between the multiple interferometers in the QWs module. Second, because the number of local measurements scales quadratically with $\mathcal{N}$, the measurement error, especially for low-probability sites, induces a considerable offset in the state reconstruction. These two main sources of error limit the total number of practically viable steps in our experiment.

Conclusion.—To summarize, we have developed and implemented photonic QWs capable of full density-matrix reconstruction which has been achieved effectively by neural-network tomography. In this framework, we have demonstrated how introducing classical noise enables the dynamical buildup of a topologically non-trivial mixed state characterized by a non-zero-quantized Uhlmann phase. The general mechanism of restoring dynamically broken symmetries of the time-evolving states by noise is applicable to other symmetry classes of topological insulators and superconductors out of equilibrium [19]. In addition, the experimentally observed agreement between the MCD as a response property and the Uhlmann phase as a topological property of a mixed state here reconciles the topological properties of the time-dependent Hamiltonian and the mixed state in the noisy dynamics of a quenched QW.

Acknowledgements.—This work was supported by Innovation Program for Quantum Science and Technology (No. 2021ZD0301200), National Natural Science Foundation of China (Nos. 12022401, 62075207, 11874343, 12104433, 11821404, 12204468, 11874038), the Fundamental Research Funds for the Central Universities (No. WK2470000030), the CAS Youth Innovation Promotion Association (No. 2020447) and China Postdoctoral Science Foundation (No. 2021M703108). JCB acknowledges financial support from the German Research Foundation (DFG) through the Collaborative Research Centre SFB 1143 (Project No. 247310070), the Cluster of Excellence ct.qmat (Project No. 390858490), and the DFG Project 419241108.

* These authors contributed equally.
† xuxiaoye@ustc.edu.cn
‡ smhan@ustc.edu.cn
§ jan.budich@tu-dresden.de
¶ cfli@ustc.edu.cn

[1] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Quantized hall conductance in a two-dimensional periodic potential,* Phys. Rev. Lett. **49**, 405–408 (1982).
[2] K. von Klitzing, *The quantized hall effect,* Rev. Mod. Phys. **58**, 519–531 (1986).
[3] T. L. Hughes, E. Prodan, and B. A. Bernevig, *Inversion-symmetric topological insulators,* Phys. Rev. B **83**, 245132 (2011).
[4] D. T. Tran, A. Dauphin, A. G. Grushin, P. Zoller, and N. Goldman, *Probing topology by heating: Quantized circular dichroism in ultracold atoms,* Sci. Adv. **3**, e1701207 (2017).
[5] M. Jamotte, L. P. Gavensky, C. M. Smith, M. Di Liberto, and N. Goldman, *Quantized valley hall response from local bulk density variations,* arXiv preprint arXiv:2212.14054 (2022).
[6] C.-K. Chiu, J. C. Y. Teo, A. P. Schnyder, and S. Ryu, *Classification of topological quantum matter with symmetries,* Rev. Mod. Phys. **88**, 035005 (2016).
[7] B. Mera, A. Zhang, and N. Goldman, *Relating the topology of Dirac Hamiltonians to quantum geometry: When the quantum metric dictates Chern numbers and winding numbers,* SciPost Phys. **12**, 018 (2022).
[8] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, *Topological quantum computation,* Bull. Amer. Math. Soc. **40**, 31–38 (2003).
[9] T. Dai, Y. Ao, J. Bao, J. Mao, Y. Chi, Z. Fu, Y. You, X. Chen, C. Zhai, B. Tang, *et al.*, *Topologically protected quantum entanglement emitters,* Nat. Photon. **16**, 248–257 (2022).
[10] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 3rd ed. (Butterworth-Heinemann, New York, 1980).
[11] Y. Hu, P. Zoller, and J. C. Budich, *Dynamical buildup of a quantized hall response from nontopological states,* Phys. Rev. Lett. **117**, 126803 (2016).
[12] S. Barbarino, J. Yu, P. Zoller, and J. C. Budich, *Preparing atomic topological quantum matter by adiabatic nonunitary dynamics,* Phys. Rev. Lett. **124**, 010401 (2020).
[13] J. H. Wilson, J. C. W. Song, and G. Refael, *Remnant geometric hall response in a quantum quench,* Phys. Rev. Lett. **117**, 235302 (2016).
[14] N. Fläschner, D. Vogel, M. Tarnowski, B. S. Rem, D. S. Lhmann, M. Heyl, J. C. Budich, L. Mathey, K. Sengstock, and C. Weitenberg, *Observation of dynamical vortices after quenches in a system with topology,* Nat. Phys. **14**, 265–268 (2017).
[15] C.-R. Yi, L. Zhang, L. Zhang, R.-H. Jiao, X.-C. Cheng, Z.-Y. Wang, X.-T. Xu, W. Sun, X.-J. Liu, S. Chen, and J.-W. Pan, *Observing topological charges and dynamical bulk-surface correspondence with ultracold atoms,* Phys. Rev. Lett. **123**, 190603 (2019).
[16] T. Mizoguchi, Y. Kuno, and Y. Hatsugai, *Detecting bulk topology of quadrupolar phase from quench dynamics,* Phys. Rev. Lett. **126**, 016802 (2021).
[17] L. Zhang, W. Jia, and X.-J. Liu, *Universal topological quench dynamics for z2 topological phases,* Sci. Bull. **67**, 1236–1242 (2022).
[18] M. McGinley and N. R. Cooper, *Topology of one-dimensional quantum systems out of equilibrium,* Phys. Rev. Lett. **121**, 090401 (2018).
[19] M. McGinley and N. R. Cooper, *Classification of topological insulators and superconductors out of equilibrium,* Phys. Rev. B **99**, 075148 (2019).
[20] G. H. Reid, M. Lu, A. R. Fritsch, A. M. Piñeiro, and I. B. Spielman, *Dynamically induced symmetry breaking*

and out-of-equilibrium topology in a 1d quantum system, Phys. Rev. Lett. **129**, 123202 (2022).
[21] L. D'Alessio and M. Rigol, *Dynamical preparation of floquet chern insulators,* Nat. Commun. **6**, 8336 (2015).
[22] M. D. Caio, N. R. Cooper, and M. J. Bhaseen, *Quantum quenches in chern insulators,* Phys. Rev. Lett. **115**, 236403 (2015).
[23] M. D. Caio, N. R. Cooper, and M. J. Bhaseen, *Hall response and edge current dynamics in chern insulators out of equilibrium,* Phys. Rev. B **94**, 155104 (2016).
[24] P. Wang and S. Kehrein, *Phase transitions in the diagonal ensemble of two-band chern insulators,* New J. Phys. **18**, 053003 (2016).
[25] A. Kruckenhauser and J. C. Budich, *Dynamical equilibration of topological properties,* Phys. Rev. B **98**, 195124 (2018).
[26] M. D. Caio, G. Möller, N. R. Cooper, and M. Bhaseen, *Topological marker currents in chern insulators,* Nat. Phys. **15**, 257–261 (2019).
[27] F. Cardano, A. D'Errico, A. Dauphin, M. Maffei, B. Piccirillo, C. de Lisio, G. De Filippis, V. Cataudella, E. Santamato, L. Marrucci, M. Lewenstein, and P. Massignan, *Detection of zak phases and topological invariants in a chiral quantum walk of twisted photons,* Nat. Commun. **8**, 15516 (2017).
[28] M. Maffei, A. Dauphin, F. Cardano, M. Lewenstein, and P. Massignan, *Topological characterization of chiral models through their long time dynamics,* New J. Phys. **20**, 013023 (2018).
[29] E. J. Meier, F. A. An, A. Dauphin, M. Maffei, P. Massignan, T. L. Hughes, and B. Gadway, *Observation of the topological anderson insulator in disordered atomic wires,* Science **362**, 929–933 (2018).
[30] A. D'Errico, F. Di Colandrea, R. Barboza, A. Dauphin, M. Lewenstein, P. Massignan, L. Marrucci, and F. Cardano, *Bulk detection of time-dependent topological transitions in quenched chiral models,* Phys. Rev. Res. **2**, 023119 (2020).
[31] O. Viyuela, A. Rivas, and M. A. Martin-Delgado, *Uhlmann phase as a topological measure for one-dimensional fermion systems,* Phys. Rev. Lett. **112**, 130401 (2014).
[32] T. Kitagawa, M. S. Rudner, E. Berg, and E. Demler, *Exploring topological phases with quantum walks,* Phys. Rev. A **82**, 033429 (2010).
[33] J. K. Asbóth and H. Obuse, *Bulk-boundary correspondence for chiral symmetric quantum walks,* Phys. Rev. B **88**, 121406 (2013).
[34] Chiral symmetry is an anti-unitary symmetry in Fock space that may also be formulated as an anti-commuting unitary constraint at first-quantized level [6].
[35] A. Uhlmann, *Parallel transport and quantum holonomy along density operators,* Rep. Math. Phys. **24**, 229–240 (1986).
[36] E. Sjöqvist, A. K. Pati, A. Ekert, J. S. Anandan, M. Ericsson, D. K. L. Oi, and V. Vedral, *Geometric phases for mixed states in interferometry,* Phys. Rev. Lett. **85**, 2845–2849 (2000).
[37] Z. Huang and D. P. Arovas, *Topological indices for open and thermal systems via uhlmann's phase,* Phys. Rev. Lett. **113**, 076407 (2014).
[38] J. C. Budich and S. Diehl, *Topology of density matrices,* Phys. Rev. B **91**, 165140 (2015).
[39] B. Mera, C. Vlachou, N. Paunković, and V. R. Vieira, *Uhlmann connection in fermionic systems undergoing phase transitions,* Phys. Rev. Lett. **119**, 015702 (2017).
[40] C.-E. Bardyn, L. Wawer, A. Altland, M. Fleischhauer, and S. Diehl, *Probing the topology of density matrices,* Phys. Rev. X **8**, 011035 (2018).
[41] Y. He and C.-C. Chien, *Uhlmann holonomy against lindblad dynamics of topological systems at finite temperatures,* Phys. Rev. B **106**, 024310 (2022).
[42] X.-Y. Xu, Q.-Q. Wang, M. Heyl, J. C. Budich, W.-W. Pan, Z. Chen, M. Jan, K. Sun, J.-S. Xu, Y.-J. Han, C.-F. Li, and G.-C. Guo, *Measuring a dynamical topological order parameter in quantum walks,* Light: Sci. Appl. **9**, 7 (2020).
[43] X.-Y. Xu, Q.-Q. Wang, W.-W. Pan, K. Sun, J.-S. Xu, G. Chen, J.-S. Tang, M. Gong, Y.-J. Han, C.-F. Li, and G.-C. Guo, *Measuring the winding number in a large-scale chiral quantum walk,* Phys. Rev. Lett. **120**, 260501 (2018).
[44] A. Schreiber, K. N. Cassemiro, V. Potoček, A. Gábris, I. Jex, and C. Silberhorn, *Decoherence and disorder in quantum walks: From ballistic spread to localization,* Phys. Rev. Lett. **106**, 180403 (2011).
[45] H. Tang, L. Banchi, T.-Y. Wang, X.-W. Shang, X. Tan, W.-H. Zhou, Z. Feng, A. Pal, H. Li, C.-Q. Hu, M. S. Kim, and X.-M. Jin, *Generating haar-uniform randomness using stochastic quantum walks on a photonic chip,* Phys. Rev. Lett. **128**, 050503 (2022).
[46] G. Torlai and R. G. Melko, *Latent space purification via neural density operators,* Phys. Rev. Lett. **120**, 240503 (2018).

Device-independent entanglement quantification in the presence of losses

Shiladitya Mal[1 2 *] Ching-Hsu Chen[3] Pei-Sheng Lin[2] Chellasamy Jebarathinam[4]
Yeong-Cherng Liang[2 1 †]

[1] *Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*
[2] *Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan*
[3] *Department of Electrophysics, National Chiayi University, Chiayi 600, Taiwan*
[4] *Department of Physics and Center for Quantum Information Science, National Cheng Kung University, Tainan 70101, Taiwan*

Abstract. The basis of device-independent (DI) certification is a conclusive Bell test. Thus, as with a Bell test, losses have a negative impact on the possibility of performing realistic DI certification. Specifically, in this work, we investigate how losses affect the possibility of performing the task of DI entanglement quantification, as measured by negativity in a bipartite setting and entanglement depth in a tripartite setting. For the former, we obtain a tradeoff between the amount of CHSH Bell violations and detection efficiency for certifying any given amount of negativity. Whereas for the latter, we obtain upper bounds on the threshold efficiency required for certifying (genuine tripartite) entanglement via Sliwa's inequalities.

Keywords: Entanglement, device-independence, loophole, Bell test, multipartite entanglement

In pursuing new-age quantum technology, quantum entanglement [17] has been established as a valuable resource for many tasks, such as secure communication [14], fast computation [38, 18], and the generation of provable random bits [1]. Typically, entanglement certification is carried out assuming the state space dimension and the exact form of the measurements performed. These assumptions are, however, not necessarily justifiable in a realistic experimental scenario, thus making the conclusions drawn questionable [30, 25]. An elegant way to get around this problem is to use a Bell test [3, 2] and draw conclusions directly from the observed statistics showing Bell-nonlocality [5], i.e., to perform so-called device-independent (DI) entanglement *certification*. In fact, even entanglement *quantification* based on entanglement monotones [37] can be carried out in a DI manner [24, 20, 34, 10].

Of course, a DI certification or quantification has to meet its own challenges. For example, since it is based on a Bell test, it is subjected to [5] so-called locality or detection loopholes, which were only closed simultaneously relatively recently [16, 31, 15, 29]. In this work, we focus on the latter, i.e., how losses in a Bell test impact the task of DI entanglement quantification based on negativity [39] and entanglement depth [33]. Indeed, if the detection efficiency is too low, a common-cause strategy that includes the instructions of a no-detection could fake the violation of a Bell inequality when we only consider the post-selected events [5]. For example, in the simplest Bell scenario involving two parties and two dichotomic measurements each, also called the Clauser-Horne-Shimony-Holt [11] (CHSH) Bell scenario, a symmetric detection efficiency larger than $2\sqrt{2}-2 \approx 83\%$ is required for a conclusive Bell test. More generally, a minimum detection efficiency of $\frac{2}{3} \approx 67\%$ [13] is required to close the detection loophole in the CHSH Bell scenario.

To determine the impact of losses on DI certification, we follow the approach of [4] and analyze formally how postselection affects the various sets of correlation. To this end, the no-detection event $\emptyset$ is formally recognized as an additional outcome of a Bell experiment. Specifically, in a bipartite Bell test, if we denote by $P_0(a,b|x,y)$ the *a priori* conditional probability of Alice (Bob) observing the outcome a (b) given that she (he) performs the x- (y)-th measurement, then Alice's (respectively, Bob's) detection efficiency η_A (η_B) are defined such that:

$$P_0(a\neq\emptyset,b|x,y) \equiv \sum_{a\neq\emptyset} P_0(a,b|x,y) = \eta_A P_0(b|y) \quad \forall\, b,x,y,$$
$$P_0(b\neq\emptyset,a|x,y) \equiv \sum_{b\neq\emptyset} P_0(a,b|x,y) = \eta_B P_0(a|x) \quad \forall\, a,x,y. \tag{1}$$

Clearly, physically relevant detection efficiencies η_A, η_B are those bounded between 0 and 1.

Accordingly, the postselected set of correlation is obtained from the a priori correlations by

$$P_{\text{ps}}(a,b|x,y) = P(a,b|x,y,a\neq\emptyset,b\neq\emptyset) = \frac{P_0(a,b|x,y)}{\eta_A\eta_B}. \tag{2}$$

As was shown in [4], a direct consequence of Eq. (7) is that if there are losses, i.e., when $\eta_A,\eta_B<1$, the set $\mathcal{L}_{\text{ps}}(\eta_A,\eta_B)$ of $P_{\text{ps}}(a,b|x,y)$ that admits a local-hidden-variable description strictly contains the usual set of Bell-local correlations $\mathcal{L}_0 \equiv \mathcal{L}_{\text{ps}}(1,1)$ with no losses. For simplicity, we consider hereafter only a situation with symmetric detection efficiencies, namely, when $\eta_A=\eta_B=\eta$.

Then, in our work, we show that the same observation holds for various sets of correlations relevant for DI certification. Examples of which include the set of correlations generated from a quantum state having an amount of negativity bounded by N (which we denote by $\mathcal{N}_\eta^{\leq N}$ when the efficiency is η) and those generated from an n-partite quantum state that is k-producible (we denote the corresponding set by $\mathcal{Q}_\eta^{n,k}$). As an illustration, we show in Fig. 3 how $\mathcal{N}_{\eta=0.95}^{\leq 0.25}$ has expanded compared with $\mathcal{N}_{\eta=1}^{\leq 0.25}$ when we consider **the first level** of the semidefinite programming (SDP) hierarchy introduced in [24].

*shiladitya.27@gmail.com
†ycliang@mail.ncku.edu.tw

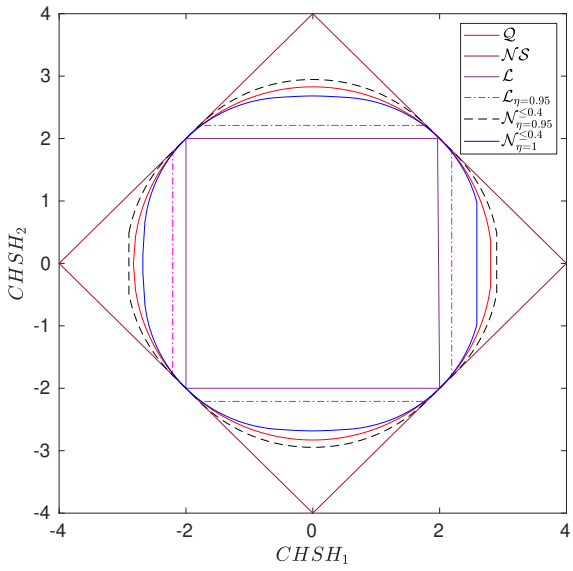


Figure 1: A projection plot showing how postselections affects the various sets of correlations assuming a symmetric detection efficiency of $\eta = 95\%$. Moving outward, we have the Bell polytope $\mathcal{L}$ (violate), the postselected local polytope $\mathcal{L}_{\text{PS}}$ (magenta), an outer approximation of the set $\mathcal{N}^{\leq 0.4}_{\eta=1}$ of quantum correlations having a negativity upper bounded by 0.4 (blue), the set $\mathcal{Q}$ of quantum correlations (red), an outer approximation of the set $\mathcal{N}^{\leq 0.4}_{\eta=0.95}$ of postselected quantum correlations having a negativity upper bounded by 0.4 (black), and the no-signaling polytope (brown). On the horizontal axis and the vertical axis, we have, respectively, the value of the Bell expression $\text{CHSH}_1 \equiv \sum_{x,y=0}^{1}(-1)^{(x+1)(y+1)}E(A_xB_y)$ and $\text{CHSH}_2 \equiv \sum_{x,y=0}^{1}(-1)^{xy}E(A_xB_y)$ where $E(A_xB_y)$ is the expectation value of the product of the ± 1 measurement outcomes. All outer approximations were computed using level 1 of the SDP hierarchy proposed in [24].

To determine the minimum value of η that one needs to certify a certain amount of negativity (in the CHSH Bell scenario), we incorporated the constraints of Eq. (6) and Eq. (7) into the SDP hierarchy introduced in [24] for lower-bounding the underlying negativity from any given correlation. In the presence of losses, we take the postselected correlations P_{ps} and the resulting CHSH value $\mathcal{S}_{\text{CHSH}}$ as our observable quantity, but in determining the underlying negativity, we use the a priori correlation P_0, which is related to the former via Eq. (7). The corresponding results are shown in Fig. 4.

As a second example of DI entanglement quantification, we consider the tripartite Bell scenario where each party can perform two binary-outcome measurements. Here, we are interested to lower bound the entanglement depth [33] of the underlying system. When there is no loss, this is known to be possible from the strength of the violation of certain tripartite Bell inequalities, as exemplified in [20]. For example, the

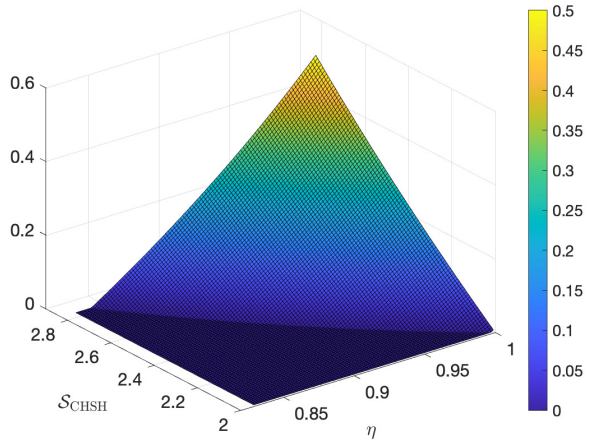


Figure 2: Three-dimensional surface plot showing a lower bound on the amount of negativity certifiable for any detection efficiency η and any given violation $\mathcal{S}_{\text{CHSH}}$ of the CHSH Bell inequality. The plot consists of 100^2 points formed by 100 values of $\mathcal{S}_{\text{CHSH}}$ sampled uniformly between 2 and $2\sqrt{2}$ and 100 values of η sampled uniformly between $2(\sqrt{2}-1)$ and 1. Here, $\mathcal{S}_{\text{CHSH}}$ could be any of the CHSH Bell expression obtained from CHSH_1 via a relabeling. All computations were done using level 2 of the SDP hierarchy proposed in [24].

tripartite Mermin Bell inequality [23] $\mathcal{I}_2$ reads as

$$-2 \overset{\mathcal{L}_0}{\leq} S_2 \equiv E(A_1B_1C_1) + E(A_2B_2C_1) + E(A_2B_1C_2) - E(A_1B_2C_2) \overset{\mathcal{L}_0}{\leq} 2 \tag{3}$$

where $E(A_xB_yC_z)$ is the tripartite expectation value. If we observe a quantum violation $|S_2| > 2$, we can immediately conclude that the state is entangled; if $|S_2| > 2\sqrt{2}$, we can even conclude that the underlying state is not 2-producible, and hence having an entanglement depth of 3, i.e., exhibiting genuine tripartite entanglement.

In this regard, it is known [32] that the complete set of facet-defining Bell inequalities in this Bell scenario falls into 46 classes $\{\mathcal{I}_i\}_{i=1}^{46}$ after taking into account the freedom of relabeling [the Mermin inequality of Eq. (23) is the 2nd among these 46]. Moreover, the corresponding 2-producible quantum bounds were determined in [35]. Again, with the help of the SDP hierarchy from [24] and equations analogous to those given in Eq. (6) and Eq. (7) for the tripartite scenario, we can determine, for any given symmetric detection efficiency η and any given Bell inequality $\mathcal{I}_i$, an upper bound on the corresponding *maximal* quantum value of S_i for the postselected local $\mathcal{L}_{\text{ps}}$ and postselected 2-producible $\mathcal{Q}_\eta^{3,2}$ correlations. In turn, this allows us to determine an *upper bound* on the minimal detection efficiency required to violate the local bound or the 2-producible bound for any given Bell inequality in this simplest tripartite Bell scenario. Our results pertaining to the upper bound of each S_i [e.g., the local upper bound of 2, or the 2-producible upper bound of $2\sqrt{2}$ of S_2 in Eq. (23)] are shown in Table 3. Accordingly, our results pertaining to the lower bound of each S_i [e.g., the local lower bound of -2, or the 2-producible lower bound of $-2\sqrt{2}$ of S_2 in Eq. (23)] are

S_i	$\eta^*_{k=1}$	$\eta^*_{k=2}$	S_i	$\eta^*_{k=1}$	$\eta^*_{k=2}$
S_1	-	-	S_{24}	80	84.51
S_2	75	75.01	S_{25}	87.51	96.7
S_3	87.87	-	S_{26}	80.18	82.94
S_4	82.84	-	S_{27}	87	94.18
S_5	82.93	95.34	S_{28}	81.28	94.1
S_6	84.42	-	S_{29}	83.61	-
S_7	81.82	88.78	S_{30}	83.8	-
S_8	81.82	87.66	S_{31}	89.2	96.99
S_9	87.2	-	S_{32}	87.1	93.29
S_{10}	-	-	S_{33}	78.8	81.15
S_{11}	86.67	-	S_{34}	86	91.82
S_{12}	85.69	-	S_{35}	88.9	95.60
S_{13}	87.33	-	S_{36}	83.1	98.45
S_{14}	85.35	-	S_{37}	83.98	-
S_{15}	83.55	94.08	S_{38}	83.9	-
S_{16}	83.22	92.78	S_{39}	78.6	81.05
S_{17}	86.4	-	S_{40}	87.6	93.24
S_{18}	84.99	98.13	S_{41}	85.3	99.43
S_{19}	84.95	97.6	S_{42}	79.8	84.98
S_{20}	83.53	-	S_{43}	86.4	-
S_{21}	83.73	94.08	S_{44}	83.4	-
S_{22}	81.28	89.87	S_{45}	83.6	-
S_{23}	93.59	-	S_{46}	89.4	-

Table 1: $\mathcal{S}_i$s in 1st and 4th column denotes the i-th class of Sliwa's inequality bounded from above. $\eta^*_{k=1}$ is an upper bound on the critical efficiency required for faithfully detecting non-separability while $\eta^*_{k=2}$ is an upper bound on the critical efficiency required for detecting non-2-producibility, i.e., an entanglement depth of 3.

shown in Table 4.

Acknowledgements This work is supported by the National Science and Technology Council (formerly Ministry of Science and Technology), Taiwan (Grants No. 107-2112-M-006-005-MY2, 109-2112-M006-010-MY3).

S_i	$\eta^*_{k=1}$	$\eta^*_{k=2}$	S_i	$\eta^*_{k=1}$	$\eta^*_{k=2}$
S_1	-	-	S_{24}	-	-
S_2	75.1	75.01	S_{25}	-	-
S_3	87.9	-	S_{26}	-	-
S_4	82.9	-	S_{27}	-	-
S_5	-	-	S_{28}	82.9	-
S_6	84.1	-	S_{29}	82.9	-
S_7	81.9	88.78	S_{30}	82.9	-
S_8	81.9	87.66	S_{31}	86.7	-
S_9	87.2	-	S_{32}	86.7	-
S_{10}	-	-	S_{33}	-	-
S_{11}	86.7	-	S_{34}	-	-
S_{12}	-	-	S_{35}	83.2	91.93
S_{13}	87.4	-	S_{36}	82.9	-
S_{14}	85.4	-	S_{37}	82.9	-
S_{15}	-	-	S_{38}	82.9	-
S_{16}	83.3	92.78	S_{39}	-	-
S_{17}	86.4	-	S_{40}	-	-
S_{18}	-	-	S_{41}	84.1	-
S_{19}	-	-	S_{42}	78.1	96.98
S_{20}	82.9	-	S_{43}	79.6	-
S_{21}	-	-	S_{44}	82.9	-
S_{22}	-	-	S_{45}	82.9	-
S_{23}	-	-	S_{46}	-	-

Table 2: $\mathcal{S}_i$s in 1st and 4th column denotes the i-th class of Sliwa's inequality bounded from bellow. $\eta^*_{k=1}$ is an upper bound on the critical efficiency required for faithfully detecting non-separability while $\eta^*_{k=2}$ is an upper bound on the critical efficiency required for detecting non-2-producibility, i.e., an entanglement depth of 3.

QAOA-MC: Markov chain Monte Carlo enhanced by Quantum Alternating Operator Ansatz

Yuichiro Nakano[1] * Hideaki Hakoshima[2] † Kosuke Mitarai[1 2] ‡ Keisuke Fujii [1 2 3] §

[1] *Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan.*

[2] *Center for Quantum Information and Quantum Biology, Osaka University, 560-0043, Japan.*

[3] *Center for Quantum Computing, RIKEN, Wako Saitama 351-0198, Japan.*

Abstract. The main advantage of quantum computation is its capability to sample from classically intractable probability distributions. One promising approach to leverage this advantage is the quantum-enhanced Markov chain Monte Carlo (quantum-enhanced MCMC) method, which utilizes outputs from quantum circuits as proposal distributions. In this work, we propose the utilization of the Quantum Alternating Operator Ansatz (QAOA) for quantum-enhanced MCMC and present a strategy to optimize its parameters in order to improve the convergence speed. By optimizing the parameters of the QAOA circuit, we can achieve a quadratic speedup in convergence.

Keywords: MCMC, Quantum-enhanced MCMC, QAOA

1 Introduction

The quantum-enhanced Markov chain Monte Carlo (quantum-enhanced MCMC) method is an algorithm that fully exploits each sampling outcome from Noisy Intermediate-Scale Quantum (NISQ) devices [1]. It uses samples from a quantum circuit as the proposal distribution in the Metropolis-Hastings method [2]. Markov chain Monte Carlo (MCMC) method [3] is a very powerful algorithm that can sample from an arbitrary probability distribution $\pi(\boldsymbol{x})$. The Metropolis-Hastings method is a typical method in MCMC. This procedure is as follows: First, propose the next state according to a proposal distribution $Q(\boldsymbol{x}'|\boldsymbol{x})$. Second, decide whether to accept the proposal based on an acceptance probability

$$A(\boldsymbol{x}'|\boldsymbol{x}) = \min\left(1, \frac{\pi(\boldsymbol{x}')}{\pi(\boldsymbol{x})}\frac{Q(\boldsymbol{x}|\boldsymbol{x}')}{Q(\boldsymbol{x}'|\boldsymbol{x})}\right). \tag{1}$$

If the proposal is rejected, the next state remains the same as the state prior to the proposal. The efficiency of the algorithm is determined by the proposal distribution. Utilizing a quantum computer for the proposal distribution, including those that are difficult to simulate on classical computers, can improve the convergence speed of MCMC compared to existing methods. The quantum-enhanced MCMC uses a distribution defined by a classically intractable quantum state generated by a circuit U as the proposal. By imposing the symmetry constraint $U = U^\top$ on the circuit U, the proposal distribution Q satisfies $Q(\boldsymbol{x}'|\boldsymbol{x}) = |\langle \boldsymbol{x}'| U |\boldsymbol{x}\rangle|^2 = |\langle \boldsymbol{x}'| U |\boldsymbol{x}\rangle|^2 = Q(\boldsymbol{x}|\boldsymbol{x}')$ and the ratio of Q in Eq. (1) is equal to 1. This allows $A(\boldsymbol{x}'|\boldsymbol{x})$ to be efficiently calculated by a classical computer, even though Q cannot be performed efficiently even with quantum computers. The circuit proposed in Ref. [4] is based on the time evolution governed by a time-independent Hamiltonian. However, this circuit still has some issues. Firstly, the implementation of the time evolution on NISQ devices faces the challenge of infeasible circuit depth, which depends on the choice of the time parameter. Secondly, the circuit's parameters are selected heuristically, and the strategy to construct a quantum circuit that improves the convergence speed of MCMC remains unclear.

In this work [5], we propose a new MCMC method called Quantum Alternating Operator Ansatz Monte Carlo (QAOA-MC) based on the quantum-enhanced MCMC. This algorithm utilizes a fixed-depth variational quantum circuit in the form of the so-called Quantum Alternating Operator Ansatz (QAOA) [6] as the proposal distribution. We thereby aim to suppress the increase in circuit depth regardless of the parameter choice. Furthermore, we develop a systematic strategy to optimize the circuit to improve the convergence speed by examining the relationship between the absolute spectral gap [7] and the acceptance rate (AR) [8] of the proposal distribution. Our numerical experiments demonstrate that QAOA-MC achieves a nearly quadratic speed-up in the convergence speed compared to the proposal using the uniform distribution. Our results indicate an acceleration of MCMC using NISQ devices and contribute to promoting the use of current NISQ devices.

2 Our Method: QAOA-MC

We propose a new MCMC method called Quantum Alternating Operator Ansatz Monte Carlo: QAOA-MC. This overview is depicted in Fig. 1(a). More details can be found in Ref. [5].

2.1 Variational quantum circuit

We apply the Quantum Alternating Operator Ansatz (QAOA) [6] to the structure of the circuit that generates MCMC proposals. Specifically, our circuit is defined as follows:

$$U(\boldsymbol{\beta}, \boldsymbol{\gamma}) = V(\boldsymbol{\beta}, \boldsymbol{\gamma})^\top V(\boldsymbol{\beta}, \boldsymbol{\gamma}), \tag{2}$$

*u830977g@ecs.osaka-u.ac.jp

†hakoshima.hideaki.qiqb@osaka-u.ac.jp

‡mitarai.kosuke.es@osaka-u.ac.jp

§fujii.keisuke.es@osaka-u.ac.jp

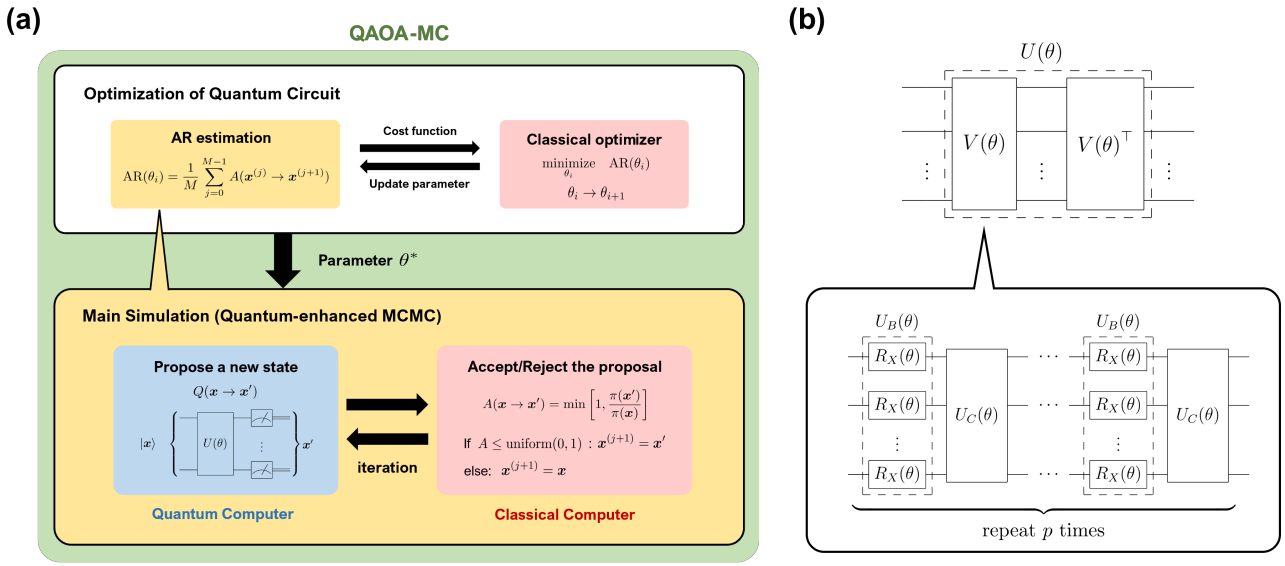


Figure 1: **(a)** The overview of the QAOA-MC algorithm. This algorithm can be summarized as follows. First, we optimize the parameter θ using the AR estimator obtained by MCMC ("Main Simulation"). After obtaining the optimized parameter θ^*, we proceed to the "main simulation" phase. We run the main simulation, using the optimized parameter θ^*. **(b)** The circuit of QAOA-MC. We arrange $V(\theta)$ and $V(\theta)^\top$ in a way that ensures the symmetry constraint: $U = U^\top$ (upper part). The lower part of the circuit $V(\theta)$ consists of two components, $U_B(\theta)$ and $U_C(\theta)$, which are repeated p times.

where

$$V(\boldsymbol{\beta}, \boldsymbol{\gamma}) = U_C(\gamma_p)U_B(\beta_p)\cdots U_C(\gamma_1)U_B(\beta_1), \quad (3)$$

$$U_B(\beta) = \exp(-iH_{\mathrm{mix}}\beta), \quad U_C(\gamma) = \exp(-i\alpha H_{\mathrm{prob}}\gamma). \quad (4)$$

Here, p represents a hyperparameter that determines the circuit depth, and α denotes the normalization factor of H_{prob} with respect to H_{mix}. The n-qubits Hamiltonians H_{mix} and H_{prob} are defined as:

$$H_{\mathrm{mix}} = \sum_{j=1}^{n} X_j, \quad (5)$$

$$H_{\mathrm{prob}} = -\sum_{\langle j,k\rangle} J_{jk} Z_j Z_k - \sum_{j=1}^{n} h_j Z_j, \quad (6)$$

where X_j and Z_j represent the Pauli operators acting on the j-th qubit. H_{prob} represents the target Hamiltonian from which we aim to sample the Boltzmann distribution. This circuit is shown in Fig. 1(b). It should be noted that the circuit defined above satisfies $U = U^\top$ due to its construction.

2.2 Optimization of circuit

The proposal distribution generated by the proposed circuit (Eq. (2)) can be optimized to achieve faster convergence. The objective function should be a computable quantity that reflects the convergence speed of MCMC. After some numerical experiments, we find that the MCMC acceptance rate can be used as the objective function. The acceptance rate (AR) [8] is defined by the following equation:

$$\mathrm{AR} = \sum_{\boldsymbol{x},\boldsymbol{x}'} \pi(\boldsymbol{x})Q(\boldsymbol{x}'|\boldsymbol{x})A(\boldsymbol{x}'|\boldsymbol{x}). \quad (7)$$

This formula includes $\pi(\boldsymbol{x})$, making it difficult to be calculated directly. However, it can be efficiently estimated by performing MCMC on $\pi(\boldsymbol{x})$. We estimate AR using samples generated from M MCMC steps as follows:

$$\mathrm{AR} \approx \frac{1}{M}\sum_{j=0}^{M-1} A(\boldsymbol{x}^{(j+1)}|\boldsymbol{x}^{(j)}), \quad (8)$$

where $\boldsymbol{x}^{(j)}$ represents the state at the j-th step of the MCMC. After experimenting with the Boltzmann distributions for various Ising models, we have discovered a relationship between AR and the absolute spectral gap δ. We now consider a single-parameter circuit $U(\theta)$, where the parameters $\{\boldsymbol{\beta}, \boldsymbol{\gamma}\}$ in Eq. (2) are set as

$$\theta = \beta_1 = \cdots = \beta_p = \gamma_1 = \cdots = \gamma_p.$$

Although there is generally no correlation between AR and δ, a correlation exists for small θ where δ increases as AR decreases. This trend continues until the AR reaches a local minimum, which often corresponds to a local maximum value of δ. Based on these observations, we optimize $U(\theta)$ by searching for a small θ that achieves the locally minimal AR.

3 Numerical Experiments & Discussion

In our numerical experiments, we analyze the absolute spectral gap δ of the MCMC transition probability ma-

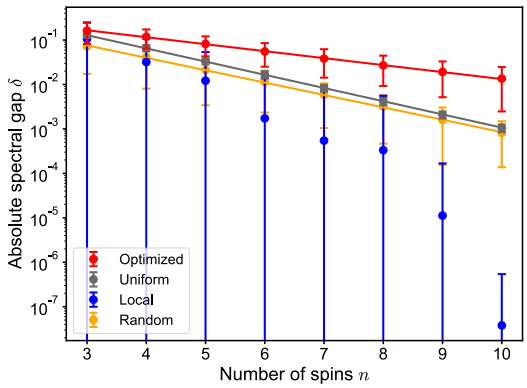


Figure 2: Relationship between model size n and average convergence speed $\langle\delta\rangle$. "Optimized" is QAOA-MC, "Uniform" is the uniform update, "Local" is the local update, and "Random" uses the randomly chosen parameter $\theta \in [0, 2\pi]$ in our circuit.

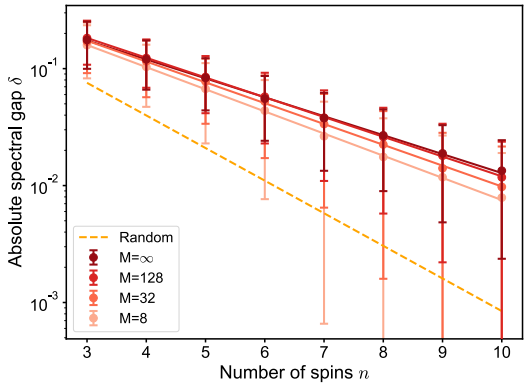


Figure 3: Relationship between model size n and average convergence speed $\langle\delta\rangle$ when M varies. The dotted line is the result of "random" in Fig. 2.

trix $P = [p_{jk}]_{2^n \times 2^n}$, where

$$p_{jk} = \begin{cases} Q(\boldsymbol{x}_j|\boldsymbol{x}_k)A(\boldsymbol{x}_j|\boldsymbol{x}_k) & (j \neq k) \\ 1 - \sum_{j \neq k} Q(\boldsymbol{x}_j|\boldsymbol{x}_k)A(\boldsymbol{x}_j|\boldsymbol{x}_k) & (j = k) \end{cases}. \quad (9)$$

The absolute spectral gap [7] is defined as the absolute difference between the two largest eigenvalues of P. This quantity can serve as the metric for evaluating the convergence speed of MCMC. A larger value of δ corresponds to faster convergence. The target distribution is the Boltzmann distribution for a spin glass, given by

$$\mu(\boldsymbol{x}) = \frac{1}{Z} \exp\left(-\frac{E(\boldsymbol{x})}{T}\right), \quad Z = \sum_{\boldsymbol{x}} \exp\left(-\frac{E(\boldsymbol{x})}{T}\right). \quad (10)$$

The energy function of the spin glass is given by

$$E(\boldsymbol{x}) = -\sum_{j>k=1}^{n} J_{jk} x_j x_k - \sum_{j=1}^{n} h_j x_j, \quad (11)$$

where $x_j \in \{1, -1\}$ is the spin variable of the j-th site.

First, we investigate the performance of QAOA-MC. We generate 500 random instances for each $3 \leq n \leq 10$ and calculate δ for each $\mu(\boldsymbol{x})$. Each instance has random coefficients $\{J_{jk}\}$ and $\{h_j\}$ drawn from a standard normal distribution. In this experiment, we optimize θ using AR calculated precisely using Eq. (7). we compare our proposal to three other proposal distributions: local update, uniform update, and the "random circuit". This "random circuit" corresponds to a distribution defined by our circuit (Eq. (2)) with a randomly chosen parameter $\theta \in [0, 2\pi]$ allowing us to verify the improvement of convergence speed through optimization. Figure 2 shows the relationship between the model size n and the average convergence speed $\langle\delta\rangle$. We fit $\langle\delta\rangle$ by 2^{-kn} with a parameter k and show the result as the straight lines in Fig. 2. QAOA-MC ("optimized") has a scaling factor k approximately 1/1.89 times that of the uniform update, which corresponds to an approximately quadratic acceleration with respect to $\langle\delta\rangle$.

We also examine the impact of MCMC estimation of AR on the performance of QAOA-MC. Since QAOA-MC uses MCMC estimates to obtain AR in practice, the objective function contains statistical errors that could adversely affect the convergence performance. AR is estimated from M samples obtained through MCMC using Eq. (8). We set M to 8, 32, 128, and ∞ (where AR is calculated directly from the target distribution using Eq. (7)) and optimize θ. Figure 3 displays the relationship between M and the resulting $\langle\delta\rangle$. As M becomes smaller, the standard deviation of $\langle\delta\rangle$ increases, and the scaling factor k deteriorates at the same time. This is because decreasing M results in a less accurate AR estimate.

4 Conclusion

In this work, we proposed a new MCMC method called QAOA-MC, based on the quantum-enhanced MCMC. We introduced the use of a QAOA-type circuit to implement the algorithm with shallow circuits and developed a systematic strategy to optimize the circuit to improve the MCMC convergence speed. Numerical experiments demonstrated that QAOA-MC achieved an approximately quadratic speed-up in the absolute spectral gap compared to using a uniform distribution for sampling from the Boltzmann distribution in a spin glass system.

Acknowledgement

This work is supported by MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant No. JPMXS0118067394 and JPMXS0120319794, and JST COI-NEXT Grant No. JPMJPF2014. KM is supported by JST PRESTO Grant No. JPMJPR2019.

References

[1] J. Preskill, "Quantum computing in the NISQ era and beyond", Quantum **2**, 79 (2018).

[2] W. K. Hastings, “Monte carlo sampling methods using markov chains and their applications”, Biometrika **57**, 97–109 (1970).

[3] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of state calculations by fast computing machines”, The Journal of Chemical Physics **21**, 1087–1092 (1953).

[4] D. Layden, G. Mazzola, R. V. Mishmash, M. Motta, P. Wocjan, J.-S. Kim, and S. Sheldon, “Quantum-enhanced markov chain monte carlo”, quant-ph/2203.12497, 10 . 48550 / ARXIV . 2203 . 12497 (2022).

[5] Y. Nakano, H. Hakoshima, K. Mitarai, and K. Fujii, “QAOA-MC: markov chain monte carlo enhanced by quantum alternating operator ansatz”, quant-ph/2305.08789 (2023).

[6] S. Hadfield, Z. Wang, B. O’gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, “From the quantum approximate optimization algorithm to a quantum alternating operator ansatz”, Algorithms **12**, 34 (2019).

[7] D. Levin and Y. Peres, *Markov chains and mixing times* (Oct. 2017).

[8] G. O. Roberts, A. Gelman, and W. R. Gilks, “Weak convergence and optimal scaling of random walk metropolis algorithms”, The Annals of Applied Probability **7**, 110–120 (1997).

Quantum state testing beyond the polarizing regime and quantum triangular discrimination

Yupan Liu[1] *

[1] *Graduate School of Mathematics, Nagoya University*

Abstract. The complexity class Quantum Statistical Zero-Knowledge (QSZK) captures computational difficulties of quantum state testing with respect to the trace distance for efficiently preparable mixed states (Quantum State Distinguishability Problem, QSDP), as introduced by Watrous (FOCS 2002). However, this class faces the same parameter issue as its classical counterpart, because of error reduction for the QSDP (the polarization lemma), as demonstrated by Sahai and Vadhan (JACM, 2003).

In this paper, we introduce quantum analogues of *triangular discrimination*, which is a symmetric version of the χ^2 divergence, and investigate *the quantum state testing problems* for quantum triangular discrimination and quantum Jensen-Shannon divergence (a symmetric version of the quantum relative entropy). These new QSZK-complete problems allow us to improve the parameter regime for testing quantum states in trace distance. Additionally, we prove that the quantum state testing for trace distance with negligible errors is in PP while the same problem without error is in BQP. This indicates that the length-preserving polarization for the QSDP implies that QSZK is in PP. The full version of this paper is available from arXiv:2303.01952.

Keywords: quantum statistical zero-knowledge, polarization lemma, quantum triangular discrimination, quantum state testing

The *quantum state testing* is generally about telling whether one quantum mixed state is close to the other, often referred to as the quantum property testing on the equality of a pair of mixed states. This problem is the quantum generalization of the classical question on testing whether two probability distributions are close, known as the *distribution testing*. These problems typically focus on the number of samples (sample complexity) needed to distinguish two mixed states[1]. However, in this paper, we will concentrate on understanding the computational complexity of these problems when mixed states that has *an efficient description*.

The Quantum State Distinguishability problem (QSDP) with respect to the *trace distance*, first introduced by Watrous [26], is a well-known example in this area of study. It is a crucial computational (promise) problem in both quantum complexity theory and quantum cryptography, and is closely related to the investigation of quantum statistical zero-knowldge (QSZK). The input to this problem consists of the description of two efficient quantum circuits Q_0 and Q_1, which specify two corresponding mixed states ρ_0 and ρ_1[2]. *Yes* instances are those in which the trace distance between the two mixed states is at least α, while *no* instances are those in which the distance is at most β, where $0 \leq \beta < \alpha \leq 1$. Any input quantum circuits that do not fit into either of these categories are considered outside the promise. In this paper, we extend the parameters α and β to efficiently computable functions and denote this parameterized version as (α, β)-QSDP, which is the quantum analogue for the Statistical Difference Problem introduced by Sahai and Vadhan [19].

In [19], error reduction for the QSDP, also known as the *polarization lemma*, demonstrates how to polarize the statistical distance between two distributions. Put it differently, for any constants α and β such that $\alpha^2 > \beta$, the lemma constructs new distributions such that they are either very far apart for *yes* instances or very close for *no* instances, which reduces errors on both sides. The polarization lemma is used to establish the SZK containment of (α, β)-SDP provided $\alpha^2 > \beta$ [19], and an analogue of the direct product lemma for the Hellinger affinity leads to error reduction for StoqMA on *no* instances [16]. Interestingly, this polarization lemma techniques works almost straightforwardly on the trace distance as noted in [26].

Sahai and Vadhan left an open problem of reducing error parameters α and β that do not coincide with the requirements of the polarization lemma technique. Specifically, the parameter regime $\alpha > \beta > \alpha^2$ which is referred to as the *non-polarizing* regime. This issue also arises in the quantum counterpart QSDP. Recently, Berman, Degwekar, Rothblum, and Vasudevan [4] made significant progress in addressing this problem by examining the limitations of existing approaches to polarization. As a result, they extended the SZK containment for SDP beyond the regime of constants α and β where $\alpha^2 - \beta > 0$ as originally stated in the polarization lemma[3].

Theorem 1 (Informal version of [4]) *(α, β)-SDP is in SZK provided that $\alpha^2 - \beta \geq 1/\mathrm{poly}(n)$ or (α, β) in the non-polarizing regime $\alpha > \beta > \alpha^2$ satisfies $\alpha - \beta \geq 1/\mathrm{poly}(n)$ and certain criteria on SD and TD[4].*

*yupan.liu.e6@math.nagoya-u.ac.jp

[1] See [18] for a comprehensive survey on the quantum property testing.

[2] In order to prepare these mixed states, we first apply quantum circuits Q_0 and Q_1 on all-zero states, and then tracing out the ancillary qubits. This will outputs mixed states ρ_0 and ρ_1 respectively.

[3] By inspecting the polarization lemma in [19], SDP is in SZK for $\alpha^2 - \beta \geq 1/O(\log n)$ as elaborated in [4].

[4] To be specific, pairs of efficient Boolean circuits (C_0, C_1) and (C_0', C_1') such pairs of the corresponding distributions (p_0, p_1) and (p_0', p_1') satisfy $\mathrm{SD}(p_0, p_1) > \mathrm{SD}(p_0', p_1') > \mathrm{SD}^2(p_0, p_1)$ and

The proof of Theorem 1 entails a series of ingenious reductions to two distribution testing problems, with respect to the Jensen-Shannon divergence and the triangular discrimination, respectively. The prior works focus on these two divergences since they captured the limitation of two known approaches to polarization. In particular, the original polarization lemma [19] focuses on reducing errors in *yes*-instances and *no*-instances alternately, while the triangular discrimination directly admits the latter. In addition, the entropy extraction approach [11] is essentially based on the Jensen-Shannon divergence[5].

The focus of this work is to address an analogous issue in the quantum scenario by examining the limitations of current techniques on demonstrating QSZK containments, achieved through investigating quantum counterparts of classical divergences related to the problem.

1 Main results

Quantum state testing beyond the polarizing regime. We introduce two quantum state testing problems, the Quantum Jensen-Shannon Divergence Problem (QJSP) and the Measured Quantum Triangular Discrimination Problem (measQTDP). QJSP corresponds to the quantum Jensen-Shannon divergence defined in [17], while measQTDP involves a quantum analogue of the triangular discrimination that will be explained later. The QSZK containments of these problems, as stated in Theorem 2, also result in the improved QSZK containments for the Quantum State Distinguishability problem[6].

Theorem 2 (Improved QSZK containments of QSDP – informal)

(1) (α,β)-QJSP *is in* QSZK *if* $\alpha-\beta \geq 1/\mathrm{poly}(n)$. *Consequently,* (α,β)-QSDP *is also in* QSZK *if* $\alpha^2 - \sqrt{2\ln 2}\beta \geq 1/\mathrm{poly}(n)$.

(2) (α,β)-measQTDP *is in* QSZK *if* $\alpha-\beta \geq 1/O(\log n)$. *This containment further implies that* (α,β)-QSDP *is in* QSZK *for certain instances*[7] *with* $\alpha^2 \leq \beta \leq \alpha$ *and* $\alpha-\beta \geq 1/O(\log(n))$.

In fact, both QJSP and measQTDP are QSZK-complete. The measured quantum triangular discrimination ($\mathrm{QTD}^{\mathrm{meas}}$) exposes the limitation of the quantum polarization lemma technique [19, 26], and its QSZK containment exhibits a natural inverse-logarithmic promise gap. Notably, another quantum analogue QTD does not achieve a similar result. Likewise, the quantum Jensen-Shannon divergence captures the limitation of the quantum entropy extraction approach [3], but our implications on QSZK containments are slightly weaker than the classical counterpart in Theorem 1, as quantum analogues of the triangular discrimination behave differently from the classical equivalent.

Easy regimes for the class QSZK. For the $(1-\epsilon,\epsilon)$-Statistical Difference Problem, if the error parameter ϵ is negligible, then this problem falls into the class PP. SDP instances within these parameter regimes are unlikely to be SZK-hard because of the oracle separation between SZK and PP [5]. We show a similar phenomenon on the $(1-\epsilon,\epsilon)$-Quantum State Distinguishability problem, and these instances are even easier to solve if there is no error, as stated in Theorem 3.

Theorem 3 (Easy regimes for QSZK, informal) *For a negligible error* $\epsilon \leq 2^{-n/2-1}$, $(1-\epsilon,\epsilon)$-$\overline{\mathrm{QSDP}}$ *is in* PP. *Moreover,* $(1,0)$-$\overline{\mathrm{QSDP}}$ *is in* BQP_1 *if there is no error.*

Here the improved SZK-hardness and QSZK-hardness follow from skillfully applying the polarization lemma for the relevant distance to the original proof, see Theorem 3.14 in [4]. We then observe that $\frac{1}{2}\mathrm{HS}^2(\rho_0,\rho_1) = \frac{1}{2}(\mathrm{Tr}(\rho_0^2)+\mathrm{Tr}(\rho_1^2)) - \mathrm{Tr}(\rho_0\rho_1)$. Hence, the remaining results are mainly derived from a *hybrid* algorithm based on the SWAP test [7], namely tossing two random coins and performing the SWAP test on the corresponding states.

In essence, the phenomenon that parameter regimes with negligible errors are easier to solve is not unique to the class QSZK. Analogous phenomena can also be observed in other quantum complexity classes, such as QMA(2) [14] and StoqMA [2]. Nevertheless, it is worth noting that these similar results in other classes do not always necessitate the *length-preserving* property[8]. Considering that SZK is a subset of QSZK, Theorem 3 can be interpreted as an indication that problems that can effectively distinguish quantum states may not remain QSZK-hard when the acceptance probability deviates *negligibly* from 0 or 1.

2 Proof techniques

The QSZK completeness of the aforementioned QJSP and measQTDP crucially relies on the relationships between quantum analogues of common classical f-divergences[9]. We start by reviewing and defining these quantum analogues. The most widely used quantum distances are the trace distance (td) with the Bures distance (B, essentially the fidelity), which are quantum counterparts of the statistical distance (SD) and the (squared) Hellinger distance (H), respectively. Other commonly

$\mathrm{TD}(p_0,p_1) > \mathrm{TD}(p_0',p_1')$.

[5] This connection arises from the fact that the Jensen-Shannon divergence can be interpreted as the (conditional) entropy difference, as indicated implicitly in Vadhan's PhD thesis [22].

[6] The reader may feel confused with Theorem 5.4 in [23] which builds upon the techniques in [19]. However, it was claimed in [12] that the proof in [19] does extend to the parameter regime of $\alpha^2 - \beta \geq 1/\mathrm{poly}(n)$, but this claim was later retracted, see [10].

[7] In particular, the QSZK containment of QSDP via a reduction to measQTDP holds for pairs of mixed states (ρ_0,ρ_1) and (ρ_0',ρ_1') satisfying the conditions $\mathrm{td}(\rho_0,\rho_1) > \mathrm{td}(\rho_0',\rho_1') > \mathrm{td}^2(\rho_0,\rho_1)$ and $\mathrm{QTD}^{\mathrm{meas}}(\rho_0,\rho_1) > \mathrm{QTD}^{\mathrm{meas}}(\rho_0',\rho_1')$.

[8] A polarization lemma for some (quantum) distance is *length-preserving* if the output length of the pair of resulting (quantum) circuits is as same as the output length of the original pair of (quantum) circuits.

[9] An f-divergence is a function $\mathrm{D}_f(p_0\|p_1)$ that measures the difference between two probability distributions p_0 and p_1, and this divergence is defined as $\mathrm{D}_f(p_0\|p_1) := \mathbb{E}_{x\sim p_1} f(p_0(x)/p_1(x))$.

	Classical	Quantum	Usages related to QSZK
SD vs. H^2	$H^2 \le SD \le \sqrt{2}H$ [13]	$\frac{1}{2}B^2 \le td \le B$ [8]	A polarization lemma for the trace distance [26]
SD vs. JS	$1 - H_2\left(\frac{1-SD}{2}\right) \le JS_2 \le SD$ [8, 21]	$1 - H_2\left(\frac{1-td}{2}\right) \le QJS_2 \le td$ [6, 8]	QJSP is QSZK-hard This work
SD vs. TD	$SD^2 \le TD \le SD$ [21]	$td^2 \le QTD^{meas} \le QTD \le td$ This work	MEASQTDP is QSZK-hard This work
JS vs. TD	$\frac{1}{2}TD \le JS \le \ln 2 \cdot TD$ [21]	$\frac{1}{2}QTD^2 \le QJS \le QTD$ This work	None
TD vs. H^2	$H^2 \le TD \le 2H^2$ [15]	$\frac{1}{2}B^2 \le QTD^{meas} \le B^2$ $\frac{1}{2}B^2 \le QTD \le B$ This work	A polarization lemma for the QTD^{meas} This work

Table 1: A comparison between classical and quantum distances (or divergences) with usages related to QSZK

used f-divergences are the KL divergence (also known as the relative entropy) and the χ^2-divergence, which are unbounded, so we instead focus on their symmetrized versions, the Jensen-Shannon divergence (JS) and the triangular discrimination (TD), respectively.

The relationship between two quantum analogues of the Jensen-Shannon divergence constitutes a specific instance of the renowned Holevo's bound, namely the measured quantum Jensen-Shannon divergence is at most the quantum Jensen-Shannon divergence (QJS). To the best of our knowledge, there is no known quantum analogue of triangular discrimination. We thus initiate the definition of the quantum triangular discrimination (QTD) and the measured quantum triangular discrimination (QTD^{meas}), based on their connection to the quantum analogues of χ^2-divergence [20]. We further examine their relationship with other aforementioned quantum distances and divergences, as stated in Theorem 4.

Theorem 4 (Inequalities on quantum analogues of the triangular discrimination – informal) *For any quantum states ρ_0 and ρ_1, we know that*

(1) $td^2(\rho_0,\rho_1) \le QTD^{meas}(\rho_0,\rho_1) \le QTD(\rho_0,\rho_1) \le td(\rho_0,\rho_1)$*;*

(2) $\frac{1}{2}QTD^2(\rho_0,\rho_1) \le QJS(\rho_0,\rho_1) \le QTD(\rho_0,\rho_1)$*;*

(3) $\frac{1}{2}B^2(\rho_0,\rho_1) \le QTD^{meas}(\rho_0,\rho_1) \le B^2(\rho_0,\rho_1)$ *and* $\frac{1}{2}B^2(\rho_0,\rho_1) \le QTD(\rho_0,\rho_1) \le B(\rho_0,\rho_1)$*.*

We summarize our new results and known inequalities in Table 1, as well as how we utilize these inequalities in our proof. In addition, we highlight that the quantum triangular discrimination behaves differently from its classical counterpart since the triangular discrimination is a constant multiplicative error approximation of the Jensen-Shannon divergence. This difference breaks down the quantum equivalent of the ingenious reduction from TDP to JSP presented in [4], leading to a slightly worse parameter in the improved QSZK containment for the Quantum State Distinguishability problem.

Leveraging inequalities in Table 1, we then proceed to prove that both QJSP and MEASQTDP are QSZK-complete. The QSZK containment of MEASQTDP utilizes a *new* polarization lemma for the measured quantum triangular discrimination, and the QSZK containment of QJSP is demonstrated through a reduction to the Quantum Entropy Difference problem [3] using a nice property of the quantum conditional entropy on a classical-quantum state. We therefore explore the limitations of current techniques for showing QSZK containments. Additionally, the QSZK-hardness of these problems is straightforwardly analogous to their classical counterparts [4] because of the corresponding inequalities in Table 1.

3 Discussion and open problems

Improved inequalities on the quantum triangular discrimination. We observe that the second inequality in Theorem 4 is not *a tight bound*. Numerical simulations indicate that the tight bound is $QTD^2(\rho_0,\rho_1) \le QJS_2(\rho_0,\rho_1) \le QTD(\rho_0,\rho_1)$ for any mixed states ρ_0 and ρ_1, with improved constant factors. This bound can be saturated by choosing mixed states ρ_0 and ρ_1 with an orthogonal support, as these instances sufficient for $QJS(\rho_0,\rho_1)$ and $QTD(\rho_0,\rho_1)$ equal to 1[10]. Furthermore, numerical simulations also suggest that the triangular inequality holds for the square root of QTD, namely $\sqrt{QTD(\rho_0,\rho_1)} + \sqrt{QTD(\rho_1,\rho_2)} \ge \sqrt{QTD(\rho_0,\rho_2)}$ for any mixed states ρ_0, ρ_1 and ρ_2. This indicates that the square root of quantum triangular discrimination is a *metric*, with the same property also holding for triangular discrimination [15].

Applications of the quantum triangular discrimination. Is there any other application of the (measured) quantum triangular discrimination besides its use in QSZK as demonstrated in this paper? For instance, Amir Yehudayoff [27] has utilized triangular discrimination to obtain a sharper lower bound on the communication complexity of the point chasing problem. Can we expect an analogous implication in the quantum world?

[10] In fact, this is a sufficient and necessary condition for QJS as well as the trace distance. By the first inequality in Theorem 4, we apply this condition to QTD.

References

[1] Rochisha Agarwal, Soorya Rethinasamy, Kunal Sharma, and Mark M Wilde. Estimating distinguishability measures on quantum computers. *arXiv preprint arXiv:2108.08406*, 2021.

[2] Dorit Aharonov, Alex B Grilo, and Yupan Liu. StoqMA vs. MA: the power of error reduction. *arXiv preprint arXiv:2010.02835*, 2020. 2

[3] Avraham Ben-Aroya, Oded Schwartz, and Amnon Ta-Shma. Quantum expanders: Motivation and constructions. *Theory of Computing*, 6:47–79, 2010. 2, 3

[4] Itay Berman, Akshay Degwekar, Ron D Rothblum, and Prashant Nalini Vasudevan. Statistical difference beyond the polarizing regime. In *Theory of Cryptography Conference*, pages 311–332. Springer, 2019. 1, 2, 3

[5] Adam Bouland, Lijie Chen, Dhiraj Holden, Justin Thaler, and Prashant Nalini Vasudevan. On the power of statistical zero knowledge. *SIAM Journal on Computing*, 49(4):FOCS17–1, 2019. 2

[6] Jop Briët and Peter Harremoës. Properties of classical and quantum Jensen-Shannon divergence. *Physical review A*, 79(5):052311, 2009. 3

[7] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Physical Review Letters*, 87(16):167902, 2001. 2

[8] Christopher A Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999. 3

[9] András Gilyén and Alexander Poremba. Improved quantum algorithms for fidelity estimation. *arXiv preprint arXiv:2203.15993*, 2022.

[10] Oded Goldreich. Errata (3-Feb-2019). http://www.wisdom.weizmann.ac.il/~/oded/entropy.html, 2019. 2

[11] Oded Goldreich and Salil Vadhan. Comparing entropies in statistical zero knowledge with applications to the structure of SZK. In *Proceedings. Fourteenth Annual IEEE Conference on Computational Complexity (Formerly: Structure in Complexity Theory Conference)(Cat. No. 99CB36317)*, pages 54–73. IEEE, 1999. 2

[12] Oded Goldreich and Salil P Vadhan. On the complexity of computational problems regarding distributions. *Studies in Complexity and Cryptography*, 6650:390–405, 2011. 2

[13] Thomas Kailath. The divergence and bhattacharyya distance measures in signal selection. *IEEE transactions on communication technology*, 15(1):52–60, 1967. 3

[14] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum Merlin-Arthur proof systems: Are multiple Merlins more helpful to Arthur? *Chicago Journal of Theoretical Computer Science*, 2009:3, 2009. 2

[15] Lucien Le Cam. *Asymptotic methods in statistical decision theory*. Springer Science & Business Media, 1986. 3

[16] Yupan Liu. StoqMA meets distribution testing. In *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021. 1

[17] Ana P Majtey, Pedro W Lamberti, and Domingo P Prato. Jensen-Shannon divergence as a measure of distinguishability between mixed quantum states. *Physical Review A*, 72(5):052310, 2005. 2

[18] Ashley Montanaro and Ronald de Wolf. A survey of quantum property testing. *Theory of Computing*, pages 1–81, 2016. 1

[19] Amit Sahai and Salil Vadhan. A complete problem for statistical zero knowledge. *Journal of the ACM (JACM)*, 50(2):196–249, 2003. 1, 2

[20] Kristan Temme, Michael James Kastoryano, Mary Beth Ruskai, Michael Marc Wolf, and Frank Verstraete. The χ^2-divergence and mixing times of quantum markov processes. *Journal of Mathematical Physics*, 51(12):122201, 2010. 3

[21] Flemming Topsøe. Some inequalities for information divergence and related measures of discrimination. *IEEE Transactions on information theory*, 46(4):1602–1609, 2000. 3

[22] Salil Pravin Vadhan. *A study of statistical zero-knowledge proofs*. PhD thesis, Massachusetts Institute of Technology, 1999. 2

[23] Thomas Vidick and John Watrous. Quantum proofs. *Foundations and Trends® in Theoretical Computer Science*, 11(1-2):1–215, 2016. 2

[24] Qisheng Wang and Zhicheng Zhang. Fast quantum algorithms for trace distance estimation. *arXiv preprint arXiv:2301.06783*, 2023.

[25] Qisheng Wang, Zhicheng Zhang, Kean Chen, Ji Guan, Wang Fang, Junyi Liu, and Mingsheng Ying. Quantum algorithm for fidelity estimation. *IEEE Transactions on Information Theory*, 69(1):273–282, 2023.

[26] John Watrous. Limits on the power of quantum statistical zero-knowledge. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 459–468. IEEE, 2002. 1, 2, 3

[27] Amir Yehudayoff. Pointer chasing via triangular discrimination. *Combinatorics, Probability and Computing*, 29(4):485–494, 2020. 3

Intra-atomic frequency comb based photonic quantum memory using single-atom-cavity setup

Chanchal[1] * G. P. Teja[1] † Sandeep K. Goyal[1] ‡

[1] *Department of Physical Sciences, Indian Institute of Science Education and Research, Mohali, Punjab 140306, India*

Abstract. On-demand and efficient storage of photons is an essential element in quantum information processing and long-distance quantum communication. Most of the quantum memory protocols require bulk systems to store photons and pose challenge for integration with photonic chip platforms. Here, we present a protocol for quantum memory using only a single-atom-cavity setup. A single atom containing a frequency comb coupled to an optical cavity can store photons efficiently. We also discuss how Rubidium and Cesium atoms coupled to nanophotonic waveguide cavities can serve as promising candidates to realize this scheme. This provides a possible realization of an on-chip quantum memory.

Keywords: Quantum memory, single atom, frequency comb, cavity

1 Introduction

A photonic quantum memory is a device that can store and re-emit photons on demand [1–3]. It is an essential component in quantum information processing applications such as quantum networks [4,5], quantum repeaters [6] and long range quantum communication [7]. In a typical atomic ensemble based quantum memory, a weak light pulse is absorbed as a delocalized atomic excitation over all the atoms in the ensemble. This is retrieved using a set of controlled pulses that emit the photon at a desired time [2,3].

To gain scalability and the practical advantage in quantum information processing, many efforts are being devoted towards integrated photonic chips [8,9]. On-chip single photon sources, on-chip beamsplitters and on-chip photon detectors are already available on integrated platform [10], while on-chip quantum memory is still a work in progress and is highly sought after device [11,12].

Here, we present a scheme for storing weak light pulses and single photons using a single atom coupled to an optical cavity proposed in [13]. The trapped atom contains an intra-atomic frequency comb (I-AFC). This joint single-atom-cavity setup results in a photon-echo, similar to the I-AFC based quantum memory protocol [14]. One can also achieve robust and efficient storage for polarization and time-bin qubits using this setup. Further, Cesium and Rubidium atoms coupled to nanophotonic waveguide cavities can serve as promising candidates for the implementation of this quantum memory protocol [13]. One of the biggest advantages of the proposed scheme is that it paves a way for the possible realization of an on-chip quantum memory. Furthermore, since this protocol requires only a frequency comb coupled to a cavity, it can also be implemented using the quantum dots inside a cavity [15,16]. On-demand single-photon sources have already been realized using quantum dots [17,18]. Combining these two can pave the way for efficient on-chip photonic quantum computation.

*ph18004@iisermohali.ac.in
†teja4477@gmail.com
‡skgoyal@iisermohali.ac.in

2 Quantum memory using single-atom I-AFC coupled to a cavity

Consider an atom that contains a frequency comb, coupled to a high finesse single-mode optical cavity [Fig. 1(a)]. The Hamiltonian for such atom-cavity system consists of three parts, the free Hamiltonian of the single-mode cavity, the free Hamiltonian of the atom and the interaction between the two systems, which reads [13]

$$\begin{aligned} H =& H_{\rm c} + H_{\rm a} + H_{\rm int} \\ =& \hbar\omega_c \hat{a}^\dagger \hat{a} + \sum_{n=1}^{N} \hbar\omega_n^e \left|e_n\right\rangle\!\left\langle e_n\right| + \sum_{n=1}^{N} \hbar\omega_n^g \left|g_n\right\rangle\!\left\langle g_n\right| \\ & - \hbar\left[\sum_n g_n \left|e_n\right\rangle\!\left\langle g_n\right| \hat{a} + \sum_n g_n^* \left|g_n\right\rangle\!\left\langle e_n\right| \hat{a}^\dagger\right], \end{aligned} \tag{1}$$

where $\hat{a}$ is the photon annihilation operator for the cavity mode. $|g_n\rangle$ and $|e_n\rangle$ denote the n-th ground state and the excited state, respectively, with coupling strength $g_n = \frac{d_n}{\hbar}\sqrt{\frac{\hbar\omega_c}{2\epsilon_0 V}}$. The d_n is the transition dipole moment between $|g_n\rangle \leftrightarrow |e_n\rangle$ transition and ω_c is the resonance frequency of the cavity. Here, it is assumed that a particular ground state $|g_n\rangle$ is coupled only to a single excited state $|e_n\rangle$.

Solving the dynamics in the frequency domain for this atom-cavity system using the standard input-output formalism [19] yields the following expression for the output field mode $\hat{a}_{\rm out}$ in terms of the input field mode $\hat{a}_{\rm in}$ [13]

$$\hat{a}_{\rm out}(\omega) = \left[1 - \frac{\kappa}{{\rm i}(\omega + \Delta_c) + \mathcal{D}(\omega) + \frac{\kappa}{2}}\right] \hat{a}_{\rm in}(\omega). \tag{2}$$

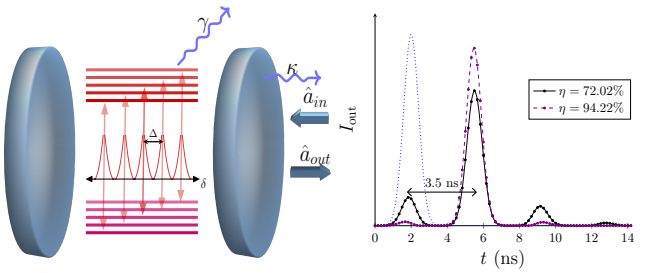
Here

$$\mathcal{D}(\omega) = \sum_n \frac{\sigma_{nn}^g |g_n|^2}{\left[i(\omega + \delta_n) + \frac{\gamma}{2}\right]}, \tag{3}$$

is the I-AFC propagator [14], and $\Delta_c = \omega_c - \omega_L$, $\delta_n = (\omega_n^e - \omega_n^g) - \omega_L$ are the detunings with respect to the input

light. γ is the spontaneous decay rate of the atom in free space and κ is the decay rate of the cavity field. Inverse Fourier-transform of Eq. (2) yields the output field in time $\hat{a}_{\rm out}(t)$.

In Fig. 1(b), we plot the output intensity $I_{\rm out} = \left\langle \hat{a}^\dagger_{\rm out}(t)\hat{a}_{\rm out}(t)\right\rangle$ as a function of time which we get by solving Eq. (2) numerically. Here we have considered a Gaussian input pulse of spectral width $2\pi \times 270$ MHz. The I-AFC associated with the atom has seven teeth with uniform comb spacing $\Delta = 2\pi \times 300$ MHz, tooth width $\gamma = 7.5$ MHz and the detuning $\Delta_c = 0$. The solid curve and the dashed curve in this figure correspond to two different values of the cavity decay rates κ. In this figure, we can clearly see that the first prominent output pulse of light is at time $t = 2$ ns which is due to the immediate reflection from the cavity. The second prominent output pulse occurs at $t \sim 5.5$ ns which is due to the emission from the cavity. There is a delay of 3.5 ns which is approximately $2\pi/\Delta$ due to the interaction of light with the setup. Hence the atom-cavity setup behaves like an I-AFC.



(a) (b)

Figure 1: (a) Schematic diagram for an I-AFC coupled to a cavity. Here, the I-AFC is interacting with a single cavity mode with decay rate κ. $\hat{a}_{\rm in}$ and $\hat{a}_{\rm out}$ represent the input and output field mode operators. γ is the spontaneous decay rate of the atom into free space. (b) Photon-echo after a delay of 3.5 ns for an ideal I-AFC coupled to a cavity. The input field in Eq. (2) is a gaussian a pulse given by $e^{-\omega^2/(2b^2)}$ with $b = 2\pi \times 270$ MHz. The two photon echoes shown in dashed purple and solid black curve correspond to the cavity decay rate 7 and 4 GHz, respectively, with the corresponding efficiencies 94.22% and 72.02%, respectively. The blue (dotted) curve shows the corresponding input field intensity.

2.1 Factors affecting quantum memory

Eqs. (2) and (3) suggest that the output field from the atom-cavity setup also depends on the cavity parameters g_n, κ and Δ_c, hence they can affect the quality of the memory. To quantify the quality of the quantum memory we consider the efficiency for this protocol, which is defined as

$$\eta = \frac{\int_{\pi/\Delta}^{3\pi/\Delta} \left\langle \hat{a}^\dagger_{\rm out}(t)\hat{a}_{\rm out}(t)\right\rangle dt}{\int \left\langle \hat{a}^\dagger_{\rm in}(t)\hat{a}_{\rm in}(t)\right\rangle dt}. \tag{4}$$

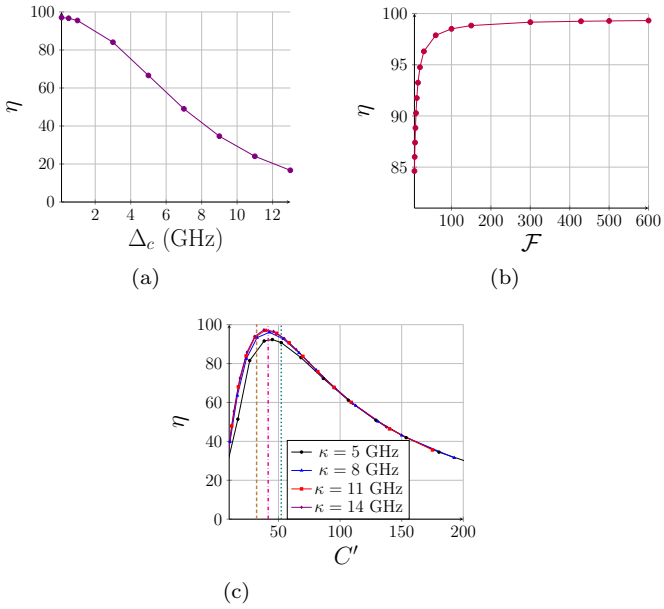


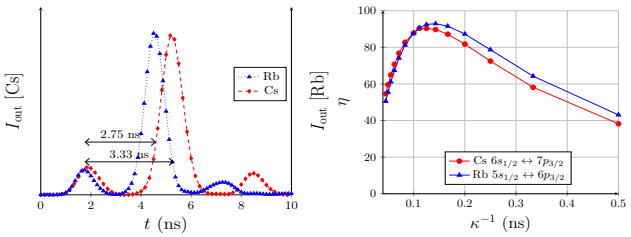
Figure 2: Effect of various parameters on the efficiency (η) of quantum memory in the single-atom-cavity setup for an ideal comb. In (a) and (b), we plot the variation of η as a function of the cavity detuning $\Delta_c = \omega_c - \omega_L$ and the comb finesse ($\mathcal{F}$), respectively for $(g', \kappa) = (1.8, 11)$ GHz with uniform comb spacing $\Delta = 2\pi \times 300$ MHz. (c) shows the plot of the variation of η for an ideal comb with comb spacing $\Delta = 2\pi \times 300$ MHz as a function of cooperativity, $C' = g'^2/(\kappa\gamma)$ for different values of κ.

In Fig. 2(a), we plot the variation of the efficiency η as a function of the cavity detuning Δ_c while keeping g', κ constant. As expected, it shows a drop in the efficiency as the cavity detuning increases. The efficiency also depends on the comb finesse $\mathcal{F} \equiv \Delta/\gamma$. In Fig. 2(b), we plot the efficiency as a function of comb finesse for the ideal comb with fixed comb spacing, $\Delta = 2\pi \times 300$ MHz by changing the peak width γ while keeping the cavity parameters fixed. This plot shows that the efficiency saturates to $\sim 100\%$ asymptotically for asymptotic values of the finesse.

For an ideal I-AFC, since all the peaks are identical, i.e., $d_n \equiv d$, we may write $g_n = g$. We define $g' = g\sqrt{\sigma^g_{nn}}$ as the effective coupling constant and define the cooperativity parameter for the atom-cavity system to be $C' = g'^2/(\kappa\gamma)$. In order to understand the effect of the cavity parameters g' and κ on the efficiency, in Fig. 2(c), we plot the variation of η as a function of the cooperativity C' for various values of κ and keeping $\Delta_c = 0$. Fig. 2(c) shows that the efficiency first increases, reaches an optimum value, and then starts decreasing again and there exists an optimum value of C' for every given value of κ, which maximizes the efficiency. The maximum efficiencies are obtained in the range $C' \sim$ 35-45 for all the values of κ. (For more details see [13]).

3 Realizing the quantum memory using Rb and Cs atoms

For realistic systems such as Rb and Cs atoms, the frequency combs obtained are usually non-uniform with unequal peak heights which affects the storage process [14, 20]. Here, we discuss the possibilities for experimental implementation of the single atom based quantum memory protocol in realistic systems such as Rb and Cs atoms coupled to nanophotonic waveguide cavity and show that the current scheme can be implemented with the existing experimental techniques.



(a) (b)

Figure 3: (a) Photon-echo for the I-AFC in Rb and Cs atoms. (b) Variation of efficiency as a function of $1/\kappa$ in Rb and Cs atoms.

One of the requirements to achieve efficient quantum memory in I-AFC-cavity setup is a cavity with high coupling strength g of the order of GHz (see Fig. 2). This, in turn requires a cavity with low mode volume of the order of $(\sim \mu\text{m})^3$. Such strong coupling can be achieved using the nano-cavities [21, 22] where mode volume $V \sim \lambda^3$ have already been realized. This tight confinement using nano-photonic cavities gives an additional advantage of potential integration with nano-photonics. Trapping in such low mode volumes results in the atom-cavity strong coupling of the order of $g \sim$ GHz along with the quality factor $Q = \omega_c/\kappa \sim 10^5$ [21].

Table 1: Rb and Cs parameters used in numerical calculations. λ is the wavelength of transition B is the magnetic field used in obtaining I-AFC. V, κ and Q are the mode volume, decay rate and quality factor of the cavity, respectively.

Atom	Transition	λ (nm)	B (T)	$V(\mu m)^3$	κ (GHz)	Q
Rb	$5s_{1/2} \leftrightarrow 6p_{3/2}$	420.3	0.15	20	~ 7	10^5
Cs	$6s_{1/2} \leftrightarrow 7p_{3/2}$	455.66	0.1	20	~ 8	10^5

Here we consider Cs and Rb atoms as examples to realize this quantum memory protocol. The parameters such as the atomic transitions used in the Cs and Rb atoms, the wavelength, applied magnetic field strength and so on for Rb and Cs atoms used for numerical calculations are given in Table 1 [13]. In Fig. 3(a) we show the photon-echo from Rb and Cs atoms calculated numerically by solving Eq. (2). The maximum efficiencies for Rb and Cs atoms are found to be 92.9% and 90.36%, respectively, for the parameters specified in Table. 1. The lesser value of the efficiencies in the case of Rb and Cs atoms is due to the inherent non-uniformity present in the frequency combs.

4 Conclusion

On-chip photonic quantum memories are essential for scalable and integrated photonic quantum information processing. Here, we have presented a scheme to store photons using only a single atom coupled to an optical cavity. The atom exhibits an I-AFC which enables the joint atom-cavity system to store photons. This provides us with a possibility to realize an on-chip quantum memory suitable for integrated photonic chips. The proposed setup is capable of storing time-multiplexed photons, along with their polarization degree of freedom efficiently, hence providing multi-mode photonic quantum memory. Although this scheme is presented trapped atoms, this can very well work with quantum dots and quantum defect centers. Since deterministic single photon sources have already been realized using quantum dots, combining it with the on-chip quantum memory can provide a robust integrated platform for photonic quantum computation.

References

[1] A. I. Lvovsky, B. C. Sanders, and W. Tittel, "Optical quantum memory," *Nat. Photonics*, vol. 3, no. 12, pp. 706–714, 2009.

[2] C. Simon, M. Afzelius, J. Appel, A. B. de la Giroday, S. Dewhurst, N. Gisin, C. Hu, F. Jelezko, S. Kröll, J. Müller, *et al.*, "Quantum memories-a review based on the european integrated project qubit applications (qap) p.," *The European Physical Journal D*, vol. 58, no. 1, 2010.

[3] K. Heshami, D. G. England, P. C. Humphreys, P. J. Bustard, V. M. Acosta, J. Nunn, and B. J. Sussman, "Quantum memories: emerging applications and recent advances," *J. Mod. Opt.*, vol. 63, no. 20, pp. 2005–2028, 2016.

[4] H. J. Kimble, "The quantum internet," *Nature*, vol. 453, no. 7198, pp. 1023–1030, 2008.

[5] C. Simon, "Towards a global quantum network," *Nat. Photonics*, vol. 11, no. 11, pp. 678–680, 2017.

[6] N. Sangouard, C. Simon, H. de Riedmatten, and N. Gisin, "Quantum repeaters based on atomic ensembles and linear optics," *Rev. Mod. Phys.*, vol. 83, pp. 33–80, Mar 2011.

[7] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, "Long-distance quantum communication with atomic ensembles and linear optics," *Nature*, vol. 414, no. 6862, pp. 413–418, 2001.

[8] B. Hacker, S. Welte, G. Rempe, and S. Ritter, "A photon–photon quantum gate based on a single atom in an optical resonator," *Nature*, vol. 536, no. 7615, pp. 193–196, 2016.

[9] S. Freer, S. Simmons, A. Laucht, J. T. Muhonen, J. P. Dehollain, R. Kalra, F. A. Mohiyaddin, F. E. Hudson, K. M. Itoh, J. C. McCallum, *et al.*, "A single-atom quantum memory in silicon," *Quantum Science and Technology*, vol. 2, no. 1, p. 015009, 2017.

[10] R. Uppu, F. T. Pedersen, Y. Wang, C. T. Olesen, C. Papon, X. Zhou, L. Midolo, S. Scholz, A. D. Wieck, A. Ludwig, *et al.*, "Scalable integrated single-photon source," *Science advances*, vol. 6, no. 50, p. eabc8268, 2020.

[11] T. Zhong, J. M. Kindem, J. G. Bartholomew, J. Rochman, I. Craiciu, E. Miyazono, M. Bettinelli, E. Cavalli, V. Verma, S. W. Nam, *et al.*, "Nanophotonic rare-earth quantum memory with optically controlled retrieval," *Science*, vol. 357, no. 6358, pp. 1392–1395, 2017.

[12] C. Liu, T.-X. Zhu, M.-X. Su, Y.-Z. Ma, Z.-Q. Zhou, C.-F. Li, and G.-C. Guo, "On-demand quantum storage of photonic qubits in an on-chip waveguide," *Phys. Rev. Lett.*, vol. 125, p. 260504, Dec 2020.

[13] Chanchal, G. P. Teja, and S. K. Goyal, "Intra-atomic frequency-comb-based photonic quantum memory using single-atom-cavity setup," *Phys. Rev. A*, vol. 107, p. 012614, Jan 2023.

[14] G. P. Teja, C. Simon, and S. K. Goyal, "Photonic quantum memory using an intra-atomic frequency comb," *Phys. Rev. A*, vol. 99, p. 052314, May 2019.

[15] J. An, A. Franceschetti, and A. Zunger, "The excitonic exchange splitting and radiative lifetime in pbse quantum dots," *Nano Lett.*, vol. 7, no. 7, pp. 2129–2135, 2007.

[16] P. Zhang, G. Song, and L. Yu, "Optical trapping of single quantum dots for cavity quantum electrodynamics," *Photonics Res.*, vol. 6, no. 3, pp. 182–185, 2018.

[17] D. Heinze, D. Breddermann, A. Zrenner, and S. Schumacher, "A quantum dot single-photon source with on-the-fly all-optical polarization control and timed emission," *Nat. Commun.*, vol. 6, no. 1, pp. 1–6, 2015.

[18] F. Liu, A. J. Brash, J. OHara, L. M. Martins, C. L. Phillips, R. J. Coles, B. Royall, E. Clarke, C. Bentham, N. Prtljaga, *et al.*, "High purcell factor generation of indistinguishable on-chip single photons," *Nat. Nanotechnol.*, vol. 13, no. 9, pp. 835–840, 2018.

[19] C. W. Gardiner and M. J. Collett, "Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation," *Phys. Rev. A*, vol. 31, pp. 3761–3774, Jun 1985.

[20] G. P. Teja and S. K. Goyal, "Studying the effect of fluctuating environment on intra-atomic frequency comb based quantum memory," *Sci. Rep.*, vol. 11, p. 11439, Jun 2021.

[21] D. van Oosten and L. Kuipers, "Trapping a single atom with a fraction of a photon using a photonic crystal nanocavity," *Phys. Rev. A*, vol. 84, p. 011802, Jul 2011.

[22] J. D. Thompson, T. Tiecke, N. P. de Leon, J. Feist, A. Akimov, M. Gullans, A. S. Zibrov, V. Vuletić, and M. D. Lukin, "Coupling a single trapped atom to a nanoscale optical cavity," *Science*, vol. 340, no. 6137, pp. 1202–1205, 2013.

Operational Interpretation of the Sandwiched Rényi Divergence of Order 1/2 to 1 as Strong Converse Exponents

Ke Li[1] * Yongsheng Yao[1 2] †

[1] *Institute for Advanced Study in Mathematics, Harbin Institute of Technology, Harbin 150001, China*
[2] *School of Mathematics, Harbin Institute of Technology, Harbin 150001, China*

Abstract. We provide the sandwiched Rényi divergence of order $\alpha \in (\frac{1}{2}, 1)$, as well as its induced quantum information quantities, with an operational interpretation in the characterization of the exact strong converse exponents of quantum tasks. Specifically, we consider (a) smoothing of the max-relative entropy, (b) quantum privacy amplification, and (c) quantum information decoupling. We solve the problem of determining the exact strong converse exponents for these three tasks, with the performance being measured by the fidelity or purified distance. The results are given in terms of the sandwiched Rényi divergence of order $\alpha \in (\frac{1}{2}, 1)$. This is the first time to find the precise operational meaning for the sandwiched Rényi divergence with Rényi parameter in the interval $\alpha \in (\frac{1}{2}, 1)$.

Keywords: sandwiched Rényi divergence, strong converse exponent, quantum shannon theory

1 Introduction

Rényi's information divergence, defined for two probability densities, is a fundamental information quantity which has played importance roles in a diversity of fields, ranging from information theory, to probability theory, and to thermodynamics and statistical physics. Its quantum generalization, due to the noncommutativity nature of density matrices, can take infinitely many possible forms. To find which one is the correct quantum generalization is significant and nontrivial.

The sandwiched Rényi divergence is one of the proper quantum generalization of Rényi's information divergence. For two density matrices ρ and σ, it is defined as [1, 2]

$$D^*_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2\alpha}}\rho\sigma^{\frac{1-\alpha}{2\alpha}}\right)^\alpha, \qquad (1)$$

where $\alpha \in (0,1)\cup(1,\infty)$ is a real parameter. Since its discovery, several operational interpretations for this quantity have been found. On the one hand, with $\alpha \in (1,\infty)$, it characterizes the strong converse exponent for quantum hypothesis testing [3, 4], for classical communication over classical-quantum channels [5], for classical data compression with quantum side information [6], and for entanglement-assisted or quantum-feedback-assisted communication over quantum channels [7, 8, 9]. On the other hand, with $\alpha \in (1,2)$ or $\alpha \in (1,\infty)$, the sandwiched Rényi divergence also characterizes the direct error exponent for the smoothing of the max-relative entropy [10, 11], for quantum privacy amplification [12, 10], for quantum information decoupling and state merging [11], and for quantum channel simulation [13].

The other proper quantum generalization of Rényi's information divergence that has found operational interpretations, is Petz's Rényi divergence [14]

$$D_\alpha(\rho\|\sigma) := \frac{1}{\alpha-1}\log \mathrm{Tr}\left(\rho^\alpha\sigma^{1-\alpha}\right), \qquad (2)$$

for $\alpha \in (0,1)\cup(1,\infty)$. With $\alpha \in (0,1)$, Petz's Rényi divergence characterizes the direct error exponent for quantum hypothesis testing in both the symmetric setting [15, 16, 17] and the asymmetric setting [18, 19, 20], and for classical data compression with quantum side information [6, 21]. It is also believed to provide the solution to the long-standing open problem of determining the reliability function of classical-quantum channels [22, 23, 24, 19, 25, 26, 21, 27], with $\alpha \in (0,1)$ too.

The roles that the two quantum Rényi divergences have played so far cause people to guess that the correct quantum generalization of Rényi's information divergence may be

$$\begin{cases} D^*_\alpha(\rho\|\sigma) & \text{when } \alpha \in (1,\infty), \\ D_\alpha(\rho\|\sigma) & \text{when } \alpha \in (0,1). \end{cases} \qquad (3)$$

Indeed, this has been conjectured in the literature; see, e.g. [3].

In this paper, we find an operational interpretation to the sandwiched Rényi divergence of order $\alpha \in (\frac{1}{2}, 1)$. This includes an operational interpretation to $D^*_\alpha(\rho\|\sigma)$ itself, as well as to its induced information quantities, the sandwiched Rényi conditional entropy and the regularized sandwiched Rényi mutual information, all in the interval $\alpha \in (\frac{1}{2}, 1)$. These results are obtained by considering the strong converse exponents for several quantum information tasks. Employing the purified distance (or, equivalently, the fidelity) as the measure of the performance, we determine these strong converse exponents, which are given in terms of the above-mentioned sandwiched Rényi information quantities of order $\alpha \in (\frac{1}{2}, 1)$. Specifically, our results are as follows.

- *Smoothing of the max-relative entropy* . Let ρ be a quantum state, σ be a positive semidefinite operator, and $\epsilon \in [0,1]$. The smooth max-relative entropy based on the purified distance is defined

*carl.ke.lee@gmail.com
†yongsh.yao@gmail.com

as [28]

$$D^{\epsilon}_{\max}(\rho\|\sigma) := \min\left\{\lambda\in\mathbb{R}\;\middle|\;\left(\exists\tilde{\rho}\in\mathcal{B}^{\epsilon}(\rho)\right)\,\tilde{\rho}\le 2^{\lambda}\sigma\right\}, \tag{4}$$

where $\mathcal{B}^{\epsilon}(\rho) := \{\tilde{\rho}\in\mathcal{S}_{\le}(\mathcal{H}) \mid P(\tilde{\rho},\rho)\le\epsilon\}$ is the ϵ-ball of (subnormalized) quantum states around ρ. Regarding $D^{\epsilon}_{\max}(\rho\|\sigma)$ as a function of ϵ, we introduce its inverse function

$$\begin{aligned}\epsilon(\rho\|\sigma,\lambda) :=& \min\{\epsilon \mid D^{\epsilon}_{\max}(\rho\|\sigma)\le\lambda\}\\ =& \min\left\{P(\rho,\tilde{\rho})\;\middle|\;\tilde{\rho}\in\mathcal{S}_{\le}(\mathcal{H}) \text{ and } \tilde{\rho}\le 2^{\lambda}\sigma\right\}\end{aligned} \tag{5}$$

and call it the smoothing quantity.

The smooth max-relative entropy is not only a basic tool in quantum information theory, but it also quantifies in an exact way an operational task. In [29], it is shown that $D^{\epsilon}_{\max}(\rho\|\sigma)$ is the cost of preparing the box (ρ,σ) with error ϵ, in the resource theory of asymmetric distinguishability. So, the study of $D^{\epsilon}_{\max}(\rho\|\sigma)$ is equivalent to the study of the dilation problem of asymmetric distinguishability.

The quantum asymptotic equipartition property [30, 31] states that

$$\lim_{n\to\infty}\frac{1}{n}D^{\epsilon}_{\max}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma). \tag{6}$$

This shows that the relative entropy is a sharp threshold. The large-deviation type behaviors are more conveniently stated in terms of $\epsilon(\rho\|\sigma,\lambda)$. When $r > D(\rho\|\sigma)$, the rate of exponential decay of $\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}, nr)$ towards 0 is determined recently in [10]. On the other hand, when $r < D(\rho\|\sigma)$, the smoothing quantity $\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}, nr)$ must converge to 1 exponentially fast; see [29, 32, 33] for bounds using Rényi relative entropies. The exact rate of this exponential convergence is called the strong converse exponent, which remains unknown.

We determine the strong converse exponent for the smoothing of the max-relative entropy. Our result is the following Theorem 1.

Theorem 1 *For $\rho\in\mathcal{S}(\mathcal{H})$, $\sigma\in\mathcal{P}(\mathcal{H})$ and $r\in\mathbb{R}$, we have*

$$\begin{aligned}&\lim_{n\to\infty}\frac{-1}{n}\log\left(1-\epsilon(\rho^{\otimes n}\|\sigma^{\otimes n}, nr)\right)\\ =&\sup_{\frac{1}{2}\le\alpha\le 1}\frac{1-\alpha}{\alpha}\left\{D^{*}_{\alpha}(\rho\|\sigma)-r\right\}.\end{aligned} \tag{7}$$

Proof. see [34] for full version. □

- *Quantum privacy amplification.* Consider a classical-quantum state

$$\rho_{XE} = \sum_{x\in\mathcal{X}} p_x|x\rangle\langle x|_X\otimes\rho^{x}_{E}. \tag{8}$$

Let the system X, which is also regarded as a classical random variable, represent an imperfect random number that is partially correlated with an adversary Eve's system E. In the procedure of privacy amplification, we apply a hash function $f:\mathcal{X}\to\mathcal{Z}$ on X to extract a random number Z, which is expected to be uniformly distributed and independent of the adversary's system E. The action of the hash function f can be written as a quantum operation

$$\mathcal{P}_f : \omega\mapsto\sum_{x\in\mathcal{X}}\langle x|\omega|x\rangle|f(x)\rangle\langle f(x)|. \tag{9}$$

So the resulting state of privacy amplification is

$$\mathcal{P}_f(\rho_{XE}) = \sum_{z\in\mathcal{Z}}|z\rangle\langle z|_Z\otimes\sum_{x\in f^{-1}(z)}p_x\rho^{x}_{E}. \tag{10}$$

The effect is measured by two quantities. One is the size of the extracted randomness, $\log|\mathcal{Z}|$ in bits. The other one is the security parameter, defined as

$$\mathfrak{P}^{\mathrm{pa}}(\rho_{XE},f) := \max_{\omega_E\in\mathcal{S}(E)} F^2\left(\mathcal{P}_f(\rho_{XE}),\pi_Z\otimes\omega_E\right), \tag{11}$$

where π_Z is the maximally mixed state. Since the purified distance is a function of fidelity, the security parameter employed here takes the same information as the one based on purified distance in previous works (e.g., [35, 36, 10]).

In the asymptotic setting where an arbitrary large number of copies of the state ρ_{XE} is available, we apply the hash function $f_n:\mathcal{X}^{\times n}\to\mathcal{Z}_n$ to extract private randomness from $\rho^{\otimes n}_{XE}$, for any $n\in\mathbb{N}$. It has been proven in [37, 38] that to achieve asymptotically perfect privacy amplification such that $\mathfrak{P}^{\mathrm{pa}}(\rho^{\otimes n}_{XE}, f_n)\to 1$, the rate of randomness extraction must satisfy

$$\limsup_{n\to\infty}\frac{1}{n}\log|\mathcal{Z}_n| \le H(X|E)_{\rho}. \tag{12}$$

Finer asymptotic results, including the second-order expansion based on purified distance [35] and that based on trace distance [39], as well as the large-deviation type error exponent [10], have been obtained later.

On the other hand, when the rate of randomness extraction is larger than $H(X|E)_{\rho}$, the strong converse property holds. Specifically, for any sequence of hash functions $\{f_n:\mathcal{X}^{\times n}\to\mathcal{Z}_n\}_{n\in\mathbb{N}}$, we have

$$\begin{aligned}&\liminf_{n\to\infty}\frac{1}{n}\log|\mathcal{Z}_n| > H(X|E)_{\rho}\\ \Rightarrow&\lim_{n\to\infty}\mathfrak{P}^{\mathrm{pa}}(\rho^{\otimes n}_{XE},f_n)=0,\end{aligned} \tag{13}$$

and the decay of $\mathfrak{P}^{\mathrm{pa}}(\rho^{\otimes n}_{XE},f_n)$ is exponentially fast. This can be seen, from the one-shot converse bound in terms of the smooth conditional min-entropy [40, 35] combined with the asymptotic equipartition property [30]. The work [41] proved

the strong converse property by providing a bound on the rate of exponential decay in terms of the sandwiched Rényi conditional entropy; see also recent works [42, 32] for bounds employing Petz's Rényi conditional entropy. The optimal achievable exponent of this decay is called the strong converse exponent and is defined as

$$\begin{aligned} &E_{sc}^{\mathrm{pa}}(\rho_{XE}, r) \\ &:= \inf\left\{ \limsup_{n\to\infty} \frac{-1}{n} \log \mathfrak{P}^{\mathrm{pa}}(\rho_{XE}^{\otimes n}, f_n) \;\middle|\; \liminf_{n\to\infty} \frac{1}{n} \log |\mathcal{Z}_n| \geq r \right\}. \end{aligned} \tag{14}$$

We derive the exact expression for $E_{sc}^{\mathrm{pa}}(\rho_{XE}, r)$. The result is given in the following theorem.

Theorem 2 *Let ρ_{XE} be a classical-quantum state. For any rate $r \geq 0$, we have*

$$E_{sc}^{\mathrm{pa}}(\rho_{XE}, r) = \sup_{\frac{1}{2}\leq\alpha\leq 1} \frac{1-\alpha}{\alpha}\big\{r - H_\alpha^*(X|E)_\rho\big\}. \tag{15}$$

Proof. see [34] for full version. □

- *Quantum information decoupling.* Let ρ_{RA} be a bipartite quantum state with A in the lab and R held by a referee. Quantum information decoupling is the task of removing the correlation between system A and system R, by performing quantum operations on A. We focus on decoupling strategy via discarding a subsystem [43]; other strategies, such as those of [44] and [45], can be treated similarly. A general decoupling scheme $\mathcal{D}$ consists of a catalytic system A' in a state $\sigma_{A'}$ and a unitary transformation $U : \mathcal{H}_{AA'} \to \mathcal{H}_{\bar{A}\tilde{A}}$. We write

$$\mathcal{D} := (\sigma_{A'},\ U : \mathcal{H}_{AA'} \to \mathcal{H}_{\bar{A}\tilde{A}}). \tag{16}$$

Discarding the subsystem $\tilde{A}$, the goal of quantum information decoupling is to make the resulting state on R and $\bar{A}$ close to a product form. Thus the performance of this scheme is characterized by

$$\begin{aligned} &\mathfrak{P}^{\mathrm{dec}}(\rho_{RA}, \mathcal{D}) \\ &:= \max_{\substack{\omega_R\in\mathcal{S}(R),\\ \omega_{\bar{A}}\in\mathcal{S}(\bar{A})}} F^2\big(\mathrm{Tr}_{\tilde{A}}[U(\rho_{RA}\otimes\sigma_{A'})U^*], \omega_R\otimes\omega_{\bar{A}}\big). \end{aligned} \tag{17}$$

The cost is measured by the amount of discarded qubits, namely, $\log|\tilde{A}|$.

It has been established that, when arbitrarily many copies of the state ρ_{RA} is available, asymptotically perfect decoupling can be achieved if and only if the rate of decoupling cost is at least $\frac{1}{2}I(R:A)_\rho$ [43]. With catalyst, the second-order asymptotics has been derived in [46]. Recently, in the catalytic setting too, we have conducted the exponential analysis, obtaining the best exponent for the convergence of the performance towards the perfect in case that the rate of decoupling cost is below a critical value [11].

When the rate of decoupling cost is smaller than $\frac{1}{2}I(R:A)_\rho$, the strong converse property states that for any sequence of decoupling schemes $\big\{\mathcal{D}_n = (\sigma_{A'_n},\ U_n : \mathcal{H}_{A^nA'_n} \to \mathcal{H}_{\bar{A}_n\tilde{A}_n})\big\}_{n\in\mathbb{N}}$, we have

$$\begin{aligned} &\limsup_{n\to\infty} \frac{1}{n}\log|\tilde{A}_n| < \frac{1}{2}I(R:A)_\rho \\ &\Rightarrow \lim_{n\to\infty} \mathfrak{P}^{\mathrm{dec}}(\rho_{RA}^{\otimes n}, \mathcal{D}_n) = 0. \end{aligned} \tag{18}$$

Moreover, the convergence is exponentially fast. This follows from the one-shot smooth-entropy bound [47] coupled with the asymptotic equipartition property [30, 31]; see also [48, 41] for proofs employing Rényi entropies, and [46] for discussions on the catalytic setting. The optimal rate of exponential decay of $\mathfrak{P}^{\mathrm{dec}}(\rho_{RA}^{\otimes n}, \mathcal{D}_n)$ in the strong converse domain, for a fixed rate of decoupling cost, is called the strong converse exponent. Formally, it is defined as

$$\begin{aligned} &E_{sc}^{\mathrm{dec}}(\rho_{RA}, r) := \\ &\inf\left\{ \limsup_{n\to\infty} \frac{-1}{n} \log \mathfrak{P}^{\mathrm{dec}}(\rho_{RA}^{\otimes n}, \mathcal{D}_n) \middle| \limsup_{n\to\infty} \frac{1}{n} \log|\tilde{A}_n| \leq r \right\}. \end{aligned} \tag{19}$$

We derive the exact expression for $E_{sc}^{\mathrm{dec}}(\rho_{RA}, r)$. Our result is as follows.

Theorem 3 *Let $\rho_{RA} \in \mathcal{S}(RA)$ and $r \geq 0$. We have*

$$E_{sc}^{\mathrm{dec}}(\rho_{RA}, r) = \sup_{\frac{1}{2}\leq\alpha\leq 1} \frac{1-\alpha}{\alpha}\big\{I_\alpha^{*,\mathrm{reg}}(R:A)_\rho - 2r\big\}. \tag{20}$$

Proof. see [34] for full version. □

Because these results have extended the operational significance of the sandwiched Rényi divergence to the range $\alpha \in (\frac{1}{2}, 1)$, we conclude that the quantum generalization of Rényi's information divergence is more complicated than that was conjectured in Eq. (3).

To derive our results, we employ two different methods. For the problem of smoothing the max-relative entropy, we first prove the special case in which ρ and σ commute, using the method of types [49], and then we reduce the general case to the commutative case by exploiting some new properties of the fidelity function. For the problems of quantum privacy amplification and quantum information decoupling, we prove the achievability parts and the optimality parts separately, using different ideas. The achievability parts are technically more involved. We employ an approach developed by Mosonyi and Ogawa [5], to derive a weaker bound in terms of the log-Euclidean Rényi conditional entropy (mutual information) and then improve it to obtain the final bound. On the other hand, the optimality parts are accomplished by adapting the techniques developed by Leditzky, Wilde and Datta [41].

References

[1] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum Rényi entropies: a new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, 2013.

[2] Mark M Wilde, Andreas Winter, and Dong Yang. Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Communications in Mathematical Physics*, 331(2):593–622, 2014.

[3] Milán Mosonyi and Tomohiro Ogawa. Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies. *Communications in Mathematical Physics*, 334(3):1617–1648, 2015.

[4] Masahito Hayashi and Marco Tomamichel. Correlation detection and an operational interpretation of the Rényi mutual information. *Journal of Mathematical Physics*, 57(10):102201, 2016.

[5] Milán Mosonyi and Tomohiro Ogawa. Strong converse exponent for classical-quantum channel coding. *Communications in Mathematical Physics*, 355(1):373–426, 2017.

[6] Hao-Chung Cheng, Eric P Hanson, Nilanjana Datta, and Min-Hsiu Hsieh. Non-asymptotic classical data compression with quantum side information. *IEEE Transactions on Information Theory*, 67(2):902–930, 2020.

[7] Manish K Gupta and Mark M Wilde. Multiplicativity of completely bounded p-norms implies a strong converse for entanglement-assisted capacity. *Communications in Mathematical Physics*, 334(2):867–887, 2015.

[8] Tom Cooney, Milán Mosonyi, and Mark M Wilde. Strong converse exponents for a quantum channel discrimination problem and quantum-feedback-assisted communication. *Communications in Mathematical Physics*, 344(3):797–829, 2016.

[9] Ke Li and Yongsheng Yao. Strong converse exponent for entanglement-assisted communication. *arXiv preprint arXiv:2209.00555*, 2022.

[10] Ke Li, Yongsheng Yao, and Masahito Hayashi. Tight exponential analysis for smoothing the max-relative entropy and for quantum privacy amplification. *arXiv preprint arXiv:2111.01075*, 2021.

[11] Ke Li and Yongsheng Yao. Reliability function of quantum information decoupling via the sandwiched Rényi divergence. *arXiv preprint arXiv:2111.06343*, 2021.

[12] Masahito Hayashi. Precise evaluation of leaked information with secure randomness extraction in the presence of quantum attacker. *Communications in Mathematical Physics*, 333(1):335–350, 2015.

[13] Ke Li and Yongsheng Yao. Reliable simulation of quantum channels. *arXiv preprint arXiv:2112.04475*, 2021.

[14] Dénes Petz. Quasi-entropies for finite quantum systems. *Reports on Mathematical Physics*, 23(1):57–65, 1986.

[15] Koenraad MR Audenaert, John Calsamiglia, Ramón Munoz-Tapia, Emilio Bagan, Ll Masanes, Antonio Acin, and Frank Verstraete. Discriminating states: The quantum Chernoff bound. *Physical Review Letters*, 98(16):160501, 2007.

[16] Michael Nussbaum and Arleta Szkoła. The Chernoff lower bound for symmetric quantum hypothesis testing. *The Annals of Statistics*, 37(2):1040–1057, 2009.

[17] Ke Li. Discriminating quantum states: the multiple chernoff distance. *The Annals of Statistics*, 44(4):1661–1679, 2016.

[18] Hiroshi Nagaoka. The converse part of the theorem for quantum Hoeffding bound. *arXiv preprint quant-ph/0611289*, 2006.

[19] Masahito Hayashi. Error exponent in asymmetric quantum hypothesis testing and its application to classical-quantum channel coding. *Physical Review A*, 76(6):062301, 2007.

[20] Koenraad MR Audenaert, Michael Nussbaum, Arleta Szkoła, and Frank Verstraete. Asymptotic error rates in quantum hypothesis testing. *Communications in Mathematical Physics*, 279(1):251–283, 2008.

[21] Joseph M Renes. Achievable error exponents of data compression with quantum side information and communication over symmetric classical-quantum channels. *arXiv preprint arXiv:2207.08899*, 2022.

[22] M Burnashev and Alexander S Holevo. On the reliability function for a quantum communication channel. *Problems of Information Transmission*, 34(2):97–107, 1998.

[23] Andreas Winter. Coding theorems of quantum information theory. *PhD Thesis, Universität Bielefeld*, 1999.

[24] Alexander S Holevo. Reliability function of general classical-quantum channel. *IEEE Transactions on Information Theory*, 46(6):2256–2261, 2000.

[25] Marco Dalai. Lower bounds on the probability of error for classical and classical-quantum channels. *IEEE Transactions on Information Theory*, 59(12):8027–8056, 2013.

[26] Hao-Chung Cheng, Min-Hsiu Hsieh, and Marco Tomamichel. Quantum sphere-packing bounds with polynomial prefactors. *IEEE Transactions on Information Theory*, 65(5):2872–2898, 2019.

[27] Hao-Chung Cheng. A simple and tighter derivation of achievability for classical communication over quantum channels. *arXiv preprint arXiv:2208.02132*, 2022.

[28] Nilanjana Datta. Max-relative entropy of entanglement, alias log robustness. *International Journal of Quantum Information*, 7(02):475–491, 2009.

[29] Xin Wang and Mark M Wilde. Resource theory of asymmetric distinguishability. *Physical Review Research*, 1(3):033170, 2019.

[30] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55(12):5840–5847, 2009.

[31] Marco Tomamichel. *Quantum information processing with finite resources: mathematical foundations*, volume 5. Springer, 2015.

[32] Salzmann Robert and Datta Nilanjana. Total insecurity of communication via strong converse for quantum privacy amplification. *arXiv preprint arXiv:2022.11090*, 2022.

[33] Mark M Wilde. On distinguishability distillation and dilution exponents. *arXiv preprint arXiv:2022.12433*, 2022.

[34] Ke Li and Yongsheng Yao. Operational interpretation of the sandwiched rényi divergence of order 1/2 to 1 as strong converse exponents. *arXiv preprint arXiv:2209.00554*, 2022.

[35] Marco Tomamichel and Masahito Hayashi. A hierarchy of information quantities for finite block length analysis of quantum tasks. *IEEE Transactions on Information Theory*, 59(11):7693–7710, 2013.

[36] Anurag Anshu, Mario Berta, Rahul Jain, and Marco Tomamichel. Partially smoothed information measures. *IEEE Transactions on Information Theory*, 66(8):5022–5036, 2020.

[37] Igor Devetak and Andreas Winter. Distillation of secret key and entanglement from quantum states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 461(2053):207–235, 2005.

[38] Renato Renner. Security of quantum key distribution. *PhD Thesis*, 2005.

[39] Yu-Chen Shen, Li Gao, and Hao-Chung Cheng. Optimal second-order rates for quantum soft covering and privacy amplification. *arXiv preprint arXiv:2202.11590*, 2022.

[40] Marco Tomamichel, Christian Schaffner, Adam Smith, and Renato Renner. Leftover hashing against quantum side information. *IEEE Transactions on Information Theory*, 57(8):5524–5535, 2011.

[41] Felix Leditzky, Mark M Wilde, and Nilanjana Datta. Strong converse theorems using Rényi entropies. *Journal of Mathematical Physics*, 57(8):082202, 2016.

[42] Yu-Chen Shen, Li Gao, and Hao-Chung Cheng. Strong converse for privacy amplification against quantum side information. *arXiv preprint arXiv:2202.10263*, 2022.

[43] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. The mother of all protocols: restructuring quantum information's family tree. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2108):2537–2563, 2009.

[44] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Quantum state merging and negative information. *Communications in Mathematical Physics*, 269(1):107–136, 2007.

[45] Berry Groisman, Sandu Popescu, and Andreas Winter. Quantum, classical, and total amount of correlations in a quantum state. *Physical Review A*, 72(3):032317, 2005.

[46] Christian Majenz, Mario Berta, Frédéric Dupuis, Renato Renner, and Matthias Christandl. Catalytic decoupling of quantum information. *Physical Review Letters*, 118(8):080503, 2017.

[47] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, 2011.

[48] Naresh Sharma. A strong converse for the quantum state merging protocol. *arXiv preprint arXiv:1404.5940*, 2014.

[49] Imre Csiszár and János Körner. *Information Theory: Coding Theorems for Discrete Memoryless Systems*. Cambridge University Press, 2011.

Computation of Green's function by local variational quantum compilation

Shota Kanasugi[1] * Shoichiro Tsutsui[2] Yuya O. Nakagawa[2] Kazunori Maruyama[1]
Hirotaka Oshima[1] Shintaro Sato[1]

[1] *Quantum Laboratory, Fujitsu Research, Fujitsu Limited., Kanagawa, Japan*
[2] *QunaSys Inc., Tokyo, Japan*

Abstract. Computation of the Green's function for large-scale quantum many-body systems is crucial in material science. It is also a representative problem for which quantum computers are expected to outperform classical computers. Here, we propose an efficient method to compute the Green's function utilizing the local variational quantum compilation, which simulates the dynamics of large-scale quantum systems using a low-depth quantum circuit constructed through optimization on a smaller-size subsystem. By performing numerical simulations and estimation of the gate count for the Fermi-Hubbard model, we demonstrate the advantage of our method compared to the standard approach based on Trotter decomposition.

Keywords: Quantum algorithm, Quantum simulation, Quantum many-body system

1 Introduction

Simulation of quantum many-body systems is one of the most promising tasks for which quantum computers are believed to outperform classical computers. An important physical quantity for studying such quantum many-body systems is the Green's function, which tells us a variety of crucial information such as particle density, quasiparticle energy dispersion, density of states (DOS), and dynamical response under external fields. There are some algorithms proposed to compute the Green's function using either long-term fault-tolerant quantum computers (FTQCs) or noisy intermediate-scale quantum (NISQ) computers. The implementation of the FTQC techniques is, however, a long-term goal since they require an enormous number of very high-fidelity qubits as well as gate operations. The NISQ techniques, on the other hand, are expected to be executable on near-term quantum computers, whereas they are not applicable to the simulation of large-scale quantum systems requiring many quantum gates. The lack of feasible techniques is a significant problem for practical applications such as material design, which requires computation of the Green's function for large-scale quantum systems of sizes that are difficult to deal with on classical or NISQ computers. In this sense, it is desirable to devise a more sophisticated approach that can bridge the gap between the methods for FTQCs and NISQ devices.

In this work, to bridge such a gap between the FTQC and NISQ techniques, we propose a new approach to compute the Green's function on quantum computers based on the local variational quantum compilation (LVQC) algorithm [1]. This method enables us to compute the Green's function for large-scale quantum many-body systems even in the near-term quantum computing era, in which only smaller-scale quantum devices such as NISQ devices are available. Specifically, the LVQC approximates the time evolution operator for large-scale quantum many-body systems as a low-depth quantum circuit by performing an optimization procedure only on a smaller-size subsystem. Since the formulation of the LVQC relies mainly on the existence of the Lieb-Robinson (LR) bound [2], which is a theoretical upper bound of the speed of propagation of the information in quantum many-body systems, it is applicable to the broad class of quantum many-body systems with local interactions. Although the LVQC algorithm was originally devised for spin systems [1], simulating the Green's function of fermionic systems (e.g., strongly correlated electron systems) is important in practical applications such as material design. Thus, we develop an extended version of the LVQC algorithm that can be applied not only to spin systems but also to fermionic systems. Our proposal is to utilize the approximate time evolution operator obtained by this extended LVQC algorithm for computing the Green's function. Leveraging the LVQC algorithm, we can significantly reduce the number of quantum gates needed to simulate Green's function. In addition, the LVQC can be used to efficiently implement a time evolution operator on not only NISQ devices but also on early-FTQCs or FTQCs through optimization on NISQ or classical computers. Therefore, our method is valid not only for NISQ devices but also for larger-scale quantum computers such as FTQCs.

We demonstrate the validity of our method by performing a numerical simulation for the Fermi-Hubbard model, which is the simplest model of interacting fermions but is essential to study the nature of strongly correlated electron systems, using a quantum circuit simulator. We confirm that our method well reproduces the exact Green's function for up to a 4×4 lattice size (32 qubits) close to a limitation that can be treated by classical computers. Furthermore, based on numerical results and a formal estimation of gate count, we show that our LVQC-based method computes the Green's function more accurately and efficiently compared to a standard method based on the Trotter decomposition.

*kanasugi.shota@fujitsu.com

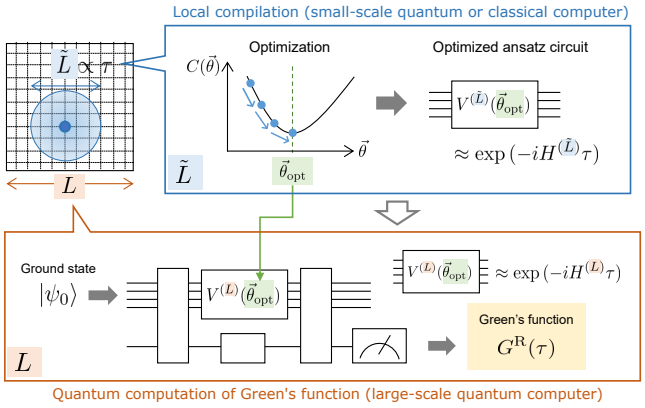


Figure 1: Overview of the LVQC approach to computing the Green's function.

In the following, we briefly summarize our proposal and results. Details of this work are shown in Ref. [3].

2 Our Proposal

Here, we summarize our proposal. We aim to calculate the retarded Green's function at zero temperature, which is defined as

$$G^{\mathrm{R}}_{a,b}(t) = -i\Theta(t)\,\langle\psi_0|\,\{e^{iHt}c_a e^{-iHt}, c_b^\dagger\}\,|\psi_0\rangle\,, \quad (1)$$

where t denotes time, $\Theta(t)$ is the Heaviside step function, $|\psi_0\rangle$ is the ground state of the target Hamiltonian H, and c_a ($c_b^\dagger$) is fermionic annihilation (creation) operator with a (b) denoting the fermionic mode. Assuming that the ground state $|\psi_0\rangle$ can be prepared on a quantum computer with high accuracy, the Green's function in Eq. (1) can be evaluated using a quantum computer if we can implement the time evolution operator e^{-iHt} as a quantum circuit [4]. However, it is generally difficult to efficiently implement the time evolution operator for a large-scale quantum system on a quantum computer, and hence some sophisticated approximation method is needed. Our proposal is to compute the Green's function by utilizing an approximate time evolution circuit constructed by the LVQC algorithm. To simulate the Green's function of a large class of quantum many-body systems using the LVQC, we also performed a theoretical formulation of the LVQC algorithm for fermionic systems that was not provided in the original paper [1].

The overview of our protocol is illustrated in Fig. 1. When simulating the Green's function for a L-site system with the Hamiltonian $H^{(L)}$ using the ansatz $V^{(L)}(\vec{\theta})$, our quantum-classical hybrid algorithm proceeds as follows.

1. Define a local subsystem consisting of $\tilde{L}(< L)$ sites and specify the local Hamiltonian $H^{(\tilde{L})}$, its time evolution operator $U^{(\tilde{L})}(\tau) = e^{-iH^{(\tilde{L})}\tau}$ with τ being a fixed time, and the local ansatz $V^{(\tilde{L})}(\vec{\theta})$. Here, the compilation size $\tilde{L} \propto \tau$ is determined by the LR bound of the target system.

2. Optimize the parameters $\vec{\theta}$ to minimize a cost function, which describes the difference between $U^{(\tilde{L})}(\tau)$ and $V^{(\tilde{L})}(\vec{\theta})$, by using the local $\tilde{L}$-sites system. Then, we obtain an optimal parameter $\vec{\theta}_{\mathrm{opt}}$ that realizes $V^{(\tilde{L})}(\vec{\theta}_{\mathrm{opt}}) \approx e^{-iH^{(\tilde{L})}\tau}$.

3. Calculate the Green's function for L-site total systems by using the ground state $|\psi_0\rangle$ and the optimized circuit $V^{(L)}(\vec{\theta}_{\mathrm{opt}})$ that approximates the time evolution operator at size L, i.e., $V^{(L)}(\vec{\theta}_{\mathrm{opt}}) \approx e^{-iH^{(L)}\tau}$. A long-time-scale dynamics at $t = n\tau$ $(n = 1, 2, 3, ...)$ can also be calculated by adopting $(V^{(L)}(\vec{\theta}_{\mathrm{opt}}))^n \approx e^{-iH^{(L)}n\tau}$.

3 Numerical Demonstration

To demonstrate the validity of our method, we performed numerical simulations for the Fermi-Hubbard model using a classical simulator. Here, we show some major results.

In this work, we investigate the accuracy of our LVQC method for various lattice sizes. As a representative result, Fig. 2 shows the Green's function and DOS of the 4×4 site lattice (32 qubits) Fermi-Hubbard model obtained by the LVQC method. Here, the DOS is an important quantity for studying the spectral properties of materials, which is calculated through the Fourier transform of the Green's function. The LVQC protocol is performed through the minimization of the cost function at the compilation size $\tilde{L} = 2 \times 2$. We note that this result is the largest-scale numerical demonstration for our method because the 4×4 site lattice is on the border of the limit of the system size in which the Green's function can be calculated by using a classical computer. In Fig. 2, we compare the results of the LVQC algorithm with the exact value shown in Ref. [5], in which the DOS is obtained by the Lanczos method. The LVQC algorithm nicely reproduces the overall peak structure of the exact DOS of the 4×4 lattice Fermi-Hubbard model. Since it is known that the accuracy of the LVQC algorithm is hardly altered with increasing the system size [1] (see also Fig. 3), we expect that the LVQC method is also valid for accurately calculating the Green's function of the classically-intractable size of lattice more than 4×4 lattice.

We also investigate the parameter dependencies of the accuracy of our LVQC method and compare it with that of a standard approach based on the Trotter decomposition. As a representative result, we here show the dependence of the accuracy on the system size L by fixing the compilation size $\tilde{L}$. Figure 3 shows the absolute error of the Green's function for the $L \times 1$ site lattice Fermi-Hubbard model as a function of the lattice size L. The LVQC results (orange diamonds) are obtained by executing the LVQC protocol with the compilation size $\tilde{L} = 2 \times 1$, while the Trotter results (blue dots) are obtained by calculating the Green's function using the approximate time evolution operator based on the Trotter decomposition. We see that the absolute error of the LVQC method is much smaller than that of the Trotter decomposition in a wide range of L. In addition, the

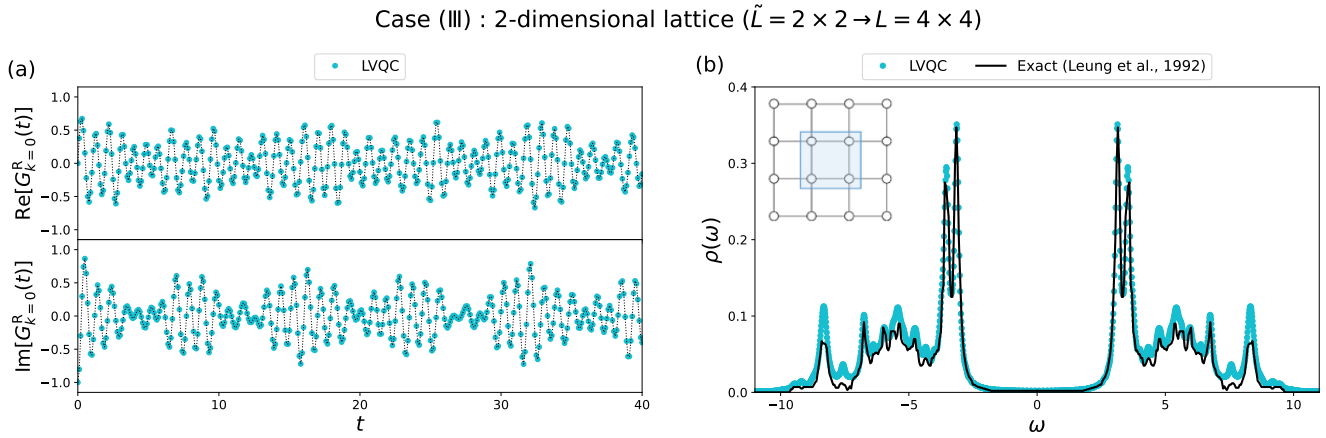


Figure 2: (a) Green's function and (b) DOS for 4×4 site lattice Fermi-Hubbard model. The cyan circles represent the results obtained by the LVQC method. The black line in panel (b) represents the exact DOS taken from Ref. [5].

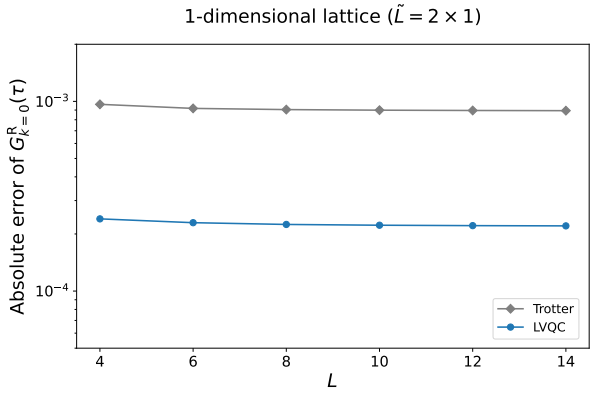


Figure 3: The absolute error of the Green's function for the $L \times 1$ site lattice Fermi-Hubbard model. The orange diamonds and blue dots represent the results obtained by the LVQC method and Trotter decomposition method, respectively.

absolute error for both LVQC and Trotter decomposition are hardly altered with increasing the system size L owing to the LR bound of the Fermi-Hubbard model. This is a remarkable result indicating that the LVQC method enables us to compute the Green's function more accurately than the standard Trotter decomposition irrespective of the system size.

4 Resource Estimation

To discuss the feasibility of the LVQC method from the viewpoint of the computational resource, we estimated the gate count needed to compute the spectral function (i.e., the Fourier transform of the Green's function) for a given precision. Specifically, based on the formal gate count of the quantum circuit needed to compute the Green's function and scaling of errors assessed from numerical results, we estimated the number of single-qubit rotation R_z gates and CNOT gates needed to simulate the Fermi-Hubbard model. Although a very conservative estimation, we find that the LVQC (Trotter decomposition) method requires about 7.8×10^7 (1.7×10^8) R_z gates and 2.1×10^9 (4.7×10^9) CNOT gates to simulate the 20×20 site lattice model, in which calculating the Green's function with high accuracy is not achievable by classical computers. Hence, the number of both R_z and CNOT gates required for the LVQC method is at most less than half of that for the Trotter decomposition method. This result suggests that the LVQC method is more practical than Trotter decomposition regarding gate counts even in classically-intractable size systems.

5 Summary

In summary, we proposed an efficient method to compute the Green's function on quantum computers by utilizing the LVQC algorithm. The Green's function is computed through the measurements of the time evolution circuit prepared by the LVQC protocol. By performing the numerical simulation for the Fermi-Hubbard model, we verified that our LVQC method is able to compute the Green's function more efficiently and accurately compared to a standard approach based on the Trotter decomposition. The formal estimation of the gate count also indicates that our LVQC method has a practical advantage against the Trotter decomposition.

The proposed method enables us to accurately simulate the Green's function for large-scale quantum many-body systems using fewer computational resources compared to standard approaches. This method will bring remarkable progress toward the realization of quantum advantage for practical problems in condensed matter physics, quantum chemistry, and material sciences.

References

[1] K. Mizuta, Y. O. Nakagawa, K. Mitarai, and K. Fujii. Local Variational Quantum Compilation of Large-Scale Hamiltonian Dynamics. PRX Quantum 3, 040302 (2022).

[2] E. H. Lieb and D. W. Robinson. The finite group velocity of quantum spin systems. Commun. Math. Phys. 28, 251 (1972).

[3] S. Kanasugi, S. Tsutsui, Y. O. Nakagawa, K. Maruyama, H. Oshima, and S. Sato. Computation of Green's function by local variational quantum compilation. arXiv:2303.15667.

[4] S. Endo, I. Kurata, and Y. O. Nakagawa. Calculation of the Green's function on near-term quantum computers. Phys. Rev. Res. 2, 033281 (2020).

[5] P. W. Leung, M. A. Novotny, and P. E. Oppenheimer. Density of states of the two-dimensional Hubbard model on a 4×4 lattice. Phys. Rev. B 46, 11779 (1992).

Activating strong nonlocality from local sets: An elimination paradigm

Subhendu B. Ghosh[1] * Tathagata Gupta[1] † Ardra A V[2] ‡
Anandamay Das Bhowmik[1] § Sutapa Saha[3] ¶ Tamal Guha[4] ‖ Amit Mukherjee[5] **

[1] *Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India*
[2] *School of Physics, IISER Thiruvananthapuram, Vithura, Kerala 695551, India*
[3] *Department of Astrophysics and High Energy Physics, S.N. Bose National Center for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata 700106, India*
[4] *Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong*
[5] *Department of Physics of Complex Systems, S.N. Bose National Center for Basic Sciences, Block JD, Sector III, Salt Lake, Kolkata 700106, India*

Abstract. Apart from the Bell nonlocality, quantum theory exhibits another kind of nonlocality that involves the indistinguishability of the locally preparable set of multipartite states. While activation of Bell nonlocality from local correlations via local operations and shared randomness is impossible, the activation of the latter kind of nonlocality has already been reported. This work observes that a stronger notion of such a nonlocality, that deals with elimination instead of discrimination, can be activated from locally preparable multipartite systems. Moreover, none of the transformed product states can be eliminated, even if all but one of the parties come together.

Keywords: Quantum nonlocality, activation, local operations and classical communication, local indistinguishability, local irreducibility, genuinely nonlocal product states.

The celebrated notion of Bell nonlocality excludes any local-realistic description to substitute multipartite quantum correlations. This, in turn, identifies quantum correlations to be advantageous over its classical counterpart in several practical applications. However, quantum theory admits a more elegant nonclassicality in question of state discrimination. While given a single copy every pure classical preparation can be distinctively identified from the others, only the orthogonal quantum states are perfectly distinguishable with a single copy. The volume of such distinguishable states for multipartite scenario gets further decreased under the limited measurement setting, like, local operations and classical communication (LOCC). Unlike most of the nonclassical aspects of quantum correlations obtained from the entangled quantum states, Bennett *et. al.* [1] first reported that the LOCC indistinguishablity holds even for orthogonal product quantum states and coined the term *quantum nonlocality without entanglement* for such a phenomenon. Consequently, a plethora of important works have been carried out in this direction which have significant importance to understand the complex topology of quantum state space structure.

Limits on state discrimination in quantum theory further give rise to several interesting questions in the context of state elimination, where instead of identification the main goal is to rule out one or more quantum states from an ensemble of consideration. Now, if the performed measurement preserves the orthogonality of the remaining states after elimination, then it is further helpful in the context of the state discrimination. Motivated by this fact, recently Halder *et al.* have introduced a stronger notion of quantum nonlocality for product states which forbids elimination of any of the product states of a set under orthogonality preserving local measurements (OPLM or local OPM) [2]. Consequently, these fueled a plethora of interesting studies in the recent past from the stronger perspective of state indistinguishability, i.e. irreducibility under OPLM [3, 4, 5, 6, 7].

Apart from revealing the elegant intricacies of state space structure, local indistinguishability and irreducibility of quantum states also indicate the prospect of locking of information such that unlocking requires entangled resources. This characteristic certainly has a crucial significance in various quantum cryptographic schemes, viz., secret sharing, data hiding [8, 9, 10, 11, 12]. However, in the practical settings, the complexity to retrieve a hidden information should depend on their mutual trustworthiness. Also, it might be important for one of those agents to manipulate the complexity should their mutual trustworthiness change after they have shared the secret with each other. For instance, consider three agents Alice, Bob and Charlie who agree to share a LOCC distinguishable quantum secret at first. However, in time, Charlie may distrust others and want to update the complexity of the secret, upon which the revealing of the secret must demand all of them to be in the same lab. This motivates to propose another version of quantum nonclassicality, which deals with the activation of quantum nonlocality from the locally distinguishable quantum states. The framework has

* subhendubghosh@gmail.com
† tathagatagupta@gmail.com
‡ ardraav17@iisertvm.ac.in
§ ananda.adb@gmail.com
¶ sutapa.gate@gmail.com
‖ g.tamal91@gmail.com
** amitisiphys@gmail.com

recently been reported in [13] for initially distinguishable entangled states and in [14] for product states.

Note that, the task can be trivially accomplished if the agents flag the indistinguishable ensemble with a distinguishable one and according to the trust update Charlie can discard his distinguishable share. This redundancy is termed as activation of nonlocality from a locally redundant set [13].

In the present work, we have dealt with the activation of a stronger nonlocality from a set of locally distinguishable product states, which are also free from local redundancy. Precisely speaking, besides the nonlocal aspects of state discrimination, we have further considered a stronger version, which is related to the impossibility of state elimination instead of state discrimination. Notably, the authors in [14] have introduced a similar notion for bipartite product states, however with a higher dimensional quantum states. In contrast, we have shown that such a feature is generic even in the smaller dimensional quantum systems. Further, we have extended this activation phenomenon in the multipartite scenario and have come up with the strongest possible nonlocality activation. In particular, performing OPLM on a $\mathbb{C}^{6^{\otimes 3}}$ we transform the set to tripartite orthogonal product qutrits, which is locally irreducible even if all but one players come together. Lastly, motivated by the practical situation of trust-updated secret sharing we propose a set of states in $\mathbb{C}^{3^{\otimes 2}} \otimes \mathbb{C}^6$, which can be used to activate strongest form of nonlocality by performing an OPLM only at the third party's possession. Besides, our results also draw a significant difference between the two types of quantum nonlocality – while the Bell nonlocality can not be activated from a shared local correlations [15, 16], the stronger version of nonlocality related to state identification can be activated from locally distinguishable product states. Importantly, our last example activates the strongest possible version of nonlocality without involving any communication between the parties.

In the following we present the statement of our final result which depicts the example of activation of strong quantum nonlocality without entanglement from a local set by performing a measurement on the possession of a single party. This has vivid importance in the framework of data hiding and secret sharing between all but one untrusted party. Precisely speaking, in such a scenario the particular trusted agent (personified as Charlie) has full authority to judge how trustworthy are the other parties and depending upon that he may compel others to meet him in person to decode a hidden secret. As an example consider the following set $\mathcal{G}_4$ of 27 orthogonal product states $|\zeta_i^{\pm}\rangle\,,\ i \in \{1, \cdots, 4, 6, \cdots, 9, 11, \cdots, 14\}$ and $|\zeta_j\rangle\,,\ j \in \{5, 10, 15\}$ in $\mathbb{C}^{3^{\otimes 2}} \otimes \mathbb{C}^6$,

$$
\begin{aligned}
|\zeta_1^{\pm}\rangle &= |0\rangle\,|1\rangle\,|\mathbf{0} \pm \mathbf{1} + \mathbf{4} \pm \mathbf{5}\rangle \\
|\zeta_2^{\pm}\rangle &= |0\rangle\,|2\rangle\,|\mathbf{0} \pm \mathbf{2} + \mathbf{4} \pm \mathbf{3}\rangle \\
|\zeta_3^{\pm}\rangle &= |1\rangle\,|2\rangle\,|\mathbf{0} \pm \mathbf{1} + \mathbf{4} \pm \mathbf{5}\rangle \\
|\zeta_4^{\pm}\rangle &= |2\rangle\,|1\rangle\,|\mathbf{0} \pm \mathbf{2} + \mathbf{4} \pm \mathbf{3}\rangle \\
|\zeta_5\rangle &= |0\rangle\,|0\rangle\,|\mathbf{0} - \mathbf{4}\rangle \\
|\zeta_6^{\pm}\rangle &= |1\rangle\,|0 \pm 1\rangle\,|\mathbf{0} - \mathbf{4}\rangle \\
|\zeta_7^{\pm}\rangle &= |2\rangle\,|0 \pm 2\rangle\,|\mathbf{0} - \mathbf{4}\rangle \\
|\zeta_8^{\pm}\rangle &= |2\rangle\,|0 \pm 1\rangle\,|\mathbf{1} - \mathbf{5}\rangle \\
|\zeta_9^{\pm}\rangle &= |1\rangle\,|0 \pm 2\rangle\,|\mathbf{2} - \mathbf{3}\rangle \\
|\zeta_{10}\rangle &= |1\rangle\,|1\rangle\,|\mathbf{1} - \mathbf{5}\rangle \\
|\zeta_{11}^{\pm}\rangle &= |0 \pm 1\rangle\,|0\rangle\,|\mathbf{1} - \mathbf{5}\rangle \\
|\zeta_{12}^{\pm}\rangle &= |0 \pm 2\rangle\,|0\rangle\,|\mathbf{2} - \mathbf{3}\rangle \\
|\zeta_{13}^{\pm}\rangle &= |0 \pm 1\rangle\,|1\rangle\,|\mathbf{2} - \mathbf{3}\rangle \\
|\zeta_{14}^{\pm}\rangle &= |0 \pm 2\rangle\,|2\rangle\,|\mathbf{1} - \mathbf{5}\rangle \\
|\zeta_{15}\rangle &= |2\rangle\,|2\rangle\,|\mathbf{2} - \mathbf{3}\rangle
\end{aligned}
$$

Proposition 1 *The set $\mathcal{G}_4$ is distinguishable under LOCC and free from local redundancy.*

However,

Theorem 2 *The set $\mathcal{G}_4$ can be deterministically transformed, via a single local OPM at Charlie's site, to an orthogonal set of tripartite product states which is locally irreducible in every bipartition.*

Besides its foundational interest to understand the topology of the state spaces of composite quantum systems, our work deserves significant importance from the practical perspective. It has mimicked an interesting framework of secured data hiding between several parties, where the distributor is flexible to update the distinguishibility of the secured data hidden in the correlation of the given states. Recently, we have also extended this approach of genuinely activating quantum nonlocality to show [17] generation of some novel and stronger resources, for example, local quantum state unmarkability [18].

The current work was published in Physical Review A as a Letter [19].

References

[1] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Quantum Nonlocality without Entanglement, Phys. Rev. A 59, 1070 (1999).

[2] S. Halder, M. Banik, S. Agrawal, and S. Bandyopadhyay, Strong Quantum Nonlocality without Entanglement, Phys. Rev. Lett. 122, (2019).

[3] S. Rout, A. G. Maity, A. Mukherjee, S. Halder, and M. Banik, Genuinely Nonlocal Product Bases: Classification and Entanglement-Assisted Discrimination, Phys. Rev. A 100, (2019).

[4] S. Rout, A. G. Maity, A. Mukherjee, S. Halder, and M. Banik, Multiparty Orthogonal Product States with Minimal Genuine Nonlocality, Phys. Rev. A 104, (2021).

[5] Z.-C. Zhang and X. Zhang, Strong Quantum Nonlocality in Multipartite Quantum Systems, Phys. Rev. A 99, (2019).

[6] F. Shi, M. Hu, L. Chen, and X. Zhang, Strong Quantum Nonlocality with Entanglement, Phys. Rev. A 102, (2020).

[7] P. Yuan, G. Tian, and X. Sun, Strong Quantum Nonlocality without Entanglement in Multipartite Quantum Systems, Phys. Rev. A 102, (2020).

[8] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Hiding Bits in Bell States, Phys. Rev. Lett. 86, 5807 (2001).

[9] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, Quantum Data Hiding, IEEE Trans. Inform. Theory 48, 580 (2002).

[10] T. Eggeling and R. F. Werner, Hiding Classical Data in Multipartite Quantum States, Phys. Rev. Lett. 89, (2002).

[11] W. Matthews, S. Wehner, and A. Winter, Distinguishability of Quantum States Under Restricted Families of Measurements with an Application to Quantum Data Hiding, Commun. Math. Phys. 291, 813 (2009).

[12] D. Markham and B. C. Sanders, Graph States for Quantum Secret Sharing, Phys. Rev. A 78, (2008).

[13] S. Bandyopadhyay and S. Halder, Genuine Activation of Nonlocality: From Locally Available to Locally Hidden Information, Phys. Rev. A 104, (2021).

[14] M.-S. Li and Z.-J. Zheng, Genuine Hidden Nonlocality without Entanglement: From the Perspective of Local Discrimination, New J. Phys. 24, 043036 (2022).

[15] M. Forster, S. Winkler, and S. Wolf, Distilling Nonlocality, Phys. Rev. Lett. 102, (2009).

[16] S. G. A. Brito, M. G. M. Moreno, A. Rai, and R. Chaves, Nonlocality Distillation and Quantum Voids, Phys. Rev. A 100, (2019).

[17] Tathagata Gupta, Subhendu B. Ghosh, Ardra A V, Anandamay Das Bhowmik, Sutapa Saha, Tamal Guha, Ramij Rahaman, and Amit Mukherjee. Hierarchical activation of quantum nonlocality: Stronger than local indistinguishability. Accepted in Physical Review A.

[18] S. Sen et al., Local Quantum State Marking, Phys. Rev. A 105, (2022).

[19] S. B. Ghosh, T. Gupta, A. A. V., A. Das Bhowmik, S. Saha, T. Guha, and A. Mukherjee, Activating Strong Nonlocality from Local Sets: An Elimination Paradigm, Phys. Rev. A 106, (2022).

Generalizing Hardy's paradox by means of the failure of transitivity of implications

Kai-Siang Chen[1] Shiladitya Mal[1 2] Gelo Noel M. Tabia[1 2] Yeong-Cherng Liang[1 2 *]

[1] *Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan*

[2] *Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan*

Abstract. As with a Bell inequality, Hardy's paradox manifests a contradiction between the prediction given by quantum theory and local-hidden variable theories. In this work, we generalize Hardy's arguments to an *arbitrary* Bell scenario involving two observers. Our construction reduces to that of Meng *et al.* [Phys. Rev. A. **98**, 062103 (2018)] and can be naturally interpreted as a demonstration of the failure of the transitivity of implications. Our generalization is equivalent to a ladder-proof-type argument for Hardy's paradox. Furthermore, it provably exhibits a higher success probability compared with existing proposals. Moreover, this advantage persists even if we allow imperfections in realizing zero-probability constraints in such paradoxes.

Keywords: Bell-nonlocality, Hardy's paradox, nonlocality without inequality, logical contradiction, quantum correlations, entanglement

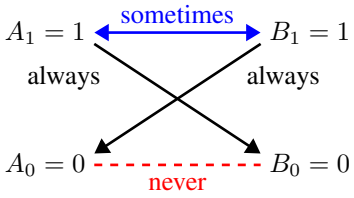
In 1964, Bell [1] showed that local-hidden-variable (LHV) models *cannot* reproduce all quantum-mechanical predictions. In particular, he demonstrated how, with the help of so-called Bell inequalities, one can experimentally falsify the predictions of LHV models. Nowadays, we know that Bell-nonlocality not only opens the door to answer fundamental questions in physics, but also serves as an important resource for device-independent quantum information [3, 14].

Interestingly, Bell inequalities are not the only way to manifest Bell-nonlocality. Hardy's proof of nonlocality [7], on the other hand, manifests the incompatibility between quantum mechanics and LHV models using logical arguments. Consider the simplest Bell scenario, i.e., one in which two observers each perform two binary-outcome measurements. Let x and y (a and b) represent, respectively, the setting/input (outcome/output) of Alice and Bob side, and A_x (B_y) denotes the outcome of Alice (Bob) when given input x (y). The probability distribution $\{P(a,b|x,y) = P(A_x, B_y)\}$ admissible in LHV models can be described by convex mixtures of local deterministic strategies $\{A_x = f_A(x,\lambda), B_y = f_B(y,\lambda)\}$, where f_A (f_B) is a deterministic function of the input x (y) and LHV λ. The Hardy paradox of [7] is encapsulated by:

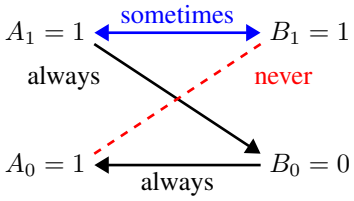
$$\begin{aligned} P(0,0|0,0) &= 0, \quad P(1,1|0,1) = 0, \\ P(1,1|1,0) &= 0, \quad P(1,1|1,1) = q > 0. \end{aligned} \tag{1}$$

In LHV models, the equality constraints of Eq. (1) imply $P(1,1|1,1) = 0$, which contradicts the inequality constraint of Eq. (1). Interestingly, Stapp [15] showed that Eq. (1) can also be interpreted as the failure of the transitivity implications (FTI), thereby demonstrating Bell-nonlocality (see also [10]). The probability $P(1,1|1,1)$ in Eq. (1) is called as the success probability, as the corresponding event facilitates (initiates) the chain or logical reasoning in Hardy's (Stapp's) arguments, see Fig. 1.

Later, inspired by Cabello's idea [4], Liang and Li [9] (see also [8]) generalized Hardy argument beyond the simplest Bell scenario, yielding a larger success probability. In this work, we propose a different generalization of Hardy's arguments. As with Stapp's work and its extension discussed in [11], our arguments can be interpreted as a demonstration of FTI. Not only that our generalization gives rise to a higher success probability than that offered by these previous proposals, it also recovers them as special cases.



(a) Hardy's argument



(b) Stapp's argument

Figure 1: Comparison of the logical reasoning between Hardy's original argument and Stapp's reformulation.

Let us first focus on the simplest Clauser-Horne-Shimony-Holt (CHSH) Bell scenario. For the original Hardy paradox, it is known [13] that the maximal success probability $P(1,1|1,1)$ attainable in quantum theory is approximately 9.02%. In [8], the authors showed that the success probability of the so-called Cabello's argument [9], defined by $q - p$ in Eq. (2) can reach $\approx 10.79\%$:

$$\begin{aligned} P(0,0|0,0) &= p, \quad P(1,1|0,1) = 0, \\ P(1,1|1,0) &= 0, \quad P(1,1|1,1) = q, \end{aligned} \tag{2}$$

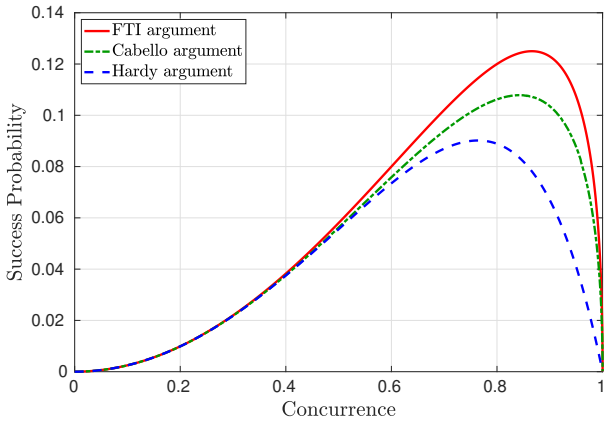
In contrast, the conditions of our FTI argument are encapsulated by:

$$\begin{aligned} P(0,0|0,0) &= 0, \quad P(1,1|0,1) = r, \\ P(1,1|1,0) &= 0, \quad P(1,1|1,1) = q, \end{aligned} \tag{3}$$

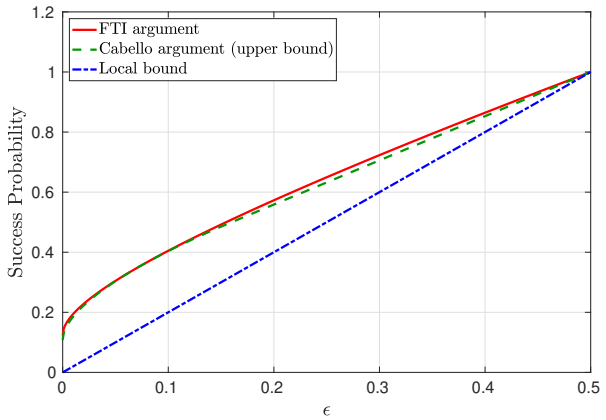
where the success probability $q - r$ can reach 12.5%, higher

*ycliang@mail.ncku.edu.tw

than that achievable in Eq. (2). In fact, this advantage holds even if we relax the zero constraints in both arguments. Moreover, we show that for most of the partially entangled 2-qubit state, the maximum success probability given by the FTI argument can be higher than the maximum ones given by Hardy's and Cabello's arguments, see Fig. 2.



(a) The maximal success probability as a function of the concurrence [16] (an entanglement measure) of the two qubit pure states used for their manifestation. From top to bottom, we give the plot of $q - r$ in Eq. (3) for our FTI argument (red, solid), $q - p$ in Eq. (2) for so-called Cabello's argument (green, dashed-dotted), and q in Eq. (1) for Hardy's argument (blue, dashed).



(b) Success probability as a function of the deviation ϵ from the zero-probability constraints. The top curve (red, solid) shows a lower bound on the success probability for our FTI argument (obtained by optimizing over two-qubit states and projective measurements); the middle curve (green, dashed) shows an upper bound on the success probability for so-called Cabello's argument obtained using level-3 of the semidefinite programming (SDP) hierarchy described in Ref. [12].

Figure 2: Comparison of the maximal success probability for various Hardy-type paradoxes in the CHSH Bell scenario.

In Figs. 3 and 4, we schematically illustrate the form of four Hardy-type arguments in an arbitrary k-input d-output Bell scenario. Note that Hardy's argument with $d = 2$-output is also known as the ladder proof of nonlocality in [2]. The generalization of Stapp's argument for the $k = d = 2$ Bell scenario to cases of arbitrary $k, d \geq 2$ has been discussed in [11], and was coined Stapp's argument therein. Here, we adopt the same terminology in the following discussion. Interestingly, the authors of [11] proved that in the k-input 2-output scenario, these generalized Stapp's argument and the ladder proof of nonlocality are equivalent. In the companion technical draft, we show the following theorem, which establishes this equivalence for arbitrary $k, d \geq 2$.

Theorem 1 *For any bipartite Bell scenario, Hardy's argument is equivalent to Stapp's argument.*

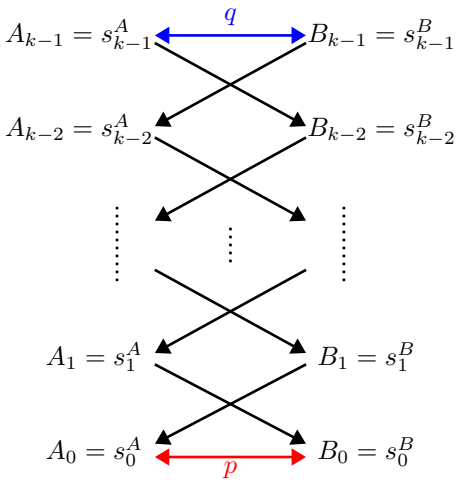


Figure 3: Logical structure of generalized Cabello's arguments in the k-input d-output Bell scenario. The generalization of Hardy's original argument to these cases is recovered by setting $p = 0$.

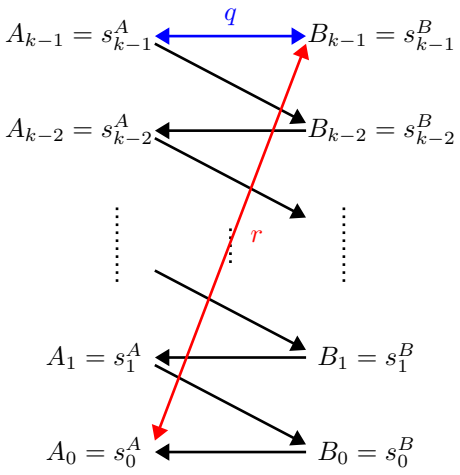
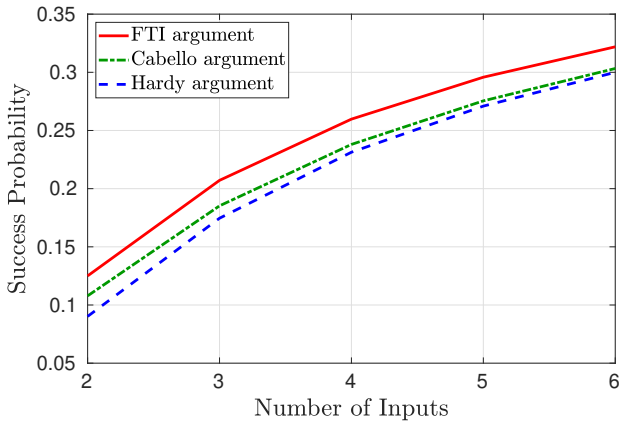


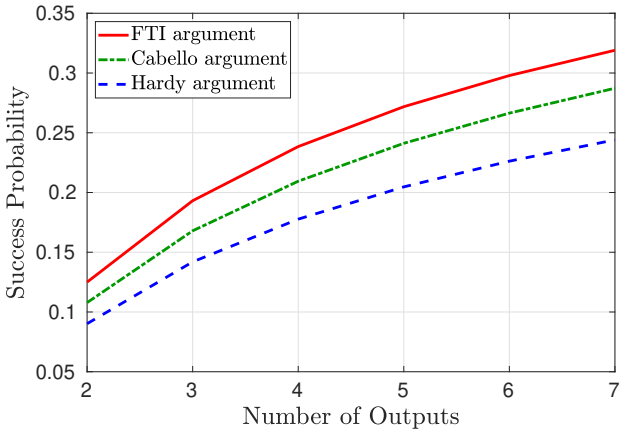
Figure 4: Logical structure of our FTI arguments in the k-input d-output Bell scenario. Generalized Stapp's arguments as introduced in [11] are recovered by setting $r = 0$.

For completeness, we show also in Fig. 5 a comparison of the success probability of our FTI argument (generalized to a general bipartite Bell scenario) against the other generalizations [2, 5, 6]. Again, our success probability is always higher than that obtained from all these other proposals.

As with [11], we propose a generalization of Hardy's paradox tailored for quantum systems of arbitrary Hilbert space dimension and that involves an arbitrary number of measurement settings. Our generalization gives a type of proof of



(a) For Bell scenarios with more inputs, we plot here the results corresponding to the ladder proof of nonlocality without inequality [2, 11] (blue, dashed), the generalization of Cabello's argument due to [5] (green, dashed-dotted), and our FTI argument (red, solid).



(b) For Bell scenarios with more outputs, we plot here the results corresponding to a generalization of Hardy's proof given by Chen *et al.* [6] (blue, dashed), and our analogous generalization of Cabello's argument Eq. (2) (green, dashed-dotted), as well as our FTI argument (red, solid).

Figure 5: Comparison of the upper bound on the success probability for three different Hardy-type paradoxes beyond the CHSH scenarios. Due to Theorem 1, generalized Stapp's arguments and generalized Hardy's arguments give the same success probability. These upper bounds were obtained by considering on level-1 of the SDP hierarchy introduced in Ref. [12].

Bell-nonlocality via the failure of the transitivity of implications. Moreover, its success probability is provably higher than all the other arguments [2, 5, 6, 11] proposed to date, even in the presence of a deviation from the hard-to-realize zero constraints. In the CHSH Bell scenario, we analytically derive the maximum success probability and corresponding measurements for any given partially entangled state. We also demonstrate that our generalization yields success probabilities at least as high, if not higher, compared to both Hardy's and Cabello's arguments.

For further details on our results, please see the attached technical draft.

Acknowledgements This work is supported by the National Science and Technology Council (formerly Ministry of Science and Technology), Taiwan (Grants No. 109-2112-M006-010-MY3).

References

[1] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics*, 1:195–200, Nov 1964.

[2] D. Boschi, S. Branca, F. De Martini, and L. Hardy. Ladder proof of nonlocality without inequalities: Theoretical and experimental results. *Phys. Rev. Lett.*, 79:2755–2758, Oct 1997.

[3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, Apr 2014.

[4] A. Cabello. Bell's theorem with and without inequalities for the three-qubit Greenberger-Horne-Zeilinger and w states. *Phys. Rev. A*, 65:032108, Feb 2002.

[5] J. L. Cereceda. Cabello's nonlocality for generalized three-qubit ghz states. *Quantum Studies: Mathematics and Foundations*, 4(3):205–215, 2017.

[6] J.-L. Chen, A. Cabello, Z.-P. Xu, H.-Y. Su, C. Wu, and L. C. Kwek. Hardy's paradox for high-dimensional systems. *Phys. Rev. A*, 88:062116, Dec 2013.

[7] L. Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Phys. Rev. Lett.*, 71:1665–1668, Sep 1993.

[8] S. Kunkri, S. K. Choudhary, A. Ahanj, and P. Joag. Nonlocality without inequality for almost all two-qubit entangled states based on Cabello's nonlocality argument. *Phys. Rev. A*, 73:022346, Feb 2006.

[9] L.-M. Liang and C.-Z. Li. Nonlocality without inequalities for two-qubit mixed states based on Cabello's nonlocality. *Phys. Lett. A*, 335(5):371–373, 2005.

[10] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman. Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Phys. Rep.*, 506(1–2):1 – 39, 2011.

[11] H.-X. Meng, J. Zhou, Z.-P. Xu, H.-Y. Su, T. Gao, F.-L. Yan, and J.-L. Chen. Hardy's paradox for multisetting high-dimensional systems. *Phys. Rev. A*, 98:062103, Dec 2018.

[12] T. Moroder, J.-D. Bancal, Y.-C. Liang, M. Hofmann, and O. Gühne. Device-independent entanglement quantification and related applications. *Phys. Rev. Lett.*, 111:030501, Jul 2013.

[13] R. Rabelo, L. Y. Zhi, and V. Scarani. Device-independent bounds for Hardy's experiment. *Phys. Rev. Lett.*, 109:180401, Oct 2012.

[14] V. Scarani. The device-independent outlook on quantum physics. *Acta Physica Slovaca*, 62(4):347, 2012.

[15] H. P. Stapp. *Mind, Matter and Quantum Mechanics*. Springer Verlag, 1993.

[16] W. K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, Mar 1998.

Unbounded Quantum Advantage in One-Way Strong Communication Complexity of a Distributed Clique Labelling Relation

Sumit Rout[1] * Nitica Sakharwade[1] † Some Sankar Bhattacharya [1] ‡
Ravishankar Ramanathan[2] § Paweł Horodecki[1] ¶

[1] *International Centre for Theory of Quantum Technologies (ICTQT), University of Gdańsk, Jana Bażynskiego 8, 80-309 Gdańsk, Poland*
[2] *Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong*

Abstract. We investigate two variants of the one-way zero-error classical and quantum communication complexities for a class of relations induced by a distributed clique labelling problem- where the receiver must output a valid answer (Communication Complexity of Relations- CCR) or all valid answers belonging to the relation (Strong-CCR). We find that CCR for a class of graphs does not provide quantum advantage, but for Strong-CCR the separation between classical and quantum communication grows unboundedly with the order of the graph. Similar result follows for the case with shared randomness. Applications include semi-device-independent dimension witnessing and the detection of Mutually Unbiased Bases.

Keywords: Quantum communication complexity, Orthogonality graphs, Orthogonal representation.

This extended abstract is based on the preprint arXiv:2305.10372 [1].

1 Introduction

Quantum Shannon theory replaces the classical carrier of information with quantum systems in Shannon's model of communication [2]. This initiated a tide of research to understand the advantage of encoding classical information in a quantum system while considering various information-theoretic scenarios. In prepare and measure scenarios, advantages in quantum communication complexity have been extensively explored[3, 4], which involve computing the minimum communication required between distant parties in order to perform the distributed computation of functions [5].

We consider a generalisation of communication complexity of functions to relation. A relation $\mathcal{R}$ over a bipartite prepare and measure scenario is a subset $\mathcal{R} \subseteq X \times Y \times B$, where X and Y are the set of possible input values of Alice and Bob respectively and B is the set of possible output values by Bob. The one-way communication complexity of a relation (CCR) $\mathcal{R}$ is the minimum communication that Alice requires to make with Bob for any input variables $x \in X$ and $y \in Y$ such that Bob's output b gives the tuple $(x, y, b) \in \mathcal{R}$. Additionally, we consider only *zero-error* scenario, *i.e.* probability $P(b|x, y) = 0$ whenever $(x, y, b) \notin \mathcal{R}$ for all $(x, y) \in X \times Y$. In [6], Raz showed an instance of an exponential gap between the classical and quantum CCR for an infinite set of inputs, while we show an exponential gap for a finite set of inputs.

Another closely related line of study has been to explore the advantage of quantum communication in tasks based on orthogonality graphs. In most cases, orthogonality graphs that lead to quantum advantage are not Kochen-Specker (KS) colourable [7], thus connecting this set of tasks to the feature of quantum contextuality [8], while the advantage in our work does not rely on contextuality.

In this work, we introduce a new task based on the communication complexity of relations (CCR) which we call Strong Communication Complexity of Relations (S-CCR), which, unlike the communication complexity of functions, may have more than one correct answer for Bob. This stronger variation of CCR enforces that Bob outputs all correct answers over different rounds of the prepare and measure scenario. The one-way S-CCR of relation $\mathcal{R}$ is the minimum communication that Alice requires to make with Bob for any input variables $x \in X$ and $y \in Y$ such that Bob's output b gives the tuple $(x, y, b) \in \mathcal{R}$ and that Bob's output b in different rounds of the prepare and measure scenario spans all valid b for each (x, y) input. Alternatively, one may define S-CCR as having a strategy with $P(b|x, y) > 0$ whenever $(x, y, b) \in \mathcal{R}$ for all $(x, y) \in X \times Y$.

The aim of this task is to be able to decipher or *reconstruct* the relation $\mathcal{R}$ from the observed statistics. In the limit when the number of iterations tends to infinity, the observed statistics can be used to get the conditional output probability distribution $\{P(b|x, y)\}_{(x,y,b)\in\mathcal{R}}$. We can define a natural payoff for S-CCR as follows:

$$\mathcal{P}_{\mathcal{R}} = \min_{(x,y,b)\in\mathcal{R}} P(b|x, y). \tag{1}$$

One way to interpret this payoff $\mathcal{P}_{\mathcal{R}}$ is through its relation to the probability of success of reconstructing the relation $\mathcal{R}$, where a higher payoff implies reconstruction with less runs.

The relations we consider here are induced by rules of a distributed *clique labelling problem (CLP)* over a class of graphs. Consider a graph $(\mathcal{G}, \mathcal{V}, \mathcal{E})$ where $\mathcal{V}$ and $\mathcal{E}$ are the set of vertices and edges in the graph $\mathcal{G}$ respectively.

*rsumitrout3@gmail.com
†nitica.sakh@gmail.com
‡somesankar@gmail.com
§ravi@cs.hku.hk
¶pawhorod@pg.edu.pl

A maximum size complete sub-graph of $\mathcal{G}$ is called a maximum clique. Let the graph have n number of maximum cliques of size ω. Any binary colouring of a maximum clique assigns value 1 to exactly one vertex of the maximum clique and 0 to rest of the vertices. Each such binary colourings of a maximum clique represents *labellings* of this clique. Given a graph $\mathcal{G}$ described above, dis-

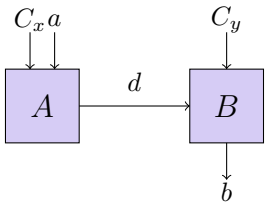


Figure 1: Given an graph $\mathcal{G}$, Alice's input is a maximum clique and a clique label, *i.e.* (C_x, a) and Bob's input is some maximum clique C_y. Bob must output a valid clique labelling b for his input clique such that $(C_x, a, C_y, b) \in \mathcal{R}_{CLP}(\mathcal{G})$. Alice can send a physical system of operational dimension d to Bob.

tributed clique labelling problem (CLP) is a prepare and measure scenario involving a referee and two spatially separated players, Alice and Bob. The referee shares the graph $\mathcal{G}$ with Alice and Bob at the beginning. The referee gives Alice the clique of size ω randomly chosen uniformly from the graph and a random possible *labelling (or binary colouring of vertices)* of the same clique, i.e., (C_x, a). The referee gives a clique C_y of size ω uniformly chosen from $\mathcal{G}$ to Bob as input. Bob must output a valid labelling of C_y which satisfies the constraints provided below, which will define the relation $\mathcal{R}_{CLP}(\mathcal{G})$ for the graph $\mathcal{G}$. We call this *Consistent Labelling of Pairwise Cliques*:

1. If Alice and Bob receive the same clique Bob's labelling should be identical to Alice's input label.
2. If Alice and Bob receive two different cliques sharing some vertices, the binary colouring of each shared vertex (0 or 1) by Bob should be identical to Alice's colouring.
3. If Alice and Bob receive two different cliques sharing some edges, the vertices belonging to an edge should not both have the binary colour 1.
4. In all other cases Bob can label the cliques independently of Alice's inputs.

Alice can use some communication (either classical or quantum) sent to Bob which we will optimise to find the communication complexity. For CCR of $\mathcal{R}_{CLP}(\mathcal{G})$, where any valid answer belonging to the relation is accepted, we show that there is no advantage in using quantum systems as carriers of information. However, S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$ where Bob's output in different runs should span over the relation, entails non-trivial quantum advantage.

Our main results consider two distinct scenarios of S-CCR depending on the availability of pre-shared correlations and direct communication resources between the two parties: (i) the spatially separated parties do not share any correlation, (ii) the communication channels can transmit systems of a fixed operational dimension. In the first scenario, we show that for the task S-CCR the separation between one-way classical and quantum communication grows with the order of the graph, specifically, the quantum complexity is $O(1)$ while the classical complexity is $\Omega(\log|\mathcal{V}|)$. We also demonstrate a quantum advantage for a relation induced by the class of *Payley graphs*. In the second scenario, we show that there exist communication tasks which imply classical channels require to be assisted by unbounded amounts of pre-shared classical correlations with a lower bound that is linear in the number of cliques while the quantum channel does not require any pre-shared resources. Additionally, we show that there exist graphs for which the task with a classical channel requires shared randomness (classical) linear in the number of cliques whereas only 1-ebit of shared entanglement assistance is sufficient.

2 Results

In the setup described earlier, Alice and Bob have access to a noiseless one-way communication channel of limited capacity and local sources of randomness (i.e. private coins) are considered to be free resources here. In the scenario when no pre-shared randomness (i.e. public coin) is allowed between Alice and Bob we calculate the necessary and sufficient classical and quantum resource required to perfectly satisfy the CCR for relation $\mathcal{R}_{CLP}(\mathcal{G})$ where graph $\mathcal{G}$ has n maximum cliques of size ω.

Theorem 1 *Given a graph $\mathcal{G}$, classical one-way zero-error communication complexity of $\mathcal{R}_{CLP}(\mathcal{G})$ is $\log_2\omega$ cbits.*

It can be shown that the quantum one-way CCR for $\mathcal{R}_{CLP}(\mathcal{G})$ is bounded from below by $\log_2\omega$ qubits. Thus, we observe no advantage of quantum communication resources over its classical analogue when considering CCR of $\mathcal{R}_{CLP}(\mathcal{G})$.

Consider some graph $\mathcal{G}$ as before, which satisfy the following conditions:

(G0): Each vertex of the graph is part of at least one maximum clique of the graph.

(G1): $\forall v, v' \in \mathcal{V}$ belonging to two different maximum cliques $\exists\ u \in \mathcal{V}$ such that u is either adjacent to v or v' but not both.

Given such graph $\mathcal{G}$, we prove a tight lower bound for classical and quantum resources required to win S-CCR for relation $\mathcal{R}_{CLP}(\mathcal{G})$. This bound is calculated for the *zero-error* scenario in which Bob should never output an outcome b such that the tuple consisting of Alice's and Bob's input, (C_x, a) and C_y respectively, and Bob's output does not belong to the relation $\mathcal{R}_{CLP}(\mathcal{G})$.

Lemma 1 *Given a graph $\mathcal{G}$ satisfying **(G0)-(G1)**, it is necessary and sufficient to communicate $\log_2|\mathcal{V}|$ cbits, where $|\mathcal{V}|$ is the order of the graph, to perform S-CCR of relation $\mathcal{R}_{CLP}(\mathcal{G})$.*

For this graph $\mathcal{G}$, one can assign complex vectors to each vertex such that these vectors are mutually orthogonal *iff* the corresponding vertices in graph $\mathcal{G}$ are adjacent. A collection of such vectors is called a *faithful orthogonal representation* of the graph $\mathcal{G}$.

Lemma 2 *Given a graph $\mathcal{G}$ satisfying **(G0)**-**(G1)** that has faithful orthogonal representation in minimum dimension $d_{\mathbb{C}}$, it is necessary and sufficient to communicate a $\log_2 d_{\mathbb{C}}$ qubits to perform S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$.*

Now using the above two results, we show that there exists an unbounded separation of quantum and classical communication resources in S-CCR of certain relations.

Theorem 2 *There exist a class of graphs such that the separation between one-way classical and quantum communication, required for zero-error reconstruction (S-CCR) of the given $\mathcal{R}_{CLP}$ induced by these graphs, is unbounded. (See [1])*

Given some graph $\mathcal{G}$ satisfying (**G0**)-(**G1**), there exist quantum advantage in S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$ if it has orthogonal representation in minimum dimension $d \leq |\mathcal{V}|$. As an example, we calculate the amount of quantum communication required for reconstruction of $\mathcal{R}_{CLP}$ for *Paley graphs*. For a Paley graph with q vertices, $\log_2 \frac{q+1}{2}$ qubit communication is sufficient for S-CCR and it yields a maximum payoff of $\frac{2}{\sqrt{q}+1}$, while $\log_2 q$ cbit communication is necessary and sufficient for non-zero payoff.

When Alice and Bob are allowed to have pre-shared correlations along with one-way direct communication resources we find that there exist graphs for which non-zero payoff while using restricted classical communication implies the presence of shared correlation. Also, allowing for pre-shared randomness removes the advantage of using quantum communication over classical communication in S-CCR of relation $\mathcal{R}_{CLP}(\mathcal{G})$. Alice needs $\log_2 \omega$ cbit or qubit communication with Bob for this task. Now we allow for fixed direct communication capacity, either quantum or classical, and compare the amount of shared randomness required to accomplish S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$.

Corollary 1 *Given a graph $\mathcal{G}$ satisfying **(G0)**-**(G1)**, with faithful orthogonal representation in minimum dimension ω, it is necessary to share randomness with n-inputs whlie communicating $\log_2 \omega$ cbit to accomplish S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$.*

We also provide a lower bound on the amount of shared randomness necessary to achieve maximum payoff for S-CCR of $\mathcal{R}_{CLP}(\mathcal{G})$ while using $\log_2 \omega$ cbits of communication. This lower bound has a connection with the existence of orthogonal arrays. Subsequently, when we consider pre-shared quantum correlations (quantum public coin) between Alice and Bob, we show that it can enhance classical communication more than classical public coin.

Theorem 3 *For classical communication with assistance from public coins, there exist graphs $\mathcal{G}$ satisfying conditions **(G0)**-**(G1)** with faithful orthogonal representation in minimum dimension ω, for which the separation between classical and quantum public coins required for perfect S-CCR of relation $\mathcal{R}_{CLP}(\mathcal{G})$ is unbounded. (See [1])*

We also discuss some applications of S-CC$\mathcal{R}_{CLP}$. The first application is the operational detection of Mutually Unbiased Bases (MUBs) from the observed statistics. When we consider some specific type of graph $\mathcal{G}$ consisting of n maximum cliques of size ω that are completely disconnected from each other, if a quantum strategy with direct communication of an ω-level system can achieve the algebraic maximum of the payoff i.e. $\frac{1}{\omega}$, then the measurements performed by Bob must be those corresponding to MUBs.

In the next application, we consider the problem of detecting the non-classical resources in both direct communication and in the shared correlation (black-box) scenario. When no public coin is available S-CC$\mathcal{R}_{CLP}$ for some graphs allows us to determine the non-classicality of the transmitted system. Second, when only a finite amount of public coins are available and 1 c-bit has been transmitted, allows us to answer whether the public coin is non-classical or otherwise.

Finally, we consider a larger class of graphs that do not have orthogonal representation in dimension ω and show that these graphs can be used to detect whether the dimension of the direct communication resource is greater than ω or otherwise.

3 Outlook

In the S-CCR problem introduced in this work, we show that there exists a class of graphs for which the separation between the dimension of quantum and classical systems necessary can be made unbounded in the absence of shared randomness between the players. In the presence of public coins, however, this separation disappears. While quantum communication does not require public coins, the amount of public coin assistance that is necessary for classical communication for accomplishing the task scales linearly with the number of cliques. Additionally, we also show that an 1 c-bit channel when assisted by $1-$*ebit* public coin performs a task that would otherwise require the assistance of an unbounded amount of classical public coin.

Finally, S-CCR task can be seen, as a qualitative simulation of the quantum statistics on demand. In fact, the relation-reconstruction condition for the strong communication complexity proposed here could bridge the gap between conventional communication complexity and sampling problems with communication [9, 10]. Precisely, in our protocol, the spatially separated parties are given some set of favourable events and it is required that the events be quantitatively simulated by classical communication so that all of them occur with nonzero probability like it is in the quantum case. Looking at the protocol from yet another angle, we can see it as a distribution of a (conditional) randomness with the help of a restricted communication channel.

References

[1] S. Rout, N. Sakharwade, S. S. Bhattacharya, R. Ramanathan, and P. Horodecki, "Unbounded quantum advantage in one-way strong communication complexity of a distributed clique labelling relation," *arXiv preprint arXiv:2305.10372*, 2023.

[2] M. M. Wilde, "From classical to quantum shannon theory," *arXiv preprint arXiv:1106.1445*, 2011.

[3] A. C.-C. Yao, "Some complexity questions related to distributive computing(preliminary report)," in *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, (New York, NY, USA), p. 209–213, Association for Computing Machinery, 1979.

[4] E. Kushilevitz and N. Nisan, *Communication Complexity*. Cambridge University Press, 1997.

[5] H. Buhrman, R. Cleve, and A. Wigderson, "Quantum vs. classical communication and computation," in *Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing*, STOC '98, (New York, NY, USA), p. 63–68, Association for Computing Machinery, 1998.

[6] R. Raz, "Exponential separation of quantum and classical communication complexity," in *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*, STOC '99, (New York, NY, USA), p. 358–367, Association for Computing Machinery, 1999.

[7] D. Saha, P. Horodecki, and M. Pawłowski, "State independent contextuality advances one-way communication," *New Journal of Physics*, vol. 21, p. 093057, Sept. 2019.

[8] S. Kochen and E. P. Specker, "The problem of hidden variables in quantum mechanics," in *The Logico-Algebraic Approach to Quantum Mechanics*, pp. 293–328, Springer Netherlands, 1975.

[9] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson, "The quantum communication complexity of sampling," *SIAM Journal on Computing*, vol. 32, no. 6, pp. 1570–1585, 2003.

[10] T. Watson, "Communication complexity with small advantage," *computational complexity*, vol. 29, p. 2, Apr 2020.

Partially Fault-tolerant Quantum Computing Architecture with Error-corrected Clifford Gates and Space-time Efficient Analog Rotations

Yutaro Akahoshi[1 2 *] Kazunori Maruyama[1 2] Hirotaka Oshima[1 2] Shintaro Sato[1 2]
Keisuke Fujii[2 3 4 5]

[1] *Quantum Laboratory, Fujitsu Research, Fujitsu Limited, 4-1-1 Kawasaki, Kanagawa 211-8588, Japan*
[2] *Fujitsu Quantum Computing Joint Research Division, Center for Quantum Information and Quantum Biology, Osaka University, 1-2 Machikaneyama, Toyonaka, Osaka, 565-8531, Japan*
[3] *Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka, 560-8531, Japan*
[4] *Center for Quantum Information and Quantum Biology, Osaka University, 560-0043, Japan*
[5] *RIKEN Center for Quantum Computing (RQC), Wako Saitama 351-0198, Japan*

Abstract. Quantum computers hold the promise of revolutionizing several computational tasks by significantly enhancing their calculation speed compared to classical computers. Although quantum computing devices has been rapidly developed in recent years, there is still a large gap between today's noisy intermediate-scale quantum (NISQ) computing and the full fault-tolerant quantum computing (FTQC) based on the quantum error correction (QEC) code due to the extremely large requirement of physical qubits for the latter. In this study, we propose a quantum computing architecture that fills the gap between NISQ and FTQC. Our architecture realizes universal computation by using noisy analog rotation gates and error-corrected Clifford gates implemented by lattice surgery. We perform direct analog rotations with a small qubit requirement and minimize remnant errors by a carefully-designed state injection protocol. Our estimation based on numerical simulations shows that, for early-FTQC devices that consist of 10^4 physical qubits with physical error probability $p = 10^{-4}$, we can perform roughly 1.72×10^7 Clifford operations and 3.75×10^4 arbitrary rotations on 64 logical qubits. Such computations cannot be realized by the existing NISQ and FTQC architectures on the same device, as well as classical computers. This extended abstract is based on a new preprint of the authors [1].

Keywords: Fault-tolerant quantum computing, Quantum error correction, Surface code, Lattice surgery, Quantum error mitigation

1 Introduction

Quantum computers are expected to provide exponential speedups of computation in certain tasks. In recent years, Quantum devices with tens to hundreds of qubits have gradually emerged and are referred to as noisy intermediate-scale quantum (NISQ) devices [2]. Unfortunately, it is still challenging to extract useful quantum advantages over the classical best approaches from NISQ devices. The main obstacle is that qubits and gate operations suffer from errors caused by undesirable interactions with the environment. Although the quantum error mitigation techniques [3] can reduce the effects of this noise, it is nontrivial whether these techniques can sufficiently mitigate errors in quantum computation involving tens to hundreds qubits within a reasonable sampling overhead.

The ultimate long-term solution to the noise problem is the realization of fault-tolerant quantum computing (FTQC) by implementing quantum error correction (QEC). However, non-Clifford gates, which are indispensable for performing universal quantum computation and providing quantum speedups, are difficult to implement fault-tolerantly. Due to the cost of the magic state distillation [4] and the Solovay-Kitaev decomposition, FTQC may require hundreds of thousands to millions of qubits [5, 6, 7, 8]. Therefore, a large gap between the NISQ and FTQC eras will exist. In the meantime, it is necessary to establish a theoretical framework that meaningfully exploits quantum computation with 10^3–10^4 qubit devices (we call these devices as "early-FTQC devices").

In this study [1], we propose a quantum computing architecture that fills the gap between NISQ and FTQC and provide evidence that a quantum device of 10^4 qubits has great potential to exhibit useful quantum advantages. In our approach, we integrate NISQ and FTQC approaches deeply. Namely, we implement quantum error correction for the logical Clifford gates and directly perform non-Clifford analog rotation gates without magic state distillation. As a drawback, the analog rotation gates suffer from unavoidable errors. In our proposal, we minimize these errors by carefully designing the state injection protocol for the analog rotation gates. Our resource estimation based on numerical simulations shows that the proposed architecture can surpass not only classical computers but also existing NISQ and FTQC approaches on a typical early-FTQC device.

2 Proposed Architecture

In this section, we briefly discuss our proposed architecture, Space-Time efficient Analog Rotation (STAR) quantum computing architecture. Below we discuss how

*yutaro.akahoshi@fujitsu.com

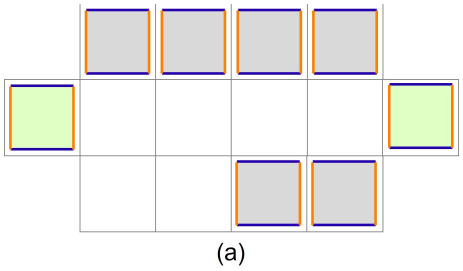


Figure 1: Example of a compact logical qubit arrangement. Gray and green squares represent data pathes and ancilla patches for the analog rotation, respectively. Other white region is necessary to perform fundamental operations of the lattice surgery. It requires $1.5n + 5$ patches to allocate n logical qubits.

to implement universal gates in the STAR architecture and give a resource estimation. Some details will be skipped due to the page restriction and can be found in the preprint [1].

2.1 Fault-Tolerant Clifford Gates

To protect quantum information from noises during computation, we employ the rotated planar surface code [9]. This QEC code has suitable features for the early-FTQC era: a relatively high threshold value against other QEC codes and a compact requirement for the number of physical qubits to encode a logical state.

Although the rotated planar surface code supports a transversal CNOT gate, it is difficult in practice due to the restricted connectivity between physical qubits. A clever way to implement logical Clifford gates even in such a situation is known as the lattice surgery [9, 10]. In the lattice surgery, logical qubits encoded in the rotated planar surface code (hereafter, we refer a logical qubit as a patch) are arranged into tiles and logical Clifford gates are implemented by combining fundamental operations: merging adjacent patches, splitting a patch into two patches, and deforming a patch. Each fundamental operation consists of stabilizer measurements and only requires nearest-neighbor connectivity.

There are mainly two schemes to perform quantum computing by the lattice surgery: one that implements explicit logical Clifford gates [9] and the other that absorbs logical Clifford gates into logical non-Clifford gates and Pauli measurements [10]. These schemes should be chosen depending on quantum circuits. The arrangement of the patches strongly depends on these schemes, and typical examples are discussed in the preprint [1]. Fig. 1 shows an example of a minimum arrangement. In Fig. 1, gray patches are data logical qubits, and green patches are reserved to generate the ancilla state of the analog rotation gate. Other white region is necessary to perform fundamental operations of the lattice surgery. This arrangement requires $1.5n + 5$ patches to allocate n data logical qubits in total.

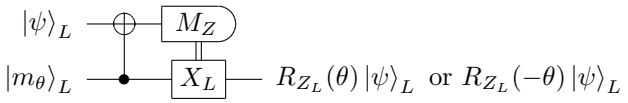


Figure 2: Quantum circuit for the analog Z rotation gate. M_Z is a destructive Z_L measurement on a logical patch.

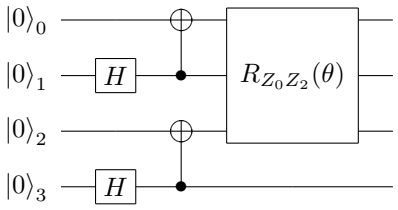


Figure 3: Quantum circuit for the injection of the ancilla state encoded in the $[[4, 1, 1, 2]]$ subsystem code.

2.2 Space-time Efficient Analog Rotation Gate

Next, let us discuss the implementation of the analog rotation gates, which is a core technology of the STAR architecture. Our strategy is as follows. We directly perform analog rotation gates without the Solovay-Kitaev decomposition and avoid magic state distillation. In comparison to the typical FTQC architecture using Clifford + T gate decomposition, this approach is advantageous in terms of the number of physical qubits and execution time. A major drawback is that the logical error rate of the analog rotation remains at $O(p)$. To minimize the effect of this remaining error, we carefully design a state injection protocol. As a result, the dominant logical error becomes a simple phase-flip channel and can be mitigated by the probabilistic error cancellation.

We implement the analog rotation gate by the gate teleportation circuit with a special ancilla state, $|m_\theta\rangle = e^{-i\frac{\theta}{2}Z}|+\rangle$, where θ can be chosen arbitrarily (Fig. 2). The output state of this circuit depends on the measurement result: namely, if the measurement result is $+1$, the output state is correctly rotated; otherwise, it is an inversely rotated state. When we obtain the inversely rotated state, we apply another rotation gate with a rotation angle 2θ for correction. This procedure is repeated until obtaining the correct state ("Repeat-Until-Success" or RUS). An average repetition number until success is estimated as $1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + \cdots = \sum_{i=1}^{\infty} \frac{n}{2^n} = 2$.

To prepare the ancilla state $|m_\theta\rangle$ within reasonable accuracy, we construct a state injection protocol utilizing the $[[4, 1, 1, 2]]$ subsystem code [11]. We first prepare the ancilla state encoded in the $[[4, 1, 1, 2]]$ subsystem code by the circuit shown in Fig. 3. Since this QEC code has a code distance of two, we can detect a single error. Thus, we discard the prepared state if the measured syndromes detect any error. Once we obtain the ancilla state that passes the post-selection, we expand it to the rotated surface code with an arbitrary code distance. After that, we perform another post-selection to remove $O(p)$ errors that occur during the expansion.

Under the circuit-level noise model, the logical error

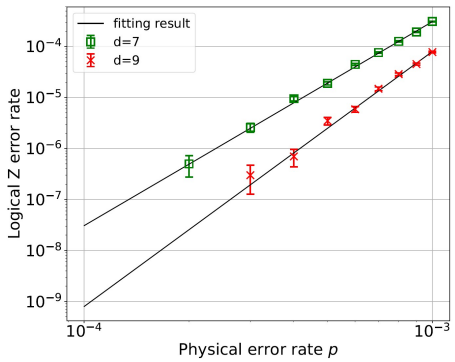


Figure 4: Logical Z error probability and its scaling behavior with $d = 7, 9$. The scaling behavior determined by a theoretically expected form, $P_L(p) \approx p^{\frac{d+1}{2}}$, is shown as solid black lines.

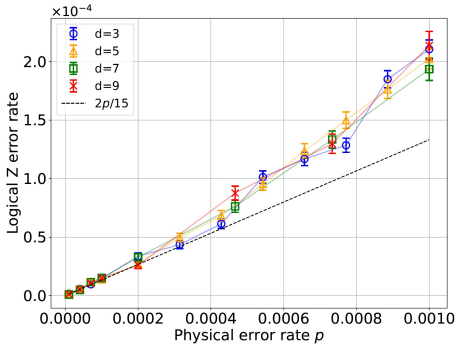


Figure 5: Logical Z error probability of the ancilla state prepared in the surface code patches with $d = 3, 5, 7, 9$. The dashed line shows the leading-order behavior expected under the circuit-level noise model, $P_{L,Z}(p) = 2p/15$.

probability of the prepared ancilla state behaves as

$$P_{Z_L}(p) = 2p/15 + O(p^2), \quad (1)$$
$$P_{X_L}(p) = O(p^2), \quad (2)$$

where p is an error probability of physical qubits. This injection protocol surpasses some previous protocols, e.g., the protocol in Ref. [12] that gives $P_L = 34p/15 + O(p^2)$. Moreover, the logical error can be regarded as the single phase-flip error if p is sufficiently small; thus, we can mitigate it by the probabilistic error cancellation [13, 14].

3 Resource Estimation

To illustrate the performance of the STAR architecture, we estimate a computational resource for early-FTQC devices based on numerical simulations. We assume a target device has $N = 10^4$ physical qubits with a physical error probability of $p = 10^{-4}$. In this extended abstract, we only consider the case where the code distance of the surface code patch is set to $d = 7$. A More detailed discussion is given in the preprint [1].

The number of the Clifford gate we can perform is estimated from the logical error probability of the surface code patch. We perform a numerical simulation on the error-correcting procedure of the surface code and determine the scaling behavior of the logical error probability. We show the logical Z error behavior in Fig. 4 as an example. As a result, with the physical error probability $p = 10^{-4}$, we estimate the available number of the Clifford gates as $N_{\text{Clifford}} \approx 1.72 \times 10^7$.

The available number of analog rotation gates can be estimated similarly. We perform a numerical simulation on the whole procedure of the state injection protocol and determine the logical error probability of the prepared ancilla state $|m_\theta\rangle$. Fig. 5 shows the logical Z error probability of $|m_\theta\rangle$ obtained by the simulation. We can see an expected behavior in Fig. 5 that the logical error probability approaches $2p/15$ by decreasing the physical error probability. As a result, the available number of the rotation gates is estimated as $N_{\text{rotation}} \approx 3.75 \times 10^4$ for $p = 10^{-4}$.

The number of logical qubits we can allocate is determined by the code distance. In general, one logical patch requires $\approx 2d^2$ physical qubits. Therefore, by employing the compact arrangement shown in Fig. 1, $(1.5n+5)\times 2d^2$ physical qubits are needed in total. For $N = 10^4$ and $d = 7$, we can allocate $n \approx 64$ logical data qubits.

The STAR architecture may be compatible with applications of quantum many-body simulations and quantum approximation optimization algorithm (QAOA) since the time-evolution operator can be implemented easily by analog rotation gates. A detailed examination of the useful applications is an important future issue.

4 Conclusion

In this work, we propose a quantum computing architecture suitable for the early-FTQC devices, the STAR architecture. In the STAR architecture, universal quantum computation is achieved by error-corrected Clifford gates and analog rotation gates. The analog rotation gate is implemented by the RUS protocol with an appropriate ancilla state. To reduce logical errors in the rotation gate, we carefully design the ancilla state injection protocol by combining the $[[4, 1, 1, 2]]$ subsystem code and post-selection. As a result, we achieve a small logical error rate of $P_L = 2p/15 + O(p^2)$ under the circuit-level noise model, which is verified numerically. Clifford operations are performed by the standard lattice surgery protocol based on the rotated surface code. Finally, we estimate a computational resource in the STAR architecture under the assumption of typical early-FTQC devices, where $N = 10^4$ physical qubits can operate with a gate fidelity of $p = 10^{-4}$. Our resource estimation shows that we can act 3.75×10^4 analog rotation gates and 1.72×10^7 Clifford gates on 64 logical qubits encoded in the $d = 7$ rotated planar surface code. Such computations cannot be realized by the existing NISQ and FTQC architectures on the same device, as well as classical computers.

We hope that our proposal and the corresponding development of quantum algorithms will bring new insights to realizing practical quantum computers in future.

References

[1] Y. Akahoshi *et al.*, Partially Fault-tolerant Quantum Computing Architecture with Error-corrected Clifford Gates and Space-time Efficient Analog Rotations. arXiv:2303.13181, (2023).

[2] J. Preskill, Quantum computing in the NISQ era and beyond. Quantum **2**, 79 (2018).

[3] S. Endo, Z. Cai, S. C. Benjamin, and X. Yuan, Hybrid quantum-classical algorithms and quantum error mitigation, Journal of the Physical Society of Japan **90**, 032001 (2021).

[4] S. Bravyi and J. Haah, Magic-state distillation with low overhead, Physical Review A **86**, 052329 (2012).

[5] C. Gidney and M. Ekera, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits, Quantum **5**, 433 (2021).

[6] N. Yoshioka *et al.*, Hunting for quantum-classical crossover in condensed matter problems, arXiv:2210.14109 (2022).

[7] M. Reiher *et al.*, Elucidating reaction mechanisms on quantum computers, in *Proceedings of the national academy of sciences* **114**, 7555 (2017).

[8] J. J. Goings *et al.*, Reliably assessing the electronic struc- ture of cytochrome p450 on today's classical comput- ers and tomorrow's quantum computers, arXiv:2202.01244 (2022).

[9] C. Horsman *et al.*, Surface code quantum computing by lattice surgery, New Journal of Physics **14**, 123011 (2012).

[10] D. Litinski, A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery, Quantum **3**, 128 (2019).

[11] D. Bacon, Operator quantum error-correcting subsystems for self-correcting quantum memories, Physical Review A **73**, 012340 (2006).

[12] L. Lao and B. Criger, Magic state injection on the rotated surface code, in *Proceedings of the 19th ACM International Conference on Computing Frontiers*, CF '22 (Association for Computing Machinery, New York, NY, USA, 2022) p. 113–120.

[13] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).

[14] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, Phys. Rev. X **8**, 031027 (2018).

Generalised Susskind-Glogower coherent states

Jean-Pierre Gazeau[1] Véronique Hussin[2] James Moran[3] * Kevin Zelaya[3]

[1] *Université Paris Cité, CNRS, Astroparticule et Cosmologie*
[2] *Département de Mathématiques et de Statistique, Université de Montréal*
[3] *Quantum Universe Center, Korea Institute for Advanced Study*
[4] *Department of Physics, Queens College of the City University of New York*

Abstract. This work is based on ***J. Math. Phys. 62 (2021)***. Susskind–Glogower coherent states, whose Fock expansion is characterised by Bessel functions, have recently attracted considerable attention for their optical properties. A shortcoming of these states is that it is not known if there exists a measure for which they resolve the identity. To this end, the modified Susskind–Glogower coherent states have been introduced as an alternative family of states that resolve the identity at the expense of modifying the expansion coefficients. In this work, the quantisation map related to the modified Susskind–Glogower coherent states is exploited, which naturally leads to a particular representation of the $\mathfrak{su}(1,1)$ Lie algebra in its discrete series where the index of the discrete series can be identified with the index of the Bessel function. We can go a step further by modifying expansion coefficients to be Bessel functions of the second kind, in which case we find a representation of the $\mathfrak{su}(2)$ Lie algebra. This enables us to define two families of generalised coherent states which exhibit interesting non-classical properties.

Keywords: Quantum optics, Generalised coherent states, Susskind-Glogower operators

1 Introduction

In the context of nonlinear f-deformed ladder operators [1], let us consider the set of operators $\{\mathbf{V}, \mathbf{V}^\dagger\}$ introduced by Susskind-Glogower [2], and recently studied in [3, 4], which are defined in terms of the conventional boson and number operators $\mathbf{a}$, $\mathbf{a}^\dagger$, $\mathbf{n} = \mathbf{a}^\dagger\mathbf{a}$, through the relations

$$\mathbf{V} = \sum_{n=1}^{\infty} |n-1\rangle\langle n| \equiv \frac{1}{\sqrt{\mathbf{n}+1}}\mathbf{a}\,,$$
$$\mathbf{V}^\dagger = \sum_{n=0}^{\infty} |n+1\rangle\langle n| \equiv \mathbf{a}^\dagger\frac{1}{\sqrt{\mathbf{n}+1}}\,. \tag{1}$$

Moreover, $\mathbb{I}$ is the identity operator in the Fock Hilbert space $\mathcal{H} = \overline{\text{Span}\{|n\rangle\}_{n=0}^{\infty}}$, i.e. the closure of all finite linear combinations of the number states $|n\rangle$, also known as Fock states [5]. The operators $\mathbf{V}$ and $\mathbf{V}^\dagger$ satisfy the commutation relation $[\mathbf{V}, \mathbf{V}^\dagger] = |0\rangle\langle 0|$, which is a projector onto the vacuum state. This property has been exploited to construct an exponential operator $D_{\text{SG}}(x) = e^{x(\mathbf{V}^\dagger - \mathbf{V})}$, with $x \in \mathbb{R}$, whose action on the vacuum state $|0\rangle$ leads to a nonlinear family of unit-norm coherent states known as *Susskind-Glogower coherent states* [6], given by

$$|\alpha\rangle_{\text{SG}} = D_{\text{SG}}(\alpha)|0\rangle = \sum_{n=0}^{\infty} \alpha^n (n+1)\frac{J_{n+1}(2r)}{r^{n+1}}|n\rangle\,, \tag{2}$$

with $\alpha \in \mathbb{C}$ and $r = |\alpha|$. Above, $J_n(z)$ is the *Bessel functions of the first kind* [7], defined by the power series

$$J_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{z}{2}\right)^{2m}}{m!\Gamma(\nu+m+1)}\,. \tag{3}$$

The Susskind-Glogower CS have been extensively discussed in the literature, where their nonclassical properties have been studied and documented. See [3, 4] for details. Nevertheless, to the best of the authors' knowledge, the resolution of the identity remains an open problem, that is, the existence of a weight function $w(\alpha)$ such that

$$\mathbb{I} = \sum_{n=0}^{\infty} |n\rangle\langle n| =^{?} \int_{\alpha\in\mathbb{C}} \frac{\mathrm{d}^2\alpha}{\pi} w(r)|\alpha\rangle_{\text{SG}}{}_{\text{SG}}\langle\alpha|\,, \tag{4}$$

is unknown or might not exist. It is not known either if this set defines a continuous frame in the sense given in [8]. A workaround for this issue was addressed in [9, 10], where the authors introduced the *modified Susskind-Glogower* coherent states, obtained after modifying the expansion coefficients of the coherent states (2) as

$$|\alpha\rangle_{\text{mSG}} = \sum_{n=0}^{\infty} \alpha^n h_n(r)|n\rangle\,, \quad h_n(r) = \sqrt{\frac{n+1}{\mathcal{N}(r)}}\frac{J_{n+1}(2r)}{r^{n+1}}\,, \tag{5}$$

where $\alpha \in \mathbb{C}$. The states $|\alpha\rangle_{\text{mSG}}$ have unit norm, while the resolution of the identity

$$\mathbb{I} = \int \frac{d^2\alpha}{\pi} w(r)|\alpha\rangle_{\text{mSG}}{}_{\text{mSG}}\langle\alpha|\,, \tag{6}$$

is fulfilled with the weight function $w(r) = \mathcal{N}(r)$, where $\mathcal{N}(r)$ stands for the normalisation constant given by

$$\mathcal{N}(r) = \frac{1}{r}\sum_{n=0}^{\infty} n[J_n(2r)]^2 = {}_1F_2\left(\begin{matrix}1/2\\2,2\end{matrix}\middle| -4r^2\right)\,. \tag{7}$$

This result ensures that the set $\{|\alpha\rangle_{\text{mSG}}\}_{\alpha\in\mathbb{C}}$ forms an overcomplete basis in the Fock space $\mathcal{H}$.

The unmodified and modified Susskind-Glogower coherent states are related through

$$|\alpha\rangle_{\text{mSG}} = \frac{1}{\sqrt{\mathcal{N}(|\alpha|)}}\frac{1}{\sqrt{\mathbf{n}+1}}|\alpha\rangle_{\text{SG}}\,. \tag{8}$$

*jamesmoran@kias.re.kr

Hence, apart from the normalisation factor, the modified Susskind-Glogower coherent states result from the action of the compact operator $1/\sqrt{\mathbf{n}+1}$ on the Susskind-Glogower coherent states. Contrary to the conventional Susskind-Glogower coherent states, it is still unknown whether a unitary operator construction exists for the modified Susskind-Glogower coherent states. Recently, some progress has been made in [11], where a non-unitary construction from a fiducial vector has been introduced.

2 Susskind-Glogower-I coherent states

We can make a further generalisation by introducing an arbitrary parameter κ into the index of the Bessel functions of the modified Susskind-Glogower coherent states, that is, we consider the change $J_{n+1}(2r) \to J_{n+\kappa}(2r)$ in (5). Thus, we introduce the family of *Susskind-Glogower-I coherent states* (SGI CS) constructed through the linear combination [12]

$$|\alpha;\kappa\rangle_{\rm I} := \sum_{n=0}^{\infty} \alpha^n h_{n;\kappa}(r)|n\rangle\,,\; h_{n;\kappa}(r) = \sqrt{\frac{\mathcal{C}_{n;\kappa}}{\mathcal{N}_\kappa(r)}}\frac{J_{n+\kappa}(2r)}{r^{n+\kappa}}\,, \tag{9}$$

where $r=|\alpha|$, $\mathcal{N}_\kappa(r)$ stands for the normalisation constant, and $\mathcal{C}_{n;\kappa}$ are unknown coefficients to be determined. The linear combination (9) must satisfy two basic properties, normalisability for $\alpha \in \mathbb{C}$ and they must satisfy a resolution of the identity. First, we focus on the resolution of the identity so that the unknown coefficients $\mathcal{C}_{n;k}$ are uniquely defined.

Making use of Bessel function integrals [13], the resolution of the identity associated with the linear combination (9) is determined from

$$\begin{aligned}\mathbb{I} &= \int_{\alpha\in\mathbb{C}} \frac{d^2\alpha}{\pi} w_\kappa(r)|\alpha;\kappa\rangle_{\rm I\,I}\langle\alpha;\kappa| \\ &= \sum_{n=0}^{\infty} 2\mathcal{C}_{n;\kappa}\mathcal{D}_\kappa \int_0^\infty dr\, r^{-(2\kappa-1)}\left[J_{n+\kappa}(2r)\right]^2 |n\rangle\langle n|\,,\end{aligned} \tag{10}$$

with the weight function fixed as $w_\kappa(r) = \mathcal{D}_\kappa\mathcal{N}_k(r)$, and $\mathcal{D}_\kappa$ is a proportionality factor independent of n. The integral in (10) is satisfied by

$$\mathcal{C}_{n;\kappa} = (n+1)_{2\kappa-1} = \frac{\Gamma(2\kappa+n)}{n!}\,,\quad \mathcal{D}_\kappa = \frac{(2\kappa-1)\Gamma(\kappa)}{2^{2\kappa-1}(1/2)_\kappa}\,. \tag{11}$$

From the convergence conditions of the integrals, we conclude that the matrix elements (10) converge for all $n=0,1,\cdots$ provided that $\kappa > \frac{1}{2}$. From (11), we have completely characterised the SGI CS given in (9). Notice that the coefficients $\mathcal{C}_{n;\kappa}$ correspond to the weighting factors of a negative binomial distribution. The latter is also a common property of the $SU(1,1)$ Perelomov CS [14, 9], but with different weighting factors.

Now, we must verify the normalisability of the SGI CS. Interestingly, an analytic expression can be determined for $\mathcal{N}_\kappa$, which is computed after expanding the Bessel functions in power series, and after arranging the resulting summations in a convenient form, we obtain

$$\mathcal{N}_\kappa(r) = \frac{\Gamma(2\kappa)}{[\Gamma(\kappa+1)]^2}\,{}_1F_2\left(\begin{matrix}1/2\\ \kappa+1,\kappa+1\end{matrix}\middle| -4r^2\right)\,, \tag{12}$$

where ${}_pF_q$ is the *generalised hypergeometric function* [15].

Therefore, the SGI CS constructed by the linear combination $|\alpha;\kappa\rangle = \sum_{n=0}^\infty c^{(I)}_{n;\kappa}(r)|n\rangle$, with expansion coefficients

$$\begin{aligned}c^{(I)}_{n;\kappa}(r) = \sqrt{\frac{(2\kappa)_n}{n!}}\,\Gamma(\kappa+1)\frac{e^{{\rm i}n\phi}J_{n+\kappa}(2r)}{r^k} \\ \times\left[{}_1F_2\left(\begin{matrix}1/2\\ \kappa+1,\kappa+1\end{matrix}\middle| -4r^2\right)\right]^{-1/2}\,,\end{aligned} \tag{13}$$

define an overcomplete and normalisable set $\{|\alpha;\kappa\rangle\}_{\alpha\in\mathbb{C}}$.

3 Susskind-Glogower-II coherent states

Now we consider an alternative set of coherent states, which are a further generalisation of the SGI CS. This is achieved by modifying the functional coefficients $h_{n,\kappa}(r)$, where $\kappa \in \mathbb{N}^+/2$, with $\mathbb{N}^+ = \{1,2,\cdots\}$, and introducing the modified Bessel functions of the second kind $K_\nu(r)$, defined as [7]

$$K_\nu(z) = \frac{\pi}{2}\frac{I_{-\nu}(z)-I_\nu(z)}{\sin\pi\nu}\,,\quad I_\nu(z) = e^{-{\rm i}\pi\nu/2}J_\nu({\rm i}z)\,, \tag{14}$$

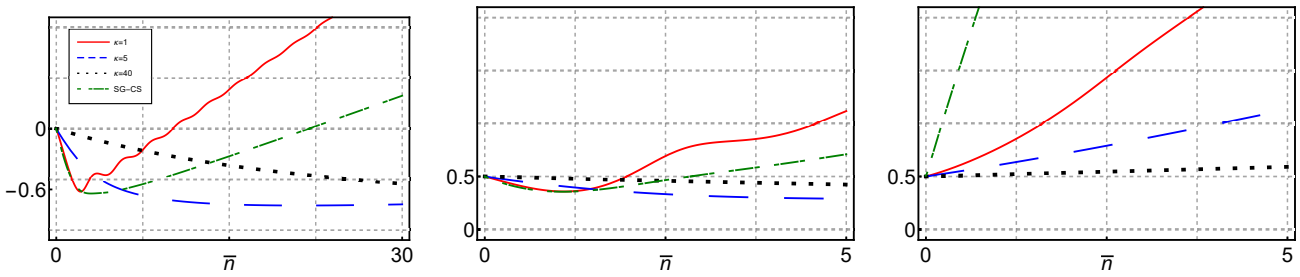
with $I_\nu(z)$ the modified Bessel function of the first kind. The function $K_\nu(z)$ behaves, in the asymptotic limit $z\to\infty$, as $K_\nu(z)\to\sqrt{\pi/(2z)}e^{-z}$. For $z\to 0$, the modified Bessel function of the second kind has branch points for all $\nu\in\mathbb{C}$ [15]. Moreover, $K_\nu(z)$ is analytic in $\mathbb{C}\backslash(-\infty,0]$. Analogously to the functions $h_{n;\kappa}(r)$ of (5), we introduce the new functions $\mathfrak{h}_{n;\kappa}$ written in terms of $K_\nu(z)$. We thus introduce the *Susskind-Glogower-II coherent states* (SGII CS) defined as

$$|z;\kappa\rangle_{\rm II} = \sum_{n=0}^{2\kappa} z^n \mathfrak{h}_{n;\kappa}(r)|n\rangle\,,\quad \mathfrak{h}_{n;\kappa}(r) = \sqrt{\frac{\mathfrak{C}_{n;\kappa}}{\mathfrak{N}_\kappa(r)}}\frac{K_{n-\kappa}(2r)}{r^{n-\kappa}}\,, \tag{15}$$

where $\in\mathbb{C}$, $r=|z|$, $\mathfrak{N}_\kappa(r)$ stands for the normalization factor, and the coefficients $\mathfrak{C}_{n;\kappa}$ are independent of z so that the set $\{|z;\kappa\rangle\}_{z\in\mathbb{C}}$ fulfils the resolution of the identity

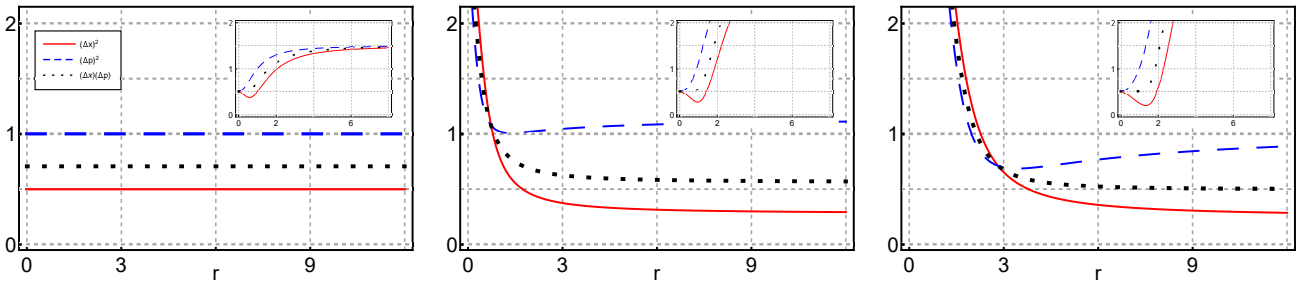
$$\mathbb{I}_{2\kappa+1} := \sum_{n=0}^{2\kappa}|n\rangle\langle n| = \int_{z\in\mathbb{C}}\frac{d^2z}{\pi}\mathfrak{w}_k(r)|z;\kappa\rangle_{\rm II\;II}\langle z;\kappa|\,, \tag{16}$$

with $\mathfrak{w}_k(r)$ the respective weight function. In this form, the set $\{|z;\kappa\rangle\}_{z\in\mathbb{C}}$ generates the $2\kappa+1$-dimensional Hilbert subspace $\mathcal{H}^{(2\kappa)} = {\rm Span}\{|n\rangle\}_{n=0}^{2\kappa}\subset\mathcal{H}$. Notice that the coherent states (15) are defined through a finite linear combination, for which the normalisation constant $\mathfrak{N}_\kappa(r)$ converges in the complex-plane as long as $\mathfrak{h}_{n;\kappa}(r)$ is free of singularities for $r\in\mathbb{R}^+\cup\{0\}$. From the asymptotic behavior previously discussed, we can guarantee the



(a) Mandel parameter (b) $(\Delta \mathbf{x})^2$ (c) $(\Delta \mathbf{p})^2$

Figure 1: Mandel parameter (a), together with the physical variances $(\Delta \mathbf{x})^2$ (b) and $(\Delta \mathbf{p})^2$ (c) for the SGI CS (9) as a function of the average number of photons for several values of κ.



(a) Mandel parameter (b) $(\Delta \mathbf{x})^2$ (c) $(\Delta \mathbf{p})^2$

Figure 2: Quadrature variances $(\Delta \mathbf{x})^2$ (solid-red), $(\Delta \mathbf{p})^2$ (dashed-blue), and the product $(\Delta \mathbf{x})(\Delta \mathbf{p})$ (dotted-black) associated with the SGII CS. The inset denotes the respective variances for the $SU(2)$ coherent states.

finite-norm condition for $n = 0, 1, \cdots, 2\kappa$. It is worth mentioning that $\mathfrak{h}_{n;\kappa}(r)$ leads to singularities either at $r = 0$ or $r \to \infty$ for $n = 2\kappa + 1$. For that reason, we have truncated the linear combination in (15) and restricted the values of κ to non-negative integers or half-integers.

The resolution of the identity (16), together with $\mathfrak{C}_{n;\kappa}$, are determined through Bessel function identities [13]. Doing the calculations shows that (16) holds for

$$\mathfrak{C}_{n;\kappa} = \binom{2\kappa}{n}, \quad \mathfrak{w}_\kappa(r) = \mathfrak{D}_\kappa \mathfrak{N}_\kappa(r), \quad \mathfrak{D}_\kappa = \frac{4(2\kappa+1)}{[\Gamma(k+1)]^2}, \tag{17}$$

where $\binom{a}{b}$ is the binomial coefficient. Thus, the expansion coefficients in the SGII CS define a binomial-like distribution weighted by a modified Bessel function of the second kind rather than the conventional binomial parameter.

Considering half-integer values of κ, that is, $\kappa = L + 1/2$ for $L \in \mathbb{N}$, we can exploit the symmetry of the modified Bessel functions of the second kind $K_\nu(z) = K_{-\nu}(z)$, together with the symmetry of the binomial coefficient $\binom{2L+1}{n}$, in order to write the SGII CS as

$$|z; L\rangle_{\text{II}} = \frac{1}{[\mathfrak{N}_L(r)]^{1/2}} \sum_{n=0}^{L} e^{\mathrm{i}n\phi} c^{(\text{II})}_{n;L}(r) \times \left[|n\rangle + e^{\mathrm{i}(2L-2n+1)\phi} |2L-n\rangle \right], \tag{18}$$

where the expansion coefficients are given by

$$c^{(\text{II})}_{n;L}(r) := \sqrt{\binom{2L+1}{n}}\, \Gamma(L-n+1/2) r^n \, {}_1F_1\left(\begin{matrix} n-L \\ 2n-2L \end{matrix} \middle| 4r \right). \tag{19}$$

Given that $\mathfrak{N}_L(r)$ is a finite sum, and using the fact that $K_{n-L-1/2}(2r)$ is an entire function for $n = 0, 1, \cdots L$, we conclude that normalisation function is well-defined on the whole complex-plane.

4 Outlook

Nonlinear coherent states may provide a useful resource in continuous variable quantum information tasks. At present it is not clear what the form of the Hamiltonians which generate the SGI CS and SGII CS described here, while for the original Susskind-Glogower coherent states, the Hamiltonian is known, but it is not clear whether they resolve the identity. Work is ongoing in both of these directions

It is also interesting to consider multiphoton representations, particularly of the SGI CS. These most closely resemble the single-mode squeezed vacuum states, and defining a two-photon representation of the SGI CS may unlock some additional analysis to compare the role of squeezing and other resources obtained from the nonlinearities, and how these resources can be used in quantum computing tasks.

References

[1] V. I. Man'ko, G. Marmo, E. C. G. Sudarshan, and F. Zaccaria, Phys. Scr. **55**, 528 (1997).

[2] L. Susskind and J. Glogower, Phys. **1**, 49 (1964).

[3] R. De J. León Montel and H. M. Moya-Cessa, Int. J. Quant Inf. **9**, 349 (2011).

[4] R. De J. León Montel, H. M. Moya-Cessa, and F. Soto-Eguibar, Rev. Mex. Fís. **57**, 133 (2011).

[5] P. A. M. Dirac, *The Principles of Quantum Mechanics*, Claredon, Oxford, 2 edition, 1935.

[6] J. Récamier, M. Gorayeb, W. L. Mochán, and J. L. Paz, Int. J. Theor. Phys. **47**, 673 (2008).

[7] A. F. Nikiforov and V. B. Uvarov, *Special Functions of Mathematical Physics: A Unified Introduction with Applications*, Birkhäuser, Germany, 1988.

[8] S. T. Ali, J. P. Antoine, and J. P. Gazeau, Ann. Phys. **222**, 1 (1993).

[9] J.-P. Gazeau, *Coherent States in Quantum Optics: An Oriented Overview*, Springer, Cham, 2019.

[10] E. M. F. Curado, S. Faci, J.-P. Gazeau, and D. Noguera, J. Opt. Soc. Am. B **38**, 3556 (2021).

[11] H. Moya-Cessa and J. Guerrero, J. Mod. Opt. **68**, 196 (2021).

[12] J.-P. Gazeau, V. Hussin, J. Moran, and K. Zelaya, J. Math. Phys. **62** (2021).

[13] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, California, 7 edition, 2007.

[14] A. Perelomov, *Generalized Coherent States and Their Applications*, Springer-Verlag, Berlin, 1986.

[15] F. W. J. Olver et al., *NIST Handbook of Mathematical Functions*, Cambridge University Press, New York, 2010.

Simplifying errors by symmetry and randomisation

arXiv:2303.02712

James Mills, Debasis Sadhukhan and Elham Kashefi

Extended abstract for AQIS 2023

I. OVERVIEW

We present a set of methods to generate less complex error channels by quantum circuit parallelisation. The resulting errors are simplified as a consequence of symmetrisation and randomisation. Initially, the case of a single error channel is analysed; this is then generalised to multiple channels. The error simplification for each method is shown to be either constant, linear, or exponential in terms of system size. Finally, some example applications are provided, along with experiments run on superconducting quantum hardware. These applications are: (1) reducing the sample complexity of matrix-inversion measurement error mitigation by error symmetrisation, (2) improving the effectiveness of noise-estimation circuit error mitigation by error randomisation, and (3) improving the predictability of noisy circuit performance by error randomisation.

II. BACKGROUND

A consequence of scaling up quantum hardware is increasing complexity of the errors. Highly complex error behaviour unpredictably degrades the quality of the output of a quantum computation. This work addresses this problem by presenting techniques that simplify the effects of errors on the output of noisy quantum circuits. Previous work on simplifying circuit noise has primarily involved the twirling of noise channels. For example, randomized compiling is a compilation technique that may be used to Pauli twirl general noise and so transform it to Pauli stochastic noise without changing the logic of the computation [1, 2]. In this work we present a framework based on using quantum circuit parallelisation to simplify the effects of stochastic Pauli errors on the output of a given computation. Where parallelisation refers to the action of running a quantum circuit in parallel across multiple sets of qubits [3–9]. Parallelisation is usually implemented to provide a linear speed-up in algorithm run-time, the key contribution of this work is to instead apply it to simplify noisy circuit errors.

We define the error complexity of a stochastic Pauli channel to be the cardinality of its set of Pauli coefficients, $|\{\mathbf{c}_P\}_{P\in\mathbf{P}^{\otimes n}}|$, i.e. the number of distinct Pauli coefficients needed to fully describe the channel. For a given error channel, coefficients for different Pauli operators that are identical are counted as a single coefficient. For example, a global depolarizing channel, D, is fully defined by a single error coefficient, so the error complexity of this channel is $|\{\mathbf{c}_D\}| = 1$. And a general stochastic Pauli channel, P, has complexity $|\{\mathbf{c}_P\}| = 4^n$. This definition of error complexity is used to refer to both the errors channels of individual parallel circuits and the effective error channels. And by error complexity reduction, we refer to the process by which an effective stochastic Pauli channel is generated with fewer coefficients than the original error channels of the parallel circuits. The techniques we propose generate less complex errors by averaging over the effects of the noise channels of the different parallel circuits. We refer to these averaged error channels, induced by the combination of circuit output results, as effective error channels. In the first approach, this is achieved by error symmetrisation; and in the second by error randomisation. For the symmetry reductions, simplification comes from coefficients of operators that can be mapped to each other by a given symmetry transformation being the same in the effective error channel. For randomisation,

simplification instead comes from the coefficients of operators which act non-trivially on the same subset of qubits becoming the same. Initially the case of a single error channel is examined, this analysis is then extended to to multiple channels. For each method, a number of different types of error complexity reduction are presented and analysed. Further reductions in error complexity are afforded by combining together different methods. In terms of system size, the error complexity reductions for the different methods are either constant, linear or exponential.

III. MAIN RESULTS

The methods presented in this work effect a reduction in the total number of coefficients required to fully describe the parallelised effective error channel. We now give informal statements of the error complexity reduction results for symmetry parallelisation, randomisation parallelisation and symmetry randomisation parallelisation. And finally we describe some applications of the error simplification techniques, along with experiments performed on superconducting hardware.

Symmetry Parallelisation

In symmetry parallelisation the error complexity reductions follow from symmetrisation of the effective channel. The same set of qubits is used for each of the different circuit parallelisations. In this case we use the term parallelisation as referring to circuits run on the same set of qubits at different times, with the different circuits still described as parallel circuits. The difference between these parallel circuits solely being the different qubit assignments on the device topology. The circuit mappings chosen result in symmetries being created in the effective error channel. The types of symmetry created were reflection, rotation, reflection and rotation, and permutation symmetry respectively.

Theorem 1. (Informal statement). *Reflection parallelisation results in an effective stochastic Pauli channel with* $O(4^n/2)$ *distinct Pauli coefficients.*

Theorem 2. (Informal statement). *Rotation parallelisation results in an effective stochastic Pauli channel with* $O(4^n/n)$ *distinct Pauli coefficients.*

Theorem 3. (Informal statement). *Reflection and rotation parallelisation results in an effective stochastic Pauli channel with* $O(4^n/2n)$ *distinct Pauli coefficients.*

Theorem 4. (Informal statement). *Permutation parallisation results in an effective stochastic Pauli channel with* $O(n^3/6)$ *distinct Pauli coefficients.*

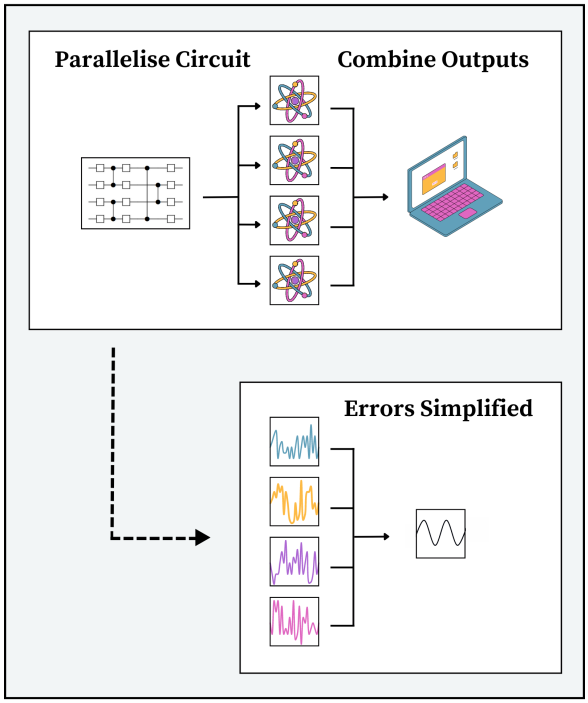


FIG. 1: *Error simplification schematic.* An input quantum circuit is run in parallel on multiple sets of noisy qubits. The computational outputs of the parallel circuits are then combined such that the effects of the errors are simplified. This simplification is achieved by the symmetrisation and randomisation of the noisy parallel circuit errors.

Randomisation parallelisation

The next approach is to use randomisation to simplify the errors in the effective channel. Here parallelisation is in space rather than time, as was the case with symmetry parallelisation. We show that by parallelising a circuit across multiple subsets of qubits with different stochastic Pauli channels, it is possible to obtain randomised effective error channels of lower complexity relative to the channels of the original circuits. To prove the randomisation parallelisation error complexity reductions we assume one of two error models. The first, referred to as $(r, 1)$, is a theoretically convenient model that allows for a reduction to global depolarizing noise. The second, referred to as $(r, 2)$, is a physically motivated model that allows for a reduction to an error channel consisting of a convex combination of depolarizing channels.

Theorem 5. (Informal statement). *Randomisation parallelisation with the (r,1) error model results in an effective stochastic Pauli channel which is convergent upon a channel with $O(1)$ Pauli coefficients, which is a global depolarizing channel with depolarizing parameter $\lambda = \frac{\eta(|\{\mathbf{P}^{\otimes n}\}|-1)}{|\{\mathbf{P}^{\otimes n}\}|}$.*

Theorem 6. (Informal statement). *Randomisation parallelisation with the (r,2) error model results in an effective stochastic Pauli channel which is convergent upon a channel with $O(2^n)$ Pauli coefficients. Each coefficient relates to a depolarizing channel acting on a subset of qubits q_j, for $j \in \{0, \ \ldots, \ 2^n - 2\}$, and the relation between Pauli coefficient η_{q_j} and depolarizing parameter λ_{q_j} is $\lambda_{q_j} = \frac{\eta_{q_j}(|\{\mathbf{P}^{\otimes n}\}|-1)}{|\{\mathbf{P}^{\otimes n}\}|}$.*

Symmetry randomisation parallelisation

Since symmetry and randomisation parallelisation apply different approaches to achieve error complexity reductions, combining the two methods results in greater reductions than that possible when each is applied individually. Only the $(r, 2)$ error model is used for the combined reductions, since no further error complexity reduction is possible for the $(r, 1)$ error model after randomisation parallelisation.

Sub-Theorem 6.1. (Informal statement). *Reflection and randomisation parallelisation using the $(r, \ 2)$ error model results in an effective stochastic Pauli channel with $O(2^n/2)$ distinct Pauli coefficients.*

Sub-Theorem 6.2. (Informal statement). *Rotation and randomisation parallelisation using the $(r, \ 2)$ error model results in an effective stochastic Pauli channel with $O(2^n/n)$ distinct Pauli coefficients.*

Sub-Theorem 6.3. (Informal statement). *Reflection, rotation and randomisation parallelisation using the $(r, \ 2)$ error model results in an effective stochastic Pauli channel with $O(2^n/2n)$ distinct Pauli coefficients.*

Sub-Theorem 6.4. (Informal statement). *Permutation and randomisation parallelisation using the $(r, \ 2)$ error model results in an effective stochastic Pauli channel with $O(n)$ distinct Pauli coefficients.*

Applications

There are many ways these methods might be usefully applied. We provide some example applications along with experiments on superconducting hardware. The first involves applying symmetry error simplification to reduce the sample cost of measurement error mitigation. In this case, the sampling overhead to perform the mitigation is reduced from 160000 to 60000 samples, with no loss of mitigation effectiveness. The second application leverages randomisation to enhance the effectiveness of noise-estimation circuit mitigation. In experiments on superconducting hardware, this the improved error mitigation effectiveness of the technique by $\sim 41\%$. The final application is in applying error randomisation to make noisy circuit performance for stable and predictable, increasing robustness to unpredictable noise accumulation and time-dependent errors. The randomisation parallelisation achieves this by making circuit noise more closely approximate depolarizing noise, and also by reducing time-dependent variation of the errors.

[1] J. J. Wallman and J. Emerson, Physical Review A **94**, 052325 (2016).
[2] A. Hashim, R. K. Naik, A. Morvan, J.-L. Ville, B. Mitchell, J. M. Kreikebaum, M. Davis, E. Smith, C. Iancu, K. P. O'Brien, I. Hincks, J. J. Wallman, J. Emerson, and I. Siddiqi, Physical Review X **11**, 041039 (2021), publisher: American Physical Society.
[3] C. Cade, L. Mineh, A. Montanaro, and S. Stanisic, Physical Review B **102**, 235122 (2020), publisher: American Physical Society.
[4] A. Broadbent and E. Kashefi, Theoretical Computer Science **410**, 2489 (2009).
[5] S. Bravyi, O. Dial, J. M. Gambetta, D. Gil, and Z. Nazario, Journal of Applied Physics **132**, 160902 (2022), publisher: American Institute of Physics.
[6] L. Liu and X. Dou, in *2021 IEEE International Symposium on High-Performance Computer Architecture (HPCA)* (2021) pp. 167–178, iSSN: 2378-203X.
[7] S. Niu and A. Todri-Sanial, "How Parallel Circuit Execution Can Be Useful for NISQ Computing?" (2021), arXiv:2112.00387 [quant-ph].
[8] S. Niu and A. Todri-Sanial, "Enabling multi-programming mechanism for quantum computing in the NISQ era," (2022), arXiv:2102.05321 [quant-ph].
[9] P. Das, S. S. Tannu, P. J. Nair, and M. Qureshi, in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, MICRO '52 (Association for Computing Machinery, New York, NY, USA, 2019) pp. 291–303.

Virtual Resource Distillation in Continuous Variables

Amalina Lai[1] * Mile Gu[1] † Ryuji Takagi[2] ‡

[1] *Nanyang Technological University*
[2] *University of Tokyo*

Abstract. Obtaining more resourceful quantum states, which are required for certain quantum processes and protocols, from less resourceful noisy states has been a keen topic of interest especially in the Noisy Intermediate-Scale Quantum (NISQ) era. Noisy states are generally states that have lost their quantum advantage to environmental factors and decoherence and may be discarded due to their lessened usefulness. While virtual resource distillation as a means to obtain more resourceful quantum states has been explored for discrete variable systems, less is known about virtual distillation in continuous variable systems, which have relevance in quantum communication and sensing. Current bounds in discrete variables will prove insufficient for the continuous variable regime. This project aims to explore a new figures of merit and provide examples for the viability of this protocol.

Keywords: probabilistic error cancellation, error mitigation, continuous variables

1 Introduction

Quantum computation often requires an input state to be put through a quantum circuit, where after which the resulting state is measured classically to obtain a measurement of interest. This quantum circuit may require some resourceful quantum state, which may be difficult to synthesize. However, measurement on this resourceful state can be simulated by manipulating measurements on a set of less resourceful (inferior) states that undergo their own quantum circuits. This is known as simulation or virtual resource distillation.

Quasiprobability sampling is a form of virtual distillation whereby a quantum circuit's measurement outcome probabilities are estimated using quasiprobability distributions. The method of quasiprobability then estimates the outcomes for quantum circuits whose quasiprobabilities may be negative, by making use of a non-negative representation of freely available quantum circuit elements that efficiently classically simulates and provides an unbiased estimator of the original quantum circuit [4]. In other words, quasiprobability sampling guarantees that its expectation value is always equal to the true mean of the distribution it attempts to simulate.

In essence, this protocol is used as a means to estimate the expectation value of a measurement on some resourceful state, by instead making measurements on free states that are available to us.

This method of quasiprobability sampling has been carried out in different situations in discrete variables, including simulating measurements on quantum magic states using free, classical stabilizer states [3]. In this manner, one might more loosely refer to this as the simulation of a quantum state, while more specifically referring the simulation of a quantum circuit involving a measurement on such a quantum state.

*nura0089@e.ntu.edu.sg
†gumile@ntu.edu.sg
‡ryuji.takagi@phys.c.u-tokyo.ac.jp

2 Virtual Resource Distillation Procedure

The sampling procedure involves first a quasiprobabilistic decomposition of the desired quantum circuit in terms of the freely available circuit elements,

$$\rho = \sum_i s_i \sigma_i \tag{1}$$

Where $s_i \in \mathbb{R}$ can be negative. Quasiprobability sampling allows the simulation of a measurement on ρ by measuring on our constituent states σ_i and then applying some post-processing on these measurement outcomes. To do so, a classical probability distribution can be derived from the quasiprobability distribution of the desired quantum circuit s_i,

$$p_i := \frac{|s_i|}{\sum_i |s_i|} \quad p_i > 0 \tag{2}$$

Sampling from this discrete, classical distribution, one obtains some i value. Consequently, one measures on the corresponding freely available circuit element ε_i to obtain a measurement outcome, m_i. However, the value of $\text{sign}(s_i)$ based on the sampling outcome i must also be tracked and included in the final simulation output, $M = \text{sign}(s_i) m_i \sum_i |s_i|$. Through repetition of this sampling process multiple times one obtains the average of N samples $\bar{M}$. This has been shown to provide an unbiased estimation of the output of the original circuit, since $\mathbb{E}[M]$ is equivalent to the original circuit's measurement outcome. However, it should be noted that the modified output of the simulation results in an increased variance of the random variable [3, 4].

3 Re-Expressing Figure of Merit In Discrete Variables

In discrete variables, the sampling procedure should output M, which is a random variable that is

bounded as $[-\max(m_i)\sum_i |s_i|, \max(m_i)\sum_i |s_i|]$, given that sign (s_i) can take on values ± 1. Depending the measurement and its measurement outcomes, the value of $\max m_i$ can vary.

Moreover, this maximum measurement outcome can be related to the operator norm of said measurement's operator, which should be Hermitian operator for a physical observable. Since the maximum measurement outcome should correspond to the largest eigenvalue, this is exactly given by the operator norm of the Hermitian operator corresponding to our measurement.

$$\max(m_i) = \|A\|_{\text{op}} \tag{3}$$

where

$$\|A\|_{\text{op}} := \inf\{c \geq 0 : \|Av\| \leq c\|v\|, \ \forall \ v \in V\} \tag{4}$$

Using Hoeffding's inequality, one can bound the probability that the average of N samples, $\bar{M}$ are within $\mathbb{E}[M]$ by some amount δ. Hoeffding's inequality [2] states that for Bernoulli (independently distributed) random variables $X_1, \ldots, X_n$, bounded as $a \leq X_i \leq b$, yields

$$\Pr(|\bar{X} - \mathbb{E}[X_i]| \geq \delta) \leq 2\exp\left[\frac{-2N^2\delta^2}{N(b-a)^2}\right] \tag{5}$$

This can be applied to our virtual resource distillation protocol, which involves independent random variables obtained through the sampling procedure elaborated above. One obtains then

$$\varepsilon := \Pr(\bar{M} - \mathbb{E}[M] \geq \delta) \tag{6}$$

$$\leq 2\exp\left[\frac{-2N\delta^2}{(2\max(m_i)\sum_i |s_i|)^2}\right] \tag{7}$$

$$= 2\exp\left[\frac{-N\delta^2}{2(\|A\|_{\text{op}}\sum_i |s_i|)^2}\right] \tag{8}$$

Thus the probability that the average of N samples, $\bar{M}$ are within $\mathbb{E}[M]$ by some amount δ is given by $1-\varepsilon$. This Hoeffding bound can be rearranged to yield the simulation cost, related to the number of samples, N,

$$N \geq \frac{2}{\delta^2}(\|A\|_{\text{op}}\sum_i |s_i|^2)\ln\left(\frac{\varepsilon}{2}\right) \tag{9}$$

Thus, we see that the simulation time cost scales quadratically with the sum of the absolute values of the quasiprobabilities and the maximum measurement outcome [3, 4]. Here, we extend the figure of merit used in prior literature, derived from Hoeffding's inequality, to be expressed in terms of the operator norm of the desired measurement operator.

4 Discrete Variable Bound that Tends to Infinity in Infinite Limit

In continuous variables, a state may still be similarly expressed in a quasiprobabilistic sum of free states, and the above sampling method may be applied. However, the analysis on the sampling cost would no longer hold in continuous variables, where the Hilbert space becomes infinite-dimensional.

In particular, the use of Hoeffding's inequality has an assumption that the random variables from the sampling method are bounded, which in general is not the case for continuous variable systems. In the case of continuous variables, the eigenvalues of a measurement can in general be unbounded, i.e. range between $(-\infty, \infty)$ [5]. As a result, generally $\|A\|_{\text{op}} \to \infty$ in the continuous variable, infinite-dimensional limit. This means that the lower bound for the number of samples N, also tends to infinity, rendering this figure of merit inadequate for the analysis of virtual resource distillation in continuous variable systems.

Thus, a new figures of merit must be determined to benchmark virtual resource distillation in the continuous variable regime.

5 Refined Figure of Merit

Consequently, in the application of virtual resource distillation to continuous variables, one cannot use Hoeffding's inequality to capture the cost of virtual resource distillation in continuous variable systems. Thus, a new figure of merit must be introduced.

To do so, we consider the quasiprobability sampling process for an arbitrary measurement it simulates. This process has a (one-shot) quasiprobability sampling distribution which gives the probability distribution for the sampling outcome, M, with mean $\mathbb{E}[M]$.

Moreover, as a probability distribution, this sampling distribution will have an associated variance, $\Delta^2 M$. This variance is positively related to not only the sum of negativities, $\sum_i |s_i|$, but the variance of the free states involved in the protocol as well. Thus, a higher value of sum negativity and/or variance of the free states, the higher the variance $\Delta^2 M$.

This can be applied together with the Central Limit Theorem which states that given a satisfactory minimum number of samples N, one can approximate the probability distribution of the average of N samples $(\bar{M})$ as a normal distribution with mean $\mu = \mathbb{E}[M]$ and standard deviation $\sigma_{CLT}^{QP} = \sqrt{\Delta^2 M/N}$, even if the original sampling distribution is not normal distribution.

This in turn allows us to determine the probability of successfully obtaining within some defined error bound δ in continuous variables. Moreover, through the Central Limit Theorem, one can also find an expression for the number of samples, N to satisfactorily obtain $\bar{M}$ within some error δ of $\mathbb{E}[M]$ with a minimum probability of success.

6 Example

By availing ourselves to non-Gaussian states, we consider how to distill a superior squeezed state from inferior squeezed states. In particular, we turn to a particular case of a non-Gaussian state, $\tau = s\sigma_0 + (1-s)\sigma_1$, composed of the convex sum of two different Gaussian states, σ_0 and σ_1 as illustrated in Figure 1. The constituent

Figure 1: Plot of the Wigner function of the non-Gaussian state τ in 3-dimensions in the phase-space. The state is a convex sum of two Gaussian states,σ_0 and σ_1 with different levels of squeezing in the x-quadrature, yielding distinct variances such that $\Delta^2 X_1^{sq} < \Delta^2 X_0^{sq}$.

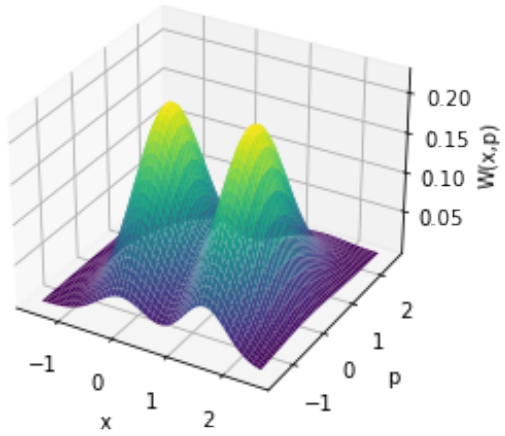
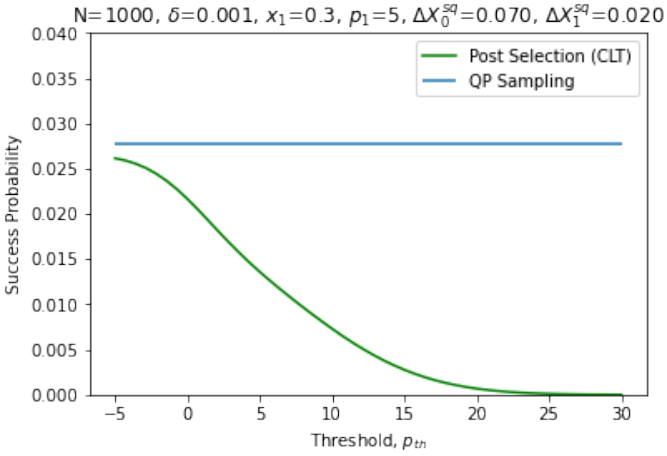


Figure 2: The ability for virtual resource distillation (green) to become the dominant protocol compared to post-selection (blue) requires a minimum amount of squeezing. In the figure, quasiprobability dominates with small Gaussian variances.



Gaussian states are not equally squeezed, with σ_0 being the less squeezed of the two. Given this non-Gaussian state together with the less squeezed Gaussian state, σ_0, one can obtain the more superior squeezed state of the two through some process. While this was introduced in the Heersink paper [1], their case was limited to both Gaussian states σ_0 and σ_1 having the same variance along the x- and p-quadratures (i.e. the two states are equally squeezed along the x-quadrature). As a result, there was not an actual distillation of a more resourceful squeezed state, since their final distilled state would have a squeezing that is roughly equivalent to the free state available to them. Thus in this project, the case was adapted and generalized to one where the variances of the two are not equal, such that we restrict ourselves to less resourceful states in an attempt to obtain a more resourceful, more squeezed Gaussian state.

Using the refined figure of merit introduced in the previous section, we can compare the probability of successfully distilling a more resourceful state via the traditional post-selection method against the new virtual resource distillation protocol put forth here. Virtual resource distillation, through analytical expressions, have been shown to provide advantage over the post-selection method, particularly when a minimum squeezing amount for the Gaussians have been reached, as seen in Figure 2. The probability of success for virtual resource distillation is greater than that for post-selection, across the range of post-selection threshold, p_{th}.

7 Conclusion

We identified problems porting this virtual resource distillation protocol from discrete variables to continuous variables, and address them by introducing a new Figure of Merit for the continuous variable regime. With this new Figure of Merit, we also show how it can be used to compare virtual resource distillation and a traditional post-selection protocol. virtual resource distillation in continuous variables has been shown here to not only be a viable protocol, but one that can potentially provide advantage against post-selection.

References

[1] J Heersink et al. “Distillation of squeezing from non-Gaussian quantum states”. In: *Physical review letters* 96.25 (2006), p. 253601.

[2] Wassily Hoeffding. “Probability inequalities for sums of bounded random variables”. In: *The collected works of Wassily Hoeffding* (1994), pp. 409–426.

[3] Mark Howard and Earl Campbell. “Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing”. In: *Physical Review Letters* 118.9 (2017). ISSN: 1079-7114. DOI: 10.1103/physrevlett.118.090501. URL: http://dx.doi.org/10.1103/PhysRevLett.118.090501.

[4] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. “Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities”. In: *Physical Review Letters* 115.7 (2015). ISSN: 1079-7114. DOI: 10.1103/physrevlett.115.070501. URL: http://dx.doi.org/10.1103/PhysRevLett.115.070501.

[5] Christian Weedbrook et al. “Gaussian quantum information”. In: *Reviews of Modern Physics* 84.2 (2012), p. 621.

Tailoring Non-Stabilizer Simulations for Analyzing Fault-Tolerant Error-Correction Codes

Mark B. Myers II[1] *

[1] *Centre for Quantum Technologies, National University of Singapore*

Abstract. As quantum devices emerge with increasing qubit counts, implementations with fault-tolerant quantum error correction (FTQEC) are of particular significance; however, such systems are difficult to verify and validate due to difficulties associated with simulating large systems under the effects of realistic noise. We aim to expand the understanding of quasi-probabilistic methods, for simulating non-stabilizer noise in quantum circuits, by investigating the method in the context of FTQEC simulations. We were able to achieve improved simulation efficiency by tailoring the method using domain knowledge, which was found to be consistent across the various non-unitary and unitary noise channels we investigated.

Keywords: quantum computation and simulation, fault-tolerant quantum error correction, stabilizer formalism, quantum noise channels, quasi-probabilistic simulations

1 Introduction

In our study we consider the method proposed by Bennink et al., which extends stabilizer simulations to include non-stabilizer channels. This method allows for any completely-positive trace-preserving channel in a circuit to be represented exactly as a quasi-probability distribution over stabilizer operations [1]. These decompositions are used in conjunction with Monte Carlo techniques to analyse various properties of a circuit. Crucially, Bennink et al. proposed that the efficiency of their method depends on both the number of non-stabilizer channels in a circuit and the negativity of the non-stabilizer channels being modeled; where negativity is a measure of how close the channel is to being a stabilizer channel. Notably, Hakkaku et al. have previously used this method to analyze the surface code under the effects of non-stabilizer noise. Their investigation focused on a single non-stabilizer channel that mixed over-rotation and bit-flip noise; wherein, they focused on more idealized noise models[2]. In this study we take a deeper look at this quasi-probability method, exploring the suitability of this method for obtaining metrics from FTQEC simulation where non-stabilizer noise is present. We highlight how Bennink's method can be tailored to FTQEC simulations for improved efficiency; whereby, we exploit several general properties of noise and error correction codes (ECCs).

2 Simulation Setup

Simulations can be employed to study and evaluate the performance of ECCs in the presence of noise. During an FTQEC simulation, various aspects of the code's performance can be analyzed, such as its ability to detect and correct specific types of errors, its code distance, or the fidelity of the ECC. When constructing an FTQEC simulation, there are five main aspects that must be considered:

- Error Correction Code - the error correction circuits
- Decoding Algorithm - the algorithm that determines where to apply correction operations
- Noise Model - the set of circuit locations where noise is applied
- Noise Channel - the type of noise being applied in the circuit
- Circuit Metric - the property being measured in the simulation

Each component of the simulation is chosen based on the research objectives. In the case of this study, we are interested in how ECCs behave under the effects of non-stabilizer noise, so we will be fixing the ECC to be the surface code, decoding algorithm to be minimum weight perfect matching, noise model to be circuit level noise, and circuit metric to be the logical infidelity; only varying the noise channels. It is important to clearly detail our choices when constructing simulations; so our approach is easier to understand and replicate.

3 Quasi-Probability Method Modifications

In our study we set out to investigate whether the quasi-probabilistic method for simulating non-stabilizer noise could be applied to FTQEC simulations to obtain valuable statistics like the logical infidelity; however, during the implementation process we found there were several properties of FTQEC simulations that lent themselves to improved simulation efficiency.

The first modification we made was related to the error correction capacity of a code; where an ECC can correct t faults. We can use this information to improve the estimator for the logical infidelity, because we can calculate the exact values for the portion of the estimator where t or fewer faults occur.

The second modification we made was related to the fact that we are predominantly focused on weak noise regimes, meaning that most of the noise-prone circuit elements will only experience an identity noise operation

*mbmyersii@u.nus.edu

in a given trial. In the base method, each identity operation still contributed to the magnitude of the trial's weight; however, by changing our sampling probabilities, we where able to reduce the average weights for each trial, and by extension reduce the variance of the estimator for the logical infidelity.

Our third modification was also related to the selection of the sampling distribution, which we optimized across all non-identity channel terms in the decomposition to reduce the variance of the estimator. The degree to which our modifications were able to reduce the variance differed slightly amongst the channels; however, it varied greatly with the channel infidelity.

4 Results and Conclusions

In this study we considered the distance 3, distance 5, and distance 7 surface codes; investigating them under the effects of the amplitude damping channel, random non-unitary channels, and random unitary channels. Across our tests, we were able to obtain the most accurate results when testing the amplitude damping and random non-unitary noise channels. Our results from the tests for unitary noise channels did not provide tight bounds on the estimators for the logical infidelity. As we increased the distance of the code the bound of our estimators increased, due to the increased number of trials necessary to achieve the same accuracy as smaller distance tests. Subsequently, we were unable to obtain tight bounds on any non-stabilizer channels for the distance 7 surface code and beyond, due to the increasing simulation costs.

This would indicate that this method is only suitable for analyzing FTQEC codes where the number of noise-prone operations is comparable to that of the distance 5 surface code, with approximately 5000 potentially noisy circuit operations. Such capabilities may be useful in some experimental applications, as long as the noise is non-unitary, or the number of noise-prone operations are limited. Crucially, we believe that these modifications were the most valuable aspect of this study, since all of the efficiency improvements we proposed will generalize to other FTQEC codes. Additionally, we believe that there is potential for further investigation about how additional information, relating to the noise channel and error correction code, could be incorporated into the simulation method for further efficiency improvements.

References

[1] Ryan S. Bennink, Erik M. Ferragut, Travis S. Humble, Jason A. Laska, James J. Nutaro, Mark G. Pleszkoch, and Raphael C. Pooser. Unbiased simulation of near-clifford quantum circuits. *Physical Review A*, 95(6), jun 2017.

[2] Shigeo Hakkaku, Kosuke Mitarai, and Keisuke Fujii. Sampling-based quasiprobability simulation for fault-tolerant quantum error correction on the surface codes under coherent noise. *Physical Review Research*, 3(4), nov 2021.

The landscape of QAOA Max-Cut Lie algebras

Martín Larocca[1] * Sujay Kazi[1 2] Marco Farinati[3] Patrick J. Coles[4] Robert Zeier[5] M. Cerezo[1] †

[1] *Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2]*New York University, New York, New York 10012, USA*
[3]*University of Burnos Aires, Buenos Aires, Argentina*
[4]*Normal Computing Corporation, New York, New York, USA*
[5]*Peter Grünberg Institute, 54245 Jülich, Germany*

Abstract. The Quantum Approximate Optimization Algorithm (QAOA) is a method for finding approximate solutions to the Max-Cut problem. This study examines Dynamical Lie algebra (DLA) for three QAOA ansatzes: the standard, orbit (orb), and multi-angle ansatzes, focusing on the reductive Lie algebra they represent and their isotypic decompositions. For the multi-angle we can fully characterize the DLA for all possible graphs, and show that this ansatz is extremely prone to barren plateaus even with a single-layered circuit. For the orb and standard ansatzes we provide an efficiently computable upper bound.

Keywords: Quantum computing, QAOA, Lie algebra, barren plateaus

1 Introduction

The quantum approximate optimization algorithm (QAOA) [1, 2] has been recently proposed as a promising candidate to make practical use of near-term quantum computers for solving combinatorial optimization problems. For example, one can find approximate solutions to the Maximum-Cut (Max-Cut) problem by encoding a graph's edges as interaction terms in an Ising-type Hamiltonian, and then using the QAOA to train a parametrized quantum circuit aimed at preparing the ground states of said Hamiltonian.

Not surprisingly, QAOA has been extensively studied in the past few years [3, 4, 5, 6, 7, 8], with many different variants of the QAOA parametrized quantum circuit ansatz being proposed [9, 10, 11, 12, 13, 14, 15]. Despite the tremendous attention QAOA has received, most of our understanding of this algorithm comes in the form of heuristics. This approach is at odds with recent advances in the field, as it has been shown that theoretically studying the underlying Dynamical Lie Algebra (DLA) [16] associated with a parametrized quantum circuit is fundamental to understand its performance. In particular, the presence of barren plateaus [17, 18, 19, 20, 21, 22] (i.e., exponentially vanishing gradients), and spurious local minima in the parameter optimization landscape can be linked to certain properties of the DLA (e.g., the dimension of the DLA is connected to the gradient scaling, such that exponential-sized DLAs are associated with barren plateaus) [20, 23, 24]. Hence, it appears that the DLA is the mathematical object that allows one to understand or predict the training landscape in QAOA.

Unfortunately, little-to-nothing is known about the DLA for QAOA Max-Cut applications. This issue gets amplified from the fact that QAOA is not immune to barren plateaus [20], indicating that a more precise understanding of its DLA is paramount. This motivates our work, which is the first comprehensive study of the DLA for various QAOA ansatzes and Max-Cut graphs.

2 Main results

We recall the Max-Cut problem and define the notation that will be used throughout this abstract. First, we recall that a graph G is defined by its vertex set $V = \{1, \ldots, n\}$ and its edge set E consisting of unordered pairs $\{k, \ell\}$ of vertices $\{k, \ell\} \in V$ with $k \neq \ell$. Given a graph G, the (optimization variant of the) Max-Cut problem, is to find a partition of its vertices into two complementary sets, such that the number of edges between those sets is as large as possible (see Fig. 1(a)). It has been shown that quantum computers can be used to find approximate solutions for the Max-Cut problem [1, 26, 8] by mapped it to the task of calculating the ground

*larocca@lanl.gov
†cerezo@lanl.gov

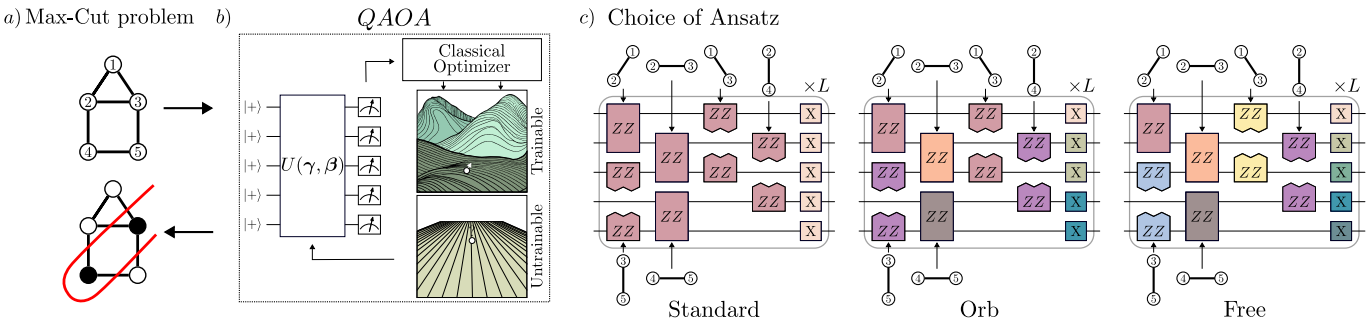


Figure 1: **Max-cut, and QAOA.** a) Given a graph, the Max-Cut problem is to determine partition of the vertices into two complementary sets, such that the number of edges between those sets is as large as possible. b) The QAOA algorithm is a hybrid quantum-classical algorithm that can be used to approximately solve the Max-Cut problem. The success of QAOA hinges on one's ability to optimizes the parameters in a quantum circuit $U(\boldsymbol{\gamma}, \boldsymbol{\beta})$ as in Eq. (3). c) While there are exists several ansatzes for $U(\boldsymbol{\gamma}, \boldsymbol{\beta})$. Here we consider the standard , orb [14] , and multi-angle (or free) [15, 25] ansatzes. In the image, a gate with ZZ indicates a two-qubit entangling gate generated by a $Z_k Z_\ell$ interaction, while the X gate indicates a single-qubit rotation around the x-axis. Boxes with the same color share the same parameter. One can see the the standard ansatz has less parameters per layer than the orb of free ansatzes.

state energy of the n-qubit Ising Hamiltonian

$$H_p = \sum_{\{k,\ell\}\in E} Z_k Z_\ell / 2 \,, \tag{1}$$

where Z_k denotes the Pauli-z operator acting on the k-th qubit. One can therefore attempt to obtain the Max-Cut variationally by defining the cost function

$$C(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \langle \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) | H_p | \psi(\boldsymbol{\gamma}, \boldsymbol{\beta}) \rangle \,, \tag{2}$$

and solving the optimization task $\arg\min_{\boldsymbol{\gamma},\boldsymbol{\beta}} C(\boldsymbol{\gamma}, \boldsymbol{\beta})$. As shown in Fig. 1(b), $|\psi(\boldsymbol{\gamma}, \boldsymbol{\beta})\rangle$ is obtained by initializing n qubits to the fiduciary state $|+\rangle^{\otimes n}$ and sending it through an L-layered parametrized quantum circuit of the form

$$U(\boldsymbol{\gamma}, \boldsymbol{\beta}) = \prod_{j=1}^{L} e^{-iH_m(\boldsymbol{\beta}_j)} e^{-iH_p(\boldsymbol{\gamma}_j)} \,. \tag{3}$$

Here, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1 \ldots, \boldsymbol{\gamma}_L)$ and $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_L)$ are vectors of trainable parameters, while $H_p(\boldsymbol{\gamma}_j)$ and $H_m(\boldsymbol{\beta}_j)$ are respectively known as the problem and mixer Hamiltonians.

Our studies are focused on three particular ansatzes, i.e., choices for $H_p(\boldsymbol{\gamma}_j)$ and $H_m(\boldsymbol{\beta}_j)$. The first is the standard ansatz as defined in the original QAOA manuscript [1]. The second and third ansatzes attempt to fix the drawbacks of the standard ansatz that it only contains a single parameter per layer. Particularly, we consider the orbit (orb) ansatz proposed in [14] where one borrows inspiration from the field of geometric quantum machine learning [27, 28] and ties the parameters in a layer according to the graph's automorphism group. Finally, we also study the multi-angle [15, 25], or free ansatz, where one assigns a single parameter to each gate in the standard ansatz. We present an example of these three ansatzes in Fig. 1(c). We will henceforth, denote the DLA of these circuits as $\mathfrak{g}_{\text{std}}$, $\mathfrak{g}_{\text{orb}}$ and $\mathfrak{g}_{\text{free}}$.

For the multi-angle QAOA ansatz we are able to fully characterize the circuit's DLA for any possible graph and prove that it only falls within one of six families. That is, we can prove the following theorem:

Theorem 1 (Mult-angle Lie algebra) *Given a connected graph, the DLA for the multi-angle QAOA ansatz $\mathfrak{g}_{\text{free}}$ fall into one of the six families depicted in Table 1.*

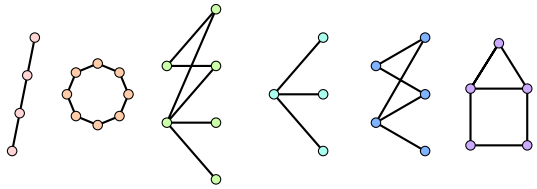
In particular, we find that for most graph families (except for the cycle and path graphs) the dimension of the DLA grows exponentially with the number of nodes n in the graph. *Such scaling implies that the multi-angle QAOA can be extremely prone to exhibiting barren plateaus, even when using a single layer of the ansatz.* For instance, we can show that the following corollary holds:

Corollary 1 *Consider a graph in the "other" category of Table 1. Then, let $\partial_\mu C(\boldsymbol{\gamma}, \boldsymbol{\beta})$ be the partial derivative of the cost. Given enough depth, we find*

$$\text{Var}_{\boldsymbol{\gamma},\boldsymbol{\beta}}[\partial_\mu C(\boldsymbol{\gamma}, \boldsymbol{\beta})] = \frac{2^n |E|}{8(2^{2n} - 4)(2^n + 2)} \,, \tag{4}$$

where $|E|$ is the number of edges.

Table 1: **Multi-angle ansatz.** Six families of connected graphs with n vertices with an example, their bipartiteness, Lie algebra (reductive decomposition), DLA (isotypic decomposition), and dimension. Here, r_{ff} denotes a $\frac{d}{2}$-dimensional free fermion representation while triv and std respectively denote the trivial and standard representations. Moreover, Dim refers to the algebra's dimension, i.e., the number of linearly independent terms in its basis.

	Path	Cycle	Even-even	Odd-odd	Even-odd	Other
Example						
Bipartite?	yes	if n is even	yes	yes	yes	no
Lie algebra	$\mathfrak{so}(2n)$	$\mathfrak{so}(2n)$ $\oplus\mathfrak{so}(2n)$	$\mathfrak{so}(2^{n-1})$ $\oplus\mathfrak{so}(2^{n-1})$	$\mathfrak{sp}(2^{n-1})$ $\oplus\mathfrak{sp}(2^{n-1})$	$\mathfrak{su}(2^{n-1})$	$\mathfrak{su}(2^{n-1})$ $\oplus\mathfrak{su}(2^{n-1})$
DLA	r_{ff} $\oplus r^*_{\text{ff}}$	$r_{(\text{ff,triv})}$ $\oplus r_{(\text{triv,ff})}$	$r_{(\text{std,triv})}$ $\oplus r_{(\text{triv,std})}$	$r_{(\text{std,triv})}$ $\oplus r_{(\text{triv,std})}$	r_{std} $\oplus r^*_{\text{std}}$	$r_{(\text{std,triv})}$ $\oplus r_{(\text{triv,std})}$
Dim	$\frac{2n(2n-1)}{2}$	$2n(2n-1)$	$2^{n-1}(2^{n-1}-1)$	$2^{n-1}(2^{n-1}+1)$	$2^{2n-2}-1$	$2^{2n-1}-2$

In the the orb and standard cases, we argue that a full characterization of the DLA for any graph is likely not possible. To make such a claim, we take a take a closer look at the DLA's symmetries [16], i.e., the set of operators that commutes with the parametrized unitary. Firstly, we show how the symmetries of the Max-Cut task get promoted to symmetries at the quantum level, and find that while the standard and orb ansatzes respect them, the multi-angle does not (and this is precisely why we can characterize it so well). Then, we show that the orb and standard ansatzes artificially introduces additional symmetries not related to the Max-Cut task. *These "hidden symmetries" arise from the fact that we use local gates in the circuit, and make the DLA graph-dependent and harder to study.* At this point, we introduce a DLA $\mathfrak{g}_{\text{Aut}}$ that solely respects the Max-Cut symmetries. Here we show that

Theorem 2 (DLA hierarchy) *For any graph G, the following chain of inclusions holds*

$$\mathfrak{g}_{\text{std}} \subseteq \mathfrak{g}_{\text{orb}} \subseteq \mathfrak{g}_{\text{Aut}} \subseteq \mathfrak{g}_{\text{free}}\,. \tag{5}$$

The power of this hierarchy resides in the fact that $\mathfrak{g}_{\text{Aut}}$ and can be characterized efficiently and therefore provide an accessible upper bound to the dimension of the orb and standard DLS. In particular, we can obtain the following result.

Theorem 3 *Given a graph G in the other family, we have*

$$\mathfrak{g}_{\text{Aut}} \oplus \mathfrak{u}(1) \oplus \mathfrak{u}(1) = \bigoplus_{\lambda} \mathfrak{u}(m_\lambda)\,. \tag{6}$$

where λ is here an index that runs over the irreps of the maximal symmetric subalgebra of $\mathfrak{u}(d)$ which respects the Max-Cut parity $\mathbb{Z}_2$ and automorphism $\text{Aut}(G)$ symmetries.

Theorem 3 shows that $\mathfrak{g}_{\text{Aut}}$ is semi-universal [29, 30].

The importance of the previous results is that we can leverage the bound in Theorem 2, along with the characterization in Theorem 3 to prove an exponential separation between the dimensions of the orb/standard DLAs, and the multi-angle DLA. For instance:

Proposition 1 *Let, G be (i) the complete graph, or (ii) G be a bipartite complete graphs. In these cases* $\dim(\mathfrak{g}_{\text{std}}), \dim(\mathfrak{g}_{\text{orb}}) \in \mathcal{O}(\text{poly}(n))$*, while* $\dim(\mathfrak{g}_{\text{free}}) \in \Omega(2^n)$.

As mentioned, Proposition 1 shows a clear separation between between the multi-angle ansatz, and the orb and standard ones.

3 Conclusions and Outlook

In this manuscript we have presented the first theoretical study of the Lie algebraic properties of the QAOA for Max-Cut. Our results have several important implications. First, we show that the large expressive power of the multi-angle ansatz leads to DLAs which are (for the vast majority of graphs) exponentially large. Thus, the multi-angle QAOA circuit is extremely prone to have barren plateaus. *For instance, we can find examples where the gradients are exponentially vanishing even when using a single layer of thie multi-angle circuit.* For the orb and standard circuit, we show that the circuits artificially introduce additional symmetries. While we provide an initial characterization of those symmetries, we open the door for studying them in future works. Finally, we present a DLA which only contains the appropriate symmetries, and which we can fully characterize. Critically, it is unknown to us if this DLA can be efficiently implemented in a quantum circuit (which opens up another research direction). Regardless, we argue that such ansatz is the appropriate one to use in QAOA for Max-Cut. In all cases, our work takes an important step forward towards characterizing the expressiveness of QAOA circuits, and paves the way towards Lie-algebraic-ansatz-design.

References

[1] E. Farhi, J. Goldstone, and S. Gutmann, arXiv preprint arXiv:1411.4028 (2014).

[2] E. Farhi and A. W. Harrow, arXiv preprint arXiv:1602.07674 (2016).

[3] R. Shaydulin, S. Hadfield, T. Hogg, and I. Safro, Quantum Information Processing **20**, 10.1007/s11128-021-03298-4 (2021).

[4] F. G. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, arXiv preprint arXiv:1812.04170 (2018).

[5] F. G. L. Brandao, R. Kueng, and D. S. França, Quantum **6**, 625 (2022).

[6] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Physical Review X **10**, 021067 (2020).

[7] S. Hadfield, T. Hogg, and E. G. Rieffel, arXiv preprint arXiv:2105.06996 (2021).

[8] M. P. Harrigan, K. J. Sung, M. Neeley, K. J. Satzinger, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, R. Barends, S. Boixo, *et al.*, Nature Physics , 1 (2021).

[9] A. Bärtschi and S. Eidenbenz, in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, 2020) pp. 72–82.

[10] Z. Wang, N. C. Rubin, J. M. Dominy, and E. G. Rieffel, Physical Review A **101**, 012320 (2020).

[11] R. Tate, B. Gard, G. Mohler, and S. Gupta, arXiv preprint arXiv:2112.11354 (2021).

[12] J. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, arXiv preprint arXiv:2301.11292 (2023).

[13] F. G. Fuchs, K. O. Lye, H. M. Nilsen, A. J. Stasik, and G. Sartor, arXiv preprint arXiv:2203.06095 (2022).

[14] F. Sauvage, M. Larocca, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2207.14413 https://doi.org/10.48550/arXiv.2207.14413 (2022).

[15] R. Herrman, P. C. Lotshaw, J. Ostrowski, T. S. Humble, and G. Siopsis, Scientific Reports **12**, 1 (2022).

[16] R. Zeier and T. Schulte-Herbrüggen, Journal of mathematical physics **52**, 113510 (2011).

[17] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nature Communications **9**, 1 (2018).

[18] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Nature Communications **12**, 1 (2021).

[19] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, PRX Quantum **3**, 010313 (2022).

[20] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, Quantum **6**, 824 (2022).

[21] K. Sharma, M. Cerezo, L. Cincio, and P. J. Coles, Physical Review Letters **128**, 180505 (2022).

[22] M. Cerezo and P. J. Coles, Quantum Science and Technology **6**, 035006 (2021).

[23] M. Larocca, N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2109.11676 (2021).

[24] L. Schatzki, M. Larocca, F. Sauvage, and M. Cerezo, arXiv preprint arXiv:2210.09974 (2022).

[25] K. Shi, R. Herrman, R. Shaydulin, S. Chakrabarti, M. Pistoia, and J. Larson, *2022 IEEE/ACM 7th Symposium on Edge Computing (SEC)*, , 414 (2022).

[26] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, Algorithms **12**, 34 (2019).

[27] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, PRX Quantum **3**, 030341 (2022).

[28] Q. T. Nguyen, L. Schatzki, P. Braccia, M. Ragone, M. Larocca, F. Sauvage, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2210.08566 (2022).

[29] I. Marvian, arXiv preprint arXiv:2302.12466 (2023).

[30] S. Kazi, M. Larocca, and M. Cerezo, arXiv preprint arXiv:2303.00728 (2023).

Sequential maximum confidence measurements

Kieran Flatt[1] * Hanwool Lee[1] † Joonwoo Bae[1] ‡

[1] *Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea*

Abstract. Sequential state discrimination is a form of state discrimination in which multiple parties aim to determine, through repeated measurement of the same quantum system, which state was prepared from some ensemble. Previous schemes have been based on unambiguous state discrimination, which is limited in its application. Here, we present a new sequential scheme based on maximum confidence measurements, which generalise unambiguous approaches. We characterise both the scenarios in which it can perform optimally, propose an alternative scheme which is applicable to all ensembles and analyse the latter's performance in some concrete examples.

Keywords: quantum information, quantum measurements, state discrimination

If a quantum system is projectively measured, it will be left in a pure state different from that it initially had. One may thus interpret that no more information can be extracted about the initial state. One surprising result, due to Bergou et al. [1], shows that this intuition is incorrect. In a sequential measurement scheme, any number of parties are able to determine which state was prepared, up to the same limits imposed in the single measurement scenario, from a given ensemble. This result gives rise to the topic of sequential state discrimination, in which any number of parties can attempt to determine which state was prepared by repeated measurement of the same system.

Bergou et al's scheme, however, is based upon unambiguous state discrimination (USD) and, thus, is restricted in the ensembles it can be applied to. Chefles' criterion, in particular, states that USD can be performed only for ensembles which form a linearly independent set [3].

In contrast with USD, maximum confidence measurements (MCMs) are state discrimination schemes with a universal applicability [2]. Consider an ensemble of N states, $\{qi, \rho_i\}_{i=1}^N$. We can construct a positive operator valued measure (POVM) such that each element π_i maximises the confidence C that a state was prepared given the relevant outcome:

$$C(i) = \mathrm{P}_{P|M}(\rho_i|\pi_i) = \frac{q_i \mathrm{P}_{M|P}(\pi_i|\rho_i)}{\mathrm{P}(\pi_i)}, \qquad (1)$$

taken from the standard Bayesian probability rule. The confidence-maximising operators may not form a complete set and so an inconclusive outcome π_0 will sometimes be necessary.Typically the rate of this outcome should be minimised. USD can then be defined as an MCM for which all confidences can reach $C(i) = 1$, and indeed the formalism will output the latter scheme for linearly independent ensembles.

In recent work, we have introduced the first analysis of *sequential maximum confidence measurements*. These have a wider range of applicability and therefore richer structure than the previous sequential state discrimination schemes. Our result characterise firstly the ensembles to which each party can obtain the same maximum confidence as would be available to a single measuring party. We then look at the most general scenario, and construct a universal sequential state discriminaton scheme, for which all parties are able to obtain some information about the initially prepared state. A trade-off is introduced between the confidence available to later parties and the inconclusive outcome rate of earlier parties. In fact, it turns out then all parties will be able to get arbitrarily close to the absolute maximum, at the cost, of course, of an arbitrarily high error rate for previous parties.

Let us quickly define the relevant terms and notation. The general scheme is depicted in Fig. 1. One party prepares a quantum system in a state chosen from some ensemble $\{qi, \rho_i\}_{i=1}^N$. This system is sent to a party who performs a measurement defined by the Kraus operators $K_i^{(1)}$ which aims to determine the initial state. This measurement leaves the system in a state $\rho_i^{(1)} \propto K_i^{(1)} \rho_i K_i^{(1)\dagger}$, and this is sent to a third party who measures with a different set of Kraus operators. The scheme can be extended in an obvious manner to M parties. The task is to construct the measurement in such a way that all parties learn something about the state. The mth party's Kraus operators are labelled $K_i^{(m)}$. Results from the theory of state discrimination tell us that POVM elements with rank greater than one will only decrease the confidence. Measurements outcomes are labelled $i = 0, 1...N$, where 0 denotes the inconclusive outcome.We call the "absolute maximum confidence" that which would be available for a single measuring party who is able to optimise her measurement with no further considerations and label it $\max[C(i)]$.

Our first result concerns the rangle of ensembles for which each party is able to get the absolute maximum confidence. We prove that this is always possible if the number of states N in the ensemble is less than or equal to the dimension of the Hilbert space. This is a weaker condition than the requirement of linear independence, which, for example, also enforces that all states are pure in the qubit case. Sequential MCMs therefore allow a

*kflatt@kaist.ac.kr
†hanwool283@kaist.ac.kr
‡joonwoo.bae@kaist.ac.kr

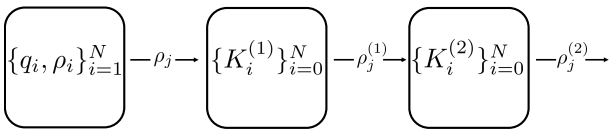


Figure 1: A multiple party state discrimination scenario includes a number of parties who each measure the same quantum system, prepared by an initial party in a state chosen from some ensemble.

wider ranger of sequential state discrimination scenarios than would otherwise be possible.

Our second result is to show that, even if the above condition does not hold, a scheme exists in which al parties are still able to gain information about the initially prepared state. In fact, they can get arbitrarily close to the maximum. Let us denote by Π_i the POVM elements which result inthe absolute maximum confidence and $K_i = \sqrt{\Pi_i}$ the associated Kraus operators. We then give all parties a weakened former of the measurement defined by the operators $\tilde{K}_i^{(m)} = \sqrt{\alpha} K_i$ for $i = 1...N$ and $\tilde{K}_0^{(m)} = (I - \alpha \sum_{i=1} \Pi_i)^{1/2}$. The weakening parameter α is directly related to the inconclusive outcome rate. The intuition for the measurement is that, as α becomes small, the resultant state will be $\tilde{\rho}_i \approx \rho_i + O(\alpha)$, and so the optimal measurement will be close to that for the original ensemble. In the small α limit, we find $\mathrm{P}(\pi_0) \approx 1 - \alpha$ and $C(i) \approx (1 - \alpha)\max[C(i)]$. The parties can therefore obtain an arbitrarily high confidence, at the cost of an arbitrarily high probability of failure for earlier parties.

We can finish our analysis by looking at a concrete example. The case we consider is that of the so-called GU states, defined as any N states symmetrically distributed around a great circle of the Bloch sphere. These ensembles have the two following properties: they form a decomposition of the identity; their absolute maximum confidence measurement is a rank-one measurement in the same direction as the state. These properties simplify the analysis and allow us to calculate the following trade-off between confidence for party m and success rate of the first $m - 1$ parties:

$$C_m(i) = \max[C(i)] \frac{1}{2} \left(1 + \Pi_{k=1}^{(m-1)} \frac{1}{2} \left(1 + \mathrm{P}_k(\pi_0) \right) \right) \quad (2)$$

where $\mathrm{P}_k(\pi_0)$ is the probability of the kth party's inconclusive outcome. Interestingly, this does not depend on the number of states in the ensemble.

We have introduce and analysed sequential maximum confidence measurements. First, we analysed the domain of applicability of absolute maximum confidence measurements, and then present a more general scheme which has universal applicability. As in the single measurement case, we find that the more general structure of MCMs allows for a more rich dynamics than USD, which is after all a special case of the former. It can also be said that our approach is more practical than the previous scheme, as USDs will in principle be impossible in the presence of detector noise and losses. These can be naturally taken into account in the wider framework of MCMs. There is scope to further explore our scheme, including the possibility of optimising over various parameters which were already assumed.

References

[1] Bergou, J., Feldman, E., Hilery M., *Phys. Rev. Lett. 111, 100501 (2013)*

[2] Croke, S. et al, *Phys Rev Lett 96, 070401 (2006)*

[3] Chefles, A., *Phys.Lett. A239 (1998)*

Recurrent generation of maximally entangled single particle states via quantum walks on cyclic graphs

Dinesh Kumar Panda[1 2 *] Colin Benjamin[1 2 †]

[1] *School of Physical Sciences, National Institute of Science Education and Research Bhubaneswar, Jatni 752050, India*

[2] *Homi Bhabha National Institute, Training School Complex, Anushaktinagar, Mumbai 400094, India*

Abstract. Maximally entangled single-particle states (MESPS) can encode more information and are robust to decoherence compared to their nonlocal counterparts. Using discrete-time quantum-walks on k-cycles where $k \in \{3, 4, 5, 8\}$ and by using either a single coin or effective-single coin or two coins in deterministic sequences, we generate MESPS for recurring time-steps, with periods 4, 6, 9, 12, and 15. For the first time, we reveal single coins like Hadamard can generate periodic MESPS on 4 or 8-cycle. This resource-saving scheme has a straightforward experimental realization and can be used in quantum cryptography with a MESPS public key.

Keywords: Entanglement, Single particle, Quantum walks, Cryptography

1 Introduction and our scheme

Quantum walks (QWs) are promising prime candidates for universal quantum computation, quantum simulations and efficient quantum algorithms [1, 2, 3, 4]. Moreover, QWs are used to explore and simulate exotic topological phases (edge states, Majorona modes etc.) [5] and to understand various complex physical and biological processes [6, 7]. QWs are achievable in labs using photons [8, 9, 10], trapped ions [11, 12], superconducting qubits [13, 14], neutral atoms [15, 16], NMR quantum information processors [17] and ultra-cold Rubidium-87 atoms [18].

Hybrid or single particle entanglement (SPE) enables encoding huge information even at a single particle level, as the entanglement is between different degrees of freedom like polarization, spatial mode and orbital angular momentum etc., belonging to the same particle [19]. Thus, SPE is resource-saving as well as it is more robust against decoherence and dephasing than its nonlocal counterparts. It has applications in analysis of states of photons, quantum liquids, and elementary particles as well as in the engineering of single photon quantum devices, and photonic quantum information. SPE is also a useful resource for quantum technologies such as quantum communication and quantum key distribution or QKD [19, 20, 21]. Exploiting QWs to generate controlled SPE is of phenomenal importance. In a recent work [22] we generate recurrent SPE via discrete time QWs on k-cycle (i.e., cyclic graph with k number of sites), using various deteministic evolution sequences with single coin, effective single coin or two coins.

2 Results

In our scheme, the QW evolve from the separable and pure initial state, $|\psi(t=0)\rangle = \cos\left(\frac{\theta}{2}\right)|0_p, 0_c\rangle + e^{i\phi}\sin\left(\frac{\theta}{2}\right)|0_p, 1_c\rangle$ with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$. Here, the QW coin space has computational basis $\{|0_c\rangle, |1_c\rangle\}$, whereas, the position space has computational basis $\{|x_p\rangle : x_p \in \{0, 1, 2, ..., k-1\}\}$. The full evolution operator of the QW is, $U_k = \hat{S}.[I_k \otimes \hat{C}_2]$ where, $\hat{S}$ and $\hat{C}_2$ are respectively be the shift and general coin operators, see [22]. Entanglement entropy (E) is used to quantify the entanglement between the coin and position degrees of freedom of the time-evolved quantum state $|\psi(t)\rangle$ [23]. We find several single-coin , effective single coin and two-coin evolution sequences which yield maximally entangled single particle states(MESPS) via the QW from the separable initial state with $\phi = \pi, \frac{\pi}{2}$. For MESPS, $E_{av} = 1$, where E_{av} is E averaged over θ. Along with general coins, we also use commonly used coin operators like Hadamard ($\hat{H}$), Fourier ($\hat{F}$), Grover ($\hat{X}$), identity ($\hat{I}$) etc., in the QW evolution.

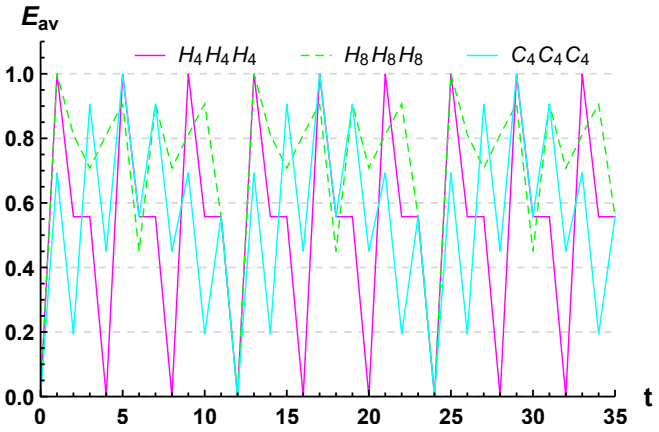


Figure 1: Average entanglement entropy E_{av} versus time steps(t) with single coin evolution sequences $H_4H_4H_4$..., $C_4C_4C_4$... for 4-cycle and $H_8H_8H_8$... for 8-cycle.

We for the first time, we report the single coin evolution sequences $H_4H_4H_4$... and $C_4C_4C_4$... which individually yields recurrent MESPS on a 4-cycle with periods 4 and 12 respectively, see Fig. 1. Apart from that, we showed that the single coin evolution sequence $H_8H_8H_8$...

*dineshkumar.quantum@gmail.com

†colin.nano@gmail.com

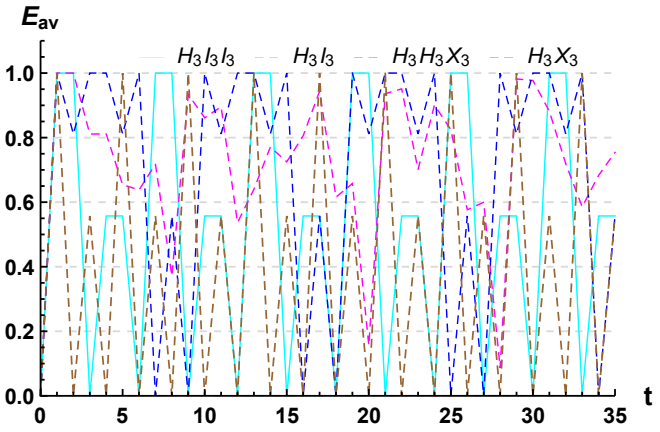


Figure 2: E_{av} versus time steps(t) with sequences: $H_3I_3I_3$..., H_3I_3..., $H_3H_3X_3$..., H_3X_3..., for 3-cycle.

generates periodic MESPS at $t = 1, 13, 25, ...$ (with period 12) on a 8-cycle. These sequences renders periodic or ordered QW too. An analytical proof for periodic QW supports the recurrent MESPS generation has been established and more than one MESPS can occur within the period of the QW, see [22].

We also see that employing the evolution sequeces $H_3H_3X_3$..., $H_3I_3I_3$..., and $H_3X_3H_3X_3$... on a 3-cycle one can obtain MESPS at all time steps up to 10, whereas on a 4-cycle their analogues give MESPS at all odd time steps $t \leq 10$, see Fig. 2. As these sequences also beget periodic QWs; thus, one obtains MESPS at larger time steps ($t > 10$) as well. Moreover, it is interesting to observe that with just evolution sequences $H_5H_5X_5$... and $H_5I_5I_5$..., one can generate MESPS for all time steps $t \leq 10$ and also at larger t, on a 5-cycle. In fact, only $H_5H_5X_5$... by itself yields MESPS at time steps $t = 1, 3, 4, 5, 6, 7, 8, 9, 10, 12, 16, ...$ with period 15.

Use in quantum cryptography–The periodically generated MESPS via our scheme can be used in quantum cryptography. We use the MESPS as the public key and we encode the message using it and the decryption requires a measurement based on the evolution sequence that generates the periodic MESPS. For details, see [22].

3 Summary

Our work [22] provides a novel scheme to generate maximally entangled single particle states or MESPS via DTQWs on both odd (3,5)- and even (4,8)-cycles, with just a single coin and with both resource-saving effective-single coin and two-coin deterministic evolution sequences. These deterministic sequences, along with generating MESPS for recurring time-steps yield the MESPS with periods 4, 6, 9, 12, and 15. For the first time, we reveal single coins like Hadamard that can generate periodic MESPS on a 4 or 8-cycle. This resource-saving scheme has a straightforward experimental realization and can be used in quantum cryptography with a MESPS public key.

One can experimentally implement our proposed scheme using linear optical elements, wherein the photon's polarization degree of freedom encodes the coin state with the position state is encoded into different time bins of the photon [24, 25]. Apart from opening a unique avenue for MESPS generation, our work significantly outperforms other existing schemes in terms of model simplicity and resource-saving architecture and periodically yields MESPS at both small and large time steps [22]. We provide a Python code for numerical experiments in [22].

Our findings naturally raise new intriguing research questions: What is the interplay between disorder and entanglement (both hybrid and nonlocal) generation for 1D or higher dimensional walkers? Can this scheme be adapted to further improve existing cryptography protocols [26, 27]? Investigations into these directions may lead to significant results which foster new local or nonlocal entanglement generation schemes and their applications.

Our presented work will significantly contribute towards state-of-art controlled (maximal) entanglement generation protocols, which is a fundamental resource in quantum computing, teleportation, and cryptography and hence, a prerequisite to constructing reliable devices for quantum information processing tasks.

References

[1] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, Phys. Rev. A 48, 1687 (1993).

[2] R. Portugal, Quantum Walks and Search Algorithms (Springer, 2019).

[3] A. Ambainis, Quantum search algorithms, SIGACT News 35, 22 (2004)

[4] V. Kendon, A random walk approach to quantum algorithms, Philos. Trans. R. Soc., A 364, 3407 (2006).

[5] Vikash Mittal, Aswathy Raj, Sanjib Dey and Sandeep K. Goyal, Persistence of topological phases in non-Hermitian quantum walks, Sci Rep 11, 10262 (2021).

[6] Kunkun Wang, Xingze Qiu, Lei Xiao, Xiang Zhan, Zhihao Bian, Wei Yi, and Peng Xue, Simulating Dynamic Quantum Phase Transitions in Photonic Quantum Walks, Phys. Rev. Lett. 122, 020501 (2019)

[7] Masoud Mohseni, Patrick Rebentrost, Seth Lloyd, and Alán Aspuru-Guzik , Environment-assisted quantum walks in photosynthetic energy transfer, J. Chem. Phys. 129, 174106 (2008).

[8] Xiao-Xu Fang, Kui An, Bai-Tao Zhang, Barry C. Sanders, He Lu, Maximal coin-position entanglement generation in a quantum walk for the third step and beyond regardless of the initial state, Phys. Rev. A 107, 012433 (2023).

[9] F. Cardano, et al., quantum walks and wave packet dynamics on a lattice with twisted photons, Science Advances 1, 15516 (2015).

[10] A. Schreiber, K. N. Cassemiro, V. Potocek, A. Gabris ,P. J. Mosley, E. Andersson, I. Jex, and C. Silberhorn, Photons walking the line: A quantum walk with adjustable coin operations, Phys. Rev. Lett. 104, 050502 (2010).

[11] H. Schmitz, R. Matjeschk, C. Schneider, J. Glueckert, M. Enderlein, T. Huber, and T. Schaetz, Quantum walk of a trapped ion in phase space, Phys. Rev. Lett. 103, 090504 (2009).

[12] F. Zahringer, et al., realization of a quantum walk with one and two trapped ions, Physical Review Letters 104, 100503 (2010).

[13] J.-Q. Zhou, L. Cai, Q.-P. Su, and C.-P. Yang, Protocol of a quantum walk in circuit QED, Phys. Rev. A 100, 012343 (2019).

[14] E. Flurin, V. Ramasesh, S. Hacohen-Gourgy, L. Martin, N. Yao, and I. Siddiqi, Observing topological invariants using quantum walks in superconducting circuits, Physical Review X 7, 031023 (2017).

[15] M. Karski, L. Forster, J.-M. Choi, A. Steffen, W. Alt, D. Meschede, and A. Widera, Quantum walk in position space with single optically trapped atoms, Science 325, 174–177 (2009).

[16] C. Robens, S. Brakhane, D. Meschede, and A. Alberti, Quantum walks with neutral atoms: Quantum interference effects of one and two particles, Laser Spectroscopy, 1-15 (2016).

[17] C. A. Ryan, M. Laforest, J. C. Boileau, and R. Laflamme, Experimental implementation of a discrete-time quantum random walk on an NMR quantum-information processor, Phys. Rev. A 72, 062317 (2005).

[18] K. Manouchehri and J. Wang, Quantum random walk with Rb atoms, Journal of Physics: Conference Series 185, 012026 (2009).

[19] S. Azzini, S. Mazzucchi, V. Moretti, D. Pastorello, and L. Pavesi, Single-Particle Entanglement, Advanced Quantum Technologies 3, 2000014 (2020).

[20] S. Adhikari, D. Home, A. S. Majumdar, A. K. Pan, A. Shenoy H., R. Srikanth, Toward secure communication using intra-particle entanglement, Quantum Inf. Process. 14, 1451 (2015).

[21] T. Pramanik, S. Adhikari, A.S. Majumdar, Dipankar Home, Alok Kumar Pan, Information transfer using a single particle path-spin hybrid entangled state, Phys. Lett. A 374, 1121 (2010).

[22] Dinesh Kumar Panda and Colin Benjamin, Recurrent generation of maximally entangled single particle states via quantum walks on cyclic graphs, arXiv:2301.04501 [quant-ph] (2023).

[23] D. Janzing, Entropy of entanglement, Compendium of Quantum Physics, 205–209 (2009).

[24] Zhi-Hao Bian, Jian Li, Xiang Zhan, Jason Twamley, and Peng Xue, Experimental implementation of a quantum walk on a circle with single photons, Phys. Rev. A 95, 052338 (2017)

[25] R. Vieira, E. P. M. Amorim, and G. Rigolin, Dynamically disordered quantum walk as a maximal entanglement generator, Phys. Rev. Lett. 111, 180503 (2013).

[26] Abhisek Panda and Colin Benjamin, Order from chaos in quantum walks on cyclic graphs, Phys. Rev. A 104, 012204 (2021).

[27] C. Vlachou et al., Quantum walk public-key cryptographic system, Int. J. Quantum Inf. 13, 1550050 (2015).

Characterization of quantum entanglement in Si quantum dot systems: Operational quasiprobability approach

Junghee Ryu[1] Hoon Ryu[1] *

[1] *Division of National Supercomputing, Korea Institute of Science and Technology Information, Daejeon 34141, Republic of Korea*

Abstract. We characterize a bipartite entanglement in a realistic silicon double quantum dot platform. Arbitrary two-qubit entangled states are generated by conducting a single-qubit rotation and a controlled-NOT operation. To quantify the entanglement, we employ a marginal operational quasiprobability (OQ) function, which serves as a reliable entanglement witness even in the presence of significant noise. We here discuss how the entanglement characteristic of the Si DQD structure are affected by charge noises which is omnipresent in semiconductor devices.

Keywords: Entanglement, Silicon quantum dot, Operational quasiprobability

1 Introduction

Quantum correlations, such as entanglement, play a crucial role in quantum information technologies, providing advantages over classical counterparts in quantum computing, quantum communication, quantum metrology, and so on. Thus, the quantification of the quantum correlations is important as it can be used to explore the potential practicality in information processing e.g., suitable states of certain quantum circuits. Especially, we here study the characterization of the quantum entanglement. To this end, we employ a marginal operational quasiprobability (OQ) function that allows negative values of the function if a given state is entangled [1]. We apply the marginal OQ method to the electron-spin qubits in a silicon (Si) double quantum dot (DQD) platform, where a single-qubit rotation and a two-qubit controlled-NOT operation are conducted sequentially in time to generate arbitrary two-qubit entangled states.

2 Method

We use a newly defined quasiprobability function for discrete systems, which is linked to experimental situations in which incompatible observables are measured consecutively. It turns out that the OQ function method identifies the nonclassicality of quantum systems in an operational way. Furthermore, for multipartite systems the marginal OQ function can be used as an entanglement witness. In principle, the OQ approach is advantageous over the entanglement verifications involving the full state tomography process in a sense that our method can be calculated with directly measurable quantities in laboratory and requires less number of measurements to characterize the entanglement.

The N-qubit OQ function is defined by applying a discrete Fourier transform on the composite expectation value $C(\mathbf{n}^1,\ldots,\mathbf{n}^N) \equiv C(\mathbf{n}^1)\otimes\cdots\otimes C(\mathbf{n}^N)$ as

$$\mathcal{W}(\mathbf{a}^1,\ldots,\mathbf{a}^N) \equiv \frac{1}{2^{NK}} \sum_{\mathbf{n}^1,\ldots,\mathbf{n}^N} (-1)^{-\mathbf{a}^1\cdot\mathbf{n}^1\cdots-\mathbf{a}^N\cdot\mathbf{n}^N} \times C(\mathbf{n}^1,\ldots,\mathbf{n}^N), \quad (1)$$

where a tuple $\mathbf{n}^i = (n_1^i, n_2^i, \ldots, n_K^i)$ represents possible measurement configurations for i-th subsystem having K measurement operators and $\mathbf{a}^i\cdot\mathbf{n}^i = \sum_k a_k^i n_k^i$. The expectation value $C(\mathbf{n}^1,\ldots,\mathbf{n}^N)$ represents the measurement configurations that are implemented in a laboratory (see Ref. [2] for more details). We consider the following formula to quantity the entanglement

$$\mathcal{N} \equiv \frac{1}{2}\sum_{\mathbf{a}} \left(|\mathcal{W}(\mathbf{a})| - \mathcal{W}(\mathbf{a})\right). \quad (2)$$

The value $\mathcal{N}$ indicates the sum of the negative components of the OQ function, thus the case of $\mathcal{N} > 0$ can be regarded as the indicator of the entanglement for given quantum systems.

Our working example is the two-qubit time responses that are generated from a Si DQD system. A 2D simulation domain of DQD structure reported in [3] encodes qubits to electron spins that are created with quantum confinement driven by biases imposed on top electrodes, see Figure 1(a). The DQD system is initialized to a $|\downarrow\downarrow\rangle$ state by filling the ground down-spin state of the left and right quantum dot with a single electron. To this end, we set the left (V_{L}) and right gate bias (V_{R}) to 555mV. For the middle gate bias (V_{M}), we consider two cases: (a) 400mV with an exchange energy (J) of 76KHz (weak interaction) and (b) 407.5mV with 18.4MHz (strong interaction) respectively, as shown in Figure 1(b). A spatial distribution of the static magnetic field that is generated by a horseshoe-shaped cobalt micro-magnet in the real case [3] is utilized as an input of simulations. The resulting Zeeman-splitting energy of the left (E_{ZL}) and right spin (E_{ZR}) turn out to be 18.31GHz and 18.45GHz respectively. All these conditions imply that we are able to implement a single-qubit rotation and a two-qubit gate operation to the initial state by controlling the middle gate bias (V_{M}).

*elec1020@kisti.re.kr

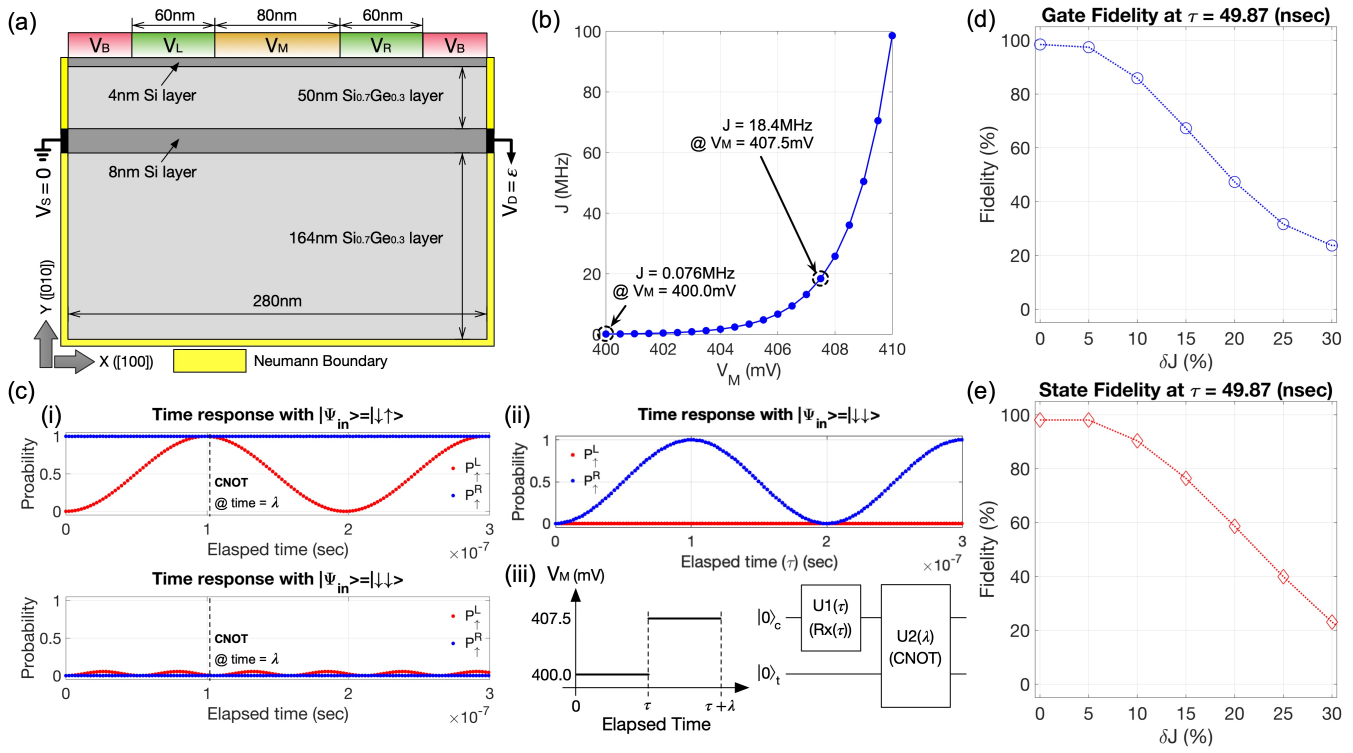


Figure 1: (a) A 2D simulation domain for our case. The real DQD structure [3] is long along the Z([001])-direction, thus it is described in a 2D manner with a periodic boundary condition along the Z-direction. (b) J given as a function V_{M} for $V_{\mathrm{R}} = V_{\mathrm{L}} = 555$mV. In our case, $J \sim$ 76KHz and 18.4MHz when $V_{\mathrm{M}} = 400$mV and 407.5mV, respectively. (c) (i) For the strong interaction ($V_M = 407.5$mV), the fastest CNOT operation can be achieved in $\sim 1.05\times10^{-7}$ (λ) seconds upon the system intialization. The gate fidelity of the CNOT operation becomes 98.35%. (ii) For the weak interaction ($V_M = 400$mV), we can make only the right spin oscillate by setting the frequency of AC pulse equal to the Zeeman-splitting energy of the right spin. (iii) A conceptual illustration for time-dependent control of V_M and resulting two-qubit unitary gate that generates the entangled states. (d, e) The fidelity of the two-qubit unitary and the corresponding output state at $\tau = 4.99\times10^{-8}$ seconds (the time spot when the output state is maximally entangled) are shown as a function of δJ, which represents the unintentional variation of J with respect to its noise-free value.

3 Results

In order to characterize the entanglement, we employ the marginal OQ method with the two measurement operators. In general, the OQ function can be constructed by using the positive operator-valued measure (POVM) measurements, but we here consider only two projective measurements defined by the Pauli matrices for cost-efficient calculations [1]. By calculating the negative values of the marginal OQ function, we can quantify the entanglement. The states are generated by the sequential application of a $R_x(\alpha)$ and a CNOT operation, thus the output can be expressed by $|\psi(\alpha)\rangle = \cos(\alpha/2)\,|00\rangle - i\sin(\alpha/2)\,|11\rangle$. The noise-driven characteristic of entanglement is also investigated by changing the noise-free exchange interaction J to $J\times(1+\delta J)$ as we treated to simulate the fidelity shown in Figure 1(c) and Figure 1(d).

Figure 2(a) shows the results of the noise-free case ($\delta J = 0$) as a function of the time τ. The blue (normalized) and green (raw) lines indicate the entanglement strength calculated with the marginal OQ method, and the red line is the one obtained with the negativity method. The maximal strength reads 0.2348 (green line) at $\tau = 4.99\times10^{-8}$ seconds. Note that there exist the intervals of τ where our method cannot characterize the entanglement precisely, which is because the marginal OQ function is constructed by using only two measurement operators for the cost-efficient calculations.

We also explore the behavior of the entanglement characteristic when the Si DQD platform suffers from the charge noises ($\delta J \neq 0$), which is omnipresent in semiconductor devices. Figure 2(b) shows the results of the noise-driven degradation in fidelity and in the marginal OQ method. The noise-driven pattern of entanglement characterization does not necessarily follow that of fidelity, and the output state of the noisy two-qubit operation still has meaningful strength of entanglement. We find that while the charge noise causes huge degradation in the state fidelity, it has a weaker effect on the entanglement resource. In a highly noisy environment, the state fidelity drops to around 20%, but more than 70% of the resource can be retained for maximally entangled Bell states. It should be noted that as shown in Figure 1(d) and 1(e), the gate and state fidelity are sensitive to charge noises, and are sharply reduced as δJ increases.

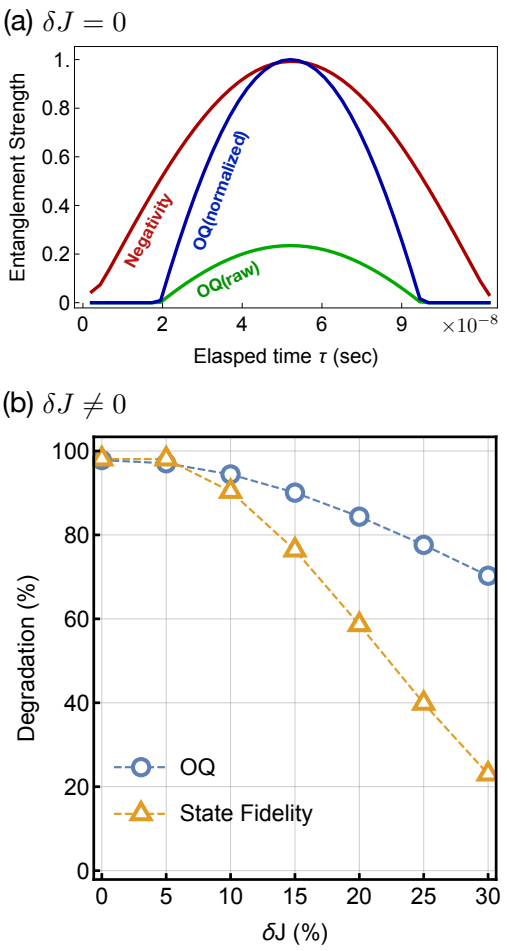


Figure 2: (a) The results of the noise-free case ($\delta J = 0$) as a function of the time τ. (b) The results of the noise-driven degradation in state fidelity and OQ method.

However, Figure 2(b) clearly shows that when the noise is too strong the results we present can be still fairly solid enough to claim the utility of the marginal OQ method as a cost-efficient indicator of entanglement strength, where the cost-efficiency of our method against the negativity method will sharply increase as the size (in qubits) of targeted quantum states increases.

4 Acknowledgements

This work has been carried out under the support of the National Research Foundation of Korea (NRF) grant (NRF-2020M3E4A1079792 and NRF-2022M3K2A1083890).

References

[1] J. Ryu *et al.* Operational quasiprobabilities for qudits *Phys. Rev. A* **88**, 052123 (2013).

[2] J. Ryu *et al.* Exploring entanglement resource in Si quantum dot systems with operational quasiprobability approach *Quantum* **6**, 827 (2022).

[3] D. M. Zajac *et al.* Resonantly driven CNOT gate for electron spins *Science* **359**, 439 (2018).

Optimal quantum metrology for two-photon absorption parameter estimation

Athena Karsa[1] * Ranjith Nair[2 3] Andy Chia[4] Kwang-Geol Lee[5]
Changhyoup Lee[1] †

[1] *Korea Research Institute of Standards and Science, Daejeon 34113, Korea*
[2] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*
[3] *Complexity Institute, Nanyang Technological University, Singapore*
[4] *Centre for Quantum Technologies, National University of Singapore, Singapore*
[5] *Department of Physics, Hanyang University, Seoul, Korea*

Abstract. Two-photon absorption (TPA) is a non-linear optical process with wide-ranging applications from spectroscopy to super-resolution imaging. Discrete-variable quantum states are optimised to maximise the quantum Fisher information (QFI) for given losses to find optimal probes for TPA parameter estimation. Our found optimal probes yield a quantum advantage compared to both the coherent state classical benchmark and the single-mode squeezed vacuum state, while also having a relatively simple structure and loss-dependent character. Further, photon counting is demonstrated to offer optimal or nearly optimal performance compared to the QFI bound and our findings help to explain the already-known behaviours of Gaussian probes.

Keywords: two-photon absorption, parameter estimation, quantum metrology, optimisation

1 Introduction

The study and use of non-classical quantum states of light is ubiquitous in quantum technologies, providing a potential resource for achieving a quantum enhancement in tasks such as metrology [1, 2], imaging [3] and spectroscopy [4]. Of particular interest are their possible role within non-linear light-matter interactions with one example, the subject of this work, being two-photon absorption (TPA). Notably, squeezed light sources have particular significance since their absorption probability scales linearly with the intensity of the light field, as opposed to the quadratic scaling observed with laser light [5, 6, 7, 8, 9, 10]. This effect may thus potentially enable non-linear spectroscopy and microscopy at low photon fluxes, in turn enabling the exploitation of TPA-based protocols for the interrogation of photosensitive samples or even living organisms [11, 12, 13, 14, 15, 16] while mitigating their degradation.

Recent years have seen a multitude of ongoing developments in quantum-enhanced protocols using states of light, owing to greater interest in quantum sensing and, alongside it, a greater wealth of experimental tools providing access to new quantum light sources and means for their utility [17, 18]. With respect to quantum-enhanced absorption measurements, first interest was in the regime of single-photon absorption [19, 20] finding that Fock states at any fixed photon number ($\bar{n}$) saturate the ultimate bound of precision given by $1/4\bar{n}$ for a single run [21, 22].

More recently interest has been targeted towards TPA [23, 24] focusing on the quantum metrological properties of Gaussian states of light, namely squeezed vacuum and coherent states, for TPA parameter estimation.

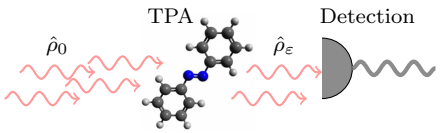


Figure 1: Schematic diagram of two-photon absorption (TPA). An input quantum state $\hat{\rho}_0$ is prepared and subsequently evolves via propagation through some TPA medium into output state $\hat{\rho}_\varepsilon$ before being measured.

Such states can straightforwardly be expressed in terms of a unitary operation acting on the vacuum, which in turn allows for a simplification of the TPA analysis which is valid only in the regime of extremely low TPA rates. While it is possible to consider this regime a valid one, since typical TPA cross-sections are generally small, it fails to allow for the study of arbitrary quantum states whose generation from the vacuum are non-trivial. Further, it fails to enable one to compare their metrological potential across all possible scales of loss, which is crucial for making accurate, practical comparisons which may also factor in experimental inefficiencies.

In this work, a general discrete-variable quantum state is optimised to maximise the QFI for fixed TPA parameter value. We propose a prospective, theoretically-optimal single-mode quantum probe; the properties of which provide insight into the loss-dependent behaviours for three particular quantum states, the coherent, squeezed vacuum and Fock state, to which our optimal probe is compared. While it turns out that for TPA the Fock state is not optimal for all losses and energies, superpositions of definite photon number states are and their underlying constituents are a direct function of the TPA loss parameter.

*athena.karsa@gmail.com
†changhyoup.lee@gmail.com

2 Quantum metrology of two-photon absorption

2.1 Limits on measurement sensitivities

Parameter estimation theory states that the error of an estimator $\hat{\varepsilon}$ of the true value of parameter ε is given by the mean-square error $\Delta^2\varepsilon = \langle(\hat{\varepsilon}-\varepsilon)^2\rangle$, where $\langle\dots\rangle$ denotes the average over all measurement results. For an unbiased estimator, its precision $\Delta\varepsilon$ is limited by the quantum Cramér-Rao bound (CRB), $\Delta^2\varepsilon \geq \frac{1}{N\mathcal{F}_{\mathrm{Q}}}$, where N is the number of repetitions of a given measurement and $\mathcal{F}_{\mathrm{Q}}$ is the quantum Fisher information (QFI).

For a given positive-operator-valued measure (POVM) $\{\hat{\Pi}_k\}$, with $\hat{\Pi}_k \geq 0$ and $\sum_k \hat{\Pi}_k = \hat{\mathbb{1}}$, one is generally able to achieve the classical Fisher information (CFI),

$$\mathcal{F}_{\mathrm{C}}(\hat{\rho}_\varepsilon, \hat{\Pi}_k) = \sum_k P_k(\varepsilon)\left(\frac{d}{d\varepsilon}\log P_k(\varepsilon)\right), \tag{1}$$

where $P_k(\varepsilon) = \mathrm{Tr}\left[\hat{\Pi}_k\hat{\rho}_\varepsilon\right]$ is the probability of measurement outcome k. Then, the CRB becomes

$$\Delta^2\varepsilon \geq \frac{1}{N\mathcal{F}_{\mathrm{C}}} \geq \frac{1}{N\mathcal{F}_{\mathrm{Q}}}, \tag{2}$$

where the second inequality is saturated for an optimal measurement choice, i.e., $\mathcal{F}_{\mathrm{Q}} = \max_{\{\hat{\Pi}_k\}}\mathcal{F}_{\mathrm{C}}$.

Given a parameter dependent state $\hat{\rho}_\varepsilon$ the corresponding QFI may be readily computed as

$$\mathcal{F}_{\mathrm{Q}} = \mathrm{Tr}\left[\hat{L}_\varepsilon^2\hat{\rho}_\varepsilon\right], \quad \partial_\varepsilon\hat{\rho}_\varepsilon = \frac{1}{2}\left(\hat{L}_\varepsilon\hat{\rho}_\varepsilon + \hat{\rho}_\varepsilon\hat{L}_\varepsilon\right), \tag{3}$$

where $\hat{L}_\varepsilon$ is symmetric logarithmic derivative (SLD),

$$\hat{L}_\varepsilon = 2\sum_{k,l}\frac{\langle l|\left(\mathcal{L}\hat{\rho}\right)|k\rangle}{\lambda_k+\lambda_l}|l\rangle\langle k|, \tag{4}$$

defined for all $\lambda_k + \lambda_l > 0$. The optimal measurement setup which saturates the quantum CRB is one constituting the POVM made up by a set of projectors over the eigenbasis of the SLD operator, $\hat{L}_\varepsilon$.

2.2 Master equation

Consider the transmission of a quantum state of light through a TPA medium, as depicted in Fig. 1, whose subsequent dynamical evolution is modelled by the Markovian Lindblad equation

$$\dot{\hat{\rho}} = \gamma\mathcal{L}\hat{\rho} = \frac{\gamma}{2}\left(2\hat{L}\hat{\rho}\hat{L}^\dagger - \hat{L}^\dagger\hat{L}\hat{\rho} - \hat{\rho}\hat{L}^\dagger\hat{L}\right). \tag{5}$$

Here, the Lindblad operator is given by the two-photon loss operator $\hat{L} = \hat{a}^2/\sqrt{2}$ and its Hermitian conjugate where $\hat{a}(\hat{a}^\dagger)$ denotes the photon annihilation(creation) operator of the field. The goal for metrological applications of TPA is to measure the absorbance $\varepsilon \equiv \gamma t$, where t is the propagation time of the illuminating field of light through the medium and γ is its loss rate.

Two-photon field dissipation, governed by Eq. (5) [25, 26, 27, 28], has a algebraic solution which yields the output density matrix after a certain loss ε (see Ref. [29]). For arbitrary input quantum states $\hat{\rho}_0$, this form can be used to compute the corresponding output states $\hat{\rho}_\varepsilon$ after TPA evolution in terms of its Fock state representation.

3 Results and Discussion

3.1 Optimal states for TPA quantum metrology

Consider an optimisation procedure whose goal is to maximise the QFI by varying the coefficients of some arbitrary quantum state $|\psi\rangle = \sum_{n=0}^{\tilde{N}} c_n|n\rangle$, with $n \geq 0$. The mean value is fixed, i.e., $\sum_n nc_n^2 = \bar{n}$ and, physicality in terms of unit trace is preserved. A nested optimisation algorithm is used whose first part searches for a global optimum, based on the evolutionary algorithm library EvoTorch [30], and the output forms the seed of a subsequent ADAM optimiser which fine-tunes the result through gradient descent.

The results of the optimisation are shown in Fig. 2 for mean photon numbers $\bar{n} = 2$ and 3. Note the reparameterisation $\Gamma = 1 - \exp(-\varepsilon) \in [0,1]$, which physically represents the probability of TPA for a two-photon Fock state input, under which, the time evolution of the density operator, Eq. (5), is recast as $\dot{\hat{\rho}} \to \frac{1}{1-\Gamma}\dot{\hat{\rho}}$. The lower panels plot the expectation values $|c_n|^2$, with maximum occupation number set to $\tilde{N} = 10$, across all losses, Γ, revealing a definite, loss-dependent behaviour on the form of the optimal probes and an underlying discrepancy between the behaviours of even and odd photon number states. Our choice of plotting $|c_n|^2$ as opposed to c_n is due to the fact that QFI is phase-invariant.

For even $\bar{n}$, the optimal quantum state takes the form

$$|\psi\rangle_{\mathrm{opt}} = \sqrt{1-\frac{\bar{n}}{\tilde{N}}}|0\rangle + \sqrt{\frac{\bar{n}}{\tilde{N}}}\left|\tilde{N}\right\rangle, \tag{6}$$

where $\tilde{N} \equiv \tilde{N}(\Gamma)$, varying across loss values such that $\tilde{N}(\Gamma=0) \to \infty$ and $\tilde{N}(\Gamma=1) \to \bar{n}$. In the limit $\varepsilon \to 0$, the QFI of $|\psi\rangle_{\mathrm{opt}}$ scales as

$$\mathcal{F}_{\mathrm{Q}}(|\psi\rangle_{\mathrm{opt}}) \simeq \frac{\bar{n}(\tilde{N}-1)}{2\varepsilon}. \tag{7}$$

In the case where $\bar{n}$ is odd, the optimal probe for short to intermediate losses is the same as in Eq. (6). In the asymptotic limit of large losses, however, the optimal probe is an equal superposition of the lowest-lying even number states available for maintaining $\bar{n}$, given by

$$|\psi\rangle_{\mathrm{opt}} = \frac{1}{\sqrt{2}}\left(|\bar{n}-1\rangle + |\bar{n}+1\rangle\right). \tag{8}$$

These findings provide further insight into prior literature's noted behaviour for coherent and squeezed vacuum states: in the limit of $\varepsilon \to 0$, the coherent state QFI and photon-counting CFI saturates at $\bar{n}^3 + \frac{1}{2}\bar{n}^2$ while the squeezed vacuum state's CFI is $\sim 10\bar{n}^2$. For relatively low energies, $\bar{n} \lesssim 10$, the squeezed vacuum is able to obtain higher precision than the coherent state even with sub-optimal photon counting. At these energies, however, the heavy-tailed distribution of the squeezed vacuum state naturally emulates our proposed optimal probe while the coherent state remains narrowly distributed around its mean. As the mean energy is progressively increased, the optimal probe becomes closer in distribution to the pure Fock state to which the coherent

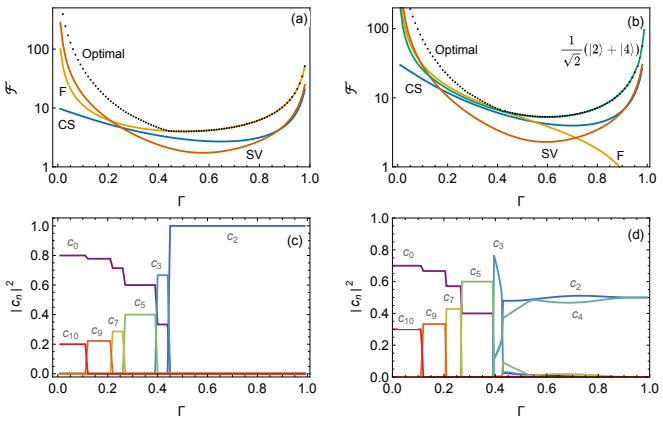


Figure 2: Panels (a) and (b) plot the QFI of the Fock state (F), coherent state (CS) and squeezed vacuum state (SV) compared to the optimal probe. Panels (c) and (d) show the associated coefficients c_n of the computationally derived optimal probe $|\psi\rangle = \sum_{n=0}^{\tilde{N}} c_n |n\rangle$ which yields maximal QFI for TPA parameter estimation, across all loss Γ. Results are given for mean values of $\bar{n} = 2$, for (a) and (c), and $\bar{n} = 3$, for (b) and (d). Also shown in (b) is the optimal probe in the limit of large loss which is an equal superposition of the lowest-lying even number states available to yield the required mean.

state's photon number distribution more closely resembles, allowing it to out-scale the squeezed vacuum. This behaviour can also be seen outside the limit of vanishing loss: intermediate to high losses show the optimal probe comprising of Fock states becomes increasingly centered around its mean; here the coherent state yields higher precision than the squeezed vacuum.

When the probe is of the form of Eq. (6) the optimal measurement is, remarkably, simple photon counting. The only non-zero off-diagonal terms in the density matrix are between the state with $\tilde{N}$ photons and the vacuum. In turn, the corresponding SLD operator is, in the Fock basis, identical to that as for the non-superposition (mixture) state except for terms corresponding to the vacuum from which no information can be obtained. This applies to all possible states with even and sub-unity means. When the mean photon number is odd, however, correlations exist between the lowest lying, non-zero even-number states which all contribute to the QFI. However, in the large loss limit where these states are truly optimal, those contributions become negligible and photon counting is very nearly optimal.

Finally, the discrepancy that exists between the behaviours of odd and even Fock states, illustrated in Fig. 3, can be explained through the decay rate of their respective contributions to the QFI as loss increases. For any state $|n\rangle$, the probability that TPA does not occur is $e^{-\frac{n(n-1)}{2}\varepsilon}$. For Fock states with $n \geq 4$, multiple transitions are available adding to the overall transition probabilities, but the decay rates are, at most, $\propto \frac{n(n-1)}{2}$. To illustrate this further, first note that the QFI takes the

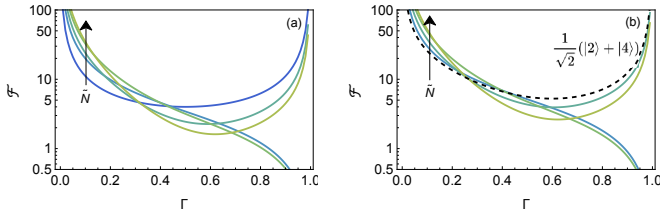


Figure 3: TPA parameter estimation QFIs for even and odd Fock states, weighted for means: (a) $\bar{n} = 2$ and (b) $\bar{n} = 3$ (states with $\tilde{N} = \bar{n}, \ldots, 6$)). Differing behaviours between even and odd $\tilde{N}$ across Γ are responsible for the crossovers seen in Gaussian states and the form of the optimal probe.

simplified form (for states diagonal in the Fock basis),

$$\mathcal{F}_{\mathrm{Q}} = \sum_k \hat{L}_k^2 \hat{\rho}_{\varepsilon,k} = \sum_k \frac{1}{h_k} g_k^2 = \sum_k \frac{1}{h_k} \left(\frac{\partial h_k}{\partial \varepsilon} \right)^2, \quad (9)$$

$$h_k = \sum_{l=0}^{k} \frac{(-1)^l}{(k-l)!l!} \frac{n!}{(n-2k)!} \frac{e^{-\frac{\varepsilon}{2}(n-2k+2l)(n-2k+2l-1)}}{F_{kl}(n-2k)}. \quad (10)$$

Then, the square of this decay rate yields a minimal pre-factor such that the limited number of transitions available can, in the large loss limit, never compensate for their respective decay rates. For intermediate losses, it is sufficient to dominate the increased decay rate allowing odd Fock states to yield higher QFIs in this region. At maximal loss ($\Gamma \to 1$) we observe a stark contrast between even and odd number state QFIs: the former again diverges $\sim \frac{n(n-1)}{2\varepsilon}$, while the latter goes to zero.

4 Conclusion

We have extended the analysis of quantum-enhanced TPA parameter estimation to consider behaviours across all scales of loss and have included results for non-Gaussian probes, specifically the Fock state. This study reveals a striking difference in behaviours of even and odd Fock states: the latter's QFI dominates in the intermediate regime while in the limit of infinite loss, tends towards zero. Here the QFI of even number states diverge, as both do in the limit of small loss.

Optimal probes for TPA parameter estimation for fixed means are discovered. They take the form of a superposition of the vacuum and, for increasing absorbance, progressively decreasing-energy odd Fock states, reducing to the pure Fock state in the large loss limit. Here photon counting is shown to saturate the QFI/ In the case of odd or sub-unity means, the optimal probe for large losses is an equal superposition of the lowest-lying even number states which maintain the mean and photon counting forms a very nearly optimal measurement.

Future research could extend this to uncover the effects of further loss arising from potential experimental inefficiencies. In addition, an analytical solution for the form of the optimal truncation number as a function of loss, as well as its precision limit, would be of value in guiding potential future implementations.

References

[1] V. Giovannetti, S. Lloyd and L. Maccone, *Advances in quantum metrology*, Nat. Photon. **5**, 222–229 (2011).

[2] E. Polino, M. Valeri, N. Spagnolo and F. Sciarrino, *Photonic quantum metrology*, AVS Quant. Sci. **2**, 024703 (2020).

[3] M. Genovese, *Real applications of quantum imaging*, J. Opt. **18**, 073002 (2016).

[4] S. Mukamel *et al.*, *Roadmap on quantum light spectroscopy*, J. Phys. B: At., Mol. Opt. Phys., **53**(7), 072002 (2020).

[5] D. N. Klyshko, *Transverse photon bunching and two-photon processes in the field of parametrically scattered light*, Sov. Phys. JTEP **56**, 753 (1982).

[6] J. Gea-Banacloche, *Two-photon absorption of nonclassical light*, Phys. Rev. Lett. **62**, 1603 (1989).

[7] J. Javanainen and P. L. Gould, *Linear intensity dependence of a two-photon transition rate*, Phys. Rev. A **41**, 5088 (1990).

[8] N. P. Georgiades, E. S. Polzik, K. Edamatsu, H. J. Kimble and A. S. Parkins, *Nonclassical excitation for atoms in a squeezed vacuum*, Phys. Rev. Lett. **75**, 3426 (1995).

[9] N. P. Georgiades, E.S. Polzik and H. J. Kimble, *Atoms as nonlinear mixers for detection of quantum correlations at ultrahigh frequencies*, Phys. Rev. A **55**, R1605 (1997).

[10] B. Dayan, A. Pe'er, A. A. Friesem and Y. Silberberg, *Two photon absorption and coherent control with broadband down-converted light*, Phys. Rev. Lett. **93**, 023005 (2004).

[11] K. E. Dorfman, F. Schlawin and S. Mukamel, *Nonlinear optical signals and spectroscopy with quantum light*, Rev. Mod. Phys. **88**, 045008 (2016).

[12] F. Schlawin, *Entangled photon spectroscopy* J. Phys. B: At. Mol. Opt. Phys. **50**, 203001 (2017).

[13] F. Schlawin, K.E. Dorfman and S. Mukamel, *Entangled two-photon absorption spectroscopy*, Acc. Chem. Res. **51**, 2207 (2018).

[14] M. Gilaberte Basset, F. Setzpfandt, F. Steinlechner, E. Beckert, T. Pertsch and M. Gräfe, *Perspectives for applications of quantum imaging*, Laser Photonics Rev. **13**, 1900097 (2019).

[15] Y. Z. Ma and B. Doughty *Nonlinear Optical Microscopy with Ultralow Quantum Light*, J. Phys. Chem. A **125**, 8765 (2021).

[16] A. Eshun, O. Varnavski, J. P. Villabona-Monsalve, R. K. Burdick, R. K. and T. Goodson III, *Entangled Photon Spectroscopy*, Acc. Chem. Res. **55**, 991 (2022).

[17] S. Pirandola, B. R. Bardhan, T. Gehring, C. Weedbrook and S. Lloyd, *Advances in photonic quantum sensing*, Nat. Photon. **12**, 724–733 (2018).

[18] A. Karsa, G. Spedalieri, Q. Zhuang, and S. Pirandola, *Quantum illumination with a generic Gaussian source*, Phys. Rev. Research **2**, 023414 (2020).

[19] A. Monras and M. G. A. Paris, *Optimal quantum estimation of loss in bosonic channels*, Phys. Rev. Lett. **98**, 160401 (2007).

[20] G. Adesso, F. Dell'Anno, S. De Siena, F. Illuminati, L. A. M. Souza, *Optimal estimation of losses at the ultimate quantum limit with non-Gaussian states*, Phys. Rev. A **79**, 040305 (2009).

[21] S. L. Braunstein, C. M. Caves and G. J. Milburn, *Generalized uncertainty relations: theory, examples, and Lorentz invariance*, Ann. of Phys. **247**, 135–173 (1996).

[22] L. Maccone, *Information-disturbance tradeoff in quantum measurements*, Phys. Rev. A **73**, 042307 (2006).

[23] C. S. Muñoz, G. Frascella, and F. Schlawin, *Quantum metrology of two-photon absorption*, Phys. Rev. Research **3**, 033250 (2021).

[24] S. Panahiyan, C. S. Muñoz, M. V. Chekhova and F. Schlawin, *Two-photon-absorption measurements in the presence of single-photon losses*, Phys. Rev. A **106**, 043706 (2022).

[25] H. D. Samaan and R. Loudon, *Quantum statistics of double-beam two-photon absorption*, J. Phys. A: Math. Gen. **8**, 1140 (1975).

[26] H. D. Samaan and R. Loudon, *Quantum statistics of single-beam two-photon absorption*, J. Phys. A: Math. Gen. **8**, 539 (1975).

[27] H. D. Samaan and R. Loudon, *Off-diagonal density matrix for single-beam two-photon absorbed light*, J. Phys. A: Math. Gen. **11**, 435 (1978).

[28] L. Gilles and P. L. Knight, *Two-photon absorption and nonclassical states of light*, Phys. Rev. A **48**, 1582 (1993).

[29] A. B. Klimov and J. L. Romero, *An algebraic solution of Lindblad-type master equations*, J. Opt. B: Quantum Semiclassical Opt. **5**, S316 (2003).

[30] N. E. Toklu, T. Atkinson, V. Micka, P. Liskowski and R. K. Srivastava, *EvoTorch: Scalable Evolutionary Computation in Python*, arViv preprint, arXiv:2302.12600 (2023).

Intrinsic randomness under general quantum measurements

Hao Dai[1 2 *] Boyang Chen[1 †] Xingjian Zhang[1 ‡] Xiongfeng Ma[1 §]

[1] *Center for Quantum Information, Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing 100084, P. R. China*

[2] *Hefei National Laboratory, University of Science and Technology of China, Hefei, Anhui 230088, P. R. China*

Abstract. Quantum measurements can produce unpredictable randomness arising from the uncertainty principle. When measuring states with projection measurements, the intrinsic randomness can be quantified by block coherence. Unlike projection measurements, there are additional and possibly hidden degrees of freedom in apparatus for generic measurements. We propose an adversary scenario to characterize the intrinsic randomness of general measurements with arbitrary input states. Interestingly, we discover that certain measurements generate nonzero randomness for all states, which suggests a new approach for designing source-independent random number generators without state characterization. Furthermore, our proposed intrinsic randomness can quantify coherence under general measurements.

Keywords: Intrinsic randomness, general measurements, quantum random number generators, quantum coherence

Introduction. — Quantum measurements can produce randomness arising from the uncertainty principle [1]. As shown in Figure 1, a measurement process is typically composed of a source characterized by a quantum state [2, 3] and a detector calibrated by a quantum measurement [4, 5]. Randomness generation can be put in an adversary scenario. Roughly speaking, intrinsic randomness represents a part of the measurement results about which the adversary, Eve, has no information. When measuring a state with projection measurements, Eve might have a certain correlation with the system. The intrinsic randomness can be quantified by the quantum coherence of the state on the measurement basis [6, 7, 8, 9, 10, 11, 12]. Yet for a general measurement described by a positive operator-valued measure (POVM), Eve might also have a correlation with the measurement device. How to quantify intrinsic randomness of the outcomes from a generic measurement is a basic problem in quantum information theory. From an experimental perspective, as noise is inevitable, randomness evaluation from POVM is also of practical interest. But unlike projection measurements, there are additional and possibly hidden degrees of freedom in apparatus for generic measurements, making it a difficult problem.

To solve the intrinsic randomness quantification problem under general measurements, we propose an adversary scenario for general measurements with arbitrary input states, based on which, we characterize the intrinsic randomness. Interestingly, we discover that under certain measurements, such as the symmetric and information-complete (SIC) measurement, all states have nonzero randomness, inspiring a new design of source-independent random number generators without state characterization. Furthermore, our results show that intrinsic randomness can quantify coherence under general measurements, which generalizes the result in the standard resource theory of state coherence.

*dhao@mail.tsinghua.edu.cn
†chen-by19@mails.tsinghua.edu.cn
‡zhang-xj18@mails.tsinghua.edu.cn
§xma@tsinghua.edu.cn

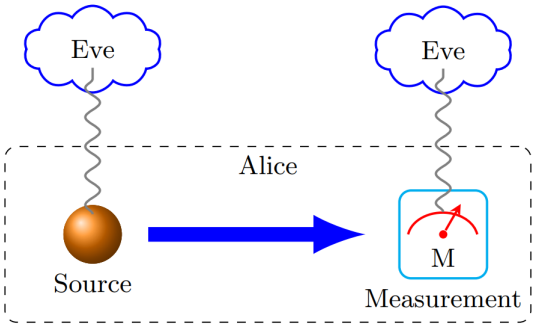


Figure 1: The source sends quantum signals in the state of ρ to the measurement device, which outputs a sequence of random numbers. Eve could have a certain correlation with the devices, where she could possess the purification of ρ on the source side and know the construction of the detection on the measurement side.

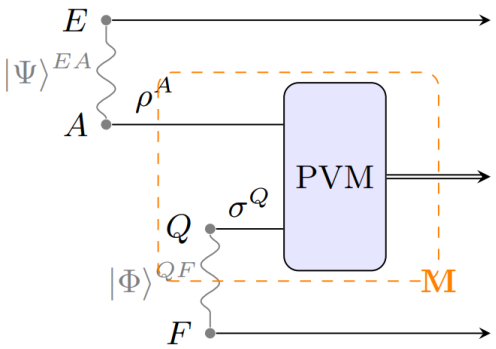


Figure 2: From Alice's perspective, the measuring process is described by POVM $\mathbf{M}$, depicted as the dashed box. Alice inputs state ρ^A and obtains classical outputs from the box. The ancillary system is generally in a mixed state σ^Q. Both the source and the ancilla could be entangled with Eve.

Process modelling. — For a POVM $\mathbf{M}$ =

$\{M_1, \cdots, M_m\}$ on d-dimensional Hilbert space $\mathcal{H}$, each element can be expressed as $M_i = A_i A_i^\dagger$, where A_i is called a POVM operator and generally not a square matrix. When measuring a state ρ, the probability of obtaining the outcome i is given by $\mathrm{tr}(M_i\rho)$ and the corresponding post-measurement state is $A_i\rho A_i^\dagger/\,\mathrm{tr}(M_i\rho)$. The set of operators $\{A_i\}$ uniquely determines the implementation of the measurement — instrument. When $\mathbf{M}$ is a projection measurement, also called projection-valued measure (PVM), $M_i = A_i$, which means that the implementation is unique. A general POVM generally corresponds to many possible implementations or different sets of operators, $\{A_i\}$. This is the challenging part of the randomness quantification of POVMs.

To quantify the amount of intrinsic randomness, the first task is to construct an adversarial perspective and examine what side information Eve may use in eavesdropping. In a quantum random number generator (QRNG), Alice first prepares a state, ρ^A, and then measures it with the POVM, $\mathbf{M}$. Apart from her knowledge of the form of these operators, Eve may correlate her system to ρ^A and $\mathbf{M}$ as well. For ρ^A, the best Eve can achieve is to hold the purified system [13]. Similarly, for $\mathbf{M}$, Eve can also hold a "purified" process which means that the measurement on the joint system is a PVM. In the most general scenario, Eve puts a state σ^Q in the measurement device and holds the purification of σ^Q. Then, Alice's device actually performs an extended measurement on both ρ^A and σ^Q, as shown in Figure 2.

From Alice's viewpoint, the measurement is calibrated to act on system A; in other words, Alice is unaware of σ^Q hidden in her measurement device. Generally, as Alice can change her input state, one would expect the measurement device to enjoy a consistent description for all possible input states on system A, or, the same POVM elements [14]. Mathematically, the consistency condition can be formulated as $\forall i$,

$$M_i = \mathrm{tr}_Q[P_i(\mathbf{1}^A \otimes \sigma^Q)]. \tag{1}$$

For simplicity, we shall omit the superscript A and Q when there is no confusion in the following discussions and denote the states ρ and σ as Alice's state and the ancillary state, respectively.

With the adversarial scenario in hand, we can see that Eve has the freedom to choose an extended PVM, $\mathbf{P}$, and the corresponding ancillary state σ satisfying the consistency condition. In the randomness analysis, we need to minimize over all possible Eve's strategies, and the intrinsic randomness is now given by

$$\begin{aligned} &R(\rho, \mathbf{M}) = \min_{\mathbf{P},\sigma} R(\rho\otimes\sigma, \mathbf{P}), \\ &\text{s.t.} \quad \forall i,\, M_i = \mathrm{tr}_Q[P_i(\mathbf{1}^A \otimes \sigma^Q)], \end{aligned} \tag{2}$$

where $R(\varrho, \mathbf{P})$ is a specific randomness function of PVMs. For example, $R(\varrho, \mathbf{P})$ can be taken as relative entropy of block coherence or block coherence of formation.

Main results. — Let us first check out a special case where the POVM is extremal, which cannot be decomposed into a linear mixture of other POVMs and includes PVMs [15]. This is an analog to a pure state, which is often considered to be decoupled from the environment.

Theorem 1 *For an extremal POVM* $\mathbf{M}$ *and a fixed input state* ρ*, all the generalized Naimark extensions give the same amount of randomness.*

Then, we can skip the minimization problem in Eq. (2) and employ any extension for the randomness function. In practice, we can take a canonical extension of $\mathbf{M}$ [16], denoted by $\mathbf{P}_c$,

$$R(\rho, \mathbf{M}) = R(\rho \otimes |0\rangle\langle 0|, \mathbf{P}_c). \tag{3}$$

A general POVM can be decomposed to extremal ones, just like a mixed state can be decomposed to pure states. In the generalized Naimark extension as shown in Figure 2, assume Eve performs measurements on her system F. Then the ancillary state can be decomposed into a mixture of pure states and the POVM can be decomposed correspondingly. As a result, Eq. (2) gives a convex-roof construction of intrinsic randomness, as presented in the following theorem.

Theorem 2 *When Eve performs a measurement on her system* F*, the intrinsic randomness of POVM outcomes is given by,*

$$\begin{aligned} R^{cf}(\rho, \mathbf{M}) = &\min_{\{\mathbf{N}^j, r_j\}} \sum_j r_j R(\rho, \mathbf{N}^j), \\ &s.t. \quad \mathbf{M} = \sum_j r_j \mathbf{N}^j, \end{aligned} \tag{4}$$

where the decomposed POVMs $\{\mathbf{N}^j\}$ *are all extremal and the randomness function* $R(\rho, \mathbf{N}^j)$ *is given by Eq.* (3)*.*

Table 1: Existing randomness evaluation and coherence measures. In the last column, we show a simple example of a specific POVM, under which no randomness or coherence should be generated. A measure cannot quantify randomness or coherence properly for general measurements if it gives a non-zero evaluation result for some states.

Ref.	randomness	coherence	measurement	$\{\mathbf{1}/2, \mathbf{1}/2\}$
[11]	✓	✓	von Neumann	not applicable
[17]	✓	×	POVM	failed
[18]	✓	×	POVM	qubit only
[19, 20]	×	✓	POVM	failed
Our work	✓	✓	POVM	succeed

Comparison with previous works. — In the literature, there are some attempts on randomness evaluation [17, 18] and coherence measures [19, 20] under POVMs, as listed in Table 1. Unfortunately, the existing measures cannot properly quantify randomness or coherence under the most general measurements. Consider a the POVM $\mathbf{M} = \{\mathbf{1}/2, \mathbf{1}/2\}$, the outcome is independent of input states. Then, it can be seen as a classical random variable taking values 0 and 1 with an equal probability. Thus, all

states should have zero randomness or coherence under this measurement. Yet for this seemingly simple example, the existing measures either fail to accord with our intuition or suffice only for the qubit case. On the other hand, our results provide a consistent quantification for the general cases.

Applications. — Quantification of intrinsic randomness has direct applications in QRNGs. After quantifying randomness for the measurement outcomes with respect to a given POVM, we find that non-random states may not exist for certain measurements. For example, there does not exist a non-random state for SIC measurement. Moreover, we can evaluate the lower bound of intrinsic randomness, $R(\mathbf{M}) = \min_\rho R(\rho, \mathbf{M})$.

Theorem 3 *For a SIC measurement* $\mathbf{M}$*, a lower bound of intrinsic randomness is given by,*

$$R(\mathbf{M}) > \log\left(\frac{d+1}{2}\right), \tag{5}$$

where d *is the dimension of the corresponding space.*

Theorem 3 inspires a new design of source-independent random number generators without state characterization.

References

[1] M. Born. Zur Quantenmechanik der Stoßvorgänge. Z. Phys, 37:863, 1926.

[2] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu. Efficient quantum state tomography. Nature Commun., 1:149, 2010.

[3] J. Haah,A. W. Harrow,Z. Ji,X. Wu, and N. Yu. Sample-optimal tomography of quantum states. IEEE Trans. Inf. Theory, 63:5628, 2017.

[4] G. M. D'Ariano, L. Maccone, and P. L. Presti. Quantum Calibration of Measurement Instrumentation. Phys. Rev. Lett., 93:250407, 2004.

[5] J. Lundeen, A. Feito, H. Coldenstrodt-Ronge, K. Pregnell, C. Silberhorn, T. Ralph, J. Eisert, M. Plenio, and I. Walmsley. Tomography of quantum detectors. Nat. Phys., 5:27, 2009.

[6] W. H. Zurek. Quantum Darwinism. Nat. Phys., 5:181, 2009.

[7] J. Åberg. Quantifying Superposition. quant-ph/0612146, 2006.

[8] T. Baumgratz, M. Cramer, and M. B. Plenio. Quantifying Coherence. Phys. Rev. Lett., 113:140401, 2014.

[9] X. Yuan, H. Zhou, Z. Cao, and X. Ma. Intrinsic randomness as a measure of quantum coherence. Phys. Rev. A, 92:022124, 2015.

[10] M. Hayashi and H. Zhu. Secure uniform random-number extraction via incoherent strategies. Phys. Rev. A, 97:012302, 2018.

[11] X. Yuan, Q. Zhao, D. Girolami, and X. Ma. Quantum Coherence and Intrinsic Randomness Adv. Quantum Technol., 2:1900053, 2019.

[12] M. Hayashi, K. Fang, and K. Wang. Finite Block Length Analysis on Quantum Coherence Distillation and Incoherent Randomness Extraction. IEEE Transactions on Information Theory, 67:3926, 2021.

[13] H.-K. Lo and H. F. Chau. Unconditional Security of Quantum Key Distribution over Arbitrarily Long Distances. science, 283:2050, 1999.

[14] P. Busch, P. J. Lahti, and P. Mittelstaedt. *The quantum theory of measurement*. Springer Berlin Heidelberg, 1996.

[15] G. M. D'Ariano, P. L. Presti, and P. Perinotti. Classical randomness in quantum measurements. J. Phys. A: Math. Gen., 38:5979, 2005.

[16] A. Peres. Neumark's theorem and quantum inseparability. Found. Phys., 20:1441, 1990.

[17] Y. Z. Law, J.-D. Bancal, V. Scarani. Quantum randomness extraction for various levels of characterization of the devices. J. Phys. A: Math. Theor., 47:424028, 2014.

[18] Z. Cao, H. Zhou, and X. Ma. Loss-tolerant measurement-device-independent quantum random number generation. New J. Phys., 17:125011, 2015.

[19] F. Bischof, H. Kampermann, and D. Bruß Resource Theory of Coherence Based on Positive-Operator-Valued Measures. Phys. Rev. Lett., 123:110402, 2019.

[20] J. Xu, L.-H. Shao, and S.-M. Fei. Coherence measures with respect to general quantum measurements. Phys. Rev. A, 102:012411, 2020.

Neural Quantum Embedding: Pushing the Limits of Quantum Supervised Learning

Tak Hur[1] * Israel F. Araujo[1] † Daniel K. Park[1 2] ‡

[1] *Department of Statistics and Data Science, Yonsei University, Seoul, Republic of Korea*
[2] *Department of Applied Statistics, Yonsei University, Seoul, Republic of Korea*

Abstract. Quantum embedding is the essential first step in quantum machine learning, and has substantial impacts on performance outcomes. In this study, we present Neural Quantum Embedding (NQE), a novel approach that efficiently optimizes quantum embedding by harnessing classical deep learning techniques. By employing NQE, we effectively enhance the lower bound of the empirical risk, leading to substantial improvements in classification performance. Moreover, NQE enhances algorithmic robustness against noise. To validate the effectiveness of NQE, we conduct experiments on IBM quantum devices for image data classification, resulting in a remarkable accuracy enhancement from 0.52 to 0.96.

Keywords: Quantum machine learning, Variational quantum classifier, Quantum feature mapping, Quantum error mitigation

1 Introduction

Quantum machine learning (QML) is a rapidly growing field that focuses on developing quantum algorithms and techniques to solve various machine learning problems. To apply QML algorithms on tasks involving classical data, the data must be mapped into a quantum state [1–18]. This process is called quantum embedding, and it is important as it can vastly affect the performance of the algorithms [19–21]. A well-designed quantum embedding can reduce the number of qubits required to represent the data, which can also reduce the computational cost of the algorithm. Furthermore, quantum embedding has the potential to represent feature-related correlations that may not be explicitly present in the classical data. These additional features can contribute to enhancing the classification accuracy of quantum machine learning algorithms.

For quantum binary classification tasks, the optimal lower bound of the empirical risk is determined by the trace distance between the two data ensembles (Section 2). As the trace distance between two quantum states is contractive under all quantum operations, having a large trace distance in the initial quantum embedding phase is crucial for successful classification. Moreover, it makes training robust against the noise as the data are separated further from the decision boundary. In this study, we present *neural quantum embedding* (NQE), which utilizes the power of classical neural network and deep learning to achieve trainable quantum embedding that is optimized for the given problems. With the numerical simulations, we demonstrate that NQE successfully improves quantum embedding by increasing the trace distance between the embedded ensembles. Experiments with IBM quantum devices further support the effectiveness of NQE on the performances of QML algorithms under noisy environments.

*takh0404@yonsei.ac.kr
†ifa@yonsei.ac.kr
‡dkd.park@yonsei.ac.kr

2 Quantum Supervised Learning

In supervised learning, the primary objective is to identify a prediction function f that minimizes the true risk $R(f) = \mathbb{E}[l(f(X), Y)]$ with respect to some loss function l, where X, Y are drawn from an unknown distribution D. Given the sample data (x_i, y_i), the goal of learning algorithms is to find the optimal function f^* that minimizes the empirical risk, $f^* = \arg\min_{f \in F} \frac{1}{N} \sum_{i=1}^{N} l(f(x_i), y_i)$ among a fixed function class F. Quantum supervised learning algorithms aim to efficiently find prediction functions with improved performance by exploiting the computational power of the quantum device.

A quantum neural network is a widely used method for quantum supervised learning. In quantum neural network, a classical input data x is first embedded into a quantum state by applying a quantum embedding circuit to an initial ground state $|x\rangle = \Phi(x)\,|0\rangle^{\otimes n}$. Then parameterized unitary $U(\theta)$ is applied to transform the embedded quantum states, and the state is measured with an observable O. The measurement outcome is used as a prediction function for supervised learning algorithms, expressed as $f(x;\theta) = \langle x|\, U^\dagger(\theta) O U(\theta)\,|x\rangle$. Then, by the gradient descent or its variant, we search for the optimal parameter θ^* that minimizes the empirical risk. For a binary classification task with label $y \in \{1, -1\}$, we can predict the label of the new data $\tilde{x}$ by the rule $y_{new} = \text{sign}[f(x_{new}; \theta^*)]$.

We can alternatively consider this procedure as a quantum two-state discrimination problem with two parameterized POVMs $E_\pm(\theta) = \frac{1}{2}(I \pm U^\dagger(\theta) O U(\theta))$. Here, we denote a probability of measuring ± 1 given an input data x as $P(E_\pm(\theta)|x)$, which can be expressed as $\langle x|\, E_\pm(\theta)\,|x\rangle$. Then, the previous decision rule on the new data becomes $y_{new} = \text{sign}[P(E_+(\theta^*)|x_{new}) - P(E_-(\theta^*)|x_{new})]$.

In such a scenario, a natural loss function is a probability of misclassification, expressed as $l(f(x;\theta), y) = P(E_{\neg y}(\theta)|x)$. Considering $2N$ numbers of balanced samples $S = \{x_i^-, -1\}_{i=1}^N \cup \{x_i^+, 1\}_{i=1}^N$, the empirical risk

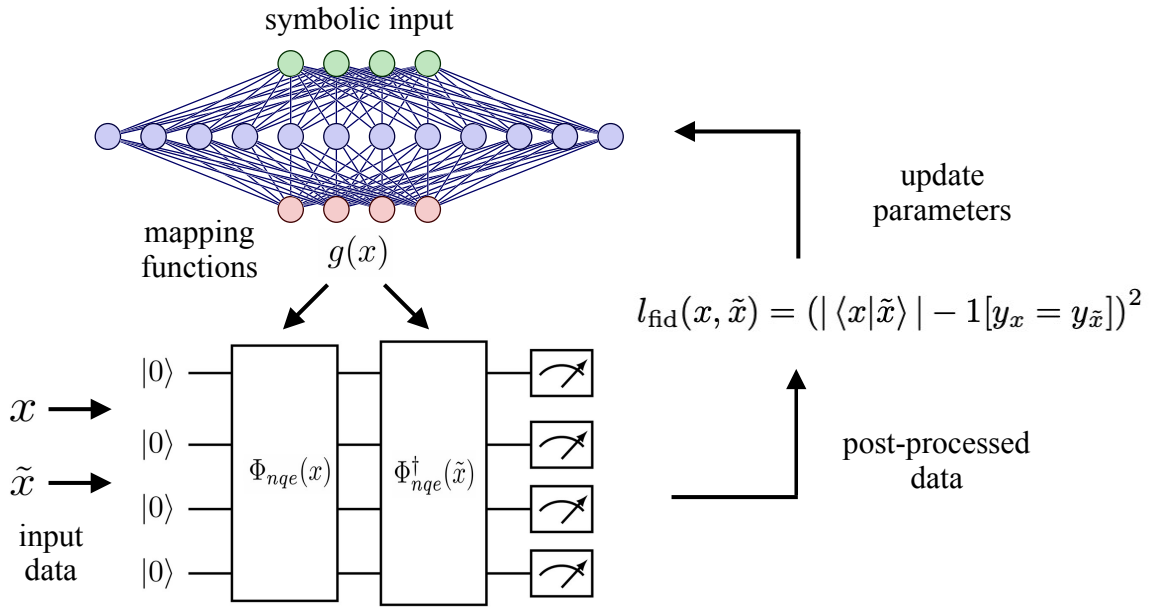


Figure 1: Overview of the NQE training. The goal of the training is to produce mapping functions that can separate the two classes of data into two orthogonal subspaces. Efficient calculation of the normal fidelity between the two state vectors produced by the feature map is performed using a quantum computer.

becomes,

$$\begin{aligned} L_s &= \frac{1}{2N}\left[\sum_{i=1}^{N} P(E_+(\theta)|x_i^-) + \sum_{i=1}^{N} P(E_-(\theta)|x_i^+)\right] \\ &= \frac{1}{2}\left[1 - \mathrm{Tr}(E_+(\theta)(\rho_+ - \rho_-))\right] \\ &\geq \frac{1}{2}[1 - D_{\mathrm{tr}}(\rho^-, \rho^+)], \end{aligned}$$

where $\rho^{\pm} = \frac{1}{N}\sum |x_i^{\pm}\rangle \langle x_i^{\pm}|$ and $D_{\mathrm{tr}}(\cdot,\cdot)$ is a trace distance operator. Here, we want to emphasize two crucial points.

1. The empirical risk is lower bounded by the trace distance between two data ensembles ρ^- and ρ^+. This is determined by the initial quantum embedding circuit, regardless of the structure of the parameterized unitaries $U(\theta)$ applied afterwards.

2. The minimum loss is achieved when $E_{\pm}(\theta)$ is a Helstrom measurement. Hence, training of a quantum neural network can be viewed as a process of finding the Helstrom measurement that optimally discriminates two data ensembles.

Thus, choosing a quantum embedding circuit that maximizes the trace distance is critical, as it pushes the lower bound of the empirical risk further below. Also, maximizing the trace distance is particularly important for NISQ (Noisy Intermediate-Scale Quantum) applications as non-unitary quantum operations strictly contracts the trace distance between two quantum states [22]. However, none of the existing quantum embeddings, including Hamiltonian, amplitude, and angle embeddings, can guarantee the effective separation of two data ensembles in the Hilbert space with a large distance. This illustrates the necessity of trainable, data-dependent embedding that can maximize the trace distance.

3 Neural Quantum Embedding

Neural Quantum Embedding utilizes a classical neural network to maximize the trace distance between two data ensembles. It can be expressed as

$$\Phi_{\mathrm{nqe}} : x \to |x\rangle = \Phi(g(x))\,|0\rangle^{\otimes n}, \tag{1}$$

where Φ is a general quantum embedding circuit and $g : \mathbb{R}^d \to \mathbb{R}^m$ is a trainable classical neural network that transforms the input data x. There is no restriction on the choice of quantum embedding circuit and the neural network structure, and we test several combinations in the numerical studies. By choosing $m < d$, we can bypass additional classical feature reduction methods, such as PCA or autoencoders, typically employed prior to quantum embedding due to the current limitations on the number of reliably controllable qubits in quantum devices.

Prior to training the quantum neural network, we train the NQE to increase the trace distance between two ensembles $D_{\mathrm{tr}}(\rho^-, \rho^+)$. Although training with the trace distance as a loss function is ideal, calculating it is computationally expensive even with the quantum computer. Instead, we used an implicit loss function derived from a fidelity measure, which is expressed as

$$l_{\mathrm{fid}}(x, \tilde{x}) = (|\langle x|\tilde{x}\rangle| - 1[y_x = y_{\tilde{x}}])^2. \tag{2}$$

This fidelity loss can be efficiently computed using the swap test [23] or directly measuring the overlap.

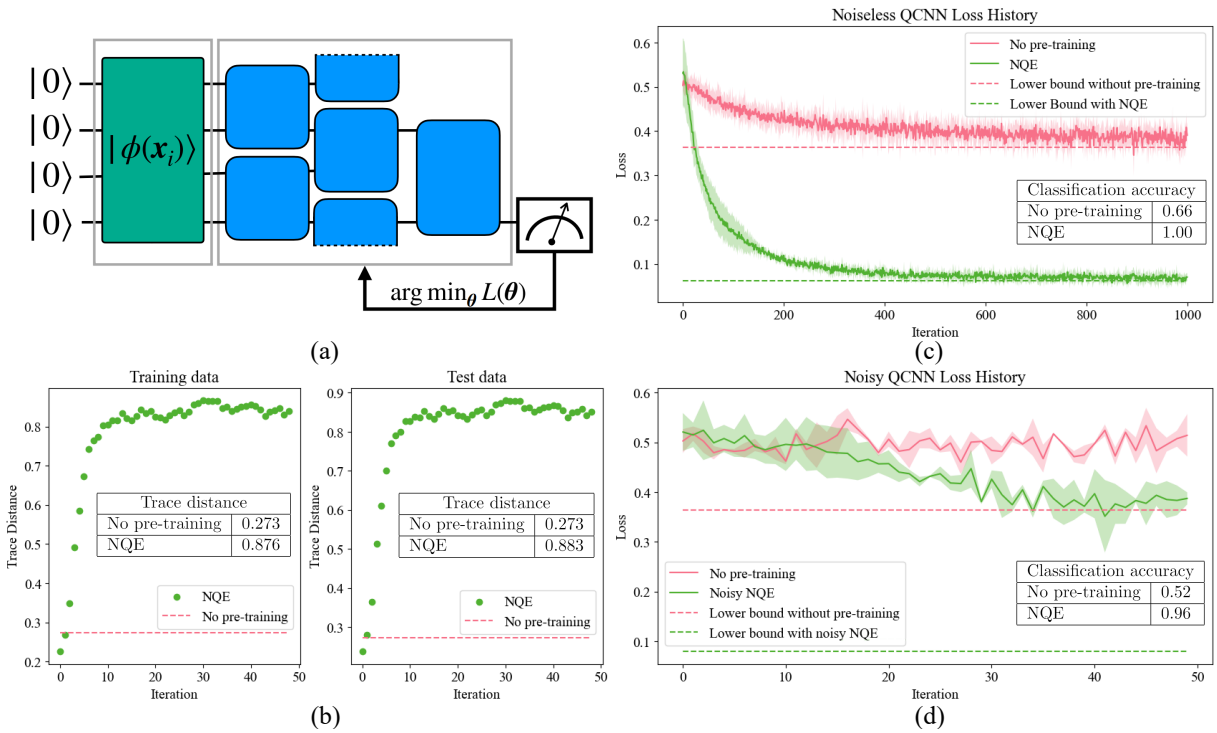


Figure 2: Summary of proof-of-principle results. (a) Structure of the QCNN circuits used in numerical studies. (b) Trace distance history during the NQE training. (c) Noiseless simulation of the QCNN Loss History and classification accuracy with and without NQE. (d) Experimental results of the QCNN loss history and classification accuracy with and without NQE, obtained through executions on IBM quantum devices.

We particularly focus on improving *ZZ feature embedding*, which is widely adopted due to its classically intractable feature mapping. It is expressed as,

$$\Phi(x) = \left[H^{\otimes n} \exp\left(i \sum_i Z_i \phi_i(x) + i \sum_{i,j} Z_i Z_j \phi_{i,j}(x) \right) \right]^L \tag{3}$$

The most commonly used functions for ϕ are $\phi_i(x) = x_i$ and $\phi_{i,j}(x) = \frac{1}{2}(\pi - x_i)(\pi - x_j)$ [19,24], but these choices are made without justifications. Although Ref. [21] numerically illustrates that the choice of ϕ can impact the performance of the QML algorithms, it fails to explain how to choose an appropriate ϕ for the problem at hand. NQE effectively solves this by replacing mapping functions with a trainable classical neural network.

4 Numerical Studies

Setup The numerical studies were performed with classes 0 and 1 of the MNIST dataset. To observe an effectiveness of NQE on performances of the QML algorithms, we used 4-qubit Quantum Convolutional Neural Network (QCNN) [25, 26] (Figure 2(a)). For the NQE, a fully connected RELU network of the dimension $4 \to 16 \to 16 \to 8$ was used. We compared NQE methods against the aforementioned function $\phi_i(x)$ and $\phi_{i,j}(x)$. The training of NQE and QCNN, and the accuracy evaluation of QCNN, were performed on the **ibmq_toronto**, **ibmq_jakarta**, and **ibm_lagos** quantum devices, respectively.

Results With the *ZZ* feature embedding, two embedded quantum ensembles are separated with the trace distance of 0.273. This sets the lower bound of the empirical risk as 0.364. In contrast, the trained NQE achieves a separation distance of 0.876, significantly reducing the empirical risk lower bound to 0.062. The trace distance history is summarized in Figure 2(b).

The training of the QCNN circuit is additional evidence supporting the effectiveness of NQE in QML. In the noiseless simulation (Figure 2(c)), we observed the empirical risk converging to the theoretical minimum in both cases, whether NQE was employed or not. This indicates that the trained QCNN circuit successfully approximates the optimal Helstrom measurements. With NQE, significant improvements were achieved, as evidenced by lower empirical risk and higher classification accuracy.

In the experiments implemented on IBM quantum devices (Figure 2(d)), the presence of noise prevents the empirical risk from converging to the theoretical minimum. However, despite the noise, the utilization of the NQE method resulted in an improved loss history. Remarkably, in some instances, the empirical risk achieved with NQE in the experiment is even lower than the theoretical limit of the previous method without NQE. Furthermore, a substantial enhancement in classification accuracy was observed, reaching 0.96 compared to the conventional method's 0.52. These results illustrate how NQE enhances the robustness of QML against noise.

References

[1] Patrick Rebentrost, Adrian Steffens, Iman Marvian, and Seth Lloyd. Quantum singular-value decomposition of nonsparse low-rank matrices. *Physical Review A*, 97(1), 2018.

[2] Leonard Wossnig, Zhikuan Zhao, and Anupam Prakash. Quantum linear system algorithm for dense matrices. *Phys. Rev. Lett.*, 120:050502, 2018.

[3] Andrew M. Childs, Robin Kothari, and Rolando D. Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, 2017.

[4] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, 2014.

[5] Carsten Blank, Daniel K Park, June-Koo Kevin Rhee, and Francesco Petruccione. Quantum classifier with tailored quantum kernel. *npj Quantum Information*, 6(1):1–7, 2020.

[6] Yoav Levine, Or Sharir, Nadav Cohen, and Amnon Shashua. Quantum Entanglement in Deep Learning Architectures. *Physical Review Letters*, 122(6):065301, 2019.

[7] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 2019.

[8] Maria Schuld and Francesco Petruccione. *Supervised Learning with Quantum Computers*. Quantum Science and Technology. Springer International Publishing, 2018.

[9] M. Schuld, M. Fingerhuth, and F. Petruccione. Implementing a distance-based classifier with a quantum interference circuit. *EPL (Europhysics Letters)*, 119(6):60002, 2017.

[10] Edwin Stoudenmire and David J Schwab. Supervised learning with tensor networks. In D. Lee, M. Sugiyama, U. Luxburg, I. Guyon, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 29, page 9, Centre Convencions Internacional Barcelona, Barcelona Spain, 2016. Curran Associates, Inc.

[11] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv preprint arXiv:1307.0411*, 2013.

[12] Daniel K. Park, Francesco Petruccione, and June-Koo Kevin Rhee. Circuit-based quantum random access memory for classical data. *Scientific Reports*, 9(1):3949, 2019.

[13] Fernando M de Paula Neto, Adenilton J da Silva, Wilson R de Oliveira, and Teresa B. Ludermir. Quantum probabilistic associative memory architecture. *Neurocomputing*, 351:101–110, 2019.

[14] Adenilton Silva, Wilson de Oliveira, and Teresa Ludermir. A Weightless Neural Node Based on a Probabilistic Quantum Memory. In *2010 Eleventh Brazilian Symposium on Neural Networks*, pages 259–264, Sao Paulo, Brazil, October 2010. IEEE. ISSN: 2375-0235.

[15] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum random access memory. *Phys. Rev. Lett.*, 100:160501, Apr 2008.

[16] Carlo A. Trugenberger. Quantum Pattern Recognition. *Quantum Information Processing*, 1(6):471–493, 2002.

[17] Dan Ventura and Tony Martinez. Quantum associative memory. *Information Sciences*, 124(1):273–296, 2000.

[18] C. A. Trugenberger. Probabilistic quantum memories. *Physical Review Letters*, 87(6), 2001.

[19] Vojtech Havlícek, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, 2019.

[20] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. Quantum embeddings for machine learning. *arXiv preprint arXiv:2001.03622*, 2020.

[21] Yudai Suzuki, Hiroshi Yano, Qi Gao, Shumpei Uno, Tomoki Tanaka, Manato Akiyama, and Naoki Yamamoto. Analysis and synthesis of feature map for kernel-based quantum classifier. *Quantum Machine Intelligence*, 2(1):9, July 2020.

[22] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, New York, NY, USA, 10th edition, 2011.

[23] Harry Buhrman, Richard Cleve, John Watrous, and Ronald de Wolf. Quantum fingerprinting. *Phys. Rev. Lett.*, 87:167902, Sep 2001.

[24] Amira Abbas, David Sutter, Christa Zoufal, Aurelien Lucchi, Alessio Figalli, and Stefan Woerner. The power of quantum neural networks. *Nature Computational Science*, 1(6):403–409, 2021.

[25] Iris Cong, Soonwon Choi, and Mikhail D. Lukin. Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278, December 2019.

[26] Tak Hur, Leeseok Kim, and Daniel K Park. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, 4(1):3, 2022.

Optimizing Gaussian Elimination-based NNA-compliant Circuit Synthesis Method by Simulated Annealing-based CNOT Gates Insertion

Zanhe QI[1] * Huan YU[1] † Shigeru YAMASHITA[1] ‡

[1] *Graduate School of Science and Engineering, Ritsumeikan University*

Abstract. Quantum circuits should be implemented on a so-called Nearest Neighbor Architecture (NNA) that only supports two-qubit operations between adjacent qubits. Thus, we usually convert a quantum circuit to an NNA-compliant circuit by inserting SWAP gates. Without that, we can usually generate a smaller NNA-compliant quantum circuit by utilizing Gaussian Elimination. This paper reveals that we can improve the Gaussian Elimination-based method by inserting CNOT gates before and/or after the target circuit in many cases. We also utilize Simulated Annealing (SA) method to get an optimal circuit. This paper shows that we can reduce the number of CNOT gates by about 19% compared to the original Gaussian Elimination-based method by inserting CNOT gates into initial circuits.

Keywords: Nearest Neighbor Architecture (NNA)-compliant, Gaussian Elimination, Inserting CNOT gates, Simulated Annealing (SA)

1 Introduction

The pursuit of practical quantum computation requires designing optimized quantum circuits that conform to quantum computers' physical limitations. This work is more difficult than conventional circuit design and requires efficient methods even before practical quantum computers are available. One frequent focus is creating Nearest Neighbor Architecture (NNA)-compliant quantum circuits, supporting operations on single qubits and physically adjacent qubits [1]. The reason for this is that most major quantum computers (for example, superconducting qubits-based ones) only support such operations [2].

Designing an NNA-compliant quantum circuit typically refers to decomposing a more large quantum circuit into single-qubit gates and CNOT gates. When a CNOT gate is non-NNA-compliant, qubits' locations are adjusted by introducing SWAP gates to make the CNOT gate NNA-compliant. Finding the smallest NNA-compliant circuit with the fewest SWAP gates is an NP-hard task, but an alternative approach called the Gaussian Elimination-based NNA-compliant circuit synthesis or GE-based method by reworking the CNOT gate-only sub-circuit to reduce the number of CNOT gates in the NNA-compliant circuit [3].

The GE-based synthesis method typically yields smaller circuits than the SWAP gate approach. Despite the GE-based synthesis result being unique for a fixed order of qubits, this method can be improved by transforming the state by inserting CNOT gates before applying it, yielding a smaller circuit. We also need to add an additional sub-circuit after the designed circuit to restore the functionality of the initial circuit. That is, we can improve the result by the GE-based synthesis method if the additional cost of the above procedure is smaller than the reduced cost of the designed circuit by the GE-based synthesis method after the above transformation. In this paper, we show that such cases indeed happen if we transform a circuit by inserting NNA-compliant CNOT gates.

*goose@ngc.is.ritsumei.ac.jp
†Y@ngc.is.ritsumei.ac.jp
‡ger@cs.ritsumei.ac.jp

2 Premilary

2.1 Representing qubit state with matrix

In a quantum circuit, a CNOT gate behaves similarly to a classical XOR gate but requires the number of inputs and outputs to be the same. The matrix representation in Fig.1 demonstrates this behavior. For example, a CNOT gate with inputs q_0 and q_1 results in outputs q_0 and $q_0 \oplus q_1$.

A matrix can be used to represent the functionality of a quantum circuit consisting only of CNOT gates, as shown in Fig.1. In this matrix, each row is determined by the state of qubits, as shown in Fig.1 (c).

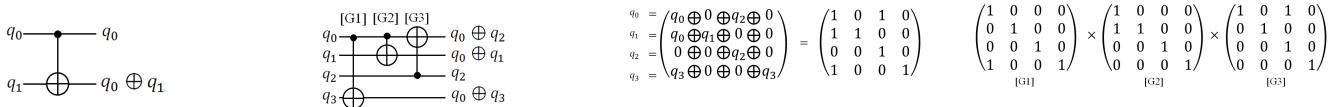
The functionality of a CNOT gate also can be represented by a matrix, as shown in Fig.1 (d). In this matrix representation, the control bits' functionality are used to modify the target bit's functionality by an XOR opeartion.

For example, the first gate (G1) in Fig.1 (d) is a CNOT gate with control bit q_0 and target bit q_3. This gate changes the matrix representation by operating the current function with XOR in 4th and 1st rows. In the resulting matrix, the first element of the fourth row corresponds to the target bit q_3 and the control bit q_0, indicated by a 1 in the leftmost matrix of Fig.1 (d).

By multiplying the three matrices in Fig.1 (d) by reversal order, we obtain the matrix depicted in Fig.1 (c), which represents the functionality of the entire circuit.

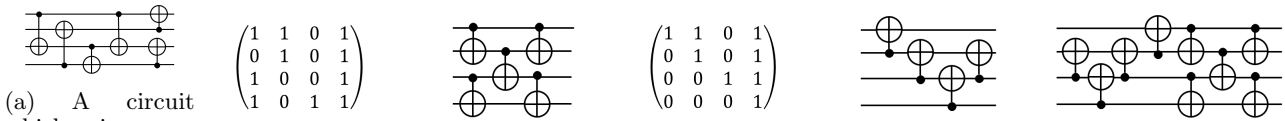
2.2 Gaussian Elimination-based NNA-compliant circuit synthesis

In this section, we want to explain how the GE-based synthesis method [3] converts the circuit as shown in Fig. 2 (a) to be NNA-compliant. First, the GE-based synthesis method takes a Boolean matrix that represents the functionality of the circuit which we want to convert to the NNA-compliant circuit, as depicted in Fig. 2 (b). It then applies Gaussian Elimination to transform the matrix into the Identity matrix. This transformation involves replacing the i-th row with the XOR operation result between the i-th and j-th rows. This operation can be viewed as a CNOT gate, where the i-th and j-th qubits serve as the target and control bits, respectively.



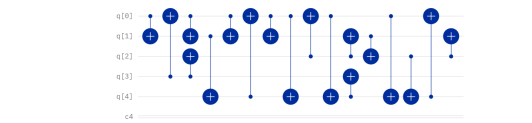
(a) A CNOT gate (b) A CNOT-based circuit (c) A matrix to represent a circuit (d) Matrices to represent CNOT gates

Figure 1: Boolean matrices to represent a circuit and CNOT gates

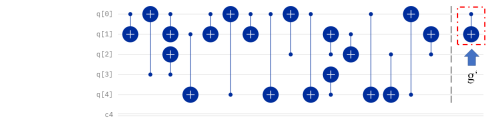


(a) A circuit which is converted to NNA-compliant (b) A matrix to represent a circuit (c) The Gaussian-down circuit (d) The upper triangular matrix (e) The Gaussian-up circuit (f) The finale circuit

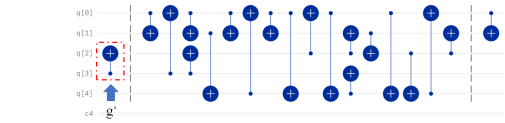
Figure 2: An example for the GE-based synthesis method



(a) The circuit which is converted to NNA-compliant



(b) The circuit after inserting the first CNOT gate



(c) The circuit after inserting the second CNOT gate

(d) The final result

Figure 3: An example for the proposed method

The GE-based synthesis process comprises two main steps. Firstly, the "Gaussian-down" step transforms the initial matrix into an upper triangular form by iteratively eliminating 1's below the diagonal, each operation corresponding to a CNOT gate, forming a "Gaussian-down circuit" (Fig.2 (c)). Secondly, the "Gaussian-up" step eliminates the remaining 1's off the diagonal, resulting in the circuit in Fig.2 (e). Finally, the NNA-compliant circuit is generated by reversing the combined Gaussian-down and Gaussian-up circuits, transforming the initial state to the final state (Fig. 2 (f)).

3 Our proposed method

In this section, we show that there is a possibility to reduce the number of CNOT gates in the designed NNA-compliant circuit by the GE-based circuit synthesis method. Our idea to do so is to try to insert NNA-compliant CNOT gates before and/or after a target circuit which we want to convert to NNA-compliant so that we can get the smaller result by the GE-based circuit synthesis method. In the following, we explain this idea by using Fig. 3.

Suppose we want to convert the circuit as shown in Fig. 3 (a) to be NNA-compliant. If we apply the GE-based circuit synthesis method to the circuit, there are 20 CNOT gates in the converted NNA-compliant circuit.

If we add only one NNA-compliant CNOT gate before the circuit as shown in Fig. 3 (a), we get the circuit as shown in Fig. 3 (b). The difference between the two circuits as shown in Fig. 3 (a) and (b) is only one CNOT gate. However, if we apply the GE-based circuit synthesis method to the circuit as shown in Fig. 3 (b), there are only 18 CNOT gates in the converted NNA-compliant circuit. This is our motivation to propose our method in this paper.

Because we can cancel the effect of the added CNOT gate by adding the same CNOT gate before the circuit which is shown in Fig. 3 (b), the above means that we can convert the circuit as shown in Fig. 3 (a) to an NNA-compliant circuit consisting of 19 CNOT gates.

Based on the above idea, we propose the following method.

Step 0: We apply the GE-based synthesis method to the target circuit, and let C be the converted NNA-compliant circuit. Let the number of CNOT gates in G be $Cost\_G$. Let $Added\_Gates$ and g be a set of gates whose initial value is an empty set. Init the initial temperature T and annealing round K.

Step 1: Add one NNA-compliant CNOT gate, g', before or after G to get G'. Apply the GE-based synthesis method to G' to get $Converted\_G'$. Let the number of CNOT gates in $Converted\_G'$ be $Cost\_G'$. Let ΔC is $Cost\_G' - Cost\_G$

Step 2: If $\Delta C < -1 - length\ (g)$ (Cause we must cancel the CNOT gates we inserted), we go to Step 1 after the following update.

- Insert g' into g
- Replace $Added\_Gates$ with g.
- Replace $Cost\_G$ with $Cost\_G'$.

- Replace G with G'.

Otherwise, if $Random\ (0,1] \leq F\ (\Delta C, T)$ ($F\ (\Delta C, T)$ is a function generating a number between (0,1], deciding to choose or not choose the worse result), we go to Step 1 after the following update.

- Insert g' into g
- Replace G with G'.

Otherwise, if do not insert any CNOT gates until n times, go to Step 3, else, go to Step 1.

Step 3: If $T \leq t$, go to Step 4. Otherwise, let $T = k * T$ (k is a number between (0,1]), and go to Step 1.

Step 4: Add the same gates in *Added_Gates* to G to get the converted circuit.

If we want to convert the circuit as shown in Fig. 3 (a), the above procedure works as follows. First at Step 0, we count the number of CNOT gates in the NNA-compliant circuit generated by the GE-based synthesis method and the number is 20; we set $Cost\_G$ to be 20 and *Added_Gates* to be an empty set and init the T and K.

Then, at Step 1 we try to insert an NNA-compliant CNOT gate, g', such that the converted NNA-compliant circuit from circuit G' has been changed by adding g'. In this example, g' is the first gate in Fig. 3 (b); we get G' as shown in Fig. 3 (b) from G as shown in Fig. 3 (a) by adding g'.

Because $Cost\_G'$ is 18 for this example, we consider that it is good to add g'. Thus we perform the following update and go to Step 1 to improve the cost further.

- Insert g' into g
- Replace G with G' as shown in Fig. 3 (b).
- Replace $Cost\_G$ to be $Cost\_G' = 18$.
- Replace *Added_Gates* with g (the g' as shown in Fig. 3 (b)).

Then at Step 1 of the second round, we find the best gate, g', to be added to the circuit as shown in Fig. 3 (b); g is the last gate in Fig. 3 (c) for this example, and so G' is the circuit as shown in Fig. 3 (c). Then $Cost\_G'$ becomes 17 for this case. The $Cost\_G'$ becomes 17 for G'. Therefore,$Cost\_G' - Cost\_G = -1$ is equal to the $-1-1$ ($length\ (g)\ is\ 1$), But $Random(1,0] \leq F(-1,T)$. So we insert g' into g and replace G with G' as shown in Fig. 3 (c), and go to Step 1 again.

Lastly, we can not insert any CNOT gate to reduce any cost many times, and the temperature is too small to continue this work. Then in Step 4, we add the gates in *Added_Gates* to the current G. In this example, *Added_Gates* contains the gates which are g in Fig. 3 (d). That is, our converted final circuit can be shown as in Fig. 3 (d) where 'Compiled NNA circuit' in the center is the current G. Thus, the total cost is 14+4 = 18 which is reduced from 20 obtained by the original GE-based synthesis method.

4 Experimental result

To verify the validity of our idea, we performed an experiment with our method and the greedy-based method we proposed before by using random circuits consisting of only CNOT gates. The results of the experiments are shown in Table. 1.

As shown in Table. 1, the number of CNOT gates is reduced by either the greedy-based or the SA-based method than the Gaussian Elimination-based method. The optimal method which is SA-based, is more effective than our improving method which is greedy-based. We also can confirm that the reduction rate by our method would increase when the number of qubits of a circuit increases.

Table 1: Experimental result

	[3]	greedy-based method		SA-based method	
#qubits	cost	cost	rate (%)	cost	rate (%)
5	18.6	16.2	12.9	15.2	17.2
6	31.5	26.8	14.9	25.7	18.4
7	40.8	33.8	17.1	32.7	19.8
8	50.3	42.4	15.7	41.1	18.2
9	75.6	64.7	14.4	63.4	16.1
10	90.3	75.0	16.9	72.8	19.3
11	106.3	86.2	18.9	85.8	19.3
12	134.9	106.8	20.8	105.9	21.5

5 Conclusion

In this paper, we considered an improved Gaussian Elimination-based NNA-compliant circuit synthesis method by inserting CNOT gates before/after the quantum circuit. We also proposed an optimization technique for this method. Experimental results confirmed that our proposed method could reduce the number of CNOT gates compared to the original Gaussian elimination-based method. Moreover, we demonstrated that the reduction rate by our method would increase when the number of qubits of a circuit increases.

For future work, we plan to explore improvements to our method on a two-dimensional lattice. Additionally, since the optimization technique of our proposed method can become computationally intensive, it would be a future challenge to investigate methods such as deep learning to reduce the computation time.

References

[1] Y. Hirata, M. Nakanishi, S. Yamashita, and Y. Nakashima An efficient method to convert arbitrary quantum circuits to ones on a linear nearest neighbor architecture. In *2009 Third International Conference on Quantum, Nano and Micro Technologies*, pages 26-33, 2009.

[2] Ibm quantum. [Online]. Available: https://quantum-computing.ibm.com/

[3] B. Nash, V. Gheorghiu, and M. Mosca Quantum circuit optimizations for nisq architectures. *Quantum Science and Technology.*, vol. 5,no. 2,p. 025010, 2020.

Incompatibility measures in multi-parameter quantum estimation under hierarchical quantum measurements

Hongzhen Chen,[1] Yu Chen,[1] and Haidong Yuan[1,*]

[1] *Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong*

(Dated: May 18, 2023)

The full articles are available on arXiv:2109.05807,2206.13095

Due to the incompatibility of the optimal measurements for different parameters, the multi-parameter quantum Cramér-Rao bound is in general not achievable. Tradeoffs among the precisions of different parameters have to be made. Quantifying such tradeoff is now one of the main subjects in quantum metrology[1–24].

Here we provides a framework to study the precision under general p-local measurements, which are the measurements that can be performed collectively on at most p copies of quantum states. This approach leads to new multi-parameter precision bounds which include the Holevo bound[25] and the Nagaoka bound[26, 27] as special cases. We also provide a systematic way to generate hierarchical analytical tradeoff relations under general p-local measurements. The obtained tradeoff relations provide a necessary condition for the saturation of the multi-parameter quantum Cramér-Rao bound under p-local measurements, which recovers the partial commutative condition[28] at $p = 1$ and the weak commutative condition at $p = \infty$. Our study thus not only provides a framework that can generate new analytical bounds on the tradeoff under general p-local measurements, but also provides a coherent picture for the existing results on the extreme cases.

The multi-parameter quantum Cramér-Rao bound is given by

$$Cov(\hat{x}) \geq \frac{1}{\nu} F_Q^{-1}, \tag{1}$$

where $Cov(\hat{x})$ is the covariance matrix for locally unbiased estimators, $\hat{x} = (\hat{x}_1, \cdots, \hat{x}_n)$, with the jk-th entry given by $Cov(\hat{x})_{jk} = E[(\hat{x}_j - x_j)(\hat{x}_k - x_k)]$, ν is the number of copies of quantum states, F_Q is the quantum Fisher information matrix. In this article, we assume F_Q is non-singular so F_Q^{-1} exists, in which case $Cov(\hat{x}) \geq \frac{1}{\nu} F_Q^{-1} > 0$ is also non-singular.

Different from the single-parameter quantum estimation, the multi-parameter quantum Cramér-Rao bound is in general not saturable. This is due to the incompatibility of the optimal measurements for different parameters. Such incompatibility is rooted in the prohibition of simultaneous

measurement of non-commutative observables and its manifested effect in multi-parameter estimation is the tradeoff on the precision limits for the estimation of different parameters.

Here we list the analytical upper bounds and the necessary condition for the saturation of QCRB under general p-local measurements.

1. For pure states, we have

$$\frac{1}{\nu}Tr[F_Q^{-1}Cov^{-1}(\hat{x})] \leq n - f(n)\|F_Q^{-\frac{1}{2}}F_{Im}F_Q^{-\frac{1}{2}}\|_F^2, \tag{2}$$

here $\|\ \|_F$ is the Frobenius norm and n is the number of parameters, $f(n) = \max\{\frac{1}{4(n-1)}, \frac{n-2}{(n-1)^2}, \frac{1}{5}\}$

2. For mixed states under p-local measurements, we have

$$\Gamma_p \leq n - f(n)\|\frac{F_Q^{-\frac{1}{2}}\bar{\mathbf{F}}_{Imp}F_Q^{-\frac{1}{2}}}{p}\|_F^2, \tag{3}$$

where $f(n) = \max\{\frac{1}{4(n-1)}, \frac{n-2}{(n-1)^2}, \frac{1}{5}\}$, $\bar{\mathbf{F}}_{Imp}$ is the imaginary part of $\bar{\mathbf{F}} = \sum_q \bar{F}_{u_q}$ with each $\bar{F}_{u_q}$ equal to either F_{u_q} or $F_{u_q}^T$, here F_{u_q} is a $n \times n$ matrix with the jk-th entry given by

$$(F_{u_q})_{jk} = \langle u_q|\sqrt{\rho_x^{\otimes p}}L_{jp}L_{kp}\sqrt{\rho_x^{\otimes p}}|u_q\rangle, \tag{4}$$

L_{jp} is the SLD of $\rho_x^{\otimes p}$ corresponding to the parameter x_j, and $\{|u_q\rangle\}$ are any set of vectors in $H_d^{\otimes p}$ that satisfies $\sum_q |u_q\rangle\langle u_q| = I_{d^p}$ with I_{d^p} denote the $d^p \times d^p$ Identity matrix.

3. For mixed states under p-local measurements, we obtain another bound as

$$\Gamma_p \leq n - \frac{1}{4(n-1)}\|\frac{C_p}{p}\|_F^2, \tag{5}$$

here

$$(C_p)_{jk} = \frac{1}{2}\|\sqrt{\rho_x^{\otimes p}}[\tilde{L}_{jp}, \tilde{L}_{kp}]\sqrt{\rho_x^{\otimes p}}\|_1, \tag{6}$$

$\tilde{L}_{jp}$ is the SLD of $\rho_x^{\otimes p}$ under the reparametrization such that the QFIM of ρ_x equals to the Identity, specifically $\tilde{L}_{jp} = \sum_q (F_Q^{-\frac{1}{2}})_{jq}L_{qp}$ with L_{qp} as the SLD of $\rho_x^{\otimes p}$ corresponding to the original parameter x_q.

4. From the above bound, we obtain a necessary condition for the saturation of the QCRB under p-local measurements, which is $\frac{C_p}{p} = 0$. For $p = 1$, this reduces to the partial commutative condition. For $p \to \infty$, we prove that

$$\lim_{p\to\infty}\frac{(C_p)_{jk}}{p} = \frac{1}{2}|Tr(\rho_x[\tilde{L}_j, \tilde{L}_k])|. \tag{7}$$

The condition, $\frac{C_p}{p} = 0$, thus reduces to the weak commutative condition, $Tr(\rho_x[\tilde{L}_j, \tilde{L}_k]) = 0$, $\forall j, k$, at $p \to \infty$. This clarifies the relation between the partial commutative condition and the weak commutative condition, which solves an open question[28].

5. We provide another simpler bound for mixed states which can be calculated with operators only on a single ρ_x.

   Given $\rho_x = \sum_{q=1}^{m} \lambda_q |\Psi_q\rangle\langle\Psi_q|$ with $\lambda_q > 0$ in the eigenvalue decomposition, under p-local measurements we have

$$\Gamma_p \leq n - \frac{1}{4(n-1)} \| \frac{T_p}{p} \|_F^2, \tag{8}$$

   where T_p is a $n \times n$ matrix with the jk-th entry given by

$$(T_p)_{jk} = \frac{1}{2} E(|\sum_{r=1}^{p} \langle\Phi_r|[\tilde{L}_j, \tilde{L}_k]|\Phi_r\rangle|), \tag{9}$$

   here $E(\cdot)$ denotes the expected value, each $|\Phi_r\rangle$ is randomly and independently chosen from the eigenvectors of ρ_x with a probability equal to the corresponding eigenvalue, i.e., each $|\Phi_r\rangle$ takes $|\Psi_q\rangle$ with probability λ_q , $q \in \{1, \cdots, m\}$. $\tilde{L}_j = \sum_\mu (F_Q^{-\frac{1}{2}})_{j\mu} L_\mu$ and $\tilde{L}_k = \sum_\mu (F_Q^{-\frac{1}{2}})_{k\mu} L_\mu$.

   For large p, this bound is almost as tight as the bound with $\frac{C_p}{p}$, the difference between $\frac{T_p}{p}$ and $\frac{C_p}{p}$ is at most of the order $O(\frac{1}{\sqrt{p}})$ with

$$\frac{(T_p)_{jk}}{p} \leq \frac{(C_p)_{jk}}{p} \leq \frac{(T_p)_{jk}}{p} + O(\frac{1}{\sqrt{p}}). \tag{10}$$

The presented framework provided a versatile tool to obtain bounds on the precision limit in multi-parameter quantum estimation under general p-local measurements, which significantly increased our knowledge on the incompatibility in multi-parameter quantum estimation. Therefore, we believe that our work makes a substantial contribution to quantum metrology and quantum information processing, and will be of interest to the general audience of AQIS.

---

* hdyuan@mae.cuhk.edu.hk

[1] Richard D. Gill and Serge Massar. State estimation for large ensembles. *Phys. Rev. A*, 61:042312, Mar 2000.

[2] F. Albarelli, M. Barbieri, M.G. Genoni, and I. Gianani. A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging. *Physics Letters A*, 384(12):126311, 2020.

[3] Angelo Carollo, Bernardo Spagnolo, Alexander A Dubkov, and Davide Valenti. On quantumness in multi-parameter quantum estimation. *Journal of Statistical Mechanics: Theory and Experiment*, 2019(9):094010, sep 2019.

[4] Xiao-Ming Lu and Xiaoguang Wang. Incorporating heisenberg's uncertainty principle into quantum multiparameter estimation. *Phys. Rev. Lett.*, 126:120503, Mar 2021.

[5] Sammy Ragy, Marcin Jarzyna, and Rafal Demkowicz-Dobrzański. Compatibility in multiparameter quantum metrology. *Phys. Rev. A*, 94:052108, Nov 2016.

[6] Yu Chen and Haidong Yuan. Maximal quantum fisher information matrix. *New Journal of Physics*, 19(6):063023, jun 2017.

[7] Jing Liu, Haidong Yuan, Xiao-Ming Lu, and Xiaoguang Wang. Quantum fisher information matrix and multiparameter estimation. *Journal of Physics A: Mathematical and Theoretical*, 53(2):023001, dec 2020.

[8] Hongzhen Chen and Haidong Yuan. Optimal joint estimation of multiple rabi frequencies. *Phys. Rev. A*, 99:032122, Mar 2019.

[9] Francesco Albarelli, Marco Barbieri, Marco G. Genoni, and Ilaria Gianani. A perspective on multiparameter quantum metrology: From theoretical tools to applications in quantum imaging. *Phys. Lett. A*, 384:126311, 2020.

[10] R. Demkowicz-Dobrzański, W. Górecki, and M. Guţă. Multi-parameter estimation beyond quantum fisher information. *J. Phys. A: Math. Theor.*, 53:363001, 2020.

[11] Jasminder S. Sidhu and Pieter Kok. Geometric perspective on quantum parameter estimation. *AVS Quantum Sci.*, 2:014701, 2020.

[12] M. D. Vidrighin, G. Donati, M. G. Genoni, X.-M. Jin, W. S. Kolthammer, M. S. Kim, A. Datta, M. Barbieri, and I. A. Walmsley. Joint estimation of phase and phase diffusion for quantum metrology. *Nat. Commun.*, 5:3532, 2014.

[13] Philip J. D. Crowley, Animesh Datta, Marco Barbieri, and I. A. Walmsley. Tradeoff in simultaneous quantum-limited phase and loss estimation in interferometry. *Phys. Rev. A*, 89:023845, Feb 2014.

[14] J.-D. Yue, Y.-R. Zhang, and H Fan. Quantum-enhanced metrology for multiple phase estimation with noise. *Sci. Rep.*, 4:5933, 2014.

[15] Yu-Ran Zhang and Heng Fan. Quantum metrological bounds for vector parameters. *Phys. Rev. A*, 90:043818, Oct 2014.

[16] Jing Liu and Haidong Yuan. Control-enhanced multiparameter quantum estimation. *Phys. Rev. A*, 96:042114, Oct 2017.

[17] Emanuele Roccia, Ilaria Gianani, Luca Mancino, Marco Sbroscia, Fabrizia Somma, Marco G Genoni, and Marco Barbieri. Entangling measurements for multiparameter estimation with two qubits. *Quantum Science and Technology*, 3(1):01LT01, nov 2017.

[18] Sholeh Razavian, Matteo G. A. Paris, and Marco G. Genoni. On the quantumness of multiparameter estimation problems for qubit systems. *Entropy*, 22(11), 2020.

[19] Alessandro Candeloro, Matteo G.A. Paris, and Marco G. Genoni. On the properties of the asymptotic incompatibility measure in multiparameter quantum estimation. *arxiv*, page 2107.13426, 2021.

[20] Koichi Yamagata, Akio Fujiwara, and Richard D. Gill. Quantum local asymptotic normality based on a new quantum likelihood ratio. *The Annals of Statistics*, 41(4):2197 – 2217, 2013.

[21] Jonas Kahn and Mădălin Guţă. Local asymptotic normality for finite dimensional quantum systems. *Communications in Mathematical Physics*, 289:597–652, Jul 2009.

[22] Yuxiang. Yang, Giulio Chiribella, and Masahito Hayashi. Attaining the ultimate precision limit in quantum state estimation. *Communications in Mathematical Physics*, 368:223–293, 2019.

[23] Jun Suzuki. Explicit formula for the holevo bound for two-parameter qubit-state estimation problem. *Journal of Mathematical Physics*, 57(4):042201, 2016.

[24] Jasminder S. Sidhu, Yingkai Ouyang, Earl T. Campbell, and Pieter Kok. Tight bounds on the simultaneous estimation of incompatible parameters. *Phys. Rev. X*, 11:011028, Feb 2021.

[25] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory.* North-Holland, Amsterdam, 1982.

[26] H. Nagaoka. A new approach to cramer–rao bounds for quantum state estimation. In Masahito Hayashi, editor, *Asymptotic theory of quantum statistical inference: Selected Papers*, Singapore, 2005. World scientific. Originally published as IEICE Technical Report, 89, 228, IT 89–42, 9–14 (1989).

[27] H. Nagaoka. A generalization of the simultaneous diagonalization of hermitian matrices and its relation to quantum estimation theory. In Masahito Hayashi, editor, *Asymptotic theory of quantum statistical inference: Selected Papers*, Singapore, 2005. World scientific.

[28] Jing Yang, Shengshi Pang, Yiyu Zhou, and Andrew N. Jordan. Optimal measurements for quantum multiparameter estimation with general states. *Phys. Rev. A*, 100:032104, Sep 2019.

Quantum error reduction with deep neural network

A.A. Zhukov[1] * W.V. Pogosow,2[1]

[1] *N. L. Dukhov All-Russia Research Institute of Automatics, Moscow, 127030, Russia*
[2] *Institute for Theoretical and Applied Electrodynamics, Moscow, 125412, Russia*

Abstract. Deep neural networks (DNN) can be applied at the post-processing stage for the improvement of the results of quantum computations on noisy intermediate-scale quantum (NISQ) processors. Here, we propose a method based on this idea, which is most suitable for digital quantum simulation characterized by the periodic structure of quantum circuits consisting of Trotter steps. A key ingredient of our approach is that it does not require any data from a classical simulator at the training stage. The network is trained to transform data obtained from quantum hardware with artificially increased Trotter steps number (noise level) towards the data obtained without such an increase. The additional Trotter steps are fictitious, i.e., they contain negligibly small rotations and, in the absence of hardware imperfections, reduce essentially to the identity gates. This preserves, at the training stage, information about relevant quantum circuit features. Two particular examples are considered that are the dynamics of the transverse-field Ising chain and XY spin chain, which were implemented on two real five-qubit IBM Q processors. A significant error reduction is demonstrated as a result of the DNN application that allows us to effectively increase quantum circuit depth in terms of Trotter steps.

Keywords: AQIS, quantum simulation, quantum computation, neural networks,machine learning, error mitigation, NISQ

Quantum information is a fast developing field that aims to utilize quantum properties, such as quantum interference and entanglement [1]. State-of-the-art quantum computers are already capable of solving many problems, which, however, are not of practical importance yet, because of relatively high quantum hardware error rates. Particularly, such processors can be useful for solving evolutionary problems. However, the simulation of the dynamics of such systems at long times requires a large number of Trotter decomposition steps of evolution operator. This leads to the fact that a large number of quantum gates are required for simulation, which means that the outcomes from the quantum computer become too noisy [2].

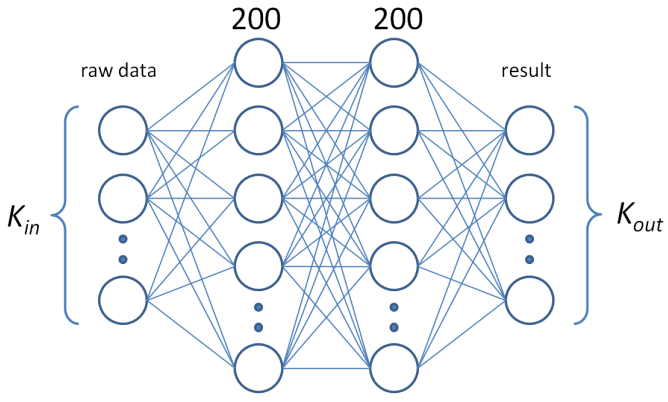


Figure 1: The feed-forward DNN architecture used in our problems for improving simulation results.

In our work, we have proposed a method for the application of classical neural networks for the improvement of the outcomes of noisy quantum computers at the post-processing stage. In contrast to other suggestions, using our approach, it is possible to get data for training a neural network without relying on a classical simulator or any other source of ideal data.

Our method is based on artificial increase of the quantum circuits depth on the training stage that can be done by incorporation of fictitious Trotter blocks formally equivalent to identity gates into the circuit (see Fig. 2). Their role is to increase noise level due to the hardware imperfections while preserving the circuit's general structure and its relevant features.

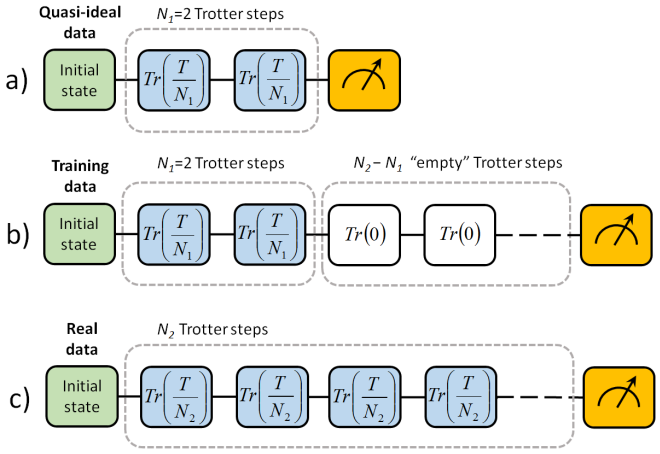


Figure 2: Schematic view of our approach: generation of quasi-ideal data with relatively shallow circuit (a); training the DNN – the data with artificially increased Trotter steps number are transformed towards quasi-ideal data (b); the trained DNN is applied to raw experimental data with the same Trotter step number as at the second stage (c).

*zugazoid@gmail.com

After being trained, the network can be applied to new data with the same Trotter step number, i.e., increased in the same way as at the training stage, but without fictitious Trotter steps. The amount of noise in this case is similar to that at the training stage. This trick allows for the effective increase of the Trotter number due to the post-processing, in the sense that errors become suppressed and results of simulations, which must have error rates below a given level, start to include data with larger Trotter step number.

We have demonstrated the basic ingredients of our approach using two examples [3]: digital quantum simulations of the dynamics of the transverse-field Ising chain and XY chain. Deep neural network with simple architectures were used at the post-processing stage. For XY chain, an additional post-selection of the results at the training stage was applied by discarding a part of the data, which does not conserve the excitation number, as required by the Hamiltonian. The proof-of-principle results obtained on a real 5-qubit IBM Athens and Bogota quantum processors show that our method allows us to increase the number of Trotter steps while maintaining the same level of errors. The significant error reduction is the main result of our demonstration. A single neural network is able to improve the data for different initial conditions.

The significant error reduction is the main result of our demonstration. A single neural network is able to improve the data for different initial conditions.

Fig. 3 shows MSE between ideal simulation data for a given Trotter step number and experimental data improved by the network with post-selection (o-shape symbols) and without post-selection (triangle-shape symbols) as well as raw data (x-shape symbols and dashed line). It is seen from this figure that neural network is able to significantly improve the quality of the data by decreasing MSE in several times for $N_2 = 4$ and 6. Post-selection leads to a further improvement. In the case for $N_2 = 8$ there is almost no improvement in the results due to the fact that the IBM Bogota quantum processor used has too high error rate (CNOT errors $\approx 2\%$) to enable quantum simulation with so long quantum circuit depth. As a result, the training data for $N_2 = 8$ is very noisy, so that the magnetization as a function of time is flat, since the probability of each qubit to be in the state 1 is close to 0.5 for any time. Such a data can not be used effectively to train a neural network.

Our method does not require a complete tomography of quantum states, which allows it to be scaled. The reason is that the network is trained to improve the data for a restricted number of quantum mean values such as spins magnetizations along different axes, order parameters, or characteristic correlators.

We believe that the proposed approach can be useful in the context of error mitigation in noisy quantum devices (especially of next generations with hardware errors decreased and qubit number increased). Particularly, it can be used in the case of periodic quantum circuits and in combination with other error reduction tools, such as post-selection or partial error correction.

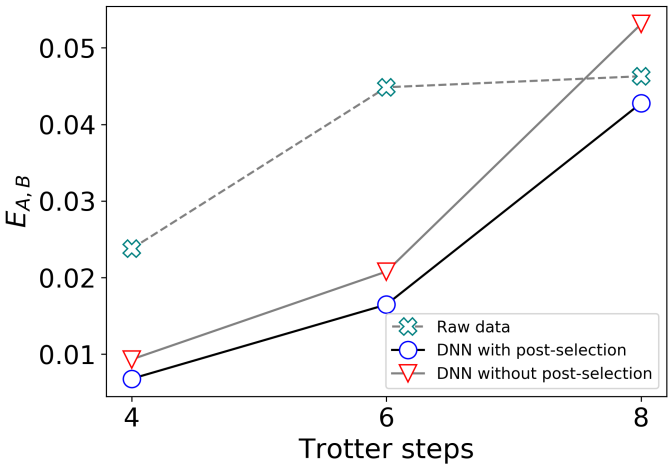


Figure 3: MSE between ideal simulation data for a given Trotter step number and experimental data from an IBM Bogota 5-qubit quantum processor improved by the network with post-selection (o-shape symbols) and without post-selection (triangle-shape symbols) as well as raw data (x-shape symbols and dashed line). The errors were averaged over time and all initial states from the computational basis.

Acknowledgements We acknowledge use of the IBM Quantum Experience for this work. The viewpoints expressed are those of the authors and do not reflect the official policy or position of IBM or the IBM Quantum Experience team.

References

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. In *Rev. Mod. Phys.* 81:865, 2009.

[2] A.A. Zhukov, S.V. Remizov, W.V. Pogosov, and Yu.E. Lozovik. Algorithmic simulation of far-from-equilibrium dynamics using quantum computer. In *Quant. Inf. Proc.* 17:1, 2018.

[3] A.A. Zhukov, W.V. Pogosov. Quantum error reduction with deep neural network applied at the post-processing stage. In *Quant. Inf. Proc.* 21:93, 2022.

Contextuality and memory cost of simulation of Majorana fermions

Susane Calegari[1] * Juani Bermejo-Vega[2] Zoltan Zimboras[3] Michał Oszmaniec[1]

[1] *Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland*
[2] *Departamento de Electromagnetismo y Física de la Materia, Universidad de Granada, 18010 Granada, Spain*
[3] *Wigner Research Centre for Physics of the Hungarian Academy of Sciences, H-1525 Budapest, P.O.Box 49, Hungary*

Abstract. We prove state-independent contextuality for Majorana fermions via graph theory and demonstrate that contextuality gives lower bounds on the memory cost of simulating restricted classes of quantum computation. Specifically, we apply these results to two models of quantum computation based on the braiding of Majorana fermions: Topological Quantum Computation (TQC) with Ising anyons and Fermionic Linear Optics (FLO), finding saturable lower bounds on the memory cost that scales, respectively, log-linearly and quadratically with the number of fermionic modes. Showing those pre-existing simulation algorithms based on the stabilizer and matchgate formalism asymptotically saturate the above bounds on the memory cost.

Keywords: Majorana Fermions, Contextuality, Memory cost, classical simulation, Pfaffian graphs

Quantifying the classical resources required to simulate quantum processes is a crucial step in understanding the distinction between quantum and classical systems. Contextuality, which has been identified as a resource in certain models of quantum computation with magic states [1], is analogous to non-locality, which is a known resource for quantum communication and cryptography [2]. Previous work by [3] has demonstrated that the presence of contextuality imposes a lower bound on the memory cost of classically simulating quantum processes [4, 5], specifically for simulating Clifford operators.

However, the techniques employed in the aforementioned study were limited to a specific class of quantum processes known as "closed subtheories." These subtheories only allow measurable quantities and their products to be measurable, thus excluding quantum processes involving local measurements that are commonly encountered in quantum communication or measurement-based quantum computation.

In this work, we address these limitations and extend the connection between contextuality and classical simulation to non-closed families of observables. We adopt the approach proposed by Abramsky and Brandenburger[6], which provides a more general proof of contextuality. To establish this connection, we consider a set of measurable observables (denoted as $\mathcal{O}$) and a set of possible outcomes (Ω) for each measurement. A context (M) is defined as a subset of $\mathcal{O}$, and $\mathcal{M}$ represents the set of all possible contexts. For each context $M \in \mathcal{M}$, a "section over" M is defined as a function $\nu : M \to \Omega$, and we denote $\nu(M)$ as the set of outcomes observed for the measurements in M. The "event" corresponding to a particular measurement outcome sequence can be represented as $\mathbb{P}_{\nu(M)} = \prod_{O\in M} \mathbb{P}^{O}_{\nu(O)}$, where $\mathbb{P}^{O}_{\nu(O)}$ denotes the projection operator associated with outcome $\nu(O)$ for observable O. A global section is a function $s : \mathcal{O} \to \Omega$ that is consistent with compatible local sections over all contexts, i.e., $s(M) = \nu(M)$ for all $M \in \mathcal{M}$. The existence of a global section implies the existence of a non-contextual value assignment (NCVA) for $\mathcal{O}$. Notably, for closed subtheories, the non-contextual constraints mentioned above are equivalent to the conventional proof of (non)contextuality that relies on the criterion $\nu(O_1)\nu(O_2) = \nu(O_1O_2)$ for commuting measurements O_1, O_2, and their product O_1O_2. However, our framework extends to non-closed subtheories where products of measurable quantities may not be measurable.

In the implementation of a quantum algorithm, one is limited to specific sets of quantum states (Σ), quantum channels ($\mathcal{T}$), and measurable observables ($\mathcal{O}$) with outcomes belonging to the set Ω. Since multiple experiments may yield the same information, we define a "subtheory" as a triple $(\Sigma, \mathcal{T}, \mathcal{O})$ without specifying its physical implementation.

To simulate quantum statistics, we employ an ontological model in the classical simulation that reproduces the Born rule probabilities of a quantum subtheory. In this classical simulation, the density matrix is represented by a probability distribution $\mu_\rho(\lambda)$ over the state space Λ, and measurements are modeled as sub-stochastic maps $\Gamma_O(\lambda', k|\lambda)$. After a measurement, the probability distribution $\mu_\rho(\lambda)$ is updated to $\mu_{\rho'}(\lambda')$ with probability $Pr(O, k|\rho, \lambda) = \sum_{\lambda,\lambda'} \Gamma_O(\lambda', k|\lambda)\mu_\rho(\lambda)$. The internal state $\lambda \in \Lambda$ contains all the information necessary to characterize the statistics of all measurements allowed in the subtheory. The lower bound on the memory required for this simulation is determined by finding a lower bound on the size of the state space Λ necessary to simulate the subtheory.

One of the main contributions of our work is establishing a connection between the memory cost of classically simulating quantum processes and state-independent contextuality. We demonstrate that the size of the ontological space Λ describing the subtheory is at least $|\Sigma|/\alpha^*$, where $\alpha^* \in \Sigma$ represents the largest subset of states with overlapping supports. In other words, $\alpha^* = |\mathcal{Z}|$ is defined as the largest subset of states for

*calegari@cft.edu.pl

which $\bigcap_{\rho\in\mathcal{Z}} \mathrm{supp}(\mu_\rho) \neq \emptyset$. Furthermore, since every measurable observable $O \in \mathcal{O}$ in the subtheory contains at least one eigenstate $\rho \in \Sigma$, we prove that if $\mathcal{O}$ is contextual, then $\bigcap_{\rho\in\Sigma} \mathrm{supp}(\mu_\rho) = \emptyset$. These results provide a lower bound on the size of the ontological space, which depends on the largest subset of states in the subtheory that satisfies the non-contextuality condition imposed by its stabilizer group.

Theorem 1 *Let Σ be the set of quantum states and $\mathcal{O}$ the set of measurable observables in a subtheory. The size of the state space Λ describing this subtheory is at least $|\Sigma|/\beta^*$, where*

$$\beta^* = \max_{\mathcal{Z}\subset\Sigma} \Big\{ |\mathcal{Z}| \;\Big|\; \mathcal{O}_{\mathcal{Z}} \text{ is non-contextual} \Big\}, \qquad (1)$$

and $\mathcal{O}_{\mathcal{Z}} \subseteq \mathcal{O}$ is the stabilizer group of the states in $\mathcal{Z}$.

In our research, we specifically focus on the simulation of the restricted model of quantum computation based on Majorana Fermions, known as "Topological quantum computation with Ising Anyons" (TQC) [7] and "Fermionic linear optics" (FLO). We establish state-independent contextuality for Majorana fermions using the sheaf-theoretic approach introduced by Abramsky and Brandenburger. In the TQC framework, initial states can be transformed into other states using braid gates and measurable observables. The allowed states in the TQC subtheory can be stabilized by pairs of Majorana operators, represented as $\mathrm{Stab}(|0\rangle^{\otimes n}) = (-im_1m_2, -im_3m_4, \ldots, -im_{2n-1}m_{2n})$. Majorana operators obey commutation rules $m_im_j + m_jm_i = 2\delta_{ij}\mathbb{I}$, where $m_i^\dagger = m_i$ for any $i, j \in [2n]$. In a simple undirected graph $G = ([2n], E)$, each Majorana mode $(\{m_j\}_{j=1}^{2n})$ corresponds to a vertex in G, and each observable $X_{i,j} \in \mathcal{O}_{\mathrm{TQC}}$ corresponds to an edge $(i, j) \in E$. For a given set of fixed-parity TQC states, Σ, the corresponding observables form $\mathcal{O} = \bigcup_{\rho\in\Sigma} \mathrm{Stab}(\rho)$. Similarly, the set of edges is $E = \bigcup_{M\in\mathcal{M}} M$, where $\mathcal{M}$ represents the set of perfect matchings that represents the stabilizer group of states in Σ. We prove the following theorem:

Theorem 2 (NCVA for TQC $\Leftrightarrow$ Pfaffian graph) *There exists a non-contextual value assignment (NCVA) for $\mathcal{O} \subseteq \mathcal{O}_{TQC}$ if and only if the graph representing $\mathcal{O}$ is Pfaffian.*

Our analysis identifies a minimal proof of contextuality represented by a set of 9 Majorana observables forming a magic square, which can be depicted as a $K_{3,3}$ graph [8]. For the TQC subtheory, $|\Sigma_{\mathrm{TQC}}| = (2n)!/n!$, and we can upper bound β^* using properties of Pfaffian graphs. This leads to the following lemma: "A classical simulation of TQC requires at least $n\log(n)$ classical bits of memory."

The TQC model lies at the intersection of two computational models: the Clifford group/stabilizer formalism model and the Fermionic Linear Optics (FLO) [9]. Quantum computation with fermionic linear optics extends the TQC model by allowing unitaries, called FLO gates, that are not restricted to the $\pi/4$ angle. Consequently, the FLO model is also contextual, and the lower bound computed for TQC can be extended to FLO. The states allowed in the FLO subtheory can be approximated within ϵ and require a state space size of at least $|\Lambda| \geq \frac{1}{\epsilon^{n^2}2^{n-2}}$. As a result, the lower bound on the memory cost of classically simulating the ϵ-approximate FLO model is $n^2\log(1/\epsilon)$.

In summary, our work establishes a connection between contextuality and memory cost in simulating quantum circuits. We develop novel techniques to derive lower bounds on the memory cost of simulating physical subtheories, focusing particularly on fermionic systems. These findings have implications for understanding Majorana Fermions through graph theory and for designing contextuality witnesses for experimental tests in these subtheories.

Comment – An earlier version of this work was initially presented at AQIS 2020, which was held exclusively in an online format, with presentations delivered through uploaded YouTube videos. Since then, significant progress has been made on this project, surpassing our initial expectations. Notably, we have successfully established a comprehensive proof of contextuality for Majorana Fermions, establishing a groundbreaking connection between contextuality and non-Pfaffian graphs. Given these remarkable advancements, we believe our work warrants consideration for a speaking opportunity at AQIS 2023.

References

[1] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, June 2014.

[2] Peter W. Shor and John Preskill. Simple Proof of Security of the BB84 Quantum Key Distribution Protocol. *Phys. Rev. Lett.*, 85(2):441–444, July 2000.

[3] Angela Karanjai, Joel J. Wallman, and Stephen D. Bartlett. Contextuality bounds the efficiency of classical simulation of quantum processes. February 2018.

[4] Adán Cabello, Mile Gu, Otfried Gühne, and Zhen-Peng Xu. Optimal Classical Simulation of State-Independent Quantum Contextuality. *Phys. Rev. Lett.*, 120(13):130401, March 2018.

[5] Matthias Kleinmann, Otfried Gühne, José R. Portillo, Jan-Ake Larsson, and Adán Cabello. Memory cost of quantum contextuality. *New J. Phys.*, 13(11):113011, November 2011.

[6] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New J. Phys.*, 13(11):113036, November 2011.

[7] Sergey Bravyi. Universal quantum computation with the $\ensuremath{\nu}=5/2$ fractional quantum Hall state. *Phys. Rev. A*, 73(4):042313, April 2006.

[8] Abu Ashik Md. Irfan, Karl Mayer, Gerardo Ortiz, and Emanuel Knill. Certified quantum measurement of Majorana fermions. *Phys. Rev. A*, 101(3):032106, March 2020.

[9] Sergey Bravyi. Lagrangian representation for fermionic linear optics. September 2004.

[10] E. B. Davies and J. T. Lewis. An operational approach to quantum probability. *Commun. Math. Phys.*, 17(3):239–260, September 1970.

[11] Masanao Ozawa. Quantum measuring processes of continuous observables. *J. Math. Phys.*, 25(1):79–87, January 1984.

[12] Nicholas Harrigan, Terry Rudolph, and Scott Aaronson. Representing probabilistic data via ontological models. July 2008.

[13] Fernando de Melo, Piotr Ćwikliński, and Barbara M. Terhal. The power of noisy fermionic quantum computation. *New J. Phys.*, 15(1):013015, January 2013.

[14] Xun Gao, Eric R. Anschuetz, Sheng-Tao Wang, J. Ignacio Cirac, and Mikhail D. Lukin. Enhancing Generative Models via Quantum Correlations. *Phys. Rev. X*, 12(2):021037, May 2022.

[15] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70(5):052328, November 2004.

[16] Jacques Hadamard. Résolution d'une question relative aux determinants. *Bull. Sci. Math.*, 17:240–246, 1893.

[17] Soumyashant Nayak. The Hadamard determinant inequality — Extensions to operators on a Hilbert space. *J. Funct. Anal.*, 274(10):2978–3002, May 2018.

A Introduction

This manuscript is based on two unpublished works. Here, we provide a brief comment on the method and the results presented in this extended abstract.

B Ontological model

To precisely state our results let us formally introduce what we mean by classically simulating a subtheory. The implementation of a quantum algorithm is always restricted to a specific set of quantum states Σ, a set of quantum channels $\mathcal{T}$, and, a set of measurable observables $\mathcal{O}$, with outcomes belonging to the set Ω, that one can use in a class of experiments. Since there may be different experiments that allow us to obtain the same information, we define a *subtheory* as a triple $(\Sigma, \mathcal{T}, \mathcal{O})$ without specifying its physical implementation.

Quantum channels are completely positive and trace-preserving maps (CPTP maps). Unitary channels, represented by $\mathcal{U}(\rho)$, are the conjugated action of the unitary $U \in U(\mathcal{H})$, $\mathcal{U}(\rho) := U\rho U^\dagger$. The quantum measurements are described by *quantum instruments*[10, 11], a set of completely positive linear maps (CP maps) $\{\mathcal{E}_k\}_k$, whose sum is trace-preserving (CPTP map), i.e., $\mathrm{tr}(\sum_k \mathcal{E}_k(\rho)) = \mathrm{tr}(\rho)$. The label k is the outcome of a measurement which occurs with probability $p_k = \mathrm{tr}(\mathcal{E}_k(\rho))$, given the condition that the density matrix is mapped to $\rho \mapsto \mathcal{E}_k(\rho)/p_k$. For von Neumann instruments $\mathcal{E}_k(\rho) = \mathbb{P}_k \rho \mathbb{P}_k$, where $\mathbb{P}_k$ is the projector corresponding to the k-th outcome of the measurement. We denote $\mathbb{P}_k^O$ or $\mathbb{P}_{(k,O)}$ as the projector associated to a von Neumann measurement of an observable O with outcome $k \in \Omega$. In this paper, we consider closed state spaces, i.e., the quantum channels and measurements can only map $\rho \mapsto \rho'$, where $\rho, \rho' \in \Sigma$.

In the framework of ontological theories, the statistics of a quantum algorithm can be reproduced by using probability distributions over a space of hidden variables, and, stochastic and sub-stochastic matrices that transform these distribution. We consider a classical state-space Λ, where $\lambda \in \Lambda$ is an internal state that encodes the information needed to determine the statistics of the subtheory [12]. The density matrix is represented, in the classical simulation, by a probability distribution $\mu_\rho(\lambda)$ over the state space. The classical counterpart of the set of quantum channels is a a set of stochastic matrices Γ_U, that map the probability distribution $\mu_\rho(\lambda)$ to $\mu_{\rho'}(\lambda') = \sum_{\lambda\in\Lambda} \Gamma_U(\lambda'|\lambda)\mu_\rho(\lambda)$. A set of von Neumann instruments, $\{\mathbb{P}_k^O\}_k$, become a set of classical instruments (sub-stochastic maps, i.e. $\sum_{\lambda'} \Gamma_{O,k}(\lambda'|\lambda) \leq 1$, for all λ, and the sum of the sub-stochastic maps over k is stochastic map), $\{\Gamma_{O,k}(\lambda'|\lambda)\}_k$. After a measurement the probability distribution $\mu_\rho(\lambda)$ is updated to $\mu_{\rho'_k}(\lambda) = \sum_{\beta\in\Lambda} \Gamma_{O,k}(\lambda|\beta)\mu_\rho(\beta)/\Pr(k|\rho, O)$, with probability $\Pr(k|\rho, O) = \sum_{\beta,\beta'} \Gamma_{O,k}(\beta'|\beta)\mu_\rho(\beta)$. The stochastic and sub-stochastic maps are updating the probability distributions on the ontological space, such that,

$$\rho \mapsto \rho' \implies \lambda \in \mathrm{supp}(\mu_\rho) \mapsto \lambda' \in \mathrm{supp}(\mu_{\rho'}). \quad (2)$$

For the subtheories considered in this work, the classical simulation can be summarized by these four objects $(\Lambda, \{\mu_{\rho_i}\}_i, \Gamma_U, \{\Gamma_{O,k}\}_k)$: the state space, the set of probability distributions over the state space, and the stochastic and sub-stochastic maps. These tools allow us to classically simulate sequences of measurements in the subtheory. The cardinality of the state-space, $|\Lambda|$, determine the size of the memory of the simulation, i.e., $\lceil \log |\Lambda| \rceil$ is an upper bound on the number of classical bits of memory required for the classical simulation.

If possible, within the subtheory, to perform a measurement that perfectly distinguishes between two different states ρ and ρ', then, we say these states are single-shot distinguishable (SSD). Accordingly, in the classical simulation of this subtheory, the support of the probability

distributions of SSD states must be disjoint, i.e.,

$$\text{SSD} \quad \Longrightarrow \quad \text{supp}(\mu_\rho) \cap \text{supp}(\mu_\sigma) = \emptyset. \tag{3}$$

In other words, no internal state can be in the support of two states that are SSD.

C Contextuality

In this section, we introduce the notion of state-independent contextuality used in our results. In Ref. [6], Abramsky and Brandenburger use sheaf theory to show that contextuality correspond exactly to obstructions to the existence of global sections. This notion is more general than the one considered in Ref. [3] and allows us to study non-closed subtheories.

We fix a set of measurable observables $\mathcal{O}$ and a set Ω of possible outcomes for each measurement. A measurement *context* is a set of commuting observables $M \subseteq \mathcal{O}$. The set of all contexts $M \subseteq \mathcal{O}$ is denoted as $\mathcal{M}$.

Definition 3 (Event) *For each context $M \in \mathcal{M}$ a section over M is a function $\nu : M \to \Omega$. We denote $\nu(M) := \{(O, \nu(O)) \,|\, O \in M\}$. Such a section describes the* event *in which the measurements in M were carried out, and the outcomes $\nu(M)$ were observed. The* event *can be represented as a projective measurement of the section over M,*

$$\mathbb{P}_{\nu(M)} = \prod_{O \in M} \mathbb{P}^O_{\nu(O)}. \tag{4}$$

For every context $M \in \mathcal{M}$ there exists a set of local sections, $\Gamma(M) = \{\nu(M)\}$. This set includes all possible outcomes for this measurement context.

Let $s : \mathcal{O} \to \Omega$ be a function that assigns a definite outcome to each observable in $\mathcal{O}$, independent of the measurement context in which it appears, and, for every subset $X \subseteq \mathcal{O}$, we define $s(X) := \{(O, s(O)) \,|\, O \in X\}$. The assignment s is a *global section* over $\mathcal{O}$ if it is consistent with a family of local section over all contexts, i.e., there exists one local section $\nu(M) \in \Gamma(M)$, such that,

$$s(M) = \nu(M), \quad \forall M \in \mathcal{M}. \tag{5}$$

Definition 4 (NCVA) *Let $s : \mathcal{O} \to \Omega$ be a global section over $\mathcal{O}$. Then, the non-contextual value assignment (NCVA) for events is defined as*

$$\lambda_s(\mathbb{P}_{\nu(M)}) = \delta_{s(M),\nu(M)} \tag{6}$$

$$= \prod_{O \in M} \delta_{s(O),\nu(O)} \tag{7}$$

If no NCVA exists for $\mathcal{O}$, then $\mathcal{O}$ is called "contextual", which provides proof of *state independent contextuality.*

We point out that, for closed subtheories, the above non-contextual constraints are equivalent to the usual proof of (non)contextuality that relies on the criterion $\nu(O_1)\nu(O_2) = \nu(O_1O_2)$ for triples of commuting measurements of the form $M = \{O_1, O_2, O_1O_2\}$. However, our framework also applies to subtheories that are non-closed, i.e., where products of measurable quantities may not be measurable in the subtheory.

D Lower bound in the memory cost

Here, we show our main results connecting a lower bound in the memory cost for classically simulating quantum processes by ontological models and contextuality. The following lemma presents a lower bound for the size of the ontological space Λ that depends on the largest subset of states in the subtheory that have overlapping supports:

Lemma 5 *Let Σ be the set of all states in a subtheory. The size of the ontological space Λ describing the subtheory is at least $|\Sigma|/\alpha^*$, where*

$$\alpha^* = \max_{\mathcal{Z} \subset \Sigma} \Big\{ |\mathcal{Z}| \;\Big|\; \bigcap_{\rho \in \mathcal{Z}} \text{supp}(\mu_\rho) \neq \emptyset \Big\}. \tag{8}$$

Proof. For every state $\rho \in \Sigma$ there exist at least one $\lambda \in \Lambda$ in support of the corresponding probability distribution μ_ρ. Therefore, we can upper bound the number of states by summing the cardinality of the support of each state

$$|\Sigma| \leq \sum_{\rho \in \Sigma} |\,\text{supp}(\mu_\rho)| \;= \sum_{\rho \in \Sigma} \left(\sum_{\lambda \in \text{supp}(\mu_\rho)} 1 \right). \tag{9}$$

One can then rewrite the right-hand side of the above equation as a sum over all $\lambda \in \Lambda$ and, for specific λ, over the states ρ that have the internal state λ in their supports. This gives us

$$|\Sigma| \leq \sum_{\lambda \in \Lambda} \left(\sum_{\{\rho : \lambda \in \text{supp}(\mu_\rho)\}} 1 \right). \tag{10}$$

Since the internal sum above can be upper bounded by α^*, we get $|\Sigma| \leq \alpha^* |\Lambda|$, which is the desired inequality. □

For any set of states that contain SSD states, $\{\rho_j\}_{j \in \mathcal{J}}$, we necessarily have $\bigcap_{j \in \mathcal{J}} \text{supp}(\mu_{\rho_j}) = \emptyset$, and hence cardinality of such a set $\mathcal{J}$ puts a lower bound on the size of Λ: $|\Lambda| \geq |\mathcal{J}|$. However, some non-SSD states must also have non-intersecting supports of $\mu_\rho(\lambda)$.

Theorem 6 *Consider a set of quantum states $\Sigma = \{\rho_j\}_{j \in \mathcal{J}}$. Let $\mathcal{O}$ be the set of all measurable observables that have at least one eigenstate in Σ. If $\mathcal{O}$ is contextual, then $\bigcap_{j \in \mathcal{J}} \text{supp}(\mu_{\rho_j}) = \emptyset$, for any simulation of this subtheory.*

Proof. We prove the theorem by contradiction i.e. we will assume that $\bigcap_{j \in \mathcal{J}} \text{supp}(\mu_{\rho_j}) \neq \emptyset$ and show that this condition implies the existence of a non-contextual value assignment for $\mathcal{O}$.

From the definition of $\mathcal{O}$, we know that for every $O \in \mathcal{O}$ exists at least one state $\rho \in \Sigma$ with eigenvalue $s(O)$, such that, $\mathbb{P}^O_{s(O)} \rho \mathbb{P}^O_{s(O)} = \rho$. By our assumption, the support of $\{\mu_{\rho_j}\}_{j \in \mathcal{J}}$ are nonempty, consequently, all states in Σ are non-SSD. This implies that all eigenstates in Σ of a given observable $O \in \mathcal{O}$ have the same eigenvalue $s(O)$,

where $s : \mathcal{O} \to \Omega$ is a well-defined function. Since the measurement of O cannot perfectly discriminate between ρ and any other state in Σ, this measurement acts as a map

$$\rho_j \mapsto \sigma_{O,j} = \frac{\mathbb{P}^O_{s(O)} \rho_j \mathbb{P}^O_{s(O)}}{\mathrm{tr}(\mathbb{P}^O_{s(O)} \rho_j)}, \quad \forall j \in \mathcal{J}, \tag{11}$$

with non-zero probability $\Pr(s(O)|\rho_j, O) = \mathrm{tr}(\mathbb{P}^O_{s(O)}\rho_j)$.

From the update rule 2, all post-measurement states $\{\sigma_j\}_{j\in\mathcal{J}}$, for this particular measurement process, are also non-SSD. Specifically, we know that $\bigcap_{j\in\mathcal{J}} \mathrm{supp}(\mu_{\rho_j}) \subset \mathrm{supp}(\mu_\rho)$ and that $\mathbb{P}^O_{s(O)}$ preserves ρ. Thus, there exists an internal state $\tilde{\lambda}$, such that, $\Gamma_{O,s(O)}(\tilde{\lambda}|\lambda) \neq 0$, and λ is mapped to $\tilde{\lambda}$,

$$\lambda \in \bigcap_{j\in\mathcal{J}} \mathrm{supp}(\mu_{\rho_j}) \mapsto \tilde{\lambda} \in \bigcap_{j\in\mathcal{J}} \mathrm{supp}(\mu_{\sigma_j}). \tag{12}$$

Consequently, $\bigcap_{j\in\mathcal{J}} \mathrm{supp}(\mu_{\sigma_j}) \neq \emptyset$ which implies that all states in $\{\sigma_j\}_{j\in\mathcal{J}}$ are non-SSD. By repeating the above arguments for two commuting observables O and O', we obtain $\mathrm{tr}(\mathbb{P}^{O'}_{s(O')}\mathbb{P}^O_{s(O)}\rho_j) \neq 0$, $\forall j \in \mathcal{J}$. Furthermore, for any commuting set of observables $M \subseteq \mathcal{O}$, and corresponding outcomes $s(M) := \{(O, s(O)|O \in M)\}$, we have

$$\mathrm{tr}(\mathbb{P}_{s(M)}\rho_j) \neq 0, \quad \forall j \in \mathcal{J}. \tag{13}$$

Thus, the set $\mathcal{O}$ is non-contextual, since the function s is a global section over $\mathcal{O}$ that is compatible with local sections given by the measurement contexts. □

From the theorem above we know that if $\mathcal{O}$ is contextual then the states in Σ cannot share an internal state, therefore that the largest subset of states, such that, $\mathcal{O}$ is non-contextual is an upper bound for α^* in Eq. 8, i.e.

$$\alpha^* \leq \max_{\mathcal{Z}\subset\Sigma} \Big\{ |\mathcal{Z}| \;\Big|\; \mathcal{O}_{\mathcal{Z}} \text{ is non-contextual} \Big\}. \tag{14}$$

The above result can be use to find a non-trivial lower bound in the memory cost of classical simulation.

E Majorana Fermions

In this work, we apply the method of finding the lower bound in the memory cost of classically simulating two models of quantum computation based on Majorana fermions: TQC with Ising anyons and FLO. Here we briefly review these two models.

Majorana fermions are described by a set Majorana operators $\{m_i\}_{i=1}^{2n}$ that satisfy the relation $\{m_i, m_j\} = 2\delta_{i,j}$ and $m_i^\dagger = m_i$. The braiding of two Majorana fermions results in a unitary $B_{i,j} = \exp(-\frac{\pi}{4} m_i m_j)$ called *braid gate* that maps a Majorana operator to another as

$$B_{i,j} m_k B_{i,j}^\dagger = \begin{cases} m_k & \text{if } k \notin \{i,j\} \\ m_j & \text{if } k = i \\ -m_i & \text{if } k = j, \end{cases} \tag{15}$$

for $i < j$.

We can use the stabilizer formalism to characterize the set of quantum states that can the prepared by TQC formalism from a free state. The stabilizer group of the initialization states $|0\rangle^{\otimes n}$ is

$$\mathrm{Stab}(|0\rangle^{\otimes n}) = (X_{1,2}, X_{3,4}, \ldots, X_{2n-1,2n}), \tag{16}$$

where $X_{i,j} = -i m_i m_j$, for $i < j$, are the measurable observables with eigenvalues ± 1. Applying any sequence of unitary gates $B_{i,j}$ maps the state $|0\rangle^{\otimes n}$ to the state $|\psi\rangle$, which is equivalent to update the stabilizer group by some permutation $\pi \in S_{2n}$ on the indexes $[2n]$,

$$\mathrm{Stab}(|\psi\rangle) = (X_{\pi(1),\pi(2)}, \ldots, X_{\pi(2n-1),\pi(2n)}). \tag{17}$$

Thus, the set of all states that can be prepared by TQC operations is

$$\Sigma_{\mathrm{TQC}} = \{ |\psi\rangle : \mathrm{Stab}(|\psi\rangle) = (X_{\pi(1),\pi(2)}, \ldots, X_{\pi(2n-1),\pi(2n)}), \pi \in S_{2n} \} \tag{18}$$

Every state in TQC have a well-defined parity, $Q|\psi\rangle = \mathrm{sgn}(\pi)|\psi\rangle$, where $Q = (-i)^n \prod_{i=1}^{2n} m_i$ is the parity operator, and $\mathrm{sgn}(\pi)$ is the sign of the permutation π. Braid gates preserve the parity, consequentely, the set of states with positive parity, Σ_+ can be generated by apply sequences of braid gates on a positive parity state, e.g. $|0\rangle^{\otimes n}$, similarly for the negative parity set Σ_-.

The braid gates, $B_{i,j}$, are a special case of unitary transformations $U = \exp\left(\sum_{i,j} h_{i,j} m_i m_j / 2\right)$, where $h_{i,j}$ is a real antisymmetric matrix. These unitary transformations are used in the context of fermionic linear optics (FLO) and called *FLO gates* [9, 13]. The conjugate action of the FLO gates acts on Majorana operators as

$$m_k \mapsto U m_k U^\dagger = \sum_{j=1}^{2n} R_{k,j} m_j, \tag{19}$$

where by choosing an appropriate matrix h one can implement an arbitrary rotation $R \in \mathrm{SO}(2n)$. A state of the subtheory is a fermionic Gaussian states and is uniquely described by its correlation matrix $C(\rho)$, a $2n \times 2n$ real anti-symmetric matrix with elements $C_{i,j}(\rho) = \frac{i}{2}\mathrm{tr}(\rho[m_i, m_j])$. Applying an FLO gate in a fermionic Gaussian state is equivalent to mapping the correlation matrix as $C \mapsto RCR^T$. Thus, any free Gaussian state ρ can be brought to block-diagonal form by a real orthogonal transformation $R \in \mathrm{SO}(2n)$

$$C(\rho) = R \bigoplus_{i=1}^{n} \begin{pmatrix} 0 & \alpha_i \\ -\alpha_i & 0 \end{pmatrix} R^T, \tag{20}$$

for $\alpha_i = \pm 1$, $\forall j$.

F Contextuality for Majorana Fermions

Any TQC operation preserves the parity of the initial states. Consequentely, it is natural to consider the fixed-parity sector $q = \pm 1$. The set of states with fixed-parity q is defined as

$$\Sigma_q = \{ |\psi\rangle : |\psi\rangle \in \Sigma_{\mathrm{TQC}} \text{ and } Q|\psi\rangle = q|\psi\rangle \}. \tag{21}$$

Consequently, the local section over every maximum context $M \in \mathcal{M}_n$ is restricted to

$$\nu(M) = \{\nu(O) : O \in \mathcal{M}, \operatorname{sgn}(M) \prod_{O \in M} \nu(M) = q\}. \quad (22)$$

The local sections are independent of the state $|\psi\rangle \in \Sigma_q$ in which the context is measured.

Theorem 7 (NCVA for TQC $\Leftrightarrow$ Pfaffian graph) *There exists a NCVA for $\mathcal{O} \subseteq \mathcal{O}_{TQC}$ if and only if the graph that represent $\mathcal{O}$ is Pfaffian.*

Proof. If exists a NCVA for $\mathcal{O}$, then exists a compatible family of local sections for all $M \in \mathcal{M}_n$ and a global section s over $\mathcal{O}$, such that,

$$\lambda_s(\mathbb{P}_{\nu(M)}) = 1, \quad \forall M \in \mathcal{M}_n. \quad (23)$$

Let the global assignment s be the orientation of the digraph $D = ([2n], \overrightarrow{E})$, then $(i,j) \in \overrightarrow{E}$ if $s(X_{i,j}) = 1$ and $(j,i) \in \overrightarrow{E}$ if $s(X_{i,j}) = -1$. Since the local sections have the restriction given by Eq. 22, the weight of any perfect matchings in D is $\operatorname{sign}(M) = q$, which means the graph is Pfaffian.

Now, let us consider the Pfaffian graph G that represents the observables in $\mathcal{O}$. If G if Pfaffian, then there exists a Pfaffian orientation D for G, such that, all perfect matchings have the same weight $\operatorname{sign}(M) = w, \forall M \in \mathcal{M}_n$. Then, the Pfaffian orientation $\overrightarrow{E}$ is a global assignment of $\mathcal{O}$. □

G Application

This method of finding the memory required for the simulation can be applied to several contextual subtheories that are relevant to quantum computing. The first example that comes to mind is the qubit stabilizer. In Ref. [14], it was shown that for any subset of qubit stabilizer states S with cardinality larger than $2^{n^2/4+7n/2}$ contain three states such that their stabilizers form a Peres-Mermin magic square. Thus, the minimum number of classical bits required to simulate this subtheory scales quadratically with the number of qubits. This proves that the Gottesman-Knill algorithm [15] is asymptotically optimal for qubit stabilizers.

Here we apply our method to two models of quantum computation based on Majorana fermions, the TQC with Ising anyons and the FLO model.

For the TQC model, the set of states allowed in the subtheory, Σ_{TQC}, have cardinality $(2n)!/n!$. The lower bound in the memory cost for a classical simulation of this subtheory can be computed by find the largest subset of states with non-contextual stabilizer group, see Eq. 14 and lemma 5.

In F, TQC subtheory was shown to be contextual for $n > 2$ using graph theory. A graph G is Pfaffian is there exist a orientation in which all perfect matchings have the same sign, i.e.,

$$\operatorname{sgn}(M) = \operatorname{sgn}(\pi_M) \prod_{i=1}^{n} a_{\pi(2i-1),\pi(2i)} = q, \quad (24)$$

where $q = \pm 1$ and $A = (a_{i,j})$ is a skew symmetric matrix assigning ± 1 values simultaneously for every observable independently of the measurement context. This condition is equivalent to a global section that assigns value to each observable, such that, for every maximum context we have the constraint

$$s(M) := \{(O, s(O)) \,|\, \operatorname{sgn}(\pi_M) \prod_{O \in M} s(O) = q\}, \quad (25)$$

where q is the parity of the states in the set $\Sigma_q \subset \Sigma_{\text{TQC}}$. Consequently, a set of states with non-contextual stabilizer group must be represented by a Pfaffian graph. This is enough to formulate the following Lemma:

Lemma 8 (Memory cost for TQC) *A classical simulation of TQC requires at least $n \log(n)$ classical bits of memory.*

Proof. From Lemma 5, we know that the classical memory require is $\log(|\Lambda|) \leq \log(|\Sigma_{\text{TQC}}|) - \log(\alpha^*)$. For TQC, $|\Sigma_{\text{TQC}}| = (2n)!/n!$, and α^* can be upper bound as shown in Eq. 14 using Pfaffian graphs, as shown in Theorem 7.

Let $\mathcal{O} \subset \mathcal{O}_{\text{TQC}}$ be a non-contextual subset of observables and A_s be a skew-symmetric matrix representing a global section over α^*. Then, Pfaffian of the matrix A_s is $\operatorname{Pf}(A) = \sum_{M \in \mathcal{M}} \operatorname{sgn}(M) = |\mathcal{M}|$, where $\mathcal{M}$ is the set of perfect matchings in $\mathcal{O}$. Since $\det(A_s) = (\operatorname{Pf}(A_s))^2$, we can use the Hadamard's inequality[16, 17] to give an upper bound to the Pfaffian

$$\begin{aligned}(\operatorname{Pf}(A_s))^2 &= |\det(A_s)| \\ &\leq \left(\prod_{i=1}^{2n} \deg(m_i)\right)^{1/2} \\ &\leq \frac{1}{2}\sum_{i=1}^{2n} \deg(m_i) = |\mathcal{O}|, \end{aligned} \quad (26)$$

, where $\deg(m_i)$ is the number of observables in $\mathcal{O}$ containing m_i. Since the cardinality of $\mathcal{O}$ is at most $\binom{2n}{2}$, we obtain the inequality $|\operatorname{Pf}(A_s)| = |\mathcal{M}| \leq 2n$. Consequently, $\alpha^* \leq 2n$. □

TQC lies in the intersection of two computational modes: the Clifford group/stabilizer formalism model and the Fermionic Linear Optics (FLO)[7]. One can see the FLO model as a generalization of TQC, as a result, the FLO model is also contextual. In this section, we apply the method for computing the lower bound in the memory cost for the classical simulation of FLO.

The states allowed in the FLO subtheory are obtained by a continuous rotation in $\mathrm{SO}(2n)$, however for a classical simulation of the model is enough to ϵ-approximate the states, covering the space of the states with an ϵ-net. The ϵ-covering net approximate the states with accuracy

$$\|C(\psi) - C(\psi')\| \leq \epsilon \implies \|\psi - \psi'\|_1 \propto \epsilon, \quad (27)$$

where the ψ is the free state that we want to ϵ-approximate with ψ'. Since the volume of the ϵ-ball is ϵ^D, where D is the dimension of the space. The number of the ϵ-balls required to cover the space of free Gaussian states is ϵ^{-D}. Since $C(\psi)$ are $2n \times 2n$ antisymmetric real matrices, the dimension $D = n(2n-1)$. Hence, the cardinality of the set of ϵ-approximate states S_{approx} is

$$|S_{\text{approx}}| \propto \left(\frac{1}{\epsilon}\right)^{\theta(n^2)}. \tag{28}$$

Since the TQC subtheory is a subset of the FLO subtheory the same upper bound on α^*, $O(2^n)$, can be used. Accordingly, the lower bound in the size of the memory required is

$$|\Lambda| \geq \frac{1}{\epsilon^{n^2} 2^{n-2}}. \tag{29}$$

Consequently, the lower bound in the memory cost of classicaly simulating the FLO model is $\Omega(n^2 \log_2 \left(\frac{1}{\epsilon}\right))$.

Noise mitigation with a quantum autoencoder

Dominick J. Joch[1] * Markus Rambach[2] Kok-Wei Bong[1] Gerardo Paz Silva[1]
Jacquiline Romero[2] Nora Tischler[1]

[1] *Centre for Quantum Dynamics and Centre for Quantum Computation and Communication Technology, Griffith University, Brisbane, Queensland 4111, Australia*
[2] *School of Mathematics and Physics, University of Queensland, QLD 4072, Australia*

Abstract. Noise present in quantum states has a significant impact on their ability to be used as a quantum information resource for many applications. A quantum autoencoder is a quantum neural network, which uses machine learning to compress quantum data. It is established that classical autoencoders can denoise classical data and theoretical work predicts the same is true for quantum data. In our work, the aim is demonstrating an optical quantum autoencoder that denoises four-dimensional quantum states (ququarts), by compressing them into a two-dimensional representation and decoding them back to the original space.

Keywords: Autoencoder, quantum machine learning, denoising, quantum optics

1 Introduction

Autoencoders are feedforward (one where the inputs to one layer all come from the previous layer) neural networks that output their input at the final layer. Their topology is typically such that they have inner bottleneck layers. Via machine learning techniques they are trained to compress input data into a smaller latent space and discarding superfluous information[1]. This mapping can be reversed to obtain the input data—so-called lossless compression.

The autoencoder is originally a classical tool, using machine learning to reduce the dimension of classical data. The same concept can be applied to quantum information with a quantum autoencoder (QAE), which falls under quantum machine learning[2]. A QAE can learn in an unsupervised manner to compress quantum information in a space with $N + K$ dimensions onto a smaller latent space of N dimensions. Thereby quantum information can be stored more efficiently in fewer resources, and it also enables one to perform noise mitigation, also called denoising.

Recent theoretical works [3], explore QAEs with different topologies, which were simulated to demonstrate mitigation of unitary noise, dephasing noise and random bit-flip errors in GHZ, W, Dicke and cluster states. It is further shown that these quantum autoencoders can denoise input states both within and outside the training data. Identifying the contributing sources of noise present in experimental preparations of quantum states is often a difficult task. By employing quantum machine learning techniques, this task can be automated to correct for noise without the need for any prior knowledge of its nature.

In this work we aim to experimentally demonstrate this novel approach to quantum noise mitigation via quantum machine learning. Prior experimental works with single photons have demonstrated encoding of two two-qubit states into two single-qubit states[2], and the encoding of qutrits to qubits[4]. As yet, the complete QAE, including the decoding step, has yet to be demonstrated in the photonic platform, as has the use of a QAE for denoising.

2 Experiment

Our QAE has a [4,2,4] topology—dimension reduction from ququarts to qubits—and we use single photons to encode quantum information. The states we use are qubit states embedded into a ququart space. These ququart states thus occupy all four modes but by design they are in a subspace that allows them to be compressed by the QAE to qubits. We choose the input states such that in the absence of noise, the QAE can be trained to implement lossless compression and decompression where the original states are recovered.

The encoding unitary is trained via a classical algorithm to minimize the probability of the photon occupying the optical modes corresponding to the junk space. We begin with a set of target training states that are corrupted by noise to obtain the input states. A parameterized unitary operator

*dominick.joch@griffithuni.edu.au

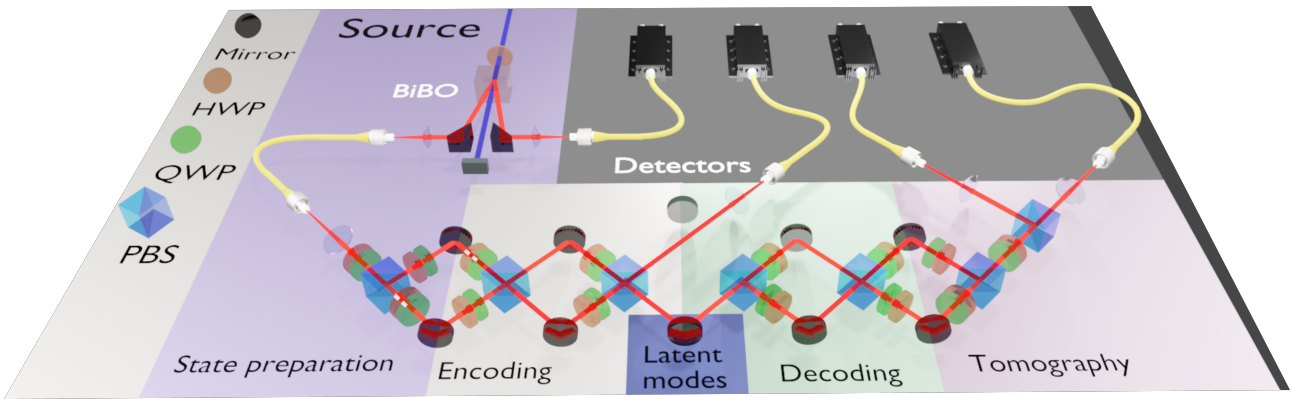


Figure 1: The experimental realization of this quantum autoencoder consists of four parts: state preparation, Encoder, Decoder and output measurement. The first pairs of waveplates and polarizing beamsplitter are used to prepare the ququart states in both the polarization and path degrees of freedom of the single photons. The next two beam splitters and four pairs of waveplates implement the parametrized encoding unitary. The decoding unitary is simply the inverse of the encoder and consists of the same optical components but in reverse order. The measurement section consists of the final interferometer and wavelates, a waveplate pair, polarizing beamsplitter and a pair of avalanche photodiode detectors.

$U(\theta_1, \theta_2, ..., \theta_n)$, is iteratively optimized to find the optimal set of parameters $\vec{\theta}$—those which minimize the occupation of the junk modes—using Adam gradient descent in our work.

In the ideal noiseless case, a set of states are encoded such that they all only occupy the qubit subspace that is the latent space, and in the decoding reverses that encoding unitary to produce the original states (lossless compression). If we consider the noisy case, when the decoder attempts to reconstruct the original input data from the compressed data, the component of the noise present in the orthogonal junk space is not reintroduced, but is instead discarded. The output state will have less noise than the input, and so a fidelity improvement with the noisless target states should be observed. For validation purposes and to test the noise mitigation properties, the trained QAE will be applied to a set of noisy test states where a significant fidelity improvement can be achieved. The QAE approach is also more generally applicable to other dimensionalities and can be extended to other physical platforms.

The experimental realization of this quantum autoencoder is shown in Fig. 1, which is conceptually equivalent to the actual setup. We encode quantum information in the polarization and path degrees of freedom of 808 nm single photons produced by by type-I spontaneous parametric downconversion. The measurement part of the setup is used to perform quantum state tomography of the ququart states. Simulations of this experimental setup show that denoising can be achieved even when imperfections are accounted for. Preliminary results indicate that the training algorithm can succesfully minimize the cost function and output states can be minimally impacted by imperfections at the output tomography. The ongoing work is to now demonstrate the fidelity improvement of the output states when noise is introduced at the input.

3 Outlook

Quantum machine learning is a field drawing increasing interest due to the power of machine learning techniques that have already been shown in classical machine learning. Though quantum encoders are still in an early stage of study and development with few experimental implementations, they are a very promising tool for quantum information processing in much the same way as their classical counterparts are for data compression and denoising. Our goal is to construct an experimental implementation of a complete quantum autoencoder and then apply it to noise corrupted input states to demonstrate the noise mitigation capabilities. This would be an important step in developing automated quantum data compression and denoising for more efficient use of quantum resources.

References

[1] J. Romero et. al. Quantum Sci. Technol. 2, 045001 (2017)

[2] Huang, Chang-Jiang and Ma, Hailan and Yin, Qi and Tang, Jun-Feng and Dong, Daoyi and Chen, Chunlin and Xiang, Guo-Yong and Li, Chuan-Feng and Guo, Guang-Can Realization of a quantum autoencoder for lossless compression of quantum data Phys. Rev. A 102, 032412 (2020)

[3] D. Bondarenko and P. Feldmann Phys. Rev. Lett. 124, 130502 (2020)

[4] A. Pepper, N. Tischler, and G. J. Pryde Phys. Rev. Lett. 122, 060501 (2019)

Uncertainty Relations in Pre- and Post-Selected Systems

Sahil,[1,2,*] Sohail,[3,†] and Sibasish Ghosh[1,2,‡]

[1]*Optics and Quantum Information Group, The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113, India*
[2]*Homi Bhabha National Institute, Training School Complex, Anushakti Nagar, Mumbai 400085, India*
[3]*Quantum Information & Computation Group, Harish-Chandra Research Institute, A CI of Homi Bhabha national Institute, Chhatnag Road, Jhunsi, Prayagraj - 211019, India.*

In this work, we derive Robertson-Heisenberg like uncertainty relation for two incompatible observables in a pre- and post-selected (PPS) system. The newly defined standard deviation and the uncertainty relation in the PPS system have physical meanings which we present here. We demonstrate two unusual properties in the PPS system using our uncertainty relation. First, for commuting observables, the lower bound of the uncertainty relation in the PPS system does not become zero even if the initially prepared state i.e., pre- selection is the eigenstate of both the observables when specific post-selections are considered. This implies that for such case, two commuting observables can disturb each other's measurement results which is in fully contrast with the Robertson-Heisenberg uncertainty relation. Secondly, unlike the standard quantum system, the PPS system makes it feasible to prepare sharply a quantum state (pre- selection) for non-commuting observables. Some applications of uncertainty and uncertainty relation in the PPS system are provided: (i) detection of mixedness of an unknown state, (ii) stronger uncertainty relation in the standard quantum system, (iii) "purely quantum uncertainty relation" that is, the uncertainty relation which is not affected (i.e., neither increasing nor decreasing) under the classical mixing of quantum states, (iv) state dependent tighter uncertainty relation in the standard quantum system, and (v) tighter upper bound for the out-of-time-order correlation function.

I. INTRODUCTION

The uncertainty relation, which Heisenberg discovered, is one of the most well-known scientific findings [1, 2]. It asserts that it is impossible to accurately measure the position and the momentum of a particle. In other words, measuring the position of a particle always affects the momentum of that particle and vice versa. Robertson developed the uncertainty relation known as "*Robertson-Heisenberg Uncertainty Relation*" (RHUR) [3] in the very later years to describe the difficulty of jointly sharp preparation of a quantum state [4] for incompatible observables. This relation not only limits the joint sharp preparation for non-commuting observables but also proved it's usefulness: to formulate quantum mechanics [5, 6], for entanglement detection [7, 8], for the security analysis of quantum key distribution in quantum cryptography [9], as a fundamental building block for quantum mechanics and quantum gravity [10], etc.

On the one side, we have the standard quantum systems where the RHUR hold while pre- and post-selected (PPS) systems, on the other side, are different kind of quantum mechanical systems that were developed by Aharonov, Bergmann, and Lebowitz (ABL) [11–13] to address the issue of temporal asymmetry in quantum mechanics. Recently, in the references [14, 15], the authors generalized the probabilities of obtaining the measurement results of an observable in a PPS system given by ABL [11].

In the later years, Aharonov, Albert, and Vaidman (AAV) [16] introduced the notion of "weak value" defined as

$$\langle A_w\rangle_\psi^\phi = \frac{\langle\phi|A|\psi\rangle}{\langle\phi|\psi\rangle}, \tag{1}$$

in a pre- and post-selected system when the observable 'A' is measured weakly. Here, $|\psi\rangle$ and $|\phi\rangle$ are pre- and post-selected states, respectively. Weak values have strange features for being complex and its real part can lie outside the max-min range of the eigenvalues of the operator of interest when the pre- and post-selections are nearly orthogonal.

In order to obtain the real and imaginary parts of the weak value of A [17, 18], first the system of interest and a pointer (ancilla) is prepared in the product state $|\psi\rangle\otimes|\xi\rangle$. Then the system-pointer is evolved under the global unitary $U = exp(-iHt)$, where $H = gA\otimes P_x$ is the von Nuemann Hamiltonian, A is the measurement operator of the system, P_x is the pointer's momentum observable, 'g' is the coupling coefficient between system and pointer and 't' is the interaction time. Now after the time evolution of the system-pointer, the system is projected to $|\phi\rangle$ and as a result, the state of the pointer collapses to the unnormalized state $|\widetilde{\xi_\phi}\rangle \approx (1 - igt\,\langle A_w\rangle_\psi^\phi P_x)\,|\xi\rangle$ in the limit $g \ll 1$, i.e., weak interaction. Now, it can be shown that the average position and momentum shifts of the pointer in state $|\xi_\phi\rangle = \frac{|\widetilde{\xi_\phi}\rangle}{\sqrt{\langle\widetilde{\xi_\phi}|\widetilde{\xi_\phi}\rangle}}$ are $\langle X\rangle_{\xi_\phi} = gt\,\mathcal{R}e(\langle A_w\rangle_\psi^\phi)$ and $\langle P_x\rangle_{\xi_\phi} = \frac{gt}{2\sigma^2}\,\mathcal{I}m(\langle A_w\rangle_\psi^\phi)$, respectively with the Gaussian pointer $\langle x|\xi\rangle = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^{1/2} e^{-x^2/4\sigma^2}$, σ is the rms width of the position distribution $|\langle x|\xi\rangle|^2$ of the pointer and thus providing the full knowledge of the weak value of A.

Recently, a lot of attention was paid to these aspects [19–40]. The measurements involving weak values are known as "Weak Measurements" or "weak PPS measurements". Since it depends on probabilistic post-selection $|\phi\rangle$, a weak value

* sahilmd@imsc.res.in
† sohail@hri.res.in
‡ sibasish@imsc.res.in

can be thought of as conditional expectation value. Moreover, when the post-selection is same as pre- selection i.e., $|\phi\rangle = |\psi\rangle$, it becomes

$$\langle A\rangle_{\psi} = \langle\psi|A|\psi\rangle\,, \tag{2}$$

the expectation value in the standard quantum system. The PPS systems can therefore be thought of as being more general than the so-called standard quantum systems.

As pre- and post-selected systems are already useful practically as well as fundamentally, then an immediate question can be asked whether there exists any uncertainty relation like the RHUR which can give the limitations on joint sharp preparation of the given pre- and post-selected states when non-commuting observables are measured.

In the present study, we demonstrate the existence of such uncertainty relations in PPS systems, which are expected as PPS systems are more generalised versions of standard ones. We first define the standard deviation of an observable in the PPS system for the given pre- and post-selections with geometrical as well as physical interpretations. After that, we derive our main result of this paper "*uncertainty relations in pre- and post-selected systems*" using the well known Cauchy-Schwarz inequality.

We provide the following physical applications of our results: (i) detection of the purity of an unknown state of any quantum systems (e.g., qubit, qutrit, two qubit, qutrit-qubit, etc) using two different definitions of the uncertainty of an observable in the PPS system, (ii) stronger uncertainty relation in the standard quantum system (i.e., the uncertainty relation that can not be made trivial or the lower bound can not be made zero for almost all possible choices of initially prepared systems) using the uncertainty relation in the PPS system, (iii) purely quantum uncertainty relations that is, the uncertainty relations which are not affected (i.e., neither increasing nor decreasing) under the classical mixing of quantum states using the uncertainty relations in PPS systems. (iv) state dependent tighter uncertainty relation in the standard system by introducing the idea of post-selection, and finally (v) tighter upper bound for the out-of-time-order correlation function. Moreover, as the RHUR has a plenty of applications, uncertainty relation in the PPS systems can also be applied in quantum optics, information, technologies, etc.

This paper is organized as follows. In sec. II, we discuss uncertainty relations in standard quantum systems. In sec. III, we derive our main results of this paper. Application of our results are given in sec. IV and finally we conclude our work in sec. V.

II. UNCERTAINTY RELATION IN STANDARD QUANTUM SYSTEM

In this section, we first interpret the standard deviation of an observable in standard quantum systems from geometrical as well as information-theoretic perspective. For establishing the standard deviation in a PPS system, we will introduce a similar interpretation. The RHUR's well-known interpretation is also provided here.

A. Standard deviation

We consider the system Hilbert space to be $\mathcal{H}$ and let $|\psi\rangle$ be a state vector in $\mathcal{H}$. Due to the probabilistic nature of the measurement outcomes of the observable A, the uncertainty in the measurement is defined as the standard deviation:

$$\langle\Delta A\rangle_{\psi} = \sqrt{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2}. \tag{3}$$

Geometric interpretation.— Standard deviation can be given a geometrical interpretation using the following Proposition.

Proposition 1. *If $|\psi\rangle \in \mathcal{H}$ is an initially prepared state of a standard quantum system and $A \in \mathcal{L}(\mathcal{H})$ is a hermitian operator, then we can decompose $A|\psi\rangle \in \mathcal{H}$ as*

$$A|\psi\rangle = \langle A\rangle_{\psi}|\psi\rangle + \langle\Delta A\rangle_{\psi}|\psi_A^{\perp}\rangle\,, \tag{4}$$

where $|\psi_A^{\perp}\rangle = \frac{1}{\langle\Delta A\rangle_{\psi}}(A - \langle A\rangle_{\psi})|\psi\rangle$, and $\langle A\rangle_{\psi} = \langle\psi|A|\psi\rangle$. Eq. (4) is sometimes known as the "Aharonov–Vaidman identity" [41].

Proof. Let $A|\psi\rangle$ and $|\psi\rangle$ are two non-orthogonal state vectors. Using Gram-Schmidt orthonormalization process, we find the unnormalized state vector $|\widetilde{\psi}_A^{\perp}\rangle \in \mathcal{H}$ orthogonal to $|\psi\rangle$ as

$$|\widetilde{\psi}_A^{\perp}\rangle = A|\psi\rangle - \frac{(\langle\psi|A)|\psi\rangle}{\langle\psi|\psi\rangle}|\psi\rangle = (A - \langle A\rangle_{\psi})|\psi\rangle\,, \tag{5}$$

where $\langle\psi|\psi\rangle = 1$ and after normalization, Eq. (5) becomes

$$A|\psi\rangle = \langle A\rangle_{\psi}|\psi\rangle + \langle\Delta A\rangle_{\psi}|\psi_A^{\perp}\rangle\,, \tag{6}$$

where $|\psi_A^{\perp}\rangle = |\widetilde{\psi}_A^{\perp}\rangle/\sqrt{\langle\widetilde{\psi}_A^{\perp}|\widetilde{\psi}_A^{\perp}\rangle}$ and $\sqrt{\langle\widetilde{\psi}_A^{\perp}|\widetilde{\psi}_A^{\perp}\rangle} = \sqrt{\langle\psi|A^2|\psi\rangle - \langle\psi|A|\psi\rangle^2} = \langle\Delta A\rangle_{\psi}$ and $\langle A\rangle_{\psi} = \langle\psi|A|\psi\rangle$. □

So, Eq. (4) can be interpreted as the unnormalized state vector $A|\psi\rangle$ which has two components and these are $\langle A\rangle_{\psi}$ along $|\psi\rangle$ and $\langle\Delta A\rangle_{\psi}$ along $|\psi_A^{\perp}\rangle$. Here we interpret $\langle\Delta A\rangle_{\psi}$ as disturbance of the state vector due to the measurement of the operator A or as the measurement error (or standard deviation) of that operator when the system is prepared in the state $|\psi\rangle$. For instance, if we set up the system in one of the eigenstates of the observable A, then from Eq. (4), it can be seen that the standard deviation of A is zero.

Information-theoretic interpretation.— From an information-theoretic approach, Eq. (3) can be written as

$$\langle\Delta A\rangle_{\psi} = \sqrt{\sum_{i=1}^{d-1}\left|\langle\psi_i^{\perp}|A|\psi\rangle\right|^2}, \tag{7}$$

where $\{|\psi\rangle, |\psi_1^\perp\rangle, |\psi_2^\perp\rangle, \cdots, |\psi_{d-1}^\perp\rangle\}$ forms an orthonormal basis such that $I = |\psi\rangle\langle\psi| + \sum_{i=1}^{d-1} |\psi_i^\perp\rangle\langle\psi_i^\perp|$ and 'd' is the dimension of the system. So, the origin of the non-zero standard deviation $\langle\Delta A\rangle_\psi$ in the standard quantum system can also be thought of due to the non-zero contributions of the unnormalized fidelities $\{|\langle\psi_i^\perp|A|\psi\rangle|\}_{i=1}^{d-1}$ which can be viewed as the spread of the information of the observable A along $\{|\psi_i^\perp\rangle\}_{i=1}^{d-1}$ directions.

B. RHUR

The well known RHUR for two non-commuting operators A and B on a Hilbert space $\mathcal{H}$ when the system is prepared in the state $|\psi\rangle$ is given by

$$\langle\Delta A\rangle_\psi^2 \langle\Delta B\rangle_\psi^2 \geq \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle\right]^2, \tag{8}$$

where $\langle\Delta A\rangle_\psi$ and $\langle\Delta B\rangle_\psi$ are the standard deviations of the operators A and B, respectively, and $[A,B] = AB - BA$ is the commutator of A and B. The derivation of Eq. (8) using the *Aharonov–Vaidman identity* can be found in [41, 42]. The stronger version is obtained by adding the "Schrödinger's term" in Eq. (8) as

$$\langle\Delta A\rangle_\psi^2 \langle\Delta B\rangle_\psi^2 \geq \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle\right]^2 + \left[\frac{1}{2}\langle\psi|\{A,B\}|\psi\rangle - \langle A\rangle_\psi\langle B\rangle_\psi\right]^2. \tag{9}$$

The RHUR is usually interpreted as the following: it puts bound on the sharp preparation of a quantum state for two non-commuting observables. Hence, a quantum state in which the standard deviations of the two non-commuting observables are both zero cannot exist.

III. MAIN RESULTS

The idea of the standard deviation or information dispersion (see preceding section) is a crucial component of the theory in a preparation-measurement situation. Pre- and post-selected systems are typical examples, therefore we define the standard deviation (uncertainty) of an observable and show that for such systems, there exist RHUR-like uncertainty relations for two non-commuting observables.

A. Standard deviation in PPS system

Geometric definition.— It is well known that when the preselection and the post-selection are same, the PPS system becomes the standard quantum system (see introduction). The following proposition generalizes Eq. (4) for the PPS system.

Proposition 2. *If a PPS system is in a pre- selected state $|\psi\rangle$ and post-selected state $|\phi\rangle$, then for a hermitian operator $A \in \mathcal{L}(\mathcal{H})$, we can decompose $A|\psi\rangle$ as*

$$A|\psi\rangle = \langle\phi|A|\psi\rangle|\phi\rangle + \langle\Delta A\rangle_\psi^\phi |\phi_{A\psi}^\perp\rangle, \tag{10}$$

where

$$\langle\Delta A\rangle_\psi^\phi = \sqrt{\langle\psi|A^2|\psi\rangle - |\langle\phi|A|\psi\rangle|^2} \tag{11}$$

and $|\phi_{A\psi}^\perp\rangle = \frac{1}{\langle\Delta A\rangle_\psi^\phi}(A|\psi\rangle - \langle\phi|A|\psi\rangle|\phi\rangle)$, a normalized state vector which is orthogonal to $|\phi\rangle$.

Proof. We assume that $A|\psi\rangle$ and $|\phi\rangle$ are two non-orthogonal state vectors. The unnormalized state vector $|\widetilde{\phi}_{A\psi}^\perp\rangle \in \mathcal{H}$ which is orthogonal to $|\phi\rangle$ is obtained using Gram-Schmidt orthonormalization process as

$$|\widetilde{\phi}_{A\psi}^\perp\rangle = A|\psi\rangle - \frac{\langle\phi|(A|\psi\rangle)}{\langle\phi|\phi\rangle}|\phi\rangle = A|\psi\rangle - \langle\phi|A|\psi\rangle|\phi\rangle, \tag{12}$$

where $\langle\phi|\phi\rangle = 1$ and after normalization, Eq. (12) becomes

$$A|\psi\rangle = \langle\phi|A|\psi\rangle|\phi\rangle + \langle\Delta A\rangle_\psi^\phi |\phi_{A\psi}^\perp\rangle,$$

where $|\phi_{A\psi}^\perp\rangle = |\widetilde{\phi}_{A\psi}^\perp\rangle / \sqrt{\langle\widetilde{\phi}_{A\psi}^\perp|\widetilde{\phi}_{A\psi}^\perp\rangle}$ and $\sqrt{\langle\widetilde{\phi}_{A\psi}^\perp|\widetilde{\phi}_{A\psi}^\perp\rangle} = \sqrt{\langle\psi|A^2|\psi\rangle - |\langle\phi|A|\psi\rangle|^2} = \langle\Delta A\rangle_\psi^\phi$. □

To define the standard deviation of the observable A in the PPS system, we now present an argument which is similar to the one used to describe the standard deviation of an observable in a standard quantum system. So, Eq. (10) can be interpreted geometrically as the unnormalized state vector $A|\psi\rangle$ which has two components $\langle\phi|A|\psi\rangle$ along the post-selection $|\phi\rangle$ and $\langle\Delta A\rangle_\psi^\phi$ along $|\phi_{A\psi}^\perp\rangle$. *Here we define $\langle\Delta A\rangle_\psi^\phi$ as the standard deviation of the observable A when the system is preselected in $|\psi\rangle$ and post-selected in $|\phi\rangle$.*

The standard deviation $\langle\Delta A\rangle_\psi^\phi$ can be realized via the weak value of the observable A as

$$\langle\Delta A\rangle_\psi^\phi = \sqrt{\langle\psi|A^2|\psi\rangle - |\langle A_w\rangle_\psi^\phi|^2|\langle\phi|\psi\rangle|^2} \tag{13}$$

$$= \sqrt{\langle\Delta A\rangle_\psi^2 + \langle A\rangle_\psi^2 - |\langle A_w\rangle_\psi^\phi|^2|\langle\phi|\psi\rangle|^2}, \tag{14}$$

where $\langle A_w\rangle_\psi^\phi$ is the weak value of the observable A defined in Eq. (1) and we have used Eq. (3) to derive Eq. (14). $|\langle\phi|\psi\rangle|^2$ is the success probability of the post-selection $|\phi\rangle$. Eq. (13) is no longer a valid expression if pre- and post-selected states are orthogonal to one another because in this situation, weak value is not defined. Then, go back to Eq. (11). It should be noted that Eq. (11) holds true whether the measurement is strong or weak.

Information-theoretic definition.— Another expression of the standard deviation $\langle\Delta A\rangle_\psi^\phi$ in the PPS system can be derived by inserting an identity operator $I = |\phi\rangle\langle\phi| + \sum_{i=1}^{d-1} |\phi_i^\perp\rangle\langle\phi_i^\perp|$, where $\{|\phi\rangle, |\phi_1^\perp\rangle, |\phi_2^\perp\rangle, \cdots, |\phi_{d-1}^\perp\rangle\}$ forms an orthonormal basis,

in the first term of the right hand side of Eq. (13) as

$$\langle \Delta A\rangle_{\psi}^{\phi}=\sqrt{\sum_{i=1}^{d-1}\left|\langle A_w\rangle_{\psi}^{\phi_i^{\perp}}\right|^2|\langle\phi_i^{\perp}|\psi\rangle|^2}. \tag{15}$$

From an information-theoretic perspective, Eq. (15) may now be understood as follows: non-zero standard deviation in the PPS system arises as a result of the non-zero contributions from the weak values $\{\langle A_w\rangle_{\psi}^{\phi_i^{\perp}}\}_{i=1}^{d-1}$ along the orthogonal post-selections $\{|\phi_i^{\perp}\rangle\}_{i=1}^{d-1}$. Note that two consecutive measurements are taken into account in a PPS system: the operator of interest A and the projection operator $\Pi_{\phi}=|\phi\rangle\langle\phi|$ which corresponds to the post-selection $|\phi\rangle$. As a result, it is hard to tell whether or not A has been measured when the weak value is zero. Because of this, it is crucial to have non-zero weak values which carry the information about the observable A. Null weak values have recently been given a useful interpretation [43]: if a successful post-selection occurs with a null weak value, then the property represented by the observable A cannot be detected by the weakly coupled quantum pointer. In other words, the pointer state remains unchanged when the weak value is zero (see introduction section). Thus, one should anticipate that the standard deviation in the PPS system should be zero if we obtain null weak values for the post-selections $\{|\phi_i^{\perp}\rangle\}_{i=1}^{d-1}$, that means the information about the observable A is not dispersed throughout the post-selections $\{|\phi_i^{\perp}\rangle\}_{i=1}^{d-1}$.

In addition to the standard deviation's geometrical and information-theoretical explanations (Eqs. (10) and (15), respectively) in the PPS system, we now study the minimum (zero) and maximum uncertainty (or standard deviation) which provide additional insights to understand the standard deviation.

Zero uncertainty.— The uncertainty $\langle \Delta A\rangle_{\psi}^{\phi}$ defined in Eq. (11) in the PPS system is zero if and only if

$$A|\psi\rangle=\langle\phi_z|A|\psi\rangle|\phi_z\rangle, \tag{16}$$

or $|\phi_z\rangle\propto A|\psi\rangle$. We have used the notation $|\phi_z\rangle$ as the post-selection for which uncertainty in PPS system becomes zero. The zero uncertainty in the PPS system can now be realised in the following way: the weak value $\langle A_w\rangle_{\psi}^{\phi_z}$ becomes non-zero i.e., $\frac{\langle\psi|A^2|\psi\rangle}{\langle\psi|A|\psi\rangle}\neq 0$ when we post-select the system to $|\phi_z\rangle$, and the weak values for all post-selections $\{|\phi_{z\,i}{}^{\perp}\rangle\}_{i=1}^{d-1}$ orthogonal to $|\phi_z\rangle$ are zero. As a result, the right side of Eq. (15) is reduced to zero. It should be noted that all post-selections orthogonal to $|\phi_z\rangle$ are "legitimate post-selections," meaning that their weak values are clearly specified. Equivalently, we can state that the information about the observable A is not dispersed along the post-selections $\{|\phi_{z\,i}{}^{\perp}\rangle\}_{i=1}^{d-1}$ as null weak values do not carry informations about the observable A (according to the above information-theoretic definition). Hence, it is guaranteed that in a particular direction there will be one and only one non-zero weak value of A in a PPS system if and only if the condition (16) is met.

Usefulness of zero uncertainty state: In this paragraph we provide the following usefulness of the zero uncertainty post-selected state $|\phi_z\rangle$.

1) In a parameter estimation scenario, where the task is to obtain the precision limit in the estimation of interaction coefficient 'g' in the interaction Hamiltonian $H=gA\otimes p$ ('p' is the pointer's momentum variable), Fisher information plays an important role whose maximum value is given by $F^{max}(g)=4\Delta^2\langle\psi|A^2|\psi\rangle$, where Δ is the standard deviation of initial distribution of the pointer state and $|\psi\rangle$ is the initially prepared state of the system [44]. In an arbitrarily post-selected state $|\phi\rangle$, Fisher information is given by $F_{\phi}(g)=4\Delta^2|\langle\phi|A|\psi\rangle|^2\leq F^{max}(g)$ [44]. Recently it was shown that in a generalized PPS system with negative quasiprobability distribution of an arbitrary quantum state, this limit can be violated [45]. Violation of such limit implies that error that occurs in estimating the unknown parameter can be reduced significantly using negative quasiprobability distribution based Fisher information compared to the usual quantum Fisher information. One can immediately see using Eq. (11) that $F_{\phi}(g)=4\Delta^2[\langle\psi|A^2|\psi\rangle-(\langle\Delta A\rangle_{\psi}^{\phi})^2]$. Now it is obvious that for the zero uncertainty post-selected state $|\phi_z\rangle$ as appeared in Eq. (16), we have $F_{\phi_z}(g)=F^{max}(g)$. Hence, to achieve the maximum Fisher information $F^{max}(g)$ in the PPS system, one must post-select the system in $|\phi_z\rangle=A|\psi\rangle/\sqrt{\langle\psi|A^2|\psi\rangle}$ which corresponds to the zero uncertainty.

2) The post-selection $|\phi_z\rangle$ alone has the ability to provide the information (e.g., $\langle\Delta A\rangle_{\psi}$ and $\langle A\rangle_{\psi}$) about the observable A. Indeed by noting that $\langle\psi|A^2|\psi\rangle=p_z(\langle A_w\rangle_{\psi}^{\phi_z})^2$ and $\langle\psi|A|\psi\rangle=p_z\langle A_w\rangle_{\psi}^{\phi_z}$, we have $\langle\Delta A\rangle_{\psi}^2=(1-p_z)p_z(\langle A_w\rangle_{\psi}^{\phi_z})^2$, where $p_z=|\langle\phi_z|\psi\rangle|^2$ is the probability of obtaining the post-selection $|\phi_z\rangle=A|\psi\rangle/\sqrt{\langle\psi|A^2|\psi\rangle}$.

Maximum uncertainty.— To achieve the maximum value of $\langle\Delta A\rangle_{\psi}^{\phi}$, the weak value $\langle A_w\rangle_{\psi}^{\phi}$ in Eq. (13) has to be zero i.e., when the post-selection $|\phi\rangle$ is orthogonal to $A|\psi\rangle$ and hence $\max(\langle\Delta A\rangle_{\psi}^{\phi})=\sqrt{\langle\psi|A^2|\psi\rangle}$. Note that, in a preparation-measurement scenario, maximum measurement error is also found to be $\sqrt{\langle\psi|A^2|\psi\rangle}$ whether the measurement of the observable A is performed in standard system (see Eq. (3)) or while performing the best estimation the operator A from the measurement of another hermitian operator [46].

B. Uncertainty relation in PPS system

After defining the standard deviation of an observable in a PPS system, interpreting it geometrically and informationally, and maintaining a parallel comparison and connection with the standard deviation in the standard system, we are now in a position to provide an uncertainty relation in a PPS system for two incompatible observables. One can formulate many different types of uncertainty relations in PPS systems (for example, [47]), but our interpretation of an uncertainty relation in a PPS system is based on the standard deviation defined in Eq. (11) or (13). Since the weak value of the observable A in the standard deviation Eq. (13) in the PPS system replaces the average value of the same observable

A in the standard deviation Eq. (3) in standard quantum system, it is not surprising that the mathematical expression of the uncertainty relation in the PPS system is a modified version of the RHUR (8), where the average values of the incompatible observables A and B in Eq. (8) will be replaced by the weak values of the respective observables when the system is pre-selected in $|\psi\rangle$ and post-selected in $|\phi\rangle$. The explicit form of the uncertainty relation in the PPS system is provided in the following theorem.

Theorem 1. *Let $A, B \in \mathcal{L}(\mathcal{H})$ be two non-commuting hermitian operators which are measured in the PPS system of our interest with $|\psi\rangle$ and $|\phi\rangle$ being pre- and post-selected states, respectively, then the product of their standard deviations satisfies*

$$\left(\langle \Delta A\rangle_{\psi}^{\phi}\right)^2\left(\langle \Delta B\rangle_{\psi}^{\phi}\right)^2 \geq \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle - \mathcal{I}m\left(W_{AB}\right)\right]^2, \tag{17}$$

where $W_{AB} = \langle\psi|A|\phi\rangle\langle\phi|B|\psi\rangle = \left(\langle A_w\rangle_{\psi}^{\phi}\right)^*\langle B_w\rangle_{\psi}^{\phi}\,|\langle\phi|\psi\rangle|^2$ (using the definition of the weak value defined in Eq. (1)).

Proof. Cauchy-Schwarz inequality for two unnormalized state vectors $|\widetilde{\phi}_{A\psi}^{\perp}\rangle$ and $|\widetilde{\phi}_{B\psi}^{\perp}\rangle$ in $\mathcal{H}$ becomes

$$\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{A\psi}^{\perp}\rangle\langle\widetilde{\phi}_{B\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle \geq \left|\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle\right|^2, \tag{18}$$

Now, as $\left|\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle\right|^2 = [\mathcal{R}e(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle)]^2 + [\mathcal{I}m(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle)]^2$ and hence

$$\left|\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle\right|^2 \geq [\mathcal{I}m(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle)]^2, \tag{19}$$

where $\mathcal{I}m(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle) = \frac{1}{2i}\left(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle - \langle\widetilde{\phi}_{B\psi}^{\perp}|\widetilde{\phi}_{A\psi}^{\perp}\rangle\right)$ and $\mathcal{R}e(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle) = \frac{1}{2}(\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle + \langle\widetilde{\phi}_{B\psi}^{\perp}|\widetilde{\phi}_{A\psi}^{\perp}\rangle)$. Now put $|\widetilde{\phi}_{A\psi}^{\perp}\rangle = A|\psi\rangle - \langle\phi|A|\psi\rangle|\phi\rangle$ defined in Eq. (12) for operator A and similarly for operator B also, then we have

$$\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle = \langle\psi|AB|\psi\rangle - \langle\psi|A|\phi\rangle\langle\phi|B|\psi\rangle\,. \tag{20}$$

Note that $\langle\widetilde{\phi}_{A\psi}^{\perp}|\widetilde{\phi}_{A\psi}^{\perp}\rangle = \left(\langle\Delta A\rangle_{\psi}^{\phi}\right)^2$ is square of the standard deviation of the observable A in the PPS system defined in Eq. (11) and similarly $\langle\widetilde{\phi}_{B\psi}^{\perp}|\widetilde{\phi}_{B\psi}^{\perp}\rangle = \left(\langle\Delta B\rangle_{\psi}^{\phi}\right)^2$ is square of the standard deviation of the observable B in the PPS system. Finally, putting these values and using Eqs. (19) and (20) in Eq. (18), it becomes Eq. (17). □

Eq. (17) is always true for any strong PPS systems [11–13] or weak PPS systems [16]. For weak PPS measurements [16], W_{AB} is expressed in terms of weak values of both the observables. If the pre- and post-selected states are same i.e., $|\phi\rangle = |\psi\rangle$, then one gets back the RHUR (8) as argued before. Eq. (17) with "Schrödinger's term" becomes

$$\left(\langle\Delta A\rangle_{\psi}^{\phi}\right)^2\left(\langle\Delta B\rangle_{\psi}^{\phi}\right)^2 \geq \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle - \mathcal{I}m\left(W_{AB}\right)\right]^2 + \left[\frac{1}{2}\langle\psi|\{A,B\}|\psi\rangle - \mathcal{R}e\left(W_{AB}\right)\right]^2. \tag{21}$$

The uncertainty relation (17) can be interpreted in the same way as we did for the RHUR (8). That is, it bounds the sharp preparation of the pair for pre- and post-selections $(|\psi\rangle, |\phi\rangle)$ for two non-commuting observables. The lower bound contains an additional term $\mathcal{I}m(W_{AB})$ compared to the RHUR (8). So even if $[A,B] \neq 0$, the bound on the right hand side of Eq. (17) can become zero implying the possibility of both the standard deviations being zero implying further the possibility of sharp preparation of a pair of pre- and post-selected states. Below, we provide the necessary and sufficient condition for such case (see *Observation 2*). Recently, the authors of the references [48, 49] confirmed that in a PPS system using the ABL- rule [11–13], it is possible to go beyond the standard lower bound in the RHUR for position and momentum observables. Not exactly, but a similar property i.e., achieving arbitrary small lower bound (which depends on the pre- and post-selections) of the product of standard deviations of two non-commuting observables in a PPS system is possible in the relations (17). We now explore two peculiar characteristics of the uncertainty relations (17) and (21) that cannot be observed in standard quantum systems.

Observation 1. If the lower bound in an uncertainty relation in any quantum system is non-zero, then we say that two incompatible observables disturb each others' measurement results. Now consider the following case. If $|\psi\rangle$ is a common eigenstate of both A and B, then $\langle\Delta A\rangle_{\psi} = 0$, $\langle\Delta B\rangle_{\psi} = 0$ implying that the measurement of one doesn't disturbs the outcome of the other. Surprisingly, this property doesn't hold in the PPS system. Note that, even if $|\psi\rangle$ is a common eigenstate of both A and B, the lower bound of the relation (21) doesn't become zero for specific post-selections which implies $\langle\Delta A\rangle_{\psi}^{\phi} \neq 0$, $\langle\Delta B\rangle_{\psi}^{\phi} \neq 0$. Hence we can say that the measurement of A is invariably disturbed by the measurement of B or vice versa in a PPS system. In the Ref. [50], Vaidman demonstrated the same property in a PPS system using the ABL- rule.

Observation 2. With two non-commuting observables in the standard quantum system, sharp preparation of a quantum state is impossible. Or equivalently, for an initially prepared state $|\psi\rangle$, it is impossible to have $\langle\Delta A\rangle_{\psi} = 0$, $\langle\Delta B\rangle_{\psi} = 0$ if $[A,B] \neq 0$. But in the PPS system, we can prepare *any* quantum state $|\psi\rangle$ which can give $\langle\Delta A\rangle_{\psi}^{\phi} = 0$, $\langle\Delta B\rangle_{\psi}^{\phi} = 0$ for a specific choice of post-selection implying sharp preparation of $|\psi\rangle$ for non-commuting observables A and B. It is easy to show that both the uncertainties $\langle\Delta A\rangle_{\psi}^{\phi}$ and $\langle\Delta B\rangle_{\psi}^{\phi}$ are zero for the common post-selection $|\phi_z\rangle$ if and only if

$$|\phi_z\rangle \propto A|\psi\rangle\,, \qquad |\phi_z\rangle \propto B|\psi\rangle\,.$$

After the normalization, we find the common post-selection condition

$$|\phi_z\rangle = \frac{A\,|\psi\rangle}{\sqrt{\langle\psi|A^2|\psi\rangle}} = \frac{B\,|\psi\rangle}{\sqrt{\langle\psi|B^2|\psi\rangle}}, \tag{22}$$

upto some phase factors.

Example. Now, consider an example of two non-commuting observables $A = \frac{1}{\sqrt{2}}(I + \sigma_x)$ and $B = \frac{1}{\sqrt{2}}(\sigma_z + \sigma_x)$ with the initially prepared state $|0\rangle$. With these specific choices, it is possible to show that condition (22) is satisfied. Recall that in order to conduct the experiment using weak values, the average values of the observables must not be zero; for this reason, we did not take into account the Pauli observables σ_x and σ_y with initially prepared state $|0\rangle$. Nonetheless, if one does not adhere to weak values, this example is still true. So, the common post-selection for this case is $|\phi_e\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and hence both the uncertainties $\langle\Delta A\rangle_0^{\phi_e}$ and $\langle\Delta B\rangle_0^{\phi_e}$ of the non-commuting observables A and B, respectively are zero for the given initially prepared state $|0\rangle$ and the conditioned post-selection $|\phi_e\rangle$ in Eq. (22). In the PPS system, it is now feasible to do the hitherto impossibly difficult task of jointly sharply preparing a quantum state for two non-commuting observables.

The aforementioned *Observations 1* and *2* demonstrate that PPS systems are capable of being even stranger than their well-known unusual results e.g., quantum Cheshire Cats [24], measurement of a component of a spin 1/2 particle which can reach $100\hbar$ [16], etc.

Comments.— The characteristics of the uncertainty relations (17) and (21) in PPS systems as compared to the RHUR (8) and Eq. (9) are substantially altered by the post-selections. These uncertainty inequalities (17) and (21) will undoubtedly have applications like the RHUR for quantum foundations, information and technologies. For instances, (i) they can be used for information extraction using commuting observables because the inequalities do not become trivial for particular choices of post-selections, (ii) one can obtain a series of uncertainty inequalities by changing the post-selections and that is advantageous for practical purposes (see stronger uncertainty relations in sec. IV), (iii) existing applications of uncertainty relations (8) and (9) in standard systems, such as entanglement detection [7], quantum metrology [51, 52], etc., can be revisited using uncertainty relations (17) and (21) in the PPS systems, (iv) PPS system based spin squeezing: spin-squeezed states are a class of states having squeezed spin variance along a certain direction, at the cost of anti-squeezed variance along an orthogonal direction. This is done by using the RHUR (8) in the standard quantum system [53–56]. Such analysis can be reintroduced in the light of PPS systems. As there is no unique definition of spin squeezing in the standard quantum systems, it is, by means of Eq. (17), also possible to define the spin squeezing non uniquely in the PPS system. A very careful analysis is required to see whether there exists some states in the PPS systems for which $\langle\Delta A\rangle_\psi^\phi = \langle\Delta B\rangle_\psi^\phi$ and inequality (17 is saturated similar to coherent spin states in the standard quantum systems.

1. *Intelligent pre- and post-selected states*

In the standard quantum system, the states for which the equality condition holds in the RHUR (8) are known as intelligent states or minimum-uncertainty states [57–59]. Minimum uncertainty states have been proposed to improve the accuracy of phase measurement in quantum interferometer [60]. Minimum-uncertainty states in the PPS systems can also be defined based on the following condition.

One can find the condition for which the inequality (17) saturates (see Appendix A) is given by

$$A|\psi\rangle - \langle\phi|A|\psi\rangle|\phi\rangle = \pm i\frac{\langle\Delta A\rangle_\psi^\phi}{\langle\Delta B\rangle_\psi^\phi}\left(B|\psi\rangle - \langle\phi|B|\psi\rangle|\phi\rangle\right). \tag{23}$$

If the sign of 'i' on the RHS of Eq. (23) is taken to be positive (negative) when the observable A appears on the LHS of the Eq. (23), then the sign of 'i' on the RHS of Eq. (23) is taken be negative (positive) when the observable B appears on the LHS of the Eq. (23). So, the pre- and post-selected states which satisfy the condition (23, can be referred as the "intelligent pre- and post-selected states". For the given pre-selection and observables in Eq. (23), one can find the post-selection which will make the Eq. (17) the most tight i.e., equality.

C. Uncertainty equality in PPS system

Recently, in the reference [61], the authors have shown that there exist variance-based uncertainty equalities from which a series of uncertainty inequalities with hierarchical structure can be obtained. It was shown that stronger uncertainty relation given by Maccone and Pati [62] is a special case of these uncertainty inequalities. Here we show such uncertainty equalities in the PPS systems. We provide interpretation of the uncertainty inequalities derived from the uncertainty equalities. Further, in Application section IV, we use uncertainty equalities in PPS systems to obtain stronger uncertainty relations and state dependent tighter uncertainty relations.

Theorem 2. *The product of standard deviations of two non-commuting hermitian operators A, $B \in \mathcal{L}(\mathcal{H})$ in a PPS system with pre- and post-selected states $|\psi\rangle$ and $|\phi\rangle$, respectively satisfies*

$$\langle\Delta A\rangle_\psi^\phi \langle\Delta B\rangle_\psi^\phi = \frac{\mp\left(\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle - \mathcal{I}m(W_{AB})\right)}{1 - \frac{1}{2}\sum_{k=1}^{d-1}\left|\langle\psi|\frac{A}{\langle\Delta A\rangle_\psi^\phi} \pm i\frac{B}{\langle\Delta B\rangle_\psi^\phi}|\phi_k^\perp\rangle\right|^2}, \tag{24}$$

where we have assumed that $\langle\Delta A\rangle_\psi^\phi$ and $\langle\Delta B\rangle_\psi^\phi$ are non-zero, and the sign should be considered such that the numerator is always real and positive. Here $\{|\phi\rangle, |\phi_k^\perp\rangle_{k=1}^{d-1}\}$ is an complete orthonormal basis in the d-dimensional Hilbert space.

Proof. Consider an orthonormal complete basis $\{|\phi\rangle, |\phi_k^\perp\rangle_{k=1}^{d-1}\}$ in the d-dimensional Hilbert space $\mathcal{H}$.

Now, define the projection operator $\Pi = I - |\phi\rangle\langle\phi|$ and the unnormalized state vector $|\xi^{\pm}\rangle = \left(\frac{A}{\langle\Delta A\rangle_{\psi}^{\phi}} \pm i\frac{B}{\langle\Delta B\rangle_{\psi}^{\phi}}\right)|\psi\rangle$. Then we have the following identity

$$\begin{aligned}\langle\xi^{\mp}|\Pi|\xi^{\mp}\rangle &= \langle\xi^{\mp}|\xi^{\mp}\rangle - \langle\xi^{\mp}|\phi\rangle\langle\phi|\xi^{\mp}\rangle \\ &= \left\{\frac{\langle\psi|A^2|\psi\rangle}{\left(\langle\Delta A\rangle_{\psi}^{\phi}\right)^2} + \frac{\langle\psi|B^2|\psi\rangle}{\left(\langle\Delta B\rangle_{\psi}^{\phi}\right)^2} \mp \frac{i\langle\psi|[A,B]|\psi\rangle}{\langle\Delta A\rangle_{\psi}^{\phi}\langle\Delta B\rangle_{\psi}^{\phi}}\right\} \\ &\quad - \left\{\frac{|\langle\phi|A|\psi\rangle|^2}{\left(\langle\Delta A\rangle_{\psi}^{\phi}\right)^2} + \frac{|\langle\phi|B|\psi\rangle|^2}{\left(\langle\Delta B\rangle_{\psi}^{\phi}\right)^2} \pm \frac{2\mathcal{I}m(W_{AB})}{\langle\Delta A\rangle_{\psi}^{\phi}\langle\Delta B\rangle_{\psi}^{\phi}}\right\} \\ &= 2 \pm 2\frac{\left(\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle - \mathcal{I}m(W_{AB})\right)}{\langle\Delta A\rangle_{\psi}^{\phi}\langle\Delta B\rangle_{\psi}^{\phi}}, \end{aligned} \tag{25}$$

where we have used Eq. (11) and $W_{AB} = \langle\psi|A|\phi\rangle\langle\phi|B|\psi\rangle$. Now, we use another expression of $\Pi = \sum_{k=1}^{d-1}|\phi_k^{\perp}\rangle\langle\phi_k^{\perp}|$ to calculate the same identity

$$\langle\xi^{\mp}|\Pi|\xi^{\mp}\rangle = \sum_{k=1}^{d-1}\left|\langle\psi|\frac{A}{\langle\Delta A\rangle_{\psi}^{\phi}} \pm i\frac{B}{\langle\Delta B\rangle_{\psi}^{\phi}}|\phi_k^{\perp}\rangle\right|^2. \tag{26}$$

So, from the Eqs. (25) and (26), we obtain the uncertainty equality (24) in the PPS system. □

Theorem 3. *The sum of the variances of two non-commuting hermitian operators $A, B \in \mathcal{L}(\mathcal{H})$ in a PPS system with pre- and post-selected states $|\psi\rangle$ and $|\phi\rangle$, respectively satisfies*

$$\begin{aligned}\left(\langle\Delta A\rangle_{\psi}^{\phi}\right)^2 + \left(\langle\Delta B\rangle_{\psi}^{\phi}\right)^2 &= \pm\left(i\langle\psi|[A,B]|\psi\rangle - 2\mathcal{I}m(W_{AB})\right) \\ &\quad + \sum_{k=1}^{d-1}|\langle\phi_k^{\perp}|(A \mp iB)|\psi\rangle|^2. \end{aligned} \tag{27}$$

Here, the '$\pm$' sign is taken suitably such that the first term in right side is always positive.

Proof. Consider an orthonormal complete basis $\{|\phi\rangle, |\phi_k^{\perp}\rangle_{k=1}^{d-1}\}$ in the d-dimensional Hilbert space $\mathcal{H}$ and hence $I - |\phi\rangle\langle\phi| = \sum_{k=1}^{d-1}|\phi_k^{\perp}\rangle\langle\phi_k^{\perp}|$. By equating the following two

$$Tr\Big((A \mp iB)|\psi\rangle\langle\psi|(A \pm iB)(I - |\phi\rangle\langle\phi|)\Big),$$

$$Tr\Big((A \mp iB)|\psi\rangle\langle\psi|(A \pm iB)(\sum_{k=1}^{d-1}|\phi_k^{\perp}\rangle\langle\phi_k^{\perp}|)\Big),$$

we have Eq. (27). □

An inequality can be obtained by discarding some of the terms in the summation corresponding to 'k' or all the terms except one term in Eq. (24) or Eq. (27). It is also possible to obtain an arbitrarily tight inequality by discarding the minimum valued term inside the summation in the denominator of Eq. (24) for a particular value of 'k'. . Note that we have to optimize the minimum $\left|\langle\psi|\frac{A}{\langle\Delta A\rangle_{\psi}^{\phi}} \pm i\frac{B}{\langle\Delta B\rangle_{\psi}^{\phi}}|\phi_k^{\perp}\rangle\right|^2$ over all possible choice of basis $\{|\phi_k^{\perp}\rangle_{k=1}^{d-1}\}$ in the subspace orthogonal to $|\phi\rangle$.

In an experiment, let's assume that a few post-selected states from $\{|\phi_k^{\perp}\rangle\}_{k=1}^{d-1}$ are not detected by the detector because of certain technical difficulties. Using such imprecise experimental data, one may still be able to obtain an uncertainty relation. In that case, the terms corresponding to the unregistered post-selections in Eq. (24) or Eq. (27) are to be eliminated.

D. Uncertainty relation for mixed pre- selection in PPS system

So far, we have only considered the pre- selected state to be pure in a PPS system. Let us now generalize the definition of the standard deviation and derive the uncertainty relations for mixed pre-selected state in the PPS system. A direct generalization of the standard deviation Eq. (11) is given by

$$\left(\langle\Delta A\rangle_{\rho}^{\phi}\right)^2 = Tr(A^2\rho) - \langle\phi|A\rho A|\phi\rangle\,. \tag{28}$$

If $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$, where $\sum_i p_i = 1$, then

$$\left(\langle\Delta A\rangle_{\rho}^{\phi}\right)^2 = \sum_i p_i\left(\langle\Delta A\rangle_{\psi_i}^{\phi}\right)^2. \tag{29}$$

Eq. (29) demonstrates the intriguing fact that, the variance of A in PPS system i.e., $\langle VarA\rangle_{\rho}^{\phi} = (\langle\Delta A\rangle_{\rho}^{\phi})^2$ does not increase under the classical mixing of quantum states. Mathematically, classical mixing of quantum states are represented by a density operator. By taking advantage of this property, one can obtain a purely quantum uncertainty relation when the pre- selection ρ is a mixed state (see sec. IV). In standard quantum systems, the variance $VarA = \langle\Delta A\rangle_{\rho}^2$ increases in general under the classical mixing of quantum states.

To realize the standard deviation in PPS system via weak value for mixed pre- selected state, we introduce another definition of the standard deviation as follows

$$\left(\langle\Delta A_w\rangle_{\rho}^{\phi}\right)^2 = Tr(A^2\rho) - |\langle A_w\rangle_{\rho}^{\phi}|^2\langle\phi|\rho|\phi\rangle\,, \tag{30}$$

where $\langle A_w\rangle_{\rho}^{\phi} = \frac{\langle\phi|A\rho|\phi\rangle}{\langle\phi|\rho|\phi\rangle}$ is the weak value of the operator A when the pre- and post-selections are ρ and $|\phi\rangle$, respectively.

Proposition 3. *The standard deviation based on weak value defined in Eq. (30) always increases under the classical mixing of quantum states that is*

$$\langle\Delta A_w\rangle_{\rho}^{\phi} \geq \langle\Delta A\rangle_{\rho}^{\phi}\,, \tag{31}$$

where the equality holds if ρ is a pure state.

Proof. Let $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$, then Eq. (30) becomes (after

TABLE I: Comparison of different properties between standard quantum systems and PPS systems.

Properties	Standard quantum systems	Pre- and post-selected systems
Standard deviation	$\langle \Delta A \rangle_\psi = \left(\langle \psi \vert A^2 \vert \psi \rangle - \langle \psi \vert A \vert \psi \rangle^2 \right)^{1/2}$	$\langle \Delta A \rangle_\psi^\phi = \left(\langle \psi \vert A^2 \vert \psi \rangle - \vert \langle A_w \rangle_\psi^\phi \vert^2 \vert \langle \phi \vert \psi \rangle \vert^2 \right)^{1/2}$
Zero standard deviation	Only if $\vert \psi \rangle$ is an eigenstate of A i.e., $\vert \psi \rangle \propto A \vert \psi \rangle$	Only if $\vert \phi \rangle \propto A \vert \psi \rangle$
Uncertainty relation	$\langle \Delta A \rangle_\psi^2 \langle \Delta B \rangle_\psi^2 \geq \left[\frac{1}{2i} \langle \psi \vert [A, B] \vert \psi \rangle \right]^2$	$\left(\langle \Delta A \rangle_\psi^\phi \right)^2 \left(\langle \Delta B \rangle_\psi^\phi \right)^2 \geq \left[\frac{1}{2i} \langle \psi \vert [A, B] \vert \psi \rangle - \mathcal{I}m \left(W_{AB} \right) \right]^2$
Joint sharp preparation	If $\vert \psi \rangle$ is the eigenstate of both A and B	If $\vert \phi \rangle = \frac{A \vert \psi \rangle}{\sqrt{\langle \psi \vert A^2 \vert \psi \rangle}} = \frac{B \vert \psi \rangle}{\sqrt{\langle \psi \vert B^2 \vert \psi \rangle}}$, up to some phase factors

using the definition of the weak value for mixed pre- selection)

$$
\begin{aligned}
&\left(\langle \Delta A_w \rangle_\rho^\phi \right)^2 \\
&= Tr(A^2 \rho) - \frac{\vert \langle \phi \vert A \rho \vert \phi \rangle \vert^2}{\langle \phi \vert \rho \vert \phi \rangle} \\
&= \sum_i p_i \langle \psi_i \vert A^2 \vert \psi_i \rangle - \frac{\vert \sum_i \sqrt{p_i} \langle \phi \vert A \vert \psi_i \rangle \sqrt{p_i} \langle \psi_i \vert \phi \rangle \vert^2}{\langle \phi \vert \rho \vert \phi \rangle} \\
&\geq \sum_i p_i \langle \psi_i \vert A^2 \vert \psi_i \rangle - \frac{\left(\sum_i p_i \vert \langle \phi \vert A \vert \psi_i \rangle \vert^2 \right) \left(\sum_i p_i \langle \phi \vert \psi_i \rangle \langle \psi_i \vert \phi \rangle \right)}{\langle \phi \vert \rho \vert \phi \rangle} \\
&= \sum_i p_i \langle \psi_i \vert A^2 \vert \psi_i \rangle - \sum_i p_i \vert \langle \phi \vert A \vert \psi_i \rangle \vert^2 \\
&= \sum_i p_i \left(\langle \Delta A \rangle_{\psi_i}^\phi \right)^2 = \left(\langle \Delta A \rangle_\rho^\phi \right)^2,
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality for the complex numbers in the first inequality and Eq. (29) in the last line. As $\langle Var A \rangle_\rho^\phi = (\langle \Delta A \rangle_\rho^\phi)^2$ does neither increase nor decrease under classical mixing of quantum states, the inequality $\langle \Delta A_w \rangle_\rho^\phi \geq \langle \Delta A \rangle_\rho^\phi$ clearly implies that under classical mixing of quantum states, the standard deviation based on weak value defined in Eq. (30) is always non-decreasing. When ρ is pure, equality automatically holds. □

It is important to note that, in general the equality in Eq. (31) does not imply that the pre- selection ρ is pure. In sec. IV A (see below), we show that only in the qubit system, equality of Eq. (31) implies that the pre- selection is a pure state. To make an equality in Eq. (31) in higher dimensional systems, we need to put conditions on the observable and post-selection (see below in sec. IV A).

The uncertainty relation (17) or (21) can be generalized for mixed pre- selection ρ also which is given by

$$
\left(\langle \Delta A \rangle_\rho^\phi \right)^2 \left(\langle \Delta B \rangle_\rho^\phi \right)^2 \geq \left[\frac{1}{2i} \langle [A, B] \rangle_\rho - \mathcal{I}m W_{AB} \right]^2, \tag{32}
$$

where $W_{AB} = \langle \phi \vert B \rho A \vert \phi \rangle$. See the derivation Eq. (32) in Appendix B. Eq. (32) holds also when the definition of standard deviation defined in Eq. (30) is considered due the *Proposition 3*.

See TABLE I for the comparison of different properties between standard quantum systems and PPS systems.

IV. APPLICATIONS

Suitably post-selected systems can offer some essential information regarding quantum systems. Below, we provide a few applications of standard deviations and uncertainty relations in PPS systems.

A. Detection of mixedness of an unknown state

Practically, partial information about a quantum state is often of great help. For example, whether an interaction has taken place with the environment, one must verify the purity of the system's state. Quantum state tomography (QST) is the most resource intensive way to verify the purity of a quantum state but here we provide some results that can be used to detect purity of the quantum state using less resources compared to the QST.

We will use the inequality (31) in *Proposition 3* to detect the mixedness of an unknown pre- selected state in a PPS system. The proofs of the following *Lemmas* are given in Appendix C.

Lemma 1. *Qubit system: In the case of a two-level quantum system (i.e., a qubit), equality in Eq. (31) holds if and only if the pre- selected state ρ is pure irrespective of choice of the observable A and the post-selected state $\vert \phi \rangle$.*

Lemma 2. *Qutrit system: If for an observable A and a complete orthonormal basis $\{\vert \phi_k \rangle\}_{k=1}^3$ (to be used as post- selected states) of any three-level quantum system (i.e., a qutrit), and the condition $\langle \phi_1 \vert A \vert \phi_2 \rangle = 0$ also holds good, then equality in Eq. (31) holds good if and only if the pre- selected state ρ is pure.*

Lemma 3. *Qubit-qubit system: Consider any two non orthogonal post-selections $\vert \phi_B \rangle$ and $\vert \phi'_B \rangle$ in the subsystem B. For any observable A, equality of $\langle \Delta (A \otimes I)_w \rangle_\rho^{\phi_{AB}}$*

and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}$ *and separately of* $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi'_{AB}}$ *and* $\langle\Delta(A\otimes I)\rangle_\rho^{\phi'_{AB}}$ *hold only when the* $2\otimes 2$ *pre- selected state* ρ *is pure, where* $|\phi_{AB}\rangle = |\phi_A\rangle|\phi_B\rangle$ *and* $|\phi'_{AB}\rangle = |\phi_A\rangle|\phi'_B\rangle$. *Two non orthogonal post-selections* $|\phi_B\rangle$ *and* $|\phi'_B\rangle$ *in the subsystem 'B' are required here due to the fact that there exists an unique* $2\otimes 2$ *mixed density matrix which satisfies the equality of Eq. (31)*

Lemma 4. *Qubit-qutrit system: If for an observable* A *and any complete orthonormal basis* $\{|\phi_A^k\rangle\}_{k=1}^3$ *(to be used as post-selected states) for a qutrit, and the condition* $\langle\phi_A^1|A|\phi_A^2\rangle = 0$ *is considered, then equality of* $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi_{AB}}$ *and* $\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}$ *and separately of* $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi'_{AB}}$ *and* $\langle\Delta(A\otimes I)\rangle_\rho^{\phi'_{AB}}$ *hold if and only if the* $3\otimes 2$ *pre- selected state* ρ *is pure.*

Extension of this method for higher dimensional systems will require more conditions to be imposed on the observable and post-selections. So it might be difficult to apply our method for higher dimensions. To overcome this difficulties, Eq. (17) or (21) can be used to detect the mixedness of the initially prepared states. Note that, Mal *et al.* have used the stronger version of the RHUR (9) to do so [63].

B. Stronger Uncertainty Relation

Motivation. If, for example, the initially prepared state of the system is an eigenstate of one of the two incompatible observables A and B, both the sides of the RHUR (8) becomes trivial (i.e., zero). For certain states, a trivial lower bound is always possible because the right side of the relation (8) contains the average of the commutator of incompatible observables. For such cases, the RHUR (8) does not capture the incompatibility of the non-commuting observables. One can think of adding *Schrödinger's term* in the RHUR but still this can be become trivial (e.g., when the prepared states is an eigenstate of either A or B). So, none of them are unquestionably appropriate to capture the incompatibility of the non-commuting observables.

It is Maccone and Pati [62] who considered a different uncertainty relation, based on the sum of the variances $\langle\Delta A\rangle_\psi^2 + \langle\Delta B\rangle_\psi^2$, that is guaranteed to be nontrivial (i.e., having non-zero lower bound) whenever the observables are incompatible on the given state $|\psi\rangle$. But there are shortcomings in the Maccone-Pati Uncertainty Relations (MPUR). It is easy to show that in two dimensional Hilbert space [64], if, for example, the initial state $|\psi\rangle$ of the system is an eigenstate of the observable A, then one finds that the first inequality $\langle\Delta A\rangle_\psi^2 + \langle\Delta B\rangle_\psi^2 \geq \pm i\langle\psi|[A,B]|\psi\rangle + \left|\langle\psi|(A\pm iB)|\psi^\perp\rangle\right|^2$ in MPUR becomes $\langle\Delta B\rangle_\psi^2 \geq \langle\Delta B\rangle_\psi^2$, where $|\psi^\perp\rangle$ is arbitrary state orthogonal to $|\psi\rangle$. Similarly, it can be shown that the second inequality $\langle\Delta A\rangle_\psi^2 + \langle\Delta B\rangle_\psi^2 \geq \frac{1}{2}\left|\langle\psi_{A+B}^\perp|(A+B)|\psi\rangle\right|^2$ in MPUR becomes $\langle\Delta B\rangle_\psi^2 \geq \frac{1}{2}\langle\Delta B\rangle_\psi^2$, where $|\psi_{A+B}^\perp\rangle = (1/\langle\Delta(A+B)\rangle_\psi)(A+B-\langle A+B\rangle_\psi)|\psi\rangle$ and $\langle\Delta(A+B)\rangle_\psi^2 = \langle(A+B)^2\rangle_\psi - \langle A+B\rangle_\psi^2$ for arbitrary dimensional Hilbert space if the initial state of the system is an eigenstate of the observable A [65]. It indicates that the first and second inequalities in MPUR for two and arbitrary dimensions, respectively, contain no information about the observable A and are therefore of no practical significance. In other words, we learn nothing new about the quantum system other than the trivial fact that $\langle\Delta B\rangle_\psi$ is non-negative. In addition, if the initially prepared state $|\psi\rangle$ is unknown, then $|\psi^\perp\rangle$ is likewise unknown in the MPUR inequalities and, so is the lower bound of MPUR. The first inequality in MPUR may be useful in a quantum system with Hilbert spaces of more than two dimensions.

Here, we demonstrate that relations (17) and (21) can be used to solve the triviality problem of the RHUR and the problem with MPUR that we have mentioned above, i.e., these uncertainty relations can provide non-trivial information about the observable A. Even if the initially prepared state (preselection) $|\psi\rangle$ is unknown, the lower bound of our stronger uncertainty relation can be calculated.

Consider the relation (17) which, using Eq. (14), becomes

$$\left(\langle\Delta A\rangle_\psi^2 + \epsilon_A\right)\left(\langle\Delta B\rangle_\psi^2 + \epsilon_B\right) \geq \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle - \mathcal{I}m(W_{AB})\right]^2, \tag{33}$$

where $\epsilon_X = \langle X\rangle_\psi^2 - |\langle X_w\rangle_\psi^\phi|^2|\langle\phi|\psi\rangle|^2$, with $X = A$ or B. Now suppose $|\psi\rangle$ is an eigenstate of A then, the Eq. (33) is nontrivial unless $|\phi\rangle = |\psi\rangle$, as, in the case when $|\phi\rangle \neq |\psi\rangle$, the inequality (33) becomes

$$\epsilon_A\left(\langle\Delta B\rangle_\psi^2 + \epsilon_B\right) \geq \left[\mathcal{I}m\left(W_{AB}\right)\right]^2. \tag{34}$$

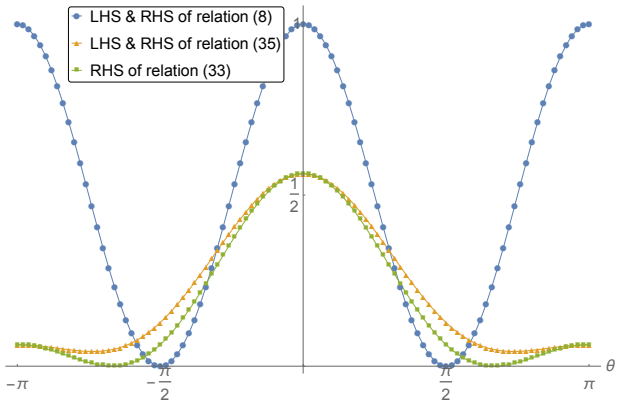


Fig. 1. Comparison between the RHURs (8) and (9), and the uncertainty relations (33) and (35). We choose $A = \sigma_x$, $B = \sigma_y$ for a spin-1/2 particle and $|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\xi}\sin(\theta/2)\,|1\rangle$, $|\phi\rangle = \cos(\omega/2)\,|0\rangle + e^{i\eta}\sin(\omega/2)\,|1\rangle$ with $\xi = 0$, $\omega = \pi/3$ and $\eta = \pi/5$. The blue curve is the LHS of the RHUR and for this particular case, it coincides with its lower bound i.e., RHS of RHUR. The orange curve is the LHS of Eq. (35) and for this particular case, it coincides with the RHS of Eq. (35). The green curve is the RHS of Eq. (33). Now, notice that, for $\theta = -\pi/2$ and $\pi/2$, the RHUR becomes trivial while for the same values of θ, the relation (33) as well as the relation (35) are nontrivial. For this particular choice of post-selection, both the relations (33) and (35) are stronger than the RHUR (8). Note that, the relation (35) is the strongest under this condition as it is non-trivial for all the values of θ. If, for the fixed values of θ and ξ, the relations (33) and (35) is trivial, then one should keep changing the values of ω and η (i.e., by choosing the post-selection suitably) until they become nontrivial which is our main goal.

Notice that, in the both sides of Eq. (34), there is a quantum state $|\phi\rangle$ which can be chosen independently in the standard quantum system. So, it is always possible to choose a suitable $|\phi\rangle$ such that the relation (33) is nontrivial. With "Schrödinger's term", the relations (33) and (34) becomes

$$\Big(\langle\Delta A\rangle_\psi^2+\epsilon_A\Big)\Big(\langle\Delta B\rangle_\psi^2+\epsilon_B\Big)\geq\left[\frac{1}{2i}\langle\psi|[\![A,B]\!]|\psi\rangle-\mathcal{I}m(W_{AB})\right]^2 + \left[\frac{1}{2}\langle\psi|\{A,B\}|\psi\rangle-\mathcal{R}e(W_{AB})\right]^2, \tag{35}$$

$$\epsilon_A\Big(\langle\Delta B\rangle_\psi^2+\epsilon_B\Big)\geq\left[\mathcal{I}m\left(W_{AB}\right)\right]^2 + \left[\frac{1}{2}\langle\psi|\{A,B\}|\psi\rangle-\mathcal{R}e(W_{AB})\right]^2, \tag{36}$$

respectively. As ϵ_A and ϵ_B can also be negative, the left hand side of relation (33) can become lower than the left hand side of relation (8). The same holds true for the right-hand side as well. So, for a fixed $|\psi\rangle$, we always want to have a nontrivial lower bound from the relations (8) and (33) which can be combined in a single uncertainty relation i.e., the *stronger uncertainty relation*

$$max\left\{\mathcal{L}_{RH},\mathcal{L}_{PPS}\right\}\geq max\left\{\mathcal{R}_{RH},\mathcal{R}_{PPS}\right\},$$

where $\mathcal{L}_{RH} = \langle\Delta A\rangle_\psi^2\langle\Delta B\rangle_\psi^2$, $\mathcal{L}_{PPS} = \Big(\langle\Delta A\rangle_\psi^2+\epsilon_A\Big)\Big(\langle\Delta B\rangle_\psi^2+\epsilon_B\Big)$, $\mathcal{R}_{RH} = \left[\frac{1}{2i}\langle\psi|[A,B]|\psi\rangle\right]^2$ and $\mathcal{R}_{PPS} = \left[\frac{1}{2i}\langle\psi|\,[A,B]\,|\psi\rangle - \mathcal{I}m\left(W_{AB}\right)\right]^2$.

In FIG. 1, comparison between the relations (8) and, both (33) and (35) is shown. Eqs. (34) and (36) capture the informations about the operator A when the initially prepared state $|\psi\rangle$ is one of the eigenstates of A, while MPUR fails to capture such informations which we have already discussed.

Moreover, even if the initial state (i.e., pre- selection) is unknown, the lower bound of the uncertainty relation (33) can be calculated experimentally and in that case we need the average value of the hermitian operator $\frac{1}{i}[A,B]$ and weak values of the operators A and B.

Sum uncertainty relation in the PPS system can also be used to obtain stronger uncertainty relation in the standard quantum system. One can easily show that

$$\Big(\langle\Delta A\rangle_\psi^2+\epsilon_A\Big)+\Big(\langle\Delta B\rangle_\psi^2+\epsilon_B\Big)\geq\pm\Big(i\,\langle\psi|[A,B]|\psi\rangle -2\mathcal{I}m(W_{AB})\Big)$$

holds in Eq. (27) in Theorem 3 by discarding the summation part which is always a positive number. This inequality remains strong against when $|\psi\rangle$ is one of the eigenstates of A by suitably choosing post-selection $|\phi\rangle$.

C. Purely quantum uncertainty relation

Motivation. In practice, it is not always possible to carry out quantum mechanical tasks with pure states because of interactions with the environment. Because the mixed initial prepared state is a classical mixture of pure quantum states, any task or measurement involves a hybrid of classical and quantum parts. In modern technologies, it is considered that quantum advantage is more effective and superior to classical advantage. Hence, a hybrid of a quantum and classical component may be less advantageous than a quantum component alone. For example, the uncertainty of an observable A in standard quantum system increases in general under classical mixing of quantum states i.e., $\langle\Delta A\rangle_\rho^2 \geq \sum_i p_i\,\langle\Delta A\rangle_{\psi_i}^2$ (where $\rho = \sum_i p_i\,|\psi_i\rangle\,\langle\psi_i|$) and this is disadvantageous in the sense that the knowledge about the observable is more uncertain than only when the average of the pure state uncertainties are considered. The uncertainty that one gets to see due to (classical) mixing of pure states is considered here as 'classical uncertainty'. The standard deviation $\langle\Delta A\rangle_\rho$ can be referred as the hybrid of classical and quantum uncertainties and hence the RHUR (8) can be considered as the hybrid uncertainty relation in the standard quantum systems.

Purely quantum uncertainty relation, a crucial component of the quantum world, may be very useful, but in order to obtain it, the classical uncertainty must be eliminated from the

hybrid uncertainty relation. To do this, we first need to determine the purely quantum uncertainty of an observable, which can be done in a number of ways, such as by eliminating the classical component of the hybrid uncertainty or by specifying a purely quantum uncertainty straight away.

Any measure of purely quantum uncertainty should have at least the following intuitive and expected property (below 'Q' represents sometimes quantum observables, sometimes states, etc for different type of quantum mechanical systems. For example, if the system is a PPS system, then 'Q' is the post-selection $|\phi\rangle$. If the system is a standard quantum system, then the term 'Q' disappears):
(i) *Quantum uncertainty should not be affected (neither increasing nor decreasing) by the classical mixing of quantum states i.e.,*

$$\mathcal{Q}(\rho, A, Q)=\sum_i p_i \mathcal{Q}(\psi_i, A, Q),\ where\ \rho =\sum_i p_i |\psi_i\rangle \langle\psi_i| .$$

Here $\mathcal{Q}(\rho, A, Q)$ is some measure of purely quantum uncertainty of the observable A for a given ρ. There might exist some other properties depending upon the nature of the system (e.g., standard systems, PPS systems, etc) but we emphasize that the most important property of a purely quantum uncertainty should be (i).

It is seen that the variance of A in PPS system i.e., $\langle VarA\rangle_\rho^\phi = (\langle\Delta A\rangle_\rho^\phi)^2$ is a purely quantum uncertainty which satisfies the property (i). Now the purely quantum mechanical uncertainty relation in this regard is Eq. (32).

As can be seen from Eq. (31) for mixed states, the second definition of the variance of A i.e., $\langle VarA_w\rangle_\rho^\phi = (\langle\Delta A_w\rangle_\rho^\phi)^2$ in the PPS system defined in Eq. (30) is a hybrid uncertainty. Hence, the uncertainty in PPS system based on weak value has both classical and quantum parts. When measurement is carried out in the PPS system and weak values are involved, classical uncertainty may be crucial in determining how much classicality (in the form of classical uncertainty) the mixed state ρ possesses. Mixed states with less classicality should have more quantumness (in the form of quantum uncertainty), and vice versa. To distinguish classical uncertainty from the hybrid uncertainty $\langle VarA_w\rangle_\rho^\phi$, we subtract the quantum uncertainty $\langle VarA\rangle_\rho^\phi$ from it i.e.,

$$C(\rho, A, \phi) = \left(\langle\Delta A_w\rangle_\rho^\phi\right)^2 - \left(\langle\Delta A\rangle_\rho^\phi\right)^2 . \qquad (37)$$

This is one of the good measures of classical uncertainty which should have some intuitive and expected properties:
(i) $\mathcal{C}(\rho, A, Q) \geq 0$ for a quantum state ρ,
(ii) $\mathcal{C}(\rho, A, Q) = 0$ when $\rho = |\psi\rangle\langle\psi|$ (absence of classical mixing),
(iii) Total classical uncertainty of disjoint systems should be the sum of individual systems's classical uncertainties:

$$\mathcal{C}(\rho, A_1\otimes I + I\otimes A_2, Q) = \mathcal{C}(\rho, A_1\otimes I, Q) + \mathcal{C}(\rho, I\otimes A_2, Q),$$

when $\rho = \rho_1 \otimes \rho_2$.

One can show that all the properties (i)-(iii) of classical uncertainty are satisfied by $C(\rho, A, \phi)$ defined in Eq. (37). Particularly, property (iii) is satisfied by taking $Q = |\phi_1\rangle|\phi_2\rangle$. Here, $|\phi_1\rangle$ and $|\phi_2\rangle$ are post-selections of the two disjoint systems, respectively.

There are some works by Luo and other authors regarding the purely quantum uncertainty relation. Initial attempt was made by Luo and Zhang [66] to obtain uncertainty relation by using skew information (introduced by Wigner and Yanase [67]) but it was found to be incorrect in general [68]. Later, another attempt was made by Luo himself [69], which is obtained by discarding the classical part from the hybrid uncertainty relation using skew information. But this uncertainty relation can't be guaranteed to be an intrinsically quantum uncertainty relation (according to property (i)) as the uncertainty they claim to be a quantum uncertainty is a product of skew information (which is a convex function under the mixing of quantum states) and a concave function under the same mixing. After that a series of successful and failed attempts was performed by modifying the works of Luo and other authors [70–73].

Instead, we have given a quantum uncertainty relation although it is based on pre- and post-selections which is different from the standard quantum mechanics but a quantumness can be seen in the relation (32).

D. State dependent tighter uncertainty relations in standard systems

The RHUR (8) or (9) is known not to be the tight one. Some existing tighter bounds are given in [61, 62, 74]. The drawback of these tighter uncertainty relations is that their lower bounds depend on the states perpendicular to the given state of the system. If the given state is unknown, then the lower bound of these uncertainty relations also remain unknown.

Here we show that by the use of arbitrary post-selected state $|\phi\rangle$, the lower bound of the RHUR based on sum uncertainties can be made arbitrarily tight and even if the given state (i.e., pre- selection here) is unknown, the lower bound of our tighter uncertainty relation can be obtained in experiments.

Theorem 4. *Let $\rho \in \mathcal{L}(\mathcal{H})$ be the density operator of the standard quantum system, then the sum of the standard deviations of two non-commuting observables A, $B \in \mathcal{L}(\mathcal{H})$ satisfies*

$$\langle\Delta A\rangle_\rho^2 + \langle\Delta B\rangle_\rho^2 \geq \pm i\, Tr([A, B]\rho) + \langle\phi|C_\pm^\dagger \rho C_\pm|\phi\rangle , \quad (38)$$

where $C_\pm = A \pm iB - \langle A \pm iB\rangle_\rho I$ and the '$\pm$' sign is taken in such a way that the first term in the right hand side is always positive.

Proof. Considering Eq. (27) for pre- selection $|\psi_j\rangle$ and mul-

tiply by 'p_j', and then after summing over 'j', we have

$$\sum_j p_j \left(\langle \Delta A\rangle^{\phi}_{\psi_j}\right)^2 + \sum_j p_j \left(\langle \Delta B\rangle^{\phi}_{\psi_j}\right)^2 = \pm i\sum_j p_j\langle\psi_j|[A,B]|\psi_j\rangle \mp 2\mathcal{I}m\Big(\sum_j p_j\langle\phi|B|\psi_j\rangle\langle\psi_j|A|\phi\rangle\Big) + \sum_{k=1}^{d-1}\sum_j p_j\langle\phi_k^{\perp}|(A\pm iB)|\psi_j\rangle\langle\psi_j|(A\mp iB)|\phi_k^{\perp}\rangle\,, \quad (39)$$

where we have used $W_{AB} = \langle\phi|B|\psi_j\rangle\langle\psi_j|A|\phi\rangle$. By using Eq. (29) for A and B when $\rho = \sum_j p_j\,|\psi_j\rangle\langle\psi_j|$, we have

$$\begin{aligned}&\left(\langle \Delta A\rangle^{\phi}_{\rho}\right)^2 + \left(\langle \Delta B\rangle^{\phi}_{\rho}\right)^2 \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) \mp 2\mathcal{I}m\Big(\langle\phi|B\rho A|\phi\rangle\Big) + \sum_{k=1}^{d-1}\langle\phi_k^{\perp}|(A\pm iB)\rho(A\mp iB)|\phi_k^{\perp}\rangle \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) + \langle\phi|(A\pm iB)\rho(A\mp iB)|\phi\rangle - \langle\phi|A\rho A|\phi\rangle \\ &\quad - \langle\phi|B\rho B|\phi\rangle + \sum_{k=1}^{d-1}\langle\phi_k^{\perp}|(A\pm iB)\rho(A\mp iB)|\phi_k^{\perp}\rangle\,, \quad (40)\end{aligned}$$

where $\mp 2\mathcal{I}m(\langle\phi|B\rho A|\phi\rangle) = \pm i(\langle\phi|B\rho A|\phi\rangle - \langle\phi|A\rho B|\phi\rangle)$ = $\langle\phi|(A\pm iB)\rho(A\mp iB)|\phi\rangle - \langle\phi|A\rho A|\phi\rangle - \langle\phi|B\rho B|\phi\rangle$ has been used. Now put $\left(\langle \Delta A\rangle^{\phi}_{\rho}\right)^2 = \mathrm{Tr}\big(A^2\rho\big) - \langle\phi|A\rho A|\phi\rangle$ defined in Eq. (28) (similarly for B also) and after subtracting $\mathrm{Tr}(A\rho)^2 + \mathrm{Tr}(B\rho)^2$ from both sides of Eq. (40) and using $|\phi\rangle\langle\phi| + \sum_{k=1}^{d-1}\left|\phi_k^{\perp}\right\rangle\left\langle\phi_k^{\perp}\right| = I$, we have

$$\begin{aligned}&\langle \Delta A\rangle^2_{\rho} + \langle \Delta B\rangle^2_{\rho} \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) + \mathrm{Tr}[(A\pm iB)(A\mp iB)\rho] - \mathrm{Tr}(A\rho)^2 - \mathrm{Tr}(B\rho)^2 \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) + \mathrm{Tr}[(A\pm iB)(A\mp iB)\rho] - \mathrm{Tr}[(A\pm iB)\rho]\,\mathrm{Tr}[(A\mp iB)\rho] \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) + \mathrm{Tr}\Big(M_{\mp}^{\dagger}M_{\mp}\rho\Big) - |\,\mathrm{Tr}(M_{\mp}\rho)|^2 \\ &= \pm i\,\mathrm{Tr}([A,B]\rho) + \mathrm{Tr}\Big[(M_{\mp} - \langle M_{\mp}\rangle_{\rho} I)^{\dagger}(M_{\mp} - \langle M_{\mp}\rangle_{\rho} I)\rho\Big], \quad (41)\end{aligned}$$

where $M_{\mp} = A \mp iB$. Now let $C_{\pm} = M_{\pm} - \langle M_{\pm}\rangle_{\rho} I$ then Eq. (41) can be rewritten as

$$\langle \Delta A\rangle^2_{\rho} + \langle \Delta B\rangle^2_{\rho} = \pm\, i\,\mathrm{Tr}([A,B]\rho) + \langle\phi|C_{\pm}^{\dagger}\rho C_{\pm}|\phi\rangle + \sum_{i}^{d-1}\langle\phi_i^{\perp}|C_{\pm}^{\dagger}\rho C_{\pm}|\phi_i^{\perp}\rangle\,,$$

where $\{|\phi\rangle, \{|\phi_i^{\perp}\rangle\}_{i=1}^{d-1}\}$ is an orthonormal basis in $\mathcal{H}$. By discarding the summation term which is always a positive number in the above equation, we obtain the inequality (38). □

Notice that, the lower bound of Eq. (38) has different non-zero values depending on different choices of the post-selections $|\phi\rangle$. The inequality (38) becomes an equality when $|\phi\rangle \propto (A \pm iB - \langle A \pm iB\rangle_{\rho} I)\,|\psi\rangle$, where $\rho = |\psi\rangle\langle\psi|$ is a pure state. In the references [61, 62, 74], the lower bound of the sum uncertainty relation depends on the state orthogonal to the initial pure state, and if the initial state is a mixed state, then the lower bound can not always be computed at least for the full rank density matrix. The reason is that we can not find a state which is orthogonal to all the eigenstates of a full rank density matrix. Moreover, if the initial density matrix is unknown then computing the lower bound will be hard. In contrast, Eq. (38) doesn't have such issues as the first and second terms in right hand side of Eq. (38) are the average values of the hermitian operators $i[A,B]$ and $(A\pm iB - \langle A\pm iB\rangle_{\rho} I)\,|\phi\rangle\langle\phi|\,(A\mp iB - \langle A\mp iB\rangle_{\rho} I)$ in the state ρ, respectively, where $\langle A\pm iB\rangle_{\rho} = \langle A\rangle_{\rho} \pm i\,\langle B\rangle_{\rho}$. All of them can be obtained in experiments even if ρ is unknown.

E. Tighter upper bound for out-of-time-order correlators

Recently Bong *et al.* [75] used the RHUR for unitary operators to give upper bound for out-of-time-order correlators (OTOC) which is defined by $F = Tr[(W_t^{\dagger}V^{\dagger}W_tV)\rho]$, where V and W_t are fixed and time dependent unitary operators, respectively. The OTOC diagnoses the spread of quantum information by measuring how quickly two commuting operators V and W fail to commute, which is quantified by $\langle|[W_t,V]|^2\rangle_{\rho} = 2(1 - Re[F])$, where $|X|^2 = X^{\dagger}X$. The OTOC has strong connection with chaos and information scrambling [76–78] and also with high energy physics [79–82]. It is known that OTOC's upper bound is essential for limiting how quickly many-body entanglement can generate [79–81]. The standard upper bound for modulus of the OTOC given by Bong *et al.* [75] is $|F| \leq cos(\theta_{VW_t} - \theta_{W_tV})$, where $\theta_{VW_t} = cos^{-1}|Tr(\rho VW_t)|$, $\theta_{W_tV} = cos^{-1}|Tr(\rho W_tV)|$.

Here, we show that uncertainty relation in PPS system for unitary operators can be used to derive tighter upper bound for the OTOC.

Theorem 5. *Let $\rho \in \mathcal{L}(\mathcal{H})$ be the system's state and $|\phi\rangle$ be any arbitrary state, then modulus of the OTOC $F = Tr[(W_t^{\dagger}V^{\dagger}W_tV)\rho]$ for fixed and time dependent unitary operators V, $W_t \in \mathcal{L}(\mathcal{H})$, respectively is upper bounded by*

$$|F| = |\,\langle W_t^{\dagger}V^{\dagger}W_tV\rangle\,| \leq cos(\theta^{\phi}_{VW_t} - \theta^{\phi}_{W_tV}), \quad (42)$$

where $\theta^{\phi}_{VW_t} = cos^{-1}||\sqrt{\rho}(VW_t)^{\dagger}\,|\phi\rangle\,||$ and $\theta^{\phi}_{W_tV} = cos^{-1}||\sqrt{\rho}(W_tV)^{\dagger}\,|\phi\rangle\,||$. Here, $||.||$ defines a vector norm.

Proof. For a given mixed state ρ and arbitrary state $|\phi\rangle$ which we consider to be pre- and post-selections, respectively, the standard deviation $\langle\Delta X\rangle^{\phi}_{\rho}$ of any operator X in the PPS system is defined as $\left(\langle\Delta X\rangle^{\phi}_{\rho}\right)^2 = Tr(XX^{\dagger}\rho) - \langle\phi|X^{\dagger}\rho X|\phi\rangle = Tr\left((\sqrt{\rho}X_0^{\phi})^{\dagger}\sqrt{\rho}X_0^{\phi}\right) = ||\sqrt{\rho}X_0^{\phi}||_F^2$, where $X_0^{\phi} = X -$

$X|\phi\rangle\langle\phi|$ and $||A||_F = \sqrt{Tr(A^\dagger A)}$ denotes the Frobenius norm of the operator A. When X is a hermitian operator, $\langle\Delta X\rangle_\rho^\phi$ becomes the standard deviation of X defined in Eq. (28). Now consider X to be unitary operators U and V. So, we can derive uncertainty relation for two unitary operators U and V using the Cauchy-Schwarz inequality for operators with Frobenius norm when the system is in pre- and post-selections ρ and $|\phi\rangle$, respectively as

$$\begin{aligned}\langle\Delta U\rangle_\rho^\phi \langle\Delta V\rangle_\rho^\phi &\geq \left|Tr\left[(\sqrt{\rho}U_0^\phi)^\dagger\sqrt{\rho}V_0^\phi\right]\right| \\ &= \left|Tr(VU^\dagger\rho) - \langle\phi|U^\dagger\rho V|\phi\rangle\right|, \end{aligned} \tag{43}$$

where $\langle\Delta U\rangle_\rho^\phi = \sqrt{1-\langle\phi|U^\dagger\rho U|\phi\rangle}$ and similarly for V also. Now, by replacing $U \to V^\dagger W_t^\dagger$ and $V \to W_t^\dagger V^\dagger$, (43) becomes

$$\begin{aligned}&|Tr(W_t^\dagger V^\dagger W_t V\rho)| \\ &\leq |\langle\phi|(V^\dagger W_t^\dagger)^\dagger\rho W_t^\dagger V^\dagger|\phi\rangle| + \langle\Delta(V^\dagger W_t^\dagger)\rangle_\rho^\phi \langle\Delta(W_t^\dagger V^\dagger)\rangle_\rho^\phi \\ &\leq ||\sqrt{\rho}V^\dagger W_t^\dagger|\phi\rangle||||\sqrt{\rho}W_t^\dagger V^\dagger|\phi\rangle|| \\ &\quad + \sqrt{1-||\sqrt{\rho}V^\dagger W_t^\dagger|\phi\rangle||^2}\sqrt{1-||\sqrt{\rho}W_t^\dagger V^\dagger|\phi\rangle||^2}, \end{aligned} \tag{44}$$

where we used the Cauchy-Schwarz inequality for vectors and $\langle\Delta(V^\dagger W_t^\dagger)\rangle_\rho^\phi = \sqrt{1-||\sqrt{\rho}V^\dagger W_t^\dagger|\phi\rangle||^2}$ and $\langle\Delta(W_t^\dagger V^\dagger)\rangle_\rho^\phi = \sqrt{1-||\sqrt{\rho}W_t^\dagger V^\dagger|\phi\rangle||^2}$, where $||\,|\chi\rangle\,|| = \sqrt{\langle\chi|\chi\rangle}$ denotes vector norm.

Now, by setting $||\sqrt{\rho}(VW_t)^\dagger|\phi\rangle|| = cos\theta_{VW_t}^\phi$ and $||\sqrt{\rho}(W_tV)^\dagger|\phi\rangle|| = cos\theta_{W_tV}^\phi$ in (44), the inequality (42) is proved. □

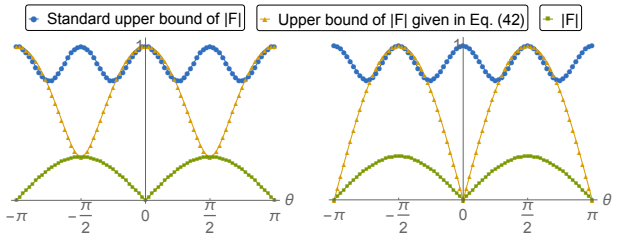


Fig. 2. For both the figures, the blue curve is the standard upper bound for $|F|$ given by Bong *et al.* [75] and the green curve is $|F|$. We have considered $V = \sigma_z$ and $W_t = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}$ for a fixed time. Initially prepared state is $|\psi\rangle = cos(\theta/2)|0\rangle + e^{i\pi/11}sin(\theta/2)|1\rangle$. Now in the left figure, the orange curve is the upper bound of $|F|$ given in Eq. (42) when the post-selection is $|\phi_1\rangle = cos(\pi/2)|0\rangle + e^{i\pi/2}sin(\pi/2)|1\rangle$. In the right figure, the orange curve is the upper bound of $|F|$ given in the Eq. (42) when the post-selection is $|\phi_2\rangle = cos(\pi/4)|0\rangle + e^{i\pi/2}sin(\pi/4)|1\rangle$. Here for two (or more) different post-selections, it is clearly seen that the upper bound given in Eq. (42) is tighter than the standard upper bound given by Bong *et al.* [75].

In Fig. 2, it is shown that by suitably choosing $|\phi\rangle$, the upper bound of $|F|$ in Eq. (42) can be made tighter than the standard upper bound given by Bong *et al.* [75]. Hence, we conclude that the tighter upper bound for the modulus of the OTOC is

$$|F| \leq min\left\{\min_\phi\{cos(\theta_{VW_t}^\phi - \theta_{W_tV}^\phi)\}, cos(\theta_{VW_t} - \theta_{W_tV})\right\}.$$

V. CONCLUSION

We have defined standard deviation of an observable in a PPS system, interpreted it geometrically as well as informationally from the perspective of weak PPS measurements and subsequently derived the Robertson-Heisenberg like uncertainty relation for two non commuting observables. Such uncertainty relations in PPS system impose limitations on the joint sharp preparation of pre- and post-selected states for two incompatible observables. We provided the necessary and sufficient condition for zero uncertainty of an observable and show its usefulness in achieving optimized Fisher information in quantum metrology. We have derived both product and sum uncertainty equalities from which a series of uncertainty inequalities can be obtained. The generalization of uncertainty relation for mixed pre- selection in PPS system has also been discussed. We have demonstrated that the PPS system can exhibit more bizarre behaviors than the usual ones. For instances, it is possible in PPS system that measurement of two compatible observables can disturb each other's measurement results i.e., the lower bound in the uncertainty relation can be made non zero by suitably choosing post-selections. A similar property in PPS system was first shown by Vaidman [50]. It is also possible that a quantum state (pre-selection) can be prepared in a PPS system for which both of the standard deviations of incompatible observables are zero although this is not possible in a standard quantum system (see section III B).

The standard deviation and uncertainty relation in the PPS system have been used to provide physical applications. (i) We have used two different definitions of the standard deviations in the PPS system to detect purity of an unknown state. (ii) The uncertainty relation in the PPS system is used to derive the stronger uncertainty relation (i.e., nontrivial for all possible choices of initially prepared states) in the standard quantum system. For two dimensional quantum system, the stronger uncertainty relation by Maccone-Pati [62] fails to provide the informations about the incompatible observables when the system state is an eigenstate of either observables. We have shown that our stronger uncertainty relation overcomes this shortcoming of Maccone-Pati uncertainty relation. (iii) Since the variance in the PPS system remains unaffected (i.e., neither increases nor decreases) by the classical mixing of quantum states, we have concluded that the uncertainty relation in the PPS system is a purely quantum uncertainty relation. In contrast, variance in the standard system increases in general under the classical mixing of quantum states. Following this observation we have provided a measure of classical uncertainty whose less value implies more purely quan-

tum uncertainty. (iv) Tighter sum uncertainty relation in the standard quantum system has been derived where the tightness depends on the post-selection. (v) Uncertainty relation in PPS system for two unitary operators has been used to provide tighter upper bound for out-of-time-order correlators.

Future directions: (i) it will be interesting if the global minimum for sum of uncertainties of non-commuting observables in the PPS system exists because that can be used to detect entanglement by suitably choosing post-selections, Similar to the work by Hofmann and Takeuchi [7]. (ii) Applications and implications of the ideas like 'zero uncertainty' and 'joint sharp preparation of a quantum state for non-commuting observables' need more attention. (iii) Tt is a matter of further study if the uncertainty relation (17) in PPS systems has applications similar to the RHUR (8), such as quantum metrology, spin squeezing, improving the accuracy of phase measurement in quantum interferometers, etc. (iv) We have derived the condition for the "intelligent pre- and post-selected states" to achieve the minimum bound of the uncertainty relation in the PPS system and intelligent pre- and post-selected states can be exploited to get highly precise phase measurements because many theoretical and experimental efforts have been made in recent years involving the minimum uncertainty states (for which the RHUR saturates) and the spin-squeezing states in the standard quantum systems (see, for example, [52, 56, 60]) for precise phase measurements.

Acknowledgment: We would like to thank Klaus Mølmer and David R. M. Arvidsson-Shukur for bringing the references [14, 15, 48, 49] and [45], respectively to our attention.

VI. APPENDICES

Appendix A

Here we derive the condition for which the inequality (17) saturates. In the Cauchy-Schwarz inequality Eq. (18), the remainder and the real term to be vanished for the equality condition of Eq. (17) i.e.,

$$|\widetilde{\phi}^{\perp}_{A\psi}\rangle - \frac{\langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{A\psi}\rangle}{\langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{B\psi}\rangle}|\widetilde{\phi}^{\perp}_{B\psi}\rangle = 0, \tag{A1}$$

$$\langle\widetilde{\phi}^{\perp}_{A\psi}|\widetilde{\phi}^{\perp}_{B\psi}\rangle + \langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{A\psi}\rangle = 0. \tag{A2}$$

Now take the inner product between $\langle\phi^{\perp}_{A\psi}|$ and Eq. (A1), and use the condition (A2), then we have

$$\langle\widetilde{\phi}^{\perp}_{A\psi}|\widetilde{\phi}^{\perp}_{A\psi}\rangle + \frac{\left(\langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{A\psi}\rangle\right)^2}{\langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{B\psi}\rangle} = 0. \tag{A3}$$

Now using $\langle\widetilde{\phi}^{\perp}_{X\psi}|\widetilde{\phi}^{\perp}_{X\psi}\rangle = \left(\langle\Delta X\rangle^{\phi}_{\psi}\right)^2$, where X={A,B}; the Eq. (A3) becomes

$$\langle\widetilde{\phi}^{\perp}_{B\psi}|\widetilde{\phi}^{\perp}_{A\psi}\rangle = \pm i\,\langle\Delta A\rangle^{\phi}_{\psi}\,\langle\Delta B\rangle^{\phi}_{\psi}\,. \tag{A4}$$

Finally, use Eqs. (A4) and (12) in Eq. (A1) to obtain the condition (23).

Appendix B

To show that the uncertainty relation (17) or (21) is also valid for mixed pre- selected state ρ, we consider the following operator

$$T = A^{\phi}_0 + (\gamma + i\epsilon)B^{\phi}_0, \tag{B1}$$

where $A^{\phi}_0 = A - A\,|\phi\rangle\,\langle\phi|$ and $B^{\phi}_0 = B - B\,|\phi\rangle\,\langle\phi|$. And γ, ϵ are some real parameters. Now for any operator T, the inequality

$$Tr(\rho TT^{\dagger}) \geq 0, \tag{B2}$$

holds. Using Eq. (B1), we have

$$\begin{aligned} Tr(\rho TT^{\dagger}) = &\left(\langle\Delta A\rangle^{\phi}_{\rho}\right)^2 + \left(\gamma^2+\epsilon^2\right)\left(\langle\Delta B\rangle^{\phi}_{\rho}\right)^2 \\ &+ \gamma\left(\langle\{A,B\}\rangle_{\rho} - 2\mathcal{R}eW_{AB}\right) \\ &- i\epsilon\left(\langle[A,B]\rangle_{\rho} - 2\mathcal{I}mW_{AB}\right) \geq 0, \end{aligned} \tag{B3}$$

where $\left(\langle\Delta A\rangle^{\phi}_{\rho}\right)^2 = Tr(\rho A^{\phi}_0 A^{\phi\dagger}_0)$ is defined in Eq. (28), $\langle[A,B]\rangle_{\rho} = Tr(\rho[A,B])$, $\langle\{A,B\}\rangle_{\rho} = Tr(\rho\{A,B\})$ and $W_{AB} = Tr(\Pi_{\phi}B\rho A)$. Now one finds the quantity $Tr(\rho TT^{\dagger})$ is minimum for $\gamma = -\frac{\langle\{A,B\}\rangle_{\rho} - 2\mathcal{R}eW_{AB}}{2\left(\langle\Delta B\rangle^{\phi}_{\rho}\right)^2}$ and $\epsilon = \frac{i(\langle[A,B]\rangle_{\rho} - 2i\mathcal{I}mW_{AB})}{2\left(\langle\Delta B\rangle^{\phi}_{\rho}\right)^2}$. Hence, $min_{\gamma,\epsilon}Tr(\rho TT^{\dagger}) \geq 0$ becomes

$$\begin{aligned} \left(\langle\Delta A\rangle^{\phi}_{\rho}\right)^2\left(\langle\Delta B\rangle^{\phi}_{\rho}\right)^2 \geq &\left[\frac{1}{2i}\langle[A,B]\rangle_{\rho} - \mathcal{I}mW_{AB}\right]^2 \\ &+ \left[\frac{1}{2}\langle\{A,B\}\rangle_{\rho} - \mathcal{R}eW_{AB}\right]^2. \end{aligned} \tag{B4}$$

By discarding the second term which is a positive number in the right hand side of Eq. (B4), the uncertainty relation (32) is achieved.

Appendix C

Here we show the proofs of all the *Lemmas* to detect mixedness of an unknown state in qubit, qutrit, qubit-qubit and qubit-qutrit systems. Let us recall the mathematical expression of the statement of the *Proposition 3* which is given by

$$\langle\Delta A_w\rangle^{\phi}_{\rho} \geq \langle\Delta A\rangle^{\phi}_{\rho}\,. \tag{C1}$$

In the following, we will use Eq. (C1) to prove all the *Lemmas*.

The general form of a mixed state is $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$ and the condition for which the equality of Eq. (C1) holds is (see proof of the *Proposition 3*)

$$\langle\phi|A|\psi_i\rangle = \lambda \langle\phi|\psi_i\rangle , \tag{C2}$$

where 'λ' is some constant which depends on the index of $|\phi\rangle$ (e.g., for $|\phi_k\rangle$, it is λ_k).

***The proof of* Lemma 1:**

Proof. We first assume that each $|\psi_i\rangle$ is distinct and hence from Eq. (C2), we have a set of equations

$$\langle\phi| (A - \lambda I) |\psi_i\rangle = 0, \tag{C3}$$

for each $|\psi_i\rangle$. Denote the unnormalized state vector $|\widetilde{\phi}_A^{\lambda}\rangle = (A - \lambda I)|\phi\rangle$. As $|\widetilde{\phi}_A^{\lambda}\rangle$ is a unnormalized state vector different from $|\phi\rangle$ and the $|\psi_i\rangle$, $\forall i$ are orthogonal to $|\widetilde{\phi}_A^{\lambda}\rangle$, it implies that $\{|\psi_i\rangle\}_{i=1}$ are confined in one dimensional Hilbert space. Hence each $|\psi_i\rangle$ is the same initially prepared state that is ρ is a pure state in a qubit system. □

***The proof of* Lemma 2:**

Proof. The qubit argument can not be generalized for the higher dimensional systems. The reason is simply because in three dimensional Hilbert space (for example) all the $|\psi_i\rangle$ can be confined in a two dimensional subspace of the Hilbert space which is orthogonal to $|\widetilde{\phi}_A^{\lambda}\rangle$. To make an "if and only if" condition, we consider the orthogonal basis $\{|\phi_k\rangle\}_{k=1}^3$ as valid post-selections. Here valid post-selections are those post-selections for which weak values are defined.

As there are three post-selections in three dimensional Hilbert space, we have three sets of equations like (C2) for the equality of the inequality (C1)

$$\{\langle\phi_1| (A - \lambda_1 I) |\psi_i\rangle = 0\}_{i=1}, \tag{C4}$$
$$\{\langle\phi_2| (A - \lambda_2 I) |\psi_i\rangle = 0\}_{i=1}, \tag{C5}$$
$$\{\langle\phi_3| (A - \lambda_3 I) |\psi_i\rangle = 0\}_{i=1}. \tag{C6}$$

Now, there are three possibilities which is implied by (C4), (C5) and (C6):
(i) The state vectors $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle = (A - \lambda_k I)|\phi_k\rangle\}_{k=1}^3$ span the whole 3-dimensional Hilbert space $\mathcal{H}$,
(ii) $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle\}_{k=1}^3$ span a 2-dimensional Hilbert space $\mathcal{H}$,
(iii) $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle\}_{k=1}^3$ span a 1-dimensional Hilbert space $\mathcal{H}$.

Below, we will show that possibility-(i) is discarded naturally whereas to discard possibility-(iii), we need a condition on observable A and post-selection $|\phi\rangle$. Then, possibility-(ii) will automatically indicate that all the $\{|\psi_i\rangle\}_{i=1}$ are same i.e., ρ is pure.

To start with possibility-(i), let's assume that possibility-(i) is true, then $\{|\psi_i\rangle\}_{i=1}$ has to be orthogonal to $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle\}_{k=1}^3$ according to (C4), (C5) and (C6) implying $|\psi_i\rangle = 0\ \forall\ i$ i.e., $\rho = 0$. So we discard this possibility.

Possibility-(iii) implies

$$\mathcal{N}_1(A - \lambda_1 I)|\phi_1\rangle = \mathcal{N}_2(A - \lambda_2 I)|\phi_2\rangle = \mathcal{N}_3(A - \lambda_3 I)|\phi_3\rangle , \tag{C7}$$

along z-axis (for example) and hence $\{|\psi_i\rangle\}_{i=1}$ span 2-dimensional xy-plane (see Fig. 4). Here $\mathcal{N}_k$ are nomalization constants. Now the inner product of (C7) with $|\phi_1\rangle$, $|\phi_2\rangle$ and $|\phi_3\rangle$, respectively gives

$$\mathcal{N}_1(\langle\phi_1|A|\phi_1\rangle - \lambda_1) = \mathcal{N}_2 \langle\phi_1|A|\phi_2\rangle = \mathcal{N}_3 \langle\phi_1|A|\phi_3\rangle , \tag{C8}$$
$$\mathcal{N}_1 \langle\phi_2|A|\phi_1\rangle = \mathcal{N}_2(\langle\phi_2|A|\phi_2\rangle - \lambda_2) = \mathcal{N}_3 \langle\phi_2|A|\phi_3\rangle , \tag{C9}$$
$$\mathcal{N}_1 \langle\phi_3|A|\phi_1\rangle = \mathcal{N}_2 \langle\phi_3|A|\phi_2\rangle = \mathcal{N}_3(\langle\phi_3|A|\phi_3\rangle - \lambda_3). \tag{C10}$$

Now, the particular choice

$$\langle\phi_1|A|\phi_2\rangle = 0 \tag{C11}$$

implies that Eq. (C8) and (C9) do not hold if $\langle\phi_1|A|\phi_3\rangle \neq 0$ and $\langle\phi_2|A|\phi_3\rangle \neq 0$, respectively. If either of Eq. (C8) and (C9) does not hold then possibility-(iii) is discarded. But, if $\langle\phi_1|A|\phi_3\rangle = 0$ and $\langle\phi_2|A|\phi_3\rangle = 0$, then we have to proceed further. Note that, by setting $\langle\phi_1|A|\phi_2\rangle = 0$ from Eq. (C11), $\langle\phi_1|A|\phi_3\rangle = 0$ and $\langle\phi_2|A|\phi_3\rangle = 0$ in Eqs. (C8), (C9) and (C10), we have

$$\lambda_k = \langle\phi_k|A|\phi_k\rangle \qquad for \qquad k = 1, 2, 3. \tag{C12}$$

Now, it is easy to see that with the values of λ_k from Eq. (C12), $\{\langle\phi_k|\widetilde{\phi}_{kA}^{\lambda_k}\rangle = 0\}_{k=1}^3$ holds. This implies that $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle\}_{k=1}^3$ can not be confined in one dimensional Hilbert space i.e., along a particular axis and in our assumption it is the z-axis. But according to Eq. (C7), $\{|\widetilde{\phi}_{kA}^{\lambda_k}\rangle\}_{k=1}^3$ are along the z-axis. Hence it shows the contradiction and we discard the possibility-(iii) when the condition $\langle\phi_1|A|\phi_2\rangle = 0$ is considered.

Finally the possibility-(ii) implies that $\{|\psi_i\rangle\}_{i=1}$ must be spanned in 1-dimensional Hilbert space $\mathcal{H}$ that is, each $|\psi_i\rangle$ is the same initially prepared state which is a pure state.

So, we conclude that if for an observable A and a complete orthonormal basis $\{|\phi_k\rangle\}_{k=1}^3$ (to be used as post-selected states) of any three-level quantum system (i.e., a qutrit), the condition $\langle\phi_1|A|\phi_2\rangle = 0$ is considered, then equality in Eq. (C1) holds good if and only if the pre- selected state ρ is pure.

□

***The proof of* Lemma 3:**

Proof. For this bipartite system, we consider the observable and the post-selection to be $A \otimes I$ and $|\phi_{AB}\rangle = |\phi_A\rangle |\phi_B\rangle$, respectively. The standard deviations defined in Eq. (28) and

(30) for the given bipartite state ρ become

$$\left(\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi_{AB}}\right)^2 = Tr[(A\otimes I)^2\rho] - \frac{|\langle\phi_{AB}|(A\otimes I)\rho|\phi_{AB}\rangle|^2}{\langle\phi_{AB}|\rho|\phi_{AB}\rangle} = Tr[A^2\rho_A] - \frac{|\langle\phi_A|A\rho_A^{\phi_B}|\phi_A\rangle|^2}{\langle\phi_A|\rho_A^{\phi_B}|\phi_A\rangle}, \quad \text{(C13)}$$

$$\left(\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}\right)^2 = Tr[(A\otimes I)^2\rho] - \langle\phi_{AB}|(A\otimes I)\rho(A\otimes I)|\phi_{AB}\rangle = Tr[A^2\rho_A] - \langle\phi_A|A\rho_A^{\phi_B}A|\phi_A\rangle, \quad \text{(C14)}$$

respectively, where $\rho_A^{\phi_B} = \langle\phi_B|\rho|\phi_B\rangle$ is the collapsed density operator of the subsystem A when a projection operator $\Pi_{\phi_B} = |\phi_B\rangle\langle\phi_B|$ is measured in the subsystem B. In a qubit-qubit system, the subsystem A is two dimensional and hence $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi_{AB}}$ from Eq. (C13) and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}$ from Eq. (C14) are equal "if and only if" $\rho_A^{\phi_B}$ is pure. Now, $\rho_A^{\phi_B}$ being pure can be from ρ being both pure and mixed. If ρ is pure then $\rho_A^{\phi_B}$ is always pure but if ρ mixed, then it is easy to see that $\rho_A^{\phi_B}$ is pure only when $\rho = \sum_{i=1}^2 p_i|\psi_A^i\rangle\langle\psi_A^i|\otimes|\phi_B^i\rangle\langle\phi_B^i|$, where $|\phi_B^1\rangle = |\phi_B\rangle$ and $\sum_{i=1}^2|\phi_B^i\rangle\langle\phi_B^i| = I$. So, let us consider another post-selection $|\phi'_B\rangle$ (which is not orthogonal to $\{|\phi_B^i\rangle\}_{i=1}^2$) and if we find $\rho_A^{\phi'_B}$ to be pure which is equivalent to the equality of $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi'_{AB}}$ and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi'_{AB}}$, then we are sure that the bipartite state ρ is a pure state (due to the virtue of qubit system discussed above), where $|\phi'_{AB}\rangle = |\phi_A\phi'_B\rangle$.

So, here is the conclusion: Consider any two non orthogonal post-selections $|\phi_B\rangle$ and $|\phi'_B\rangle$ in the subsystem B. For any observable A, equality of $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi_{AB}}$ and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}$ and separately of $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi'_{AB}}$ and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi'_{AB}}$ hold *only* when the $2\otimes 2$ pre- selected state ρ is pure. □

The proof of **Lemma 4:**

Proof. The treatment above with the condition of the qutrit system, we have the conclusion: if for an observable A and any complete orthonormal basis $\{|\phi_A^k\rangle\}_{k=1}^3$ (to be used as post-selected states) for a qutrit, the condition $\langle\phi_A^3|A|\phi_A^1\rangle = 0$ is considered, then equality of $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi_{AB}}$ and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi_{AB}}$ and separately of $\langle\Delta(A\otimes I)_w\rangle_\rho^{\phi'_{AB}}$ and $\langle\Delta(A\otimes I)\rangle_\rho^{\phi'_{AB}}$ hold if and only if the $3\otimes 2$ pre-selected state ρ is pure, where $|\phi'_{AB}\rangle = |\phi_A\phi'_B\rangle$. □

[1] W. Heisenberg, Z. Phys. 43, 172 (1927).
[2] J. A. Wheeler and H. Zurek, "Quantum Theory and Measurement", (Princeton University Press, Princeton, NJ, 1983), p. 2.
[3] H. P. Robertson, Phys. Rev. 34, 163 (1929).
[4] Consider a spin-1/2 system where we are interested in measuring σ_x and σ_z observables. Here, σ_x with eigenstates $|\uparrow_x\rangle$,$|\downarrow_x\rangle$ and σ_z with eigenstates $|0\rangle$,$|1\rangle$ are well known Pauli operators. In order to measure them, we have to prepare the system in some spin state. Let the state be $|\uparrow_x\rangle$. Now if we measure σ_x, the measurement result will be '+1' with eigenstate $|\uparrow_x\rangle$ only. We say that the preparation of the the quantum state is sharp for σ_x as the number of outcomes is only one. While the measurement results for σ_z when the system is prepared in the same state $|\uparrow_x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ are '+1' and '-1' with eigenstates $|0\rangle$ and $|1\rangle$ respectively. That is the preparation of the state $|\uparrow_x\rangle$ is not sharp for σ_z. So even if the state $|\uparrow_x\rangle$ can be prepared sharply for σ_x, we can not prepare the state $|\uparrow_x\rangle$ sharply for σ_z. This is because σ_x and σ_z do not have same eigenstates i.e., they are non-commuting (incompatible) observables. If both the observables are not sharp, then the RHUR gives a bound it puts on the sharp preparation of a quantum state for two non-commuting observables.
[5] P. Busch, T. Heinonen and P. Lahti, Phys.Rep.452,155(2007).
[6] P. J. Lahti and M. J. Maczynski, J. Math. Phys. (N.Y.) 28, 1764 (1987).
[7] Holger F. Hofmann and Shigeki Takeuchi, Phys. Rev. A 68, 032103 (2003).
[8] O. Guhne, Phys. Rev. Lett. 92, 117903 (2004).
[9] C. A. Fuchs and A. Peres, Phys. Rev. A 53, 2038 (1996).
[10] Michael J. W. Hall, Gen. Relativ. Gravit. 37, 1505 (2005).
[11] Y. Aharonov, P. G. Bergmann, and J. L. Lebowitz, Phys. Rev. B 134, 1410 (1964).
[12] Y. Aharonov and L. Vaidman, The Two-State Vector Formalism: An Updated Review, Time in Quantum Mechanics, 2008, Volume 734.
[13] A.G. Kofman, Sahel Ashhab, and Franco Nori, Physics Reports 520 (2012) 43–133.
[14] Søren Gammelmark, Brian Julsgaard, and Klaus Mølmer, Phys. Rev. Lett. 111, 160401 (2013).
[15] D. Tan, S.J. Weber, I. Siddiqi, K. Mølmer, and K.W. Murch, Phys. Rev. Lett. 114, 090403 (2015).
[16] Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60, 1351 (1988).
[17] J.S. Lundeen and K.J. Resch, Physics Letters A 334 (2005) 337–344.
[18] Richard Jozsa, Phys. Rev. A 76, 044103 (2007).
[19] Omar S. Magaña-Loaiza, M. Mirhosseini, B. Rodenburg, and R. W. Boyd, Phys. Rev. Lett. 112, 200401 (2014).
[20] O. Hosten and P. Kwiat, Science 319, 787 (2008).
[21] J. S. Lundeen, B. Sutherland, A. Patel, C. Stewart, and C. Bamber, Nature 474, 188–191(2011) .
[22] Jeff S. Lundeen and Charles Bamber, Phys. Rev. Lett. 108, 070402 (2012).
[23] G.S. Thekkadath, L. Giner, Y. Chalich, M.J. Horton, J. Banker, and J.S. Lundeen, Phys. Rev. Lett. 117, 120401 (2016).
[24] Y. Aharonov, S. Popescu, D. Rohrlich, and P. Skrzypczyk, 2013 New J. Phys. 15 113015.
[25] T. Denkmayr, H. Geppert, S. Sponar, H. Lemmel, A. Matzkin, J. Tollaksen, and Y. Hasegawa, 10.1038/ncomms5492.
[26] Debmalya Das and Arun Kumar Pati, New J. Phys. 22 063032 (2020).
[27] Sahil, Sohail, Subhrajit Modak, Sibasish Ghosh, and Arun Kumar Pati, arXiv:2107.00573 .
[28] WW. Pan, XY. Xu, Y. Kedem, QQ. Wang, Z. Chen, M. Jan, K.

Sun, JS. Xu, YJ. Han, CF. Li, and GC. Guo, Phys. Rev. Lett. 123, 150402(2019).
[29] Yusuf Turek, J. Phys. Commun. 4 075007(2020).
[30] Y. Aharonov, A. Botero, S. Popescu, B. Reznik, and J. Tollaksen, Phys. Lett. A 301 (2001) 130.
[31] J. S. Lundeen and A. M. Steinberg, Phys. Rev. Lett. 102, 020404 (2009).
[32] K. Yokota, T. Yamamoto, M. Koashi, and N. Imoto, New J. Phys. 11, 033011 (2009).
[33] Y. Aharonov and D. Rohrlich, Quantum Paradoxes: Quantum Theory for the Perplexed, Wiley–VCH.
[34] Y. Aharonov, S. Nussinov, S. Popescu, and L. Vaidman, Phys. Rev. A 87 (2012) 014105.
[35] D. R. Solli, C. F. McCormick, R. Y. Chiao, S. Popescu, and J. M. Hickmann, Phys. Rev. Lett. 92 (2004) 043601.
[36] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, Phys. Rev. Lett. 109, 100404 (2012).
[37] Arun Kumar, Pati, Uttam Singh, and Urbasi Sinha, Phys. Rev. A 92, 052120 (2015).
[38] G. Nirala, S. N. Sahoo, A. K. Pati, and U. Sinha, Phys. Rev. A 99, 022111 (2019).
[39] A. K. Pati, C. Mukhopadhyay, S. Chakraborty, and S. Ghosh, Phys. Rev. A 102, 012204 (2020).
[40] S. Goswami, S. Chakraborty, S. Ghosh, and A. S. Majumdar, Phys. Rev. A 99, 012327 (2019).
[41] Lior Goldenberg and Lev Vaidman, American of Physics 64, 1059 (1996).
[42] Leifer, M.S, Quantum Stud.: Math. Found. (2023).
[43] Q. Duprey and A. Matzkin, Phys. Rev. A 95, 032110 (2017).
[44] G. Bié Alves, B. M. Escher, R. L. de Matos Filho, N. Zagury, and L. Davidovich, Phys. Rev. A 91, 062107 (2015).
[45] D.R.M.A. Shukur, N. Y. Halpern, H. V. Lepage, A. A. Lasek, C. H. W. Barnes, and S. Lloyd, Nat. Commun. 11, 3775 (2020).
[46] Michael J. W. Hall, Phys. Rev. A 69, 052113 (2004).
[47] Arun Kumar Pati and Junde Wu, arXiv:1411.7218v1.
[48] H Bao, S. Jin, J. Duan, S. Jia, K. Mølmer, H. Shen, and Y. Xiao, Nat. Commun. 11, 5658 (2020).
[49] Maryam Khanahmadi and Klaus Mølmer, Phys. Rev. A 104, 022204 (2021).
[50] Lev Vaidman, Phys. Rev. Lett. 70, 3369 (1993).
[51] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, Phys. Rev. Lett. 96, 010401 (2006).
[52] Luca Pezzè, Augusto Smerzi, M. K. Oberthaler, Roman Schmied, and Philipp Treutlein, Rev. Mod. Phys. 90, 035005 (2018).
[53] M. Kitagawa and M. Ueda, Phys. Rev. A 47 (1993) 5138.
[54] D. J. Wineland, J. J. Bollinger, W. M. Itano, F. L. Moore, and D. J. Heinzen, Phys.Rev.A46(1992) R6797.
[55] D. J. Wineland, J. J. Bollinger, W. M. Itano, and D. J. Heinzen, Phys. Rev. A 50 (1994) 67.
[56] Jian Ma, Xiaoguang Wang, C.P. Sun, and Franco Nori, Physics Reports 509 (2011) 89–165.
[57] C. Aragone, G. Guerri, S. Salamo, and J. L. Tani, J. Phys. A 7, L149 (1974).
[58] C. Aragone, E. Chalbaud, and S. Salamo , J. Math. Phys. 17, 1963 (1976).
[59] D. A. Trifonov, J. Math. Phys. 35, 2297 (1994).
[60] Mark Hillery and Leonard Mlodinow, Phys. Rev. A 48, 1548 (1993).
[61] Yao Yao, Xing Xiao, Xiaoguang Wang, and C. P. Sun, Phys. Rev. A 91, 062113 (2015).
[62] Lorenzo Maccone and Pati K. Arun, Phys. Rev. Lett. 113, 260401 (2014).
[63] S. Mal, T. Pramanik, and A. S. Majumdar, Phys. Rev. A 87, 012105 (2013).
[64] Let $|\psi\rangle$ be the eigenstate of the observable A, then the first inequality in MPUR [62] becomes $\langle\Delta A\rangle_\psi^2 + \langle\Delta B\rangle_\psi^2 \geq \pm i\langle\psi|[A,B]|\psi\rangle + \left|\langle\psi|(A \pm iB)|\psi^\perp\rangle\right|^2 = \left|\langle\psi|B|\psi^\perp\rangle\right|^2 = \langle\Delta B\rangle_\psi^2$. Here we have used the fact that in 2D, arbitrary states $|\psi^\perp\rangle$ orthogonal to $|\psi\rangle$ and $|\psi_B^\perp\rangle$ are same upto a phase factor and hence $\left|\langle\psi|B|\psi^\perp\rangle\right| = \left|\langle\psi_B^\perp|B|\psi\rangle\right| = \langle\Delta B\rangle_\psi$ (see Eq. (4)). So, the inequality becomes $\langle\Delta B\rangle_\psi^2 \geq \langle\Delta B\rangle_\psi^2$.
[65] The second inequality in MPUR [62] is given by $\langle\Delta A\rangle_\psi^2 + \langle\Delta B\rangle_\psi^2 \geq \frac{1}{2}\left|\langle\psi_{A+B}^\perp|(A+B)|\psi\rangle\right|^2$, where $|\psi_{A+B}^\perp\rangle = (1/\langle\Delta(A+B)\rangle_\psi)(A+B-\langle A+B\rangle_\psi)|\psi\rangle$ and $\langle\Delta(A+B)\rangle_\psi^2 = \langle(A+B)^2\rangle_\psi - \langle A+B\rangle_\psi^2$. If $|\psi\rangle$ is the eigenstate of the observable A, then $\langle\Delta(A+B)\rangle_\psi = \langle\Delta B\rangle_\psi$ and using the expression of $|\psi_{A+B}^\perp\rangle$ in the above inequality, we have $\langle\Delta B\rangle_\psi^2 \geq \frac{1}{2}\langle\Delta B\rangle_\psi^2$.
[66] S.Luo and Q.Zhang, EEE Trans. Inform. Theory 50 (2004) 1778–1782.
[67] E.P. Wigner and M.M. Yanase, roc. Natl. Acad. Sci. USA 49 (1963) 910–918.
[68] K. Yanagi, S. Furuichi, and K. Kuriyama, EEE Trans. Inform. Theory 51 (2005) 4401–4404.
[69] Shunlong Luo, Phys. Rev. A 72, 042110 (2005) .
[70] D. Li, X. Li, F. Wang, H. Huang, X. Li, and L. C. Kwek, Phys. Rev. A 79, 052106 (2009).
[71] S. Furuichi, K. Yanagi, and K. Kuriyama, J. Math. Anal. Appl. 356 (2009) 179–185.
[72] Kenjiro Yanagi, 2010 J. Phys.: Conf. Ser. 201 012015.
[73] Kenjiro Yanagi, J. Math. Anal. Appl. 365 (2010) 12–18.
[74] Qiu-Cheng Song and Cong-Feng Qiao, Physics Letters A 380 (2016) 2925–2930.
[75] KW. Bong, N. Tischler, R. B. Patel, S. Wollmann, G. J. Pryde, and M. J. W. Hall, Phys. Rev. Lett. 120, 230402 (2018).
[76] Brian Swingle, Nature Physics volume 14, pages 988–990 (2018).
[77] Brian Swingle, Gregory Bentsen, Monika Schleier-Smith, and Patrick Hayden, Phys. Rev. A 94, 040302(R) (2016).
[78] N. Y. Halpern, A. Bartolotta, and J. Pollack, Commun. Phys. 2, 92 (2019).
[79] Yasuhiro Sekino and L. Susskind, JHEP10(2008)065.
[80] Lashkari, N., Stanford, D., Hastings, M., Osborne T., and Hayden P., J. High Energ. Phys. 2013, 22 (2013).
[81] Maldacena, J., Shenker, S.H., and Stanford, D., J. High Energ. Phys. 2016, 106 (2016).
[82] Hosur P., Qi XL., Roberts D.A., and Yoshida B., J. High Energy Phys. 02 (2016) 4.

The shareability of steering in two-producible states

Qiu-Cheng Song[1] * Travis J. Baker[1] † Howard M. Wiseman[1] ‡

[1] *Centre for Quantum Computation and Communication Technology (Australian Research Council), Centre for Quantum Dynamics, Griffith University, Yuggera Country, Brisbane 4111, Australia*

Abstract. Quantum steering was originally introduced as the phenomenon whereby one party (Alice) can steer the quantum system of another party (Bob) into distinct ensembles of states by performing different measurements on her subsystem. Here, we investigate steering in a network scenario involving n parties, where the global quantum state shared between them is produced using only two-party entangled states $|\psi_\alpha\rangle$ and mixing with ancillary separable states. We introduce three scenarios which can be straightforwardly implemented on standard quantum optics architecture, which we call random $\frac{n}{2}$-pair entanglement, random pair entanglement and semi-random pair entanglement (SRPE). For example, the SRPE scenario is where, among n parties, a fixed party shares the entangled state with a random party, and other $n-2$ parties are prepared in single-qubit pure states. We derive analytically necessary and sufficient steering criteria for the states in the three scenarios under different measurement settings. Strikingly, using the SRPE construction, one party can steer any one of the $n-1$ other parties, for arbitrarily large n, using only two measurements. Then, exploiting symmetry, we study various small network configurations for three- and four-parties in the three scenarios, under different measurements and parameter α in the state $|\psi_\alpha\rangle$. Motivated by these results, we investigate whether the phenomenon of collective steering could be observed in the SRPE scenario, where two parties must cooperate in order to steer a third. This is known to have applications in quantum secret sharing schemes. Using semi-definite programming techniques, we find collective steering possible, and robust to noise.

Keywords: shareability, steering, small network, collective steering

We come to the conclusion by studying the shareability of steering in three different scenarios of 2-producible multipartite entangled qubit states, which we call $\frac{n}{2}$-pair entanglement (R$\frac{n}{2}$PE), random pair entanglement (RPE), and semi-random pair entanglement (SRPE). As well as considering steerability under all projective measurements, we consider more limited measurement strategies, for which we find the necessary and sufficient steering criterion analytically for the relevant class of reduced two-qubit states, which are two-qubit X-states. Most strikingly, in the SRPE scenario, where the n-qubit state can be produced from a single entangled pair of qubits plus $n-2$ product states, one party (Alice) can simultaneously steer all $n-1$ Bobs, for arbitrary n, using only two measurements. Finally, we study the properties of small networks in the three scenarios and collecive steering in SRPE. In this extended abstract, we show the most interesting scenario we called Semi-Random Pair Entanglement (SRPE).

SRPE is the scenario that among n parties a fixed party shares the entangled state $|\psi_\alpha\rangle = \sqrt{1-\alpha}|0\rangle|0\rangle + \sqrt{\alpha}|1\rangle|1\rangle$ with a random party and other parties are prepared in single-qubit pure state $|0\rangle$ each. Fig. 1 illustrates the examples $n = 3, 4$.

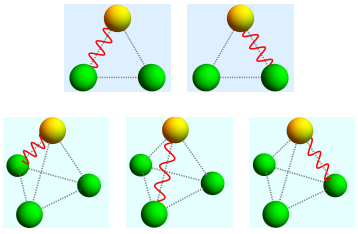


Figure 1: SRPE. e.g. $n = 3, 4$. The green ball represents an arbitrary party. The yellow denotes represents the fixed party, which randomly shares the entangled state $|\Psi_\alpha\rangle$ with an arbitrary party.

* songqiucheng19@gmail.com
† dr.travis.j.baker@gmail.com
‡ h.wiseman@griffith.edu.au

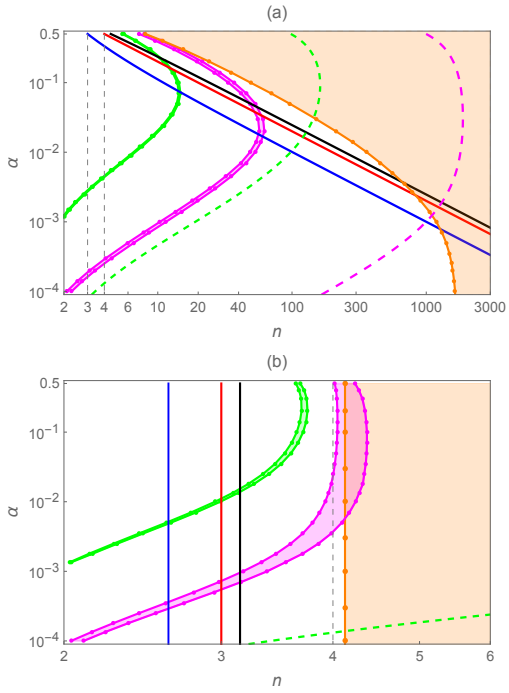


Figure 2: The steerability of the bipartite reduced SRPE state. (a) Steering from Alice to Bob. (b) Steering from Bob to Alice. The dashed green and dashed magenta curves denote the entanglement bounds with noise $\mu = 0.02, 0.002$, respectively. The blue, red, black lines represent the steering bounds for two, three and equatorial measurement schemes, respectively. The green and magenta joined dots denote the steering bounds (including lower bound and upper bound) with noise $\mu = 0.02, 0.002$ for all projective measurements, respectively. The orange joined dots denote the steering lower bound without noise for all projective measurements, but the upper bound cannot be found by the numerical method. The shaded region corresponds to the states where numerical imprecision prevents distinguishing whether the state is steerable or non-steerable.

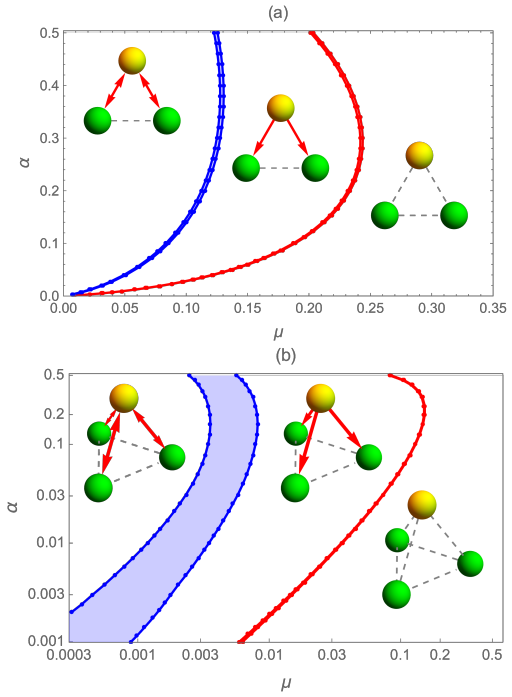


Figure 3: Small steering networks of SRPE state with noise μ. The measurement scheme (all projective measurements) is considered. Subfigure (a) shows network properties of the tripartite SRPE state with a linear scale on both axes, while subfigure (b) displaying the properties of the 4-partite SRPE state with a log scale on both axes.

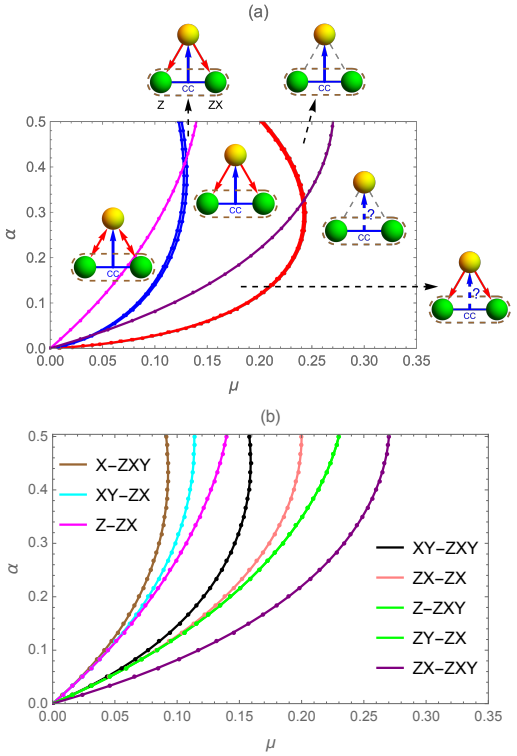


Figure 4: Collective steering of tripartite SRPE state with noise. (a) The blue and red curves represent one-way and two-way steering bounds, respectively, which are same as bounds in Fig. 3. The magenta and purple curves represent collective steering bounds which are obtained by making measurement strategies $Z-ZX$ (or $Z-ZY$) and $ZX-ZXY$ (or $ZY-ZXY$), respectively. The blue solid arrow denotes collective steering made by Bob1 and Bob2 with classical communication (CC). The blue dashed arrow with question mark indicates that the collective steering of this region is uncertain. (b) Various collective steering bounds obtained by different measurement strategies.

Towards provably optimal quantum error mitigation based on universal cost bounds

Kento Tsubouchi,[1] Takahiro Sagawa,[1,2] and Nobuyuki Yoshioka[1,3,4]

[1] *Department of Applied Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*
[2] *Quantum-Phase Electronics Center (QPEC), The University of Tokyo, Tokyo 113-8656, Japan*
[3] *Theoretical Quantum Physics Laboratory, RIKEN Cluster for Pioneering Research (CPR), Wako-shi, Saitama 351-0198, Japan*
[4] *JST, PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama, 332-0012, Japan*

I. MOTIVATION AND SUMMARY

One of the central problems in quantum technology is to establish control and understanding of unwanted noise, since an accumulation of errors may eventually spoil the practical advantage of quantum devices. Two representative approaches are quantum error correction and quantum error mitigation (QEM); the former performs measurement and feedback so that the logical information is constantly protected from external environment [1–5], while the latter is more oriented to offload the burden on the quantum device in a sense that one aims to mitigate the errors by constructing appropriate POVM elements via postprocessing. While a wide variety of QEM methods has been proposed and demonstrated in recent years [6–26], little is known about their fundamental aspects. For instance, in the context of QEC, it has been shown under local depolarizing noise that the noisy quantum state converges to the maximally mixed state exponentially with the depth L [27, 28], while the argument cannot be extended straightforwardly for QEM due to the variation of the circuit structures and the postprocessing allowed in the mitigation operations.

In this work, we address two outstanding questions: (i) the lowest sampling overhead required to perform unbiased QEM, (ii) provably optimal QEM method for sufficiently deep quantum circuits. Our strategy is to extend the applicability of quantum estimation theory, which has been argued to provide the lower bound on the sample complexity N to perform unbiased estimation on a traceless observable $\hat{X}$, given a noisy quantum state that undergo a single error channel as $\mathcal{E}(\rho)$ [29]. The technical contribution made in the present work is to bridge the quantum estimation theory and QEM by proposing a virtual quantum circuit such that (i) noisy quantum states may differ in noise strength or gate parameters (ii) noisy output state is generated from arbitrary depth of layered quantum circuit. Combining this with the quantum Cramér-Rao inequality, which states that the quantum Fisher information of the (noisy) quantum state relates the sample complexity N with the standard deviation ε of an unbiased estimator [30], we arrive at our main results stated as Theorem 1 and 2 in the following.

II. UNIVERSAL COST BOUNDS ON ERROR MITIGATION

Bounds in generic circuit.— By analyzing the quantum Fisher information of a generic layered noisy quantum circuit with depth L, we derive the following Theorem on the sample complexity for unbiased QEM:

Theorem 1. *Suppose that all noise channels are both injective and full-rank. Then, no matter how we construct the unbiased QEM that works for arbitrary quantum state, there is exponential growth with depth L in the number of copies N of the layered noisy quantum circuit required to perform unbiased estimation of $\langle \hat{X} \rangle$.*

For instance, consider the case where each layer of unitary gates is followed by the global depolarizing noise $\mathcal{E} : \rho \mapsto (1-p)\rho + pI/2^n$. By applying Theorem 1, we can show that the cost N, or the number of circuit execution of the layered noisy quantum circuit, required for the unbiased estimator of the expectation

value $\langle \hat{X} \rangle$ constructed from QEM shall satisfy

$$N \geq \frac{\mathrm{tr}[\hat{X}^2]}{2^n \varepsilon^2} \left(1 - (1-p)^L\right) \left(\frac{1}{1-p}\right)^{2L} \tag{1}$$

$$\sim \frac{\mathrm{tr}[\hat{X}^2]}{2^n \varepsilon^2} \left(\frac{1}{1-p}\right)^{2L}. \tag{2}$$

Strikingly, we can show that there exists a simple QEM technique that saturates Eq. (2) in the limit of large L. Namely, the global depolarizing noise affects the expectation value as $\langle \hat{X} \rangle^{\mathrm{noisy}} = (1-p)^L \langle \hat{X} \rangle$, and thus we achieve unbiased estimation by rescaling the measurement result as $(1-p)^{-L} \langle \hat{X} \rangle^{\mathrm{noisy}}$. Since the estimation variance on $\langle \hat{X} \rangle^{\mathrm{noisy}}$ approaches $2^{-n}\mathrm{tr}[\hat{X}^2]$ in the limit of large L, the sampling cost to estimate $\langle \hat{X} \rangle$ approaches $\frac{\mathrm{tr}[\hat{X}^2]}{2^n \varepsilon^2} \left(\frac{1}{1-p}\right)^{2L}$, which satisfies the lower bound of Theorem 1. As we numerically confirm in Fig. 1(a), the rescaling method achieves the lower bound which cannot be saturated by other unbiased methods such as the probabilistic error cancellation [7] or the generalized quantum subspace expansion [18].

Bounds in scrambling circuit.— It is noteworthy that the lower bound stated in Theorem 1 is for a generic layered quantum circuit. Since it also involves circuits that only weakly entangle qubits, the lower bound (e.g. Eq. (2)) does not depend on the qubit count n. However, if the quantum circuit scrambles the quantum state strong enough, we expect that every noise affects the measurement outcome; we must pay overhead to eliminate every local noise and thus encounter dependence on n. In fact, under local noise we can tighten the bound:

Theorem 2. *Let $U_1, U_2, ..., U_{L-1}, U_L$ be n-qubit unitary gate constructing the layered circuit, which are drawn from a set of random unitary that form unitary 2-design, and $\mathcal{E}_l$ be a local noise following the l-th unitary gate. Then, there is exponential growth with both qubit count n and depth L in the average over the number of copies N required to perform unbiased estimation of $\langle \hat{X} \rangle$ over $\{U_1, ..., U_L\}$.*

Concretely, if we consider the case of local depolarizing and amplitude damping noise, we can show that

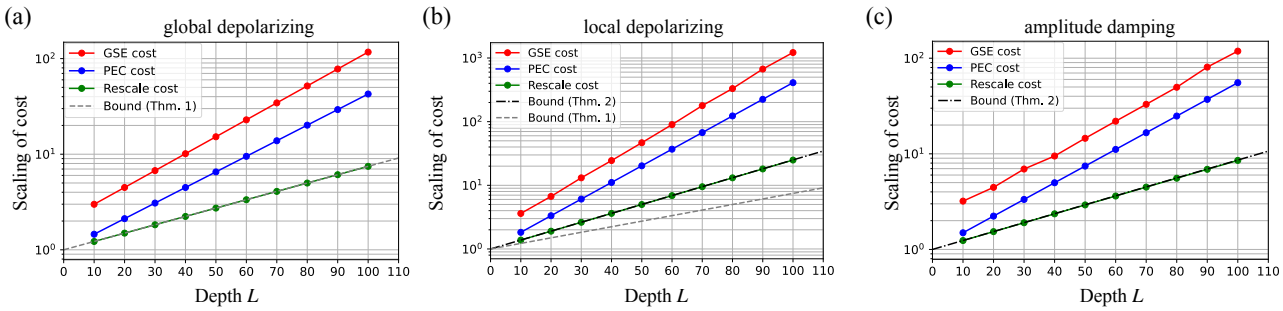


FIG. 1. Scaling of the cost to perform QEM methods for random Clifford circuit of $n = 2$ qubits under (a) global depolarizing noise, (b) local depolarizing noise, and (c) local amplitude damping noise with error rate $p = 0.01$. The red, blue, and green lines denote the sampling overhead of generalized subspace expansion [18] using power subspace, the probabilistic error cancellation as derived in Ref. [31], and the rescaling technique as explained in the main text. The rescaling factor is $(1-p)^{-L}$ and $(1-p)^{-3nL4^{n-1}/(4^n-1)}$ for global and local depolarizing noise, and $(1-p)^{-2nL4^{n-1}/(4^n-1)}$ for amplitude damping noise, respectively. Bound (Thm. 1) and Bound (Thm. 2) represent the lower bound of the cost obtained from Theorem 1 and Theorem 2, respectively. Note that GSE and the rescaling methods do not completely eliminate the errors for (b) and (c), while we confirm a significant reduction of bias.

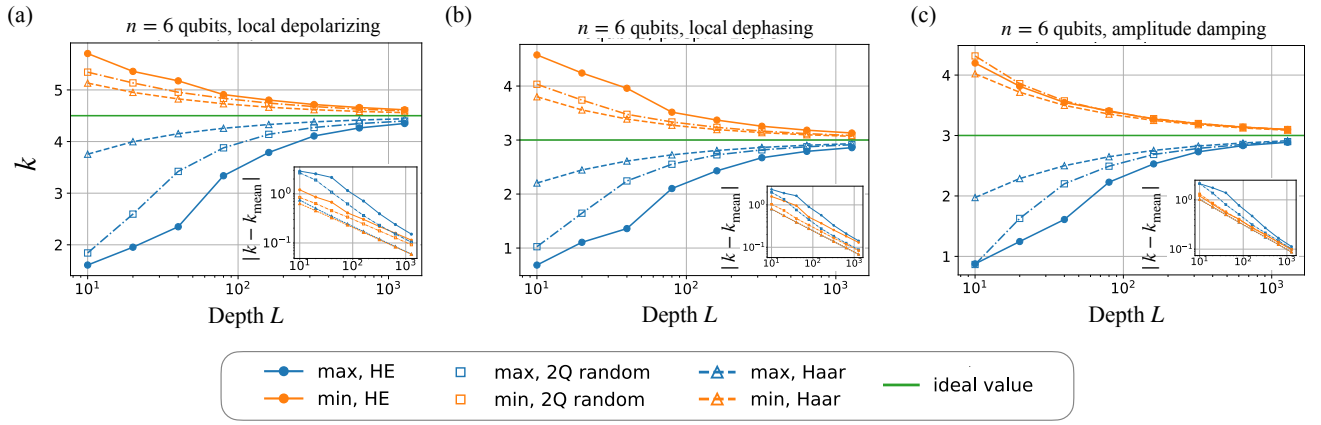


FIG. 2. Convergence of (a) local depolarizing, (b) local dephasing, and (c) amplitude damping into the global depolarizing noise under random circuits of $n = 6$ qubits with error rate $p = 0.0001$. Here, we denote by $(1-p)^{kL}$ the singular values of the unital part of the Pauli transfer matrix for the effective noise channel at each depth, where k for the maximal and minimal ones are plotted in this figure. As is highlighted in the inset, we find that all k's approach the geometric mean $k_{\rm mean}$ of the singular values for each noise channel with its fluctuation scaling as $O(1/\sqrt{L})$, implying the convergence to the global depolarizing noise. For instance, $k_{\rm mean} = 3n4^{n-1}/(4^n-1)$ for local depolarizing and $k_{\rm mean} = 2n4^n/(4^n-1)$ for both local dephasing and amplitude damping. Here, we consider three class of random circuits: hardware-efficient ansatz with linear connectivity, 2-qubit random unitary between random pairs, and Haar random unitary.

the cost bound is given as

$$\mathbb{E}[N] \geq \begin{cases} O\left(\left(1+\frac{3}{2}\frac{4^n}{4^n-1}p\right)^{nL}\right) & \text{(local dep.)} \\ O\left(\left(1+\frac{4^n}{4^n-1}p\right)^{nL}\right) & \text{(amp. damping)} \end{cases} \tag{3}$$

which we numerically confirm to be saturated at large-L regime (Fig. 1(b)(c)).

III. TOWARDS OPTIMAL QUANTUM ERROR MITIGATION

While the scaling of Eq. (3) is derived under the assumption of unitary 2-design, we conjecture that the bound shall hold for even wider class of quantum circuits and local noise. Concretely, as is presented in Fig. 2, the effect of each noise becomes indiscriminable from that of the global depolarizing noise whose error rate grows exponentially with n in the large-L regime, even when any layer of unitary gate does not constitute unitary 2-design. Note that the demonstration involves quantum circuit structure even with only linear connectivity. These results are in agreement with the phenomenological argument provided in Ref. [32] that, noise in deep layered circuits shall be modeled by global depolarizing noise with its strength fluctuating as $O(1/\sqrt{L})$.

These facts not only give us another evidence for scaling as in Eq. (3) but also imply that, although we cannot remove bias completely, we may optimally suppress the effect of noise by just rescaling the measurement results as in the case of global depolarizing noise. We have also numerically investigated other noise structure that does not satisfy the conditions in the Theorems (e.g. local dephasing) and have confirmed that such a picture holds as well. We envision that such a picture becomes even more important in the regime of early fault-tolerance quantum computing. In such a regime, one shall aim to run longer quantum circuits than in near-term regime so that the scrambling effect becomes sufficiently strong to assure the convergence to the global depolarizing noise.

[1] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, Phys. Rev. A **52**, R2493 (1995).
[2] E. Knill, R. Laflamme, and W. Zurek, Threshold accuracy for quantum computation, arXiv preprint quant-ph/9610011 (1996).
[3] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error rate, SIAM Journal on Computing **38**, 1207 (2008).
[4] D. A. Lidar and T. A. Brun, *Quantum error correction* (Cambridge university press, 2013).
[5] M. A. Nielsen and I. Chuang, *Quantum computation and quantum information* (American Association of Physics Teachers, 2002).
[6] Y. Li and S. C. Benjamin, Efficient variational quantum simulator incorporating active error minimization, Phys. Rev. X **7**, 021050 (2017).
[7] K. Temme, S. Bravyi, and J. M. Gambetta, Error mitigation for short-depth quantum circuits, Phys. Rev. Lett. **119**, 180509 (2017).
[8] A. Kandala, K. Temme, A. D. Córcoles, A. Mezzacapo, J. M. Chow, and J. M. Gambetta, Error mitigation extends the computational reach of a noisy quantum processor, Nature **567**, 491 (2019).
[9] S. Endo, S. C. Benjamin, and Y. Li, Practical quantum error mitigation for near-future applications, Phys. Rev. X **8**, 031027 (2018).
[10] E. v. d. Berg, Z. K. Minev, A. Kandala, and K. Temme, Probabilistic error cancellation with sparse pauli-lindblad models on noisy quantum processors, arXiv preprint arXiv:2201.09866 (2022).
[11] W. J. Huggins, S. McArdle, T. E. O'Brien, J. Lee, N. C. Rubin, S. Boixo, K. B. Whaley, R. Babbush, and J. R. McClean, Virtual distillation for quantum error mitigation, Phys. Rev. X **11**, 041036 (2021).
[12] B. Koczor, Exponential error suppression for near-term quantum devices, Phys. Rev. X **11**, 031057 (2021).
[13] M. Huo and Y. Li, Dual-state purification for practical quantum error mitigation, Phys. Rev. A **105**, 022427 (2022).
[14] P. Czarnik, A. Arrasmith, L. Cincio, and P. J. Coles, Qubit-efficient exponential suppression of errors, arXiv preprint arXiv:2102.06056 (2021).
[15] J. R. McClean, M. E. Kimchi-Schwartz, J. Carter, and W. A. de Jong, Hybrid quantum-classical hierarchy for mitigation of decoherence and determination of excited states, Physical Review A **95** (2017).
[16] J. R. McClean, Z. Jiang, N. C. Rubin, R. Babbush, and H. Neven, Decoding quantum errors with subspace expansions, Nature Communications **11**, 636 (2020).
[17] N. Yoshioka, T. Sato, Y. O. Nakagawa, Y.-y. Ohnishi, and W. Mizukami, Variational quantum simulation for periodic materials, Phys. Rev. Research **4**, 013052 (2022).
[18] N. Yoshioka, H. Hakoshima, Y. Matsuzaki, Y. Tokunaga, Y. Suzuki, and S. Endo, Generalized quantum subspace expansion, Phys. Rev. Lett. **129**, 020502 (2022).
[19] X. Bonet-Monroig, R. Sagastizabal, M. Singh, and T. E. O'Brien, Low-cost error mitigation by symmetry verification, Phys. Rev. A **98**, 062339 (2018).
[20] S. McArdle, X. Yuan, and S. Benjamin, Error-mitigated digital quantum simulation, Phys. Rev. Lett. **122**, 180501 (2019).
[21] Z. Cai, Quantum Error Mitigation using Symmetry Expansion, Quantum **5**, 548 (2021).
[22] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, Error mitigation with Clifford quantum-circuit data, Quantum **5**, 592 (2021).
[23] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, Learning-based quantum error mitigation, PRX Quantum **2**, 040330 (2021).
[24] R. Sagastizabal, X. Bonet-Monroig, M. Singh, M. A. Rol, C. C. Bultink, X. Fu, C. H. Price, V. P. Ostroukh, N. Muthusubramanian, A. Bruno, M. Beekman, N. Haider, T. E. O'Brien, and L. DiCarlo, Experimental error mitigation via symmetry verification in a variational quantum eigensolver, Phys. Rev. A **100**, 010302 (2019).
[25] J. Sun, X. Yuan, T. Tsunoda, V. Vedral, S. C. Benjamin, and S. Endo, Mitigating realistic noise in practical noisy intermediate-scale quantum devices, Phys. Rev. Applied **15**, 034026 (2021).
[26] S. Zhang, Y. Lu, K. Zhang, W. Chen, Y. Li, J.-N. Zhang, and K. Kim, Error-mitigated quantum gates exceeding physical fidelities in a trapped-ion system, Nature Communications **11**, 587 (2020).
[27] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, Limitations of noisy reversible computation, arXiv preprint quant-ph/9611028 (1996).
[28] M. J. Kastoryano and K. Temme, Quantum logarithmic sobolev inequalities and rapid mixing, Journal of Mathematical Physics **54**, 052202 (2013).
[29] Y. Watanabe, T. Sagawa, and M. Ueda, Optimal measurement on noisy quantum systems, Phys. Rev. Lett. **104**, 020401 (2010).
[30] S. L. Braunstein and C. M. Caves, Statistical distance and the geometry of quantum states, Physical Review Letters **72**, 3439 (1994).
[31] R. Takagi, Optimal resource cost for error mitigation, Phys. Rev. Research **3**, 033178 (2021).
[32] D. Qin, Y. Chen, and Y. Li, Error statistics and scalability of quantum error mitigation formulas, arXiv preprint arXiv:2112.06255 (2021).

Harrow-Hassidim-Lloyd algorithm without ancilla postselection

D. Babukhin[1] *

[1] *Dukhov Automatics Research Institute (VNIIA), 127055 Moscow, Russia*

Abstract. Harrow-Hassidim-Lloyd algorithm (HHL) is a quantum algorithm for solving systems of linear equations with exponential speed up. Originally, this algorithm requires postselection of ancilla qubit, which makes this algorithm probabilistic. Here we show, that under certain conditions, the HHL algorithm can work without postselection of ancilla qubit, which makes the algorithm deterministic.

Keywords: Harrow-Hassidim-Lloyd Algorithm, Systems of Linear Equations, Quantum Computing

The HHL algorithm [1] allows for solving systems of linear equations on a quantum computer with an exponential speed up due to use of the quantum phase estimation. This algorithm was applied to many practical problems, which will potentially be enhanced with use of quantum computer - quantum machine learning or solving differential equations, to name a few.

This algorithm requires postselection of an ancillary qubit, which makes the result of the algorithm probabilistic. The need for postselection increases running time to obtain the calculation result of the HHL algorithm, as well as any algorithm, based on the HHL. Postselection requires $O(\kappa)$ amplitude amplifications to make a success probability sufficiently high, where κ is a conditional number of the input matrix A. The overall complexity of the HHL is $O(log(N)s^2\kappa^2/\epsilon)$ [1], where one κ comes from the amplitude amplification step. There are methods to overcome this postselection issue, developed over the years since the HHL was invented. One solution is to use a variable-time amplitude amplification algorithm [2] to increase the probability of measuring ancilla in $|1\rangle$, which reduces the overall complexity of the HHL to $O(\kappa \log^3 \kappa)$ in conditional number. The other way is using polynomial decomposition into linear combinations of unitary operators [3], which allow reducing complexity to $O(\kappa \log \kappa)$. There are other solutions for the quantum linear system problem, based on adiabatic quantum computing.

In the work [4], we demonstrate that the HHL works without postselection, when an input matrix A and a measurement matrix M satisfy the commutator identity $K = \frac{1}{2}[[M, A_C], A_C] - \frac{1}{2}[[M, \sqrt{A_C^2 - I}], \sqrt{A_C^2 - I}] = 0$. When this relation is satisfied, the algorithm produces quantum states for two ancilla measurement outcomes ($|0\rangle$ or $|1\rangle$), in which expectation values deviate from each other only by a constant. This connection of expectation values allows using both output states to obtain an expectation value of M on the solution of the linear system. The result reduces the overall HHL complexity to $O(log(N)s^2\kappa/\epsilon)$.

We provide a particular example of a matrix, which satisfies this condition. The matrix has a following form

$$A = \begin{pmatrix} a & b & 0 & \dots & 0 & 0 & 0 \\ b & a & b & \dots & 0 & 0 & 0 \\ 0 & b & a & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a & b & 0 \\ 0 & 0 & 0 & \dots & b & a & b \\ 0 & 0 & 0 & \dots & 0 & b & a \end{pmatrix}$$

where a and b are parameters which come from a problem under the scope. This matrix satisfies $K = 0$ with an observable

$$M = X \otimes X \otimes \cdots \otimes X$$

The result is verified numerically by calculating a norm value $|K|$: in a region of parameters a and b,

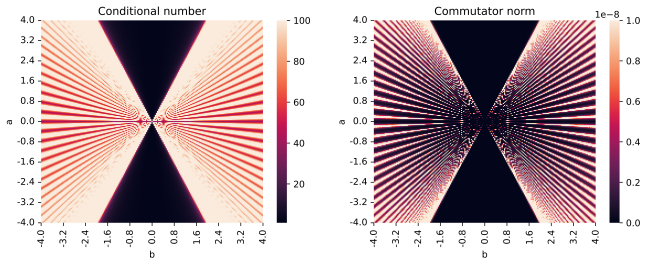


Figure 1: Heatmaps of conditional values of an input matrix A (left) and $|K|$ values (right). A dimension of the input matrix is 2^6.

where the matrix A is invertible - i.e., the HHL algorithm works - the $|K|$ value is zero, which means the HHL works postselection-free. There can be other practically-interesting classes of matrices, which satisfy the postselection-free condition and thus lead to improvement of the HHL.

References

[1] A. W. Harrow, A. Hassidim and S. Lloyd Quantum Algorithm for Linear Systems of Equations. Phys. Rev. Lett. **103**, 150502 (2009).

[2] A. Ambainis Variable time amplitude amplification and a faster quantum algorithm for solving systems of linear equations.. https://arxiv.org/abs/1010.4458.

*dv.babukhin@gmail.com

[3] A. M. Childs, R. Kothari, and R. D. Somma Quantum Algorithm for Systems of Linear Equations with Exponentially Improved Dependence on Precision. SIAM J. Comput. 46, 1920 (2017).

[4] D. V. Babukhin Harrow-Hassidim-Lloyd algorithm without ancilla postselection. Phys. Rev. A **107**, 042408 (2023).

Solving A Classification Problem Using Quantum Support Vector Machine

Seemanta Bhattacharjee[1] * MD. Muhtasim Fuad[1] † A.K.M. Fakhrul Hossain[1] ‡

[1] *Shahjalal University Of Science and Technology, Sylhet-3114, Bangladesh*

Abstract. Quantum Support Vector Machine is a kernel based approach for classification problems. We studied the applicability of quantum kernel to financial data, specifically our self-curated Dhaka Stock Exchange (DSEx) Broad Index dataset. To the best of our knowledge, this is the very first systematic research work on this dataset on the application of quantum kernel. We report empirical quantum advantage in our work where we have used several quantum kernels and proposed the best one for this dataset while providing the verification of Phase Space Terrain Ruggedness Index metric. We have estimated the resources needed to carry out these types of investigations on a larger scale for future practitioners.

Keywords: QSVM, DSEx Broad Index, PTRI verification, Quantum Resource Estimation

1 Results and Discussion

This study focused on a binary classification problem using the Dhaka Stock Exchange (DSEx) Broad Index dataset and we conducted experiments on both classical and quantum methods simultaneously across various data sizes. To the best of our knowledge, this is the first comprehensive investigation of QSVM based on the DSEx Broad Index dataset, and this work provides a robust comparison of several quantum kernels for this type of dataset. This study provides valuable insights into the performance of quantum kernels on financial datasets. These findings can inform further research on quantum machine learning applications in the finance industry.

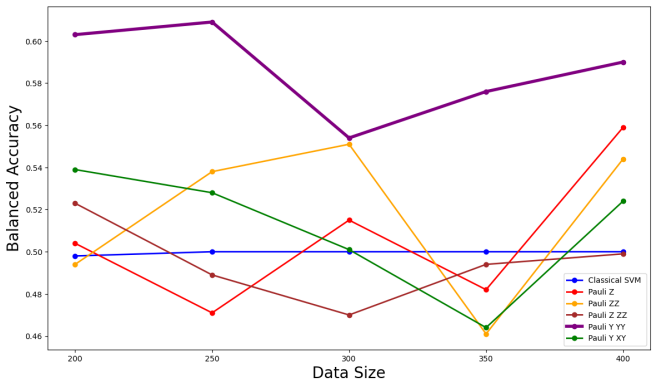


Figure 1: Comparing average Balanced Accuracy of quantum kernels with classical SVM on datasets where classical SVM performance is closest to mean performance, across varying dataset sizes and 7 features.

From figure 1, we have observed that the quantum kernel built using the Pauli Y YY feature map is the most suitable one for this particular dataset. We noticed that the kernels generated using the kernels built using Pauli Y YY feature map consistently outperformed all other quantum kernels and classical Support Vector Machine (SVM) models for every point in the feature-dataset configuration space. We have run the experiments on datasets ranging from 200 to 400 and using 5 to 7 features.

For further investigation, we have run the same experiments on datasets where the performance of classical SVMs were closest to its maximum and minimum performance respectively. Our findings show that the kernel built using Pauli Y YY feature map also outperform classical SVMs.

The Pauli Y YY feature map performs significantly better in every configuration space where the classical SVM performs worse than its average performance. By comparing the position of the data points with the horizontal "zero-advantage line" in Figure 2, we observe EQA for all of the problem instances in the configuration space.

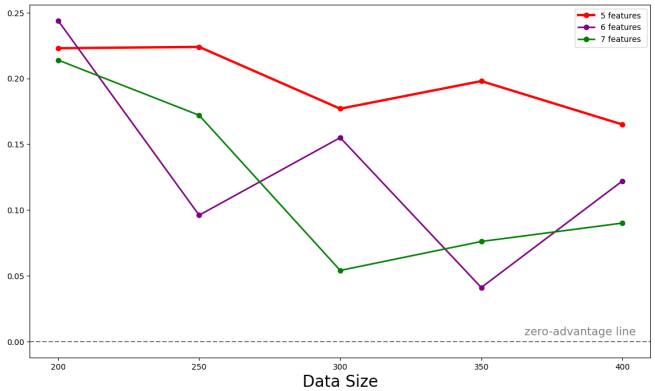


Figure 2: Difference of average Balanced Accuracy of classical SVM vs Pauli Y YY Kernel on datasets where classical SVM performance is closest to minimum performance, across varying sizes and 5 to 7 number of features.

Figure 2 implies the impact and importance of quantum approach in real world scenarios where we face the scarcity of stock market data and where the classical SVM performs worse than the average performance.

We curated the dataset through a comprehensive collection of features from various online resources, which we merged to create a cohesive dataset. We conducted Exploratory Data Analysis on the dataset to provide ini-

*babune99@gmail.com
†muhtasimfuad625@gmail.com
‡a.k.m.fakhrul.hossain@gmail.com

tial insights. We have added more economical features that usually have a direct impact on the change of stock market's trend.

This study provides valuable insights into the performance of quantum kernels on financial datasets and these findings can inform further research on quantum machine learning applications in the finance industry. Regarding the resources required for quantum kernel classification, we can minimize the number of qubits required to perform experiments with a fixed number of features using Bloch Sphere Encoding [2]. This approach has the potential to reduce the resources needed for these experiments. Additionally, exploring the impact of different hyper parameters and regularization techniques on the performance of the algorithm can further improve its accuracy and generalizability.

2 PTRI Verfication

PTRI serves to systematically identify machine learning problems where quantum kernels may exhibit empirical superiority [1]. A geophysics-inspired strategy is proposed to delineate regions of potential Empirical Quantum Advantage (EQA) within datasets, in order to facilitate the selection of the subset of problems that could benefit from a quantum kernel. To this end, one viable strategy involves analyzing the ruggedness of the manifold via PTRI. To ascertain the PTRI values for the entire configuration space, we evaluated the balanced accuracy metric, and represented the results graphically in Figure 3.

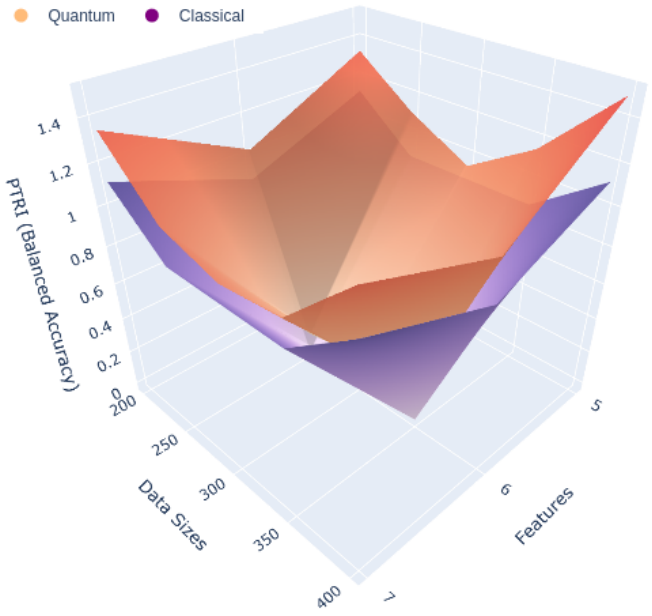


Figure 3: PTRI Scores for classical SVM and QSVM in a 15 point configuration space.

In figure 3, we presented a comparison of PTRI scores, with the purple-colored surface depicting Classical computing, and the peach-colored surface representing Quantum computing.

The data points for each configuration space coordinate were averaged over two selected datasets, resulting in a total of 15 points for each plotted configuration space. The z-axis depicts the metric, while the x- and y-axes represent the number of features and data size, respectively. We verified the PTRI metric for this dataset and observed that classical SVM's balanced accuracy did not fluctuate significantly, even as the size of the training set increased. We noted that quantum kernels performed better on smoother terrain.

3 Resource Estimation

We conducted a thorough assessment of the resources required to construct the circuit essential for developing the quantum kernel using the Y YY feature map.

Our findings provide a framework for future researchers to estimate resources for conducting similar experiments in a more comprehensive manner. The number of gates and circuit depth increases as the number of features or circuit repetition increases.

The total number of gates needed in an experiment run in quantum can be calculated by the following equation, where F defines the number of features and R defines the repetition number of circuit.

$$\text{Total} = (11 \times F - 7) \times R$$

Here, the circuit consists of four different types of gates which are H, R_x, P and C_x gates. The total number of H, R_x, P and C_x gates needed individually can be found by the following equations respectively:

$$H = F \times R$$

$$R_x = (6 \times F - 4) \times R$$

$$P = (2 \times F - 1) \times R$$

$$C_x = (2 \times F - 2) \times R$$

We have also provided the equation to calculate the quantum circuit dept for this kinds of experiments that will be a framework for quantum practitioners. The depth of a quantum circuit can be calculated using the following equation:

$$\text{Depth} = (5 \times F - 1) \times R$$

The number of qubits required is independent of the number of repetitions in the circuit. The required number of qubits precisely matches the number of features, regardless of the number of repetitions.

4 Variability and Errors

The visualization of how the balance accuracy is distributed for the Classical model is illustrated in the figure 4.

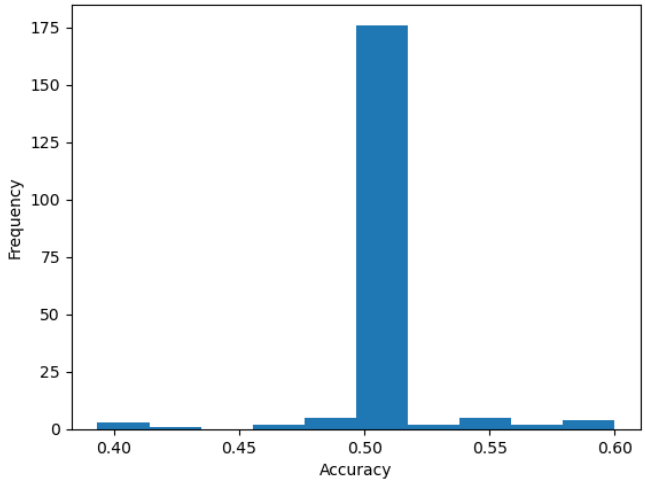


Figure 4: Examining Balanced Accuracy Variability on Configuration Space Coordinate with 200 Samples and 5 Features through 200 different experiments.

Here, to illustrate the distribution of balanced accuracy in classical SVM, we have run 200 numbers of separate experiments where the data size was 200 and number of features was 5. It showed a 2.1% standard deviation while predicting the trend in classical domain which makes this a prime candidate for the application of quantum algorithms.

References

[1] Z. Krunic, F. F. Flöther, G. Seegan, N. D. Earnest-Noble and O. Shehab. Quantum kernels for real-world predictions based on electronic health records. *IEEE Transactions on Quantum Engineering*, vol. 3, pages 1–11, 2022.

[2] J. Heredge, C. Hill, L. Hollenberg, and M. Sevior. Quantum Support Vector Machines for Continuum Suppression in B Meson Decays. *Computing and Software for Big Science.* vol. 5, 2021.

Quantum information processing with frequency-comb qubits and time-resolving detectors

Tomohiro Yamazaki[1 2 3 *] Tomoaki Arizono[2] Toshiki Kobayashi[2 3] Rikizo Ikuta[2 3] Takashi Yamamoto[2 3]

[1] *NTT Basic Research Laboratories, NTT Corporation, Atsugi, Kanagawa 243-0198, Japan*
[2] *Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan*
[3] *Center for Quantum Information and Quantum Biology, Osaka University, Toyonaka, Osaka 560-0043, Japan*

Abstract. Linear optical quantum computation using time-frequency degree of freedom has advantages in terms of error susceptibility and extensibility to high-dimensional encodings. However, time-bin and frequency-bin encodings typically require high-speed switches and electro-optic modulators, respectively, and the uses of such active devices would make scaling up difficult. Here, we propose a new scheme based on the encoding of a qubit into single-photon frequency combs. By combining qubit generators, time-resolving detectors, optical interleavers, and beam splitters, universal quantum computation can be achieved with the robustness against temporal and spectral errors. we also show the experimental feasibility of this scheme according to the error analysis.

Keywords: linear optics, frequency comb, time-resolving detector, optical interleaver

1 Introduction

Photons and their manipulations using linear optics play an indispensable role in quantum information processing [1, 2]. Photons have multiple degrees of freedom (DoF) such as polarization, path, time-frequency [3, 4], and angular-momentum [5]. The properties of qubits depends on what DoF of photons is used to form the qubits. Taking polarization qubits as an example, computational errors are caused by birefringence and the dimension of the encoding is limited to two. By construct, qubits formed by time-frequency DoF are usually less susceptible to errors because most optical components do not depend on small temporal and spectral differences. In addition, it is possible to realize high-dimensional encoding because time-frequency DoF is a continuous variable.

An encoding using time-frequency DoF is time-bin encoding in which the temporal peaks of a photon form the computation basis. Universal quantum computation can be achieved with high-speed switches and operations on another DoF such as polarization [6] or path [7, 8, 9]. Another encoding using time-frequency DoF is frequency-bin encoding in which the spectral peaks of a photon form the computation basis. Universal quantum computation can be achieved combining electro-optic modulators and pulse shapers [10, 11, 12, 13, 14, 15]. However, the use of many active devices in these approaches is prone to errors and losses and poses challenges in scaling up. Although recently proposed is to use time-resolving detectors for manipulation of frequency-bin qubits [16], the finite resolution of these detectors causes serious errors because frequency-bin qubit are susceptible to temporal shift errors.

In this study, we propose a new scheme of linear optical quantum computation (LOQC) using time-frequency DoF. We use encoding in which single-photon frequency combs form the computational basis. The state in this encoding is called the time-frequency Gottesman-Kitaev-Preskill (TFGKP) state [17, 18] from the analog of Gottesman-Kitaev-Preskill code [19] for quadrature amplitudes of light [20]. The TFGKP state is discretized in both the time and frequency domains because of its comb-shaped spectrum. Thus, it is robust against time- and spectral-shift errors. We show that universal quantum computation can be achieved using TFGKP-state generators, time-resolving detectors, optical interleavers, beam splitters. Thus, active devices such as high-speed switches and electro-optic modulators are not required. TFGKP-state generators can be efficiently realized using a cavity-enhanced nonlinear optical process [21, 22, 23, 24, 25, 26, 27]. Futhermore, in contrast to the passive scheme that use frequency-bin encoding and time-resolving detectors [16], quantum computation can be performed robustly despite the detector's finite resolutions and other temporal and spectral errors. We estimate the errors occurring in this scheme and show that the experimental requirements for fault-tolerant quantum computation are almost achievable with current state-of-the-art technologies. This work has been published in Physical Review Letters [28].

2 Result

Each frequency basis state of a TFGKP qubit is defined as a single-photon frequency comb with the same comb spacing. The basis states differ from each other by the frequency shifts from a fixed central frequency. The time basis of a TFGKP qubit is defined by the discrete Fourier transform of the frequency basis. Because of its comb-shaped spectrum, TFGKP qubit is discretized in not only frequency domain but also time domain. Figure 1 shows an example of probability distribution of a TFGKP qubit. In practice, TFGKP qubits have a finite bandwidth of the spectral envelope and a non-zero linewidth of each spectral peak. However, we can approximately consider the frequency basis as the computational basis $\{|0\rangle, |1\rangle\}$.

*tomohiro.yamazaki@ntt.com

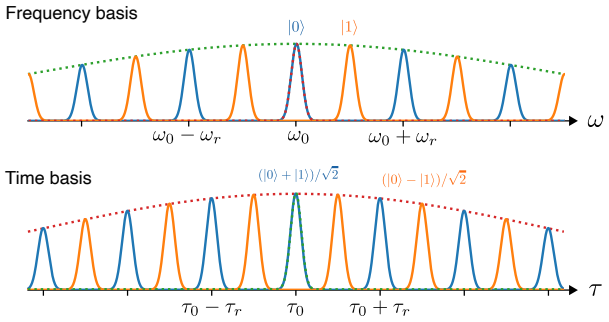


Figure 1: Probability distributions of a time-frequency Gottesmann-Kitaev-Preskill qubit in the frequency and time bases. The blue and orange lines in the frequency basis correspond to $|0\rangle$ and $|1\rangle$, and ones in the time basis correspond to $(|0\rangle + |1\rangle)/\sqrt{2}$ and $(|0\rangle - |1\rangle)/\sqrt{2}$, respectively.

Then, the time basis is approximately equivalent to the basis of $\{(|0\rangle \pm |1\rangle)/\sqrt{2}\}$.

The heralded generation of a TFGKP qubit is possible using a broadband time-frequency entangled photon pair and a cavity. When one of the two photons passes through the cavity and is detected by a time-resolving detector, the state of the other photon corresponds to a TFGKP state as the transmission spectrum of the cavity is reflected to the spectrum of the remaining photon. In particular, cavity-enhanced nonlinear optical process is suitable to the efficient generation of a TFGKP-state.

In addition to TFGKP-state generator, we use beam splitters, optical interleavers, time- and frequency-resolving detectors. We suppose the spatial beam splitters, especially 50:50 beam splitters, that are independent of time-frequency and other DoFs. Optical interleavers are spectrally periodic filter that spatially combines or separates frequency combs as shown in Fig. 2. Especially, we use 2:2 OIs that have two input and output ports. Time- and frequency-resolving detectors are used to distinguish time and spectral basis states, respectively. Although these detectors have finite resolutions, we still be able to distinguish the basis states as long as the probability distributions of the basis states are well separated. In addition, a frequency-resolving detector can be substituted by combining a 1:d OI with d detectors.

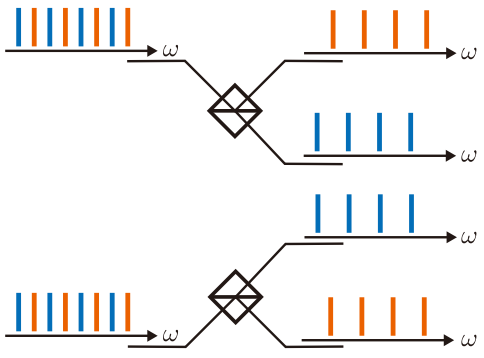


Figure 2: Graphical representation of 2:2 optical interleaver. Orange and blue spectral peaks represent $|0\rangle$ and $|1\rangle$, respectively.

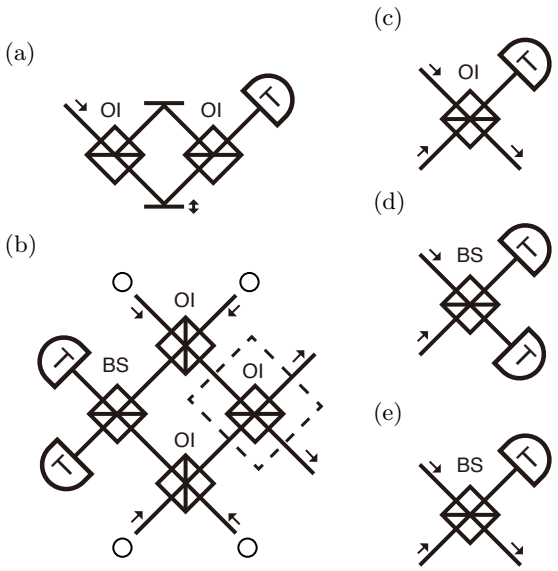


Figure 3: Concrete setups for quantum operations. All detectors, beam splitters (BSs), and optical interleavers (OIs) are time-resolving detectors, 50:50 beam splitters, and 2:2 optical interleavers, respectively. (a) Measurement in the $\cos(\frac{\theta}{2})X + \sin(\frac{\theta}{2})Y$ basis. θ is adjustable by changing the relative lengths between the two arms. (b) Bell-state generation, which succeeds when detectors detect two photons in different states. The states of input single qubits are $(|0\rangle + |1\rangle)/\sqrt{2}$. The optical interleaver enclosed by the dotted lines denotes a feed-forward operation required in 1/3 of the success cases. Its total success probability is 3/16. (c) Type-I fusion gate, which succeeds when a detector detects a photon with a probability 1/2. (d) Type-II' fusion gate, which succeeds when a detector or detectors detect $|0\rangle$ and $|1\rangle$ with a probability of 1/2. (e) Type-I' fusion gate, which succeeds when a detector detects a photon with a probability of 1/4.

Next, we show the quantum computational universality of the above toolbox. The time- and frequency-resolving detectors correspond to the measurement in the X and Z bases, respectively. An arbitrary phase gate can be realized by spatially separating each computational basis by a 1:2 OI, adding a small relative time delay, and then combining them with a 2:1 optical interleaver. As shown in Fig. 3(a), the phase gate followed by the measurement in the X basis corresponds to the measurement in the $\cos(\theta/2)X + \sin(\theta/2)Y$ basis. Fault-tolerant quantum computation can be performed with single-qubit measurements in the bases X, Z, and $(X+Y)/\sqrt{2}$ on a three-dimensional cluster state [29, 30].

Thus, the remaining part is how to generate the cluster state. This is accomplished by combing Bell-state generation setups, type-I fusion gates, type-II' fusion gates, and type-I' fusion gates that are shown in Fig. 3(b), 3(c), 3(d), 3(e), respectively. First, we generate Bell states by using Bell-state generation setups and then generate a three-qubit cluster state from two Bell states by using a

type-I' fusion gate. Once we obtain three-qubit cluster states, we can generate larger cluster states according to a similar procedure to the protocol for polarization encoding in Ref. [31] That is, we use type-I fusion gates to generate one-dimensional cluster states and use type-II' fusion gates to generation higher-dimensional cluster states. As a result, it is shown that our toolbox is sufficient for the universal quantum computation.

Finally, we consider the amount of computational errors caused by the temporal and spectral broadening. The temporal and spectral broadening can be classified to the coherent and incoherent ones, respectively. We call the insufficient separation between states corresponding to the different basis states "factor I," while we call the insufficient overlap between states corresponding to the same basis state "factor II." Factor I is related to errors on the one-qubit measurements and characterized by the total amount of coherent and incoherent broadening. On the other hand, factor II is related to errors on the entangling gates and characterized by the amount of incoherent broadening relative to that of coherent broadening. In our scheme, we use frequency-resolving measurements only for one-qubit measurements; therefore, we do not have to consider factor II on the frequency basis. On the time basis, there is an optimal amount of coherent temporal broadening owing to a tradeoff between factors I and II.

Under several experimental assumptions, we can derive the experimental requirements to make the major error probabilities lower than 0.01. In particular, we suppose the temporal resolution of detectors as the major temporal error source. From the best value among the detectors for telecom wavelengths [32], the upper bound of the spectral comb spacing and the lower bound of the finesse of TFGKP qubits are derived. These values are in the same order of magnitudes as the commercially available optical interleavers and biphoton frequency combs generated by nonlinear optical waveguide resonators [26, 27, 33]. Thus, the current state-of-the-art technologies largely meet the experimental requirements for fault-tolerant quantum computation based on our scheme.

3 Conclusion

We proposed a LOQC scheme with TFGKP state generators, time-resolving detectors, beam splitters, and optical interleavers. The discretization in both the time and frequency domains owing to TFGKP qubits leads to error robustness against both temporal and spectral errors. Furthermore, by treating the time and frequency basis asymmetrically, we realized universal quantum computation without active devices. Although we consider the case of qubits in this paper, TGKP states and their manipulations can readily be extended to the case of qudits. Thus, the considered toolbox is a good platform for realizing the recently developed field of high-dimensional LOQC using qudits [34, 35, 36]. This scheme has high error robustness and ease of operations because of its use of time–frequency DoF and passive devices, respectively. Therefore, this is a practical approach, especially for quantum computation with integrated photonic circuits [37] and quantum communication requiring multi-photon entangled states [38, 39, 40].

References

[1] E. Knill, R. Laflamme, and G. J. Milburn, "A scheme for efficient quantum computation with linear optics," *Nature*, vol. 409, pp. 46–52, Jan. 2001.

[2] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Milburn, "Linear optical quantum computing with photonic qubits," *Rev. Mod. Phys.*, vol. 79, pp. 135–174, Jan. 2007.

[3] J. M. Lukens, A. Dezfooliyan, C. Langrock, M. M. Fejer, D. E. Leaird, and A. M. Weiner, "Orthogonal Spectral Coding of Entangled Photons," *Phys. Rev. Lett.*, vol. 112, p. 133602, Apr. 2014.

[4] B. Brecht, D. V. Reddy, C. Silberhorn, and M. G. Raymer, "Photon Temporal Modes: A Complete Framework for Quantum Information Science," *Phys. Rev. X*, vol. 5, p. 041017, Oct. 2015.

[5] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the orbital angular momentum states of photons," *Nature*, vol. 412, pp. 313–316, July 2001.

[6] P. C. Humphreys, B. J. Metcalf, J. B. Spring, M. Moore, X.-M. Jin, M. Barbieri, W. S. Kolthammer, and I. A. Walmsley, "Linear Optical Quantum Computing in a Single Spatial Mode," *Phys. Rev. Lett.*, vol. 111, p. 150501, Oct. 2013.

[7] H. Takesue, "Entangling time-bin qubits with a switch," *Phys. Rev. A*, vol. 89, p. 062328, June 2014.

[8] H.-P. Lo, T. Ikuta, N. Matsuda, T. Honjo, and H. Takesue, "Entanglement generation using a controlled-phase gate for time-bin qubits," *Appl. Phys. Express*, vol. 11, p. 092801, Aug. 2018.

[9] H.-P. Lo, T. Ikuta, N. Matsuda, T. Honjo, W. J. Munro, and H. Takesue, "Quantum Process Tomography of a Controlled-Phase Gate for Time-Bin Qubits," *Phys. Rev. Applied*, vol. 13, p. 034013, Mar. 2020.

[10] J. M. Lukens and P. Lougovski, "Frequency-encoded photonic qubits for scalable quantum information processing," *Optica*, vol. 4, p. 8, Jan. 2017.

[11] M. Kues, C. Reimer, P. Roztocki, L. R. Cortés, S. Sciara, B. Wetzel, Y. Zhang, A. Cino, S. T. Chu, B. E. Little, D. J. Moss, L. Caspani, J. Azaña, and R. Morandotti, "On-chip generation of high-dimensional entangled quantum states and their coherent control," *Nature*, vol. 546, pp. 622–626, June 2017.

[12] H.-H. Lu, J. M. Lukens, N. A. Peters, O. D. Odele, D. E. Leaird, A. M. Weiner, and P. Lougovski, "Electro-Optic Frequency Beam Splitters and Tritters for High-Fidelity Photonic Quantum Information Processing," *Phys. Rev. Lett.*, vol. 120, p. 030502, Jan. 2018.

[13] H.-H. Lu, J. M. Lukens, N. A. Peters, B. P. Williams, A. M. Weiner, and P. Lougovski, "Quantum interference and correlation control of frequency-bin qubits," *Optica*, vol. 5, p. 1455, Nov. 2018.

[14] X. Lu, G. Moille, Q. Li, D. A. Westly, A. Singh, A. Rao, S.-P. Yu, T. C. Briles, S. B. Papp, and K. Srinivasan, "Efficient telecom-to-visible spectral translation through ultralow power nonlinear nanophotonics," *Nat. Photonics*, vol. 13, pp. 593–601, Sept. 2019.

[15] H.-H. Lu, E. M. Simmerman, P. Lougovski, A. M. Weiner, and J. M. Lukens, "Fully Arbitrary Control of Frequency-Bin Qubits," *Phys. Rev. Lett.*, vol. 125, p. 120503, Sept. 2020.

[16] C. Cui, K. P. Seshadreesan, S. Guha, and L. Fan, "High-Dimensional Frequency-Encoded Quantum Information Processing with Passive Photonics and Time-Resolving Detection," *Phys. Rev. Lett.*, vol. 124, p. 190502, May 2020.

[17] N. Fabre, G. Maltese, F. Appas, S. Felicetti, A. Ketterer, A. Keller, T. Coudreau, F. Baboux, M. I. Amanti, S. Ducci, and P. Milman, "Generation of a time-frequency grid state with integrated biphoton frequency combs," *Phys. Rev. A*, vol. 102, p. 012607, July 2020.

[18] N. Fabre, A. Keller, and P. Milman, "Time and frequency as quantum continuous variables," *Phys. Rev. A*, vol. 105, p. 052429, May 2022.

[19] D. Gottesman, A. Kitaev, and J. Preskill, "Encoding a qubit in an oscillator," *Phys. Rev. A*, vol. 64, p. 012310, June 2001.

[20] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, "Gaussian quantum information," *Rev. Mod. Phys.*, vol. 84, pp. 621–669, May 2012.

[21] C. Reimer, M. Kues, P. Roztocki, B. Wetzel, F. Grazioso, B. E. Little, S. T. Chu, T. Johnston, Y. Bromberg, L. Caspani, D. J. Moss, and R. Morandotti, "Generation of multiphoton entangled quantum states by means of integrated frequency combs," *Science*, vol. 351, pp. 1176–1180, Mar. 2016.

[22] P. Imany, J. A. Jaramillo-Villegas, O. D. Odele, K. Han, D. E. Leaird, J. M. Lukens, P. Lougovski, M. Qi, and A. M. Weiner, "50-GHz-spaced comb of high-dimensional frequency-bin entangled photons from an on-chip silicon nitride microresonator," *Opt. Express*, vol. 26, p. 1825, Jan. 2018.

[23] C. Reimer, S. Sciara, P. Roztocki, M. Islam, L. Romero Cortés, Y. Zhang, B. Fischer, S. Loranger, R. Kashyap, A. Cino, S. T. Chu, B. E. Little, D. J. Moss, L. Caspani, W. J. Munro, J. Azaña, M. Kues, and R. Morandotti, "High-dimensional one-way quantum processing implemented on d-level cluster states," *Nature Phys*, vol. 15, pp. 148–153, Feb. 2019.

[24] P. Imany, J. A. Jaramillo-Villegas, M. S. Alshaykh, J. M. Lukens, O. D. Odele, A. J. Moore, D. E. Leaird, M. Qi, and A. M. Weiner, "High-dimensional optical quantum logic in large operational spaces," *npj Quantum Inf*, vol. 5, p. 59, July 2019.

[25] M. Kues, C. Reimer, J. M. Lukens, W. J. Munro, A. M. Weiner, D. J. Moss, and R. Morandotti, "Quantum optical microcombs," *Nature Photon*, vol. 13, pp. 170–179, Mar. 2019.

[26] R. Ikuta, R. Tani, M. Ishizaki, S. Miki, M. Yabuno, H. Terai, N. Imoto, and T. Yamamoto, "Frequency-Multiplexed Photon Pairs Over 1000 Modes from a Quadratic Nonlinear Optical Waveguide Resonator with a Singly Resonant Configuration," *Phys. Rev. Lett.*, vol. 123, p. 193603, Nov. 2019.

[27] T. Yamazaki, R. Ikuta, T. Kobayashi, S. Miki, F. China, H. Terai, N. Imoto, and T. Yamamoto, "Massive-mode polarization entangled biphoton frequency comb," *Sci Rep*, vol. 12, p. 8964, May 2022.

[28] T. Yamazaki, T. Arizono, T. Kobayashi, R. Ikuta, and T. Yamamoto, "Linear Optical Quantum Computation with Frequency-Comb Qubits and Passive Devices," *Phys. Rev. Lett.*, vol. 130, p. 200602, May 2023.

[29] R. Raussendorf, J. Harrington, and K. Goyal, "A fault-tolerant one-way quantum computer," *Annals of Physics*, vol. 321, pp. 2242–2270, Sept. 2006.

[30] R. Raussendorf and J. Harrington, "Fault-Tolerant Quantum Computation with High Threshold in Two Dimensions," *Phys. Rev. Lett.*, vol. 98, p. 190504, May 2007.

[31] D. E. Browne and T. Rudolph, "Resource-Efficient Linear Optical Quantum Computation," *Phys. Rev. Lett.*, vol. 95, p. 010501, June 2005.

[32] B. Korzh, Q.-Y. Zhao, J. P. Allmaras, S. Frasca, T. M. Autry, E. A. Bersin, A. D. Beyer, R. M. Briggs, B. Bumble, M. Colangelo, G. M. Crouch, A. E. Dane, T. Gerrits, A. E. Lita, F. Marsili, G. Moody, C. Peña, E. Ramirez, J. D. Rezac, N. Sinclair, M. J. Stevens, A. E. Velasco, V. B. Verma, E. E. Wollman, S. Xie, D. Zhu, P. D. Hale, M. Spiropulu, K. L. Silverman, R. P. Mirin, S. W. Nam, A. G. Kozorezov, M. D. Shaw, and K. K. Berggren, "Demonstration of sub-3 ps temporal resolution with a superconducting nanowire single-photon detector," *Nat. Photonics*, vol. 14, pp. 250–255, Apr. 2020.

[33] R. Fujimoto, T. Yamazaki, T. Kobayashi, S. Miki, F. China, H. Terai, R. Ikuta, and T. Yamamoto, "Entanglement distribution using a biphoton frequency comb compatible with DWDM technology," *Opt. Express*, vol. 30, p. 36711, Sept. 2022.

[34] S. Paesani, J. F. F. Bulmer, A. E. Jones, R. Santagati, and A. Laing, "Scheme for Universal High-Dimensional Quantum Computation with Linear Optics," *Phys. Rev. Lett.*, vol. 126, p. 230504, June 2021.

[35] H. Zhang, C. Zhang, X.-M. Hu, B.-H. Liu, Y.-F. Huang, C.-F. Li, and G.-C. Guo, "Arbitrary two-particle high-dimensional Bell-state measurement by auxiliary entanglement," *Phys. Rev. A*, vol. 99, p. 052301, May 2019.

[36] Y.-H. Luo, H.-S. Zhong, M. Erhard, X.-L. Wang, L.-C. Peng, M. Krenn, X. Jiang, L. Li, N.-L. Liu, C.-Y. Lu, A. Zeilinger, and J.-W. Pan, "Quantum Teleportation in High Dimensions," *Phys. Rev. Lett.*, vol. 123, p. 070505, Aug. 2019.

[37] D. Dai, L. Liu, S. Gao, D.-X. Xu, and S. He, "Polarization management for silicon photonic integrated circuits: Polarization management for silicon photonic integrated circuits," *Laser & Photonics Reviews*, vol. 7, pp. 303–328, May 2013.

[38] K. Azuma, K. Tamaki, and H.-K. Lo, "All-photonic quantum repeaters," *Nat Commun*, vol. 6, p. 6787, Apr. 2015.

[39] F. Ewert, M. Bergmann, and P. van Loock, "Ultrafast Long-Distance Quantum Communication with Static Linear Optics," *Phys. Rev. Lett.*, vol. 117, p. 210501, Nov. 2016.

[40] J. Borregaard, H. Pichler, T. Schröder, M. D. Lukin, P. Lodahl, and A. S. Sørensen, "One-Way Quantum Repeater Based on Near-Deterministic Photon-Emitter Interfaces," *Phys. Rev. X*, vol. 10, p. 021071, June 2020.

Learning Stochastic Process with Quantum Recurrent Models

Ximing Wang[1] * Chengran Yang[2] † Mile Gu[1 2] ‡

[1] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*

[2] *Centre for Quantum Technologies, National University of Singapore, Singapore*

Keywords: Variational Quantum Algorithms, Stochastic Process, Quantum Memory

1 Introduction

Predicting future based on the observation of the past is a fundamental task in science and engineering. The predictive models are widely used in various fields, such as weather forecasting, stock market prediction, and natural language processing. While the future may depends on a long history of the past, the computational tools for prediction can only store limited information. As the result, people are interested in how to predict the future faithfully while compressing the past information as much as possible.

While the observations of the process exhibit a strong causal order, the predictive models are often designed to adapt the sequential nature of the process. With an additional assumption that the process is stationary, the predictive models can be constructed with alternating steps. At each step, the memory interact with the environment with a fixed dynamics and generate the output, and then the updated memory is been carried to the next step. Such models, take ϵ-machine [2] as an example, are study extensively in the classical region. But the construction of the minimal predictive models in the quantum region is still in its infancy.

Here we propose a variational method to construct quantum recurrent models for predicting the future of a stochastic process. By using a quantum memory and a quantum interaction between the memory and the environment, we show that, with the same amount of memory, the quantum model has a less distortion from the real process than the classical model. Also, we compared our method to a full quantum circuit model, which covers all possible distributions of the finite past and future. We show that our method is less likely to overfit the data than the full quantum circuit model as it contains much less number of tunable parameters. In addition, with a proper designed cost function, the training of the model is efficient, which makes the algorithm practical for NISQ devices.

2 Framework

Stochastic process – Stochastic processes lay at the core of information theory since the establishment of this field [1]. A stochastic process is defined as a (bi-infinite) *sequence* of random variables $\{X_t\}_{t\in\mathbb{Z}}$,taken values x_t from a finite alphabet Γ. A stochastic process is stationary if the marginal probability of any consecutive subsequences of the random variables $X_{t:t+L} := (X_t, X_{t+1}, \ldots, X_{t+L-1})$ are invariant under time translations, i.e.,

$$P(X_{t:t+L}) = P(X_{t:t+L}) \ , \tag{1}$$

for all time t, t' and any sequence length L.

Predictive models use the past information $\overleftarrow{x} := x_{-\infty:0}$ to predict the future $\overrightarrow{X} := X_{0:\infty}$ with the probability $P(\overrightarrow{X}|\overleftarrow{x})$. Storing the entire past is source-consuming and unnecessary for various processes, thus predictive models encode the past $\overleftarrow{x}$ information into classical states $s_i = f_\epsilon(\overleftarrow{x})$ and save it as its memory for further prediction. The model generates future statistics according to a set of transition rules between the classical states s_i. That is, given the memory in the state s_i, the predictive model will generate output x and update its memory to state s_j with probability $P(s_j, x|s_i)$. Repeating the transition rules allow predictive models to sequentially generate the outputs of stochastic processes. Among all classical models, the ϵ-machine requires the least amount of memory d_c – the least number of classical states,

$$d_c = \#\{s_i\}. \tag{2}$$

Reducing the amount of memory inevitably introduces the distortion of predictive models' outputs. KL-divergence has operational significance in diverse contexts. We use the KL-divergence rate D_e to quantify the distortion, which is defined as the divergence rate between two stochastic processes P, Q,

$$D_e(P||Q) := \lim_{L\to\infty} \frac{1}{L} \sum_{\overleftarrow{x}} P(\overleftarrow{x}) D_{\mathrm{KL}}(P(X_{0:L}|\overleftarrow{x})||Q(X_{0:L}|\overleftarrow{x})) \tag{3}$$

where $D_{\mathrm{KL}}(P(x)||Q(x))$ is the KL-divergence rate $D_{\mathrm{KL}}(P(x)||Q(x)) := \sum_x P(x) \log P(x)/Q(x)$.

Quantum model – A variational quantum circuit is a quantum circuit with tunable parameters, which can be described as a composition of several parameterized unitary operators, each is constructed with parameterized quantum gates. The parameters are usually trained to minimize a cost function that is chosen according to the task.

*ximing001@e.ntu.edu.sg
†cr.yang@nus.edu.sg
‡ceptryn@gmail.com

To simulate a stochastic process, we construct quantum models with a recurrent quantum circuit as shown in figure 1. Our model consists of two systems:

1. A memory system carries the information between different times steps, and

2. an output system that is measured and reset in the computational basis at each step.

With a proper training on the quantum model, the probability distribution of the measured output can approximate an information source better than any classical model with the same size of memory systems.

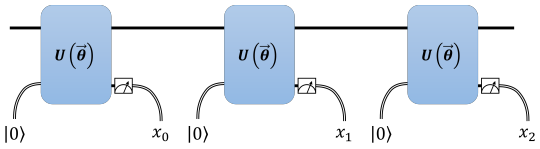


Figure 1: The recurrent quantum circuit that simulates an information source. The same unitary $U(\vec{\theta})$ is applied recurrently at each time step. This model consists of two systems, the output system O and the memory system M. While the memory system carries information between time steps, the output system is reset to $|0\rangle$ at each step.

Despite the fruitful applications of variational quantum circuits, the training of the circuits remains an obstacle due to the large size of the parameter space. It is necessary to reduce the number of parameters for better training of the quantum models, where the reduction of the parameter space can help to relax the problem of barren plateaus. A general quantum circuit $U(\vec{\theta})$ in figure 1 is unnecessarily too complex for simulating stochastic processes. To simplify the circuit, we propose a universal circuit with fewer parameters that can simulate all stochastic processes as universal quantum circuits. A simple model with 1-qubit-output is given in figure 2, which is constructed with controlled R_y gates and controlled local unitary gates acting on the memory system. This model is generated by the CS decomposition [4]. Given a unitary matrix U represented in the computational basis, the matrix can be decomposed into block matrices, where each block is corresponding to a given classical input and output on the output system. In a 1-qubit-output model, the CS decomposition decompose each block into a product of R_y rotations acting on outputs controlled by the memory and local unitary on the memory controlled by the classical input and output string. And the n-qubit-output model can be constructed decompose each controlled unitary as a $n-1$-qubit-output model. More specifically, for the simple model in figure 2, at each time step, a controlled R_y gate $\sum_s |s\rangle\langle s|^M \otimes R_y^O(\theta_s)$ is applied, and the output is measured immediately after the controlled R_y gate. This R_y gate is controlled on the computational basis $\{|s\rangle\}$ of the memory system. After the output is measured, a local unitary U_0 or U_1 is applied to the memory system based on whether the output is 0 or 1. For a 2-dimensional-memory model, a universal circuit only requires 8 parameterized gates instead of 15 parameterized gates as for universal 2-qubit unitaries.

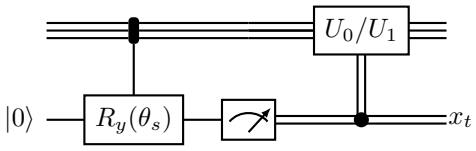


Figure 2: Universal circuit for 2-output quantum model. The unitary operators U_0 and U_1 at time t are applied conditioned on the measured output x_t. The R_y gate is controlled on the computational basis of the memory system.

3 Methods

Choice of the cost function –
As supervised learning, the training of the quantum model requires sufficient information about the original information source.

Although the diverge rate D_d is operationally meaningful, the computation of the KL-divergence is not easy, where the probability of each possible future have to be evaluated separately. The number of different futures grows exponentially with the length L of the future to be considered. That means the number of samples required to evaluate D_d also grows exponentially. In order to make the training practical, here we choose to replace the KL-divergence D_{KL} with a geometric mean fidelity inspired measure F_{GM} [6, 3], where

$$F_{GM}\left(P\left((\overrightarrow{X}')_0^L|\overleftarrow{x}\right)\middle\|P\left(\overrightarrow{X}_0^L|\overleftarrow{x}\right)\right) := \frac{\sum_{\overrightarrow{x}_0^L\in\Gamma^L} p(\overrightarrow{x}_0^L|\overleftarrow{x})q(\overrightarrow{x}_0^L|\overleftarrow{x})}{\sqrt{\sum_{\overrightarrow{x}_0^L\in\Gamma^L} p^2(\overrightarrow{x}_0^L|\overleftarrow{x})}\sqrt{\sum_{\overrightarrow{x}_0^L\in\Gamma^L} q^2(\overrightarrow{x}_0^L|\overleftarrow{x})}}. \quad (4)$$

This measure $F_{GM}\in[0,1]$ approaches to 1 if and only if $(\overrightarrow{X}')_0^L = \overrightarrow{X}_0^L$ for every possible history. This measure is also referred to as cosine similarity in the context of information retrieval [5]. As the result, by assuming the information source has a finite Markov order k, such that the causal state s is uniquely determined by the last k outputs $\overleftarrow{x}^k_{-1}$, the quantum model that minimize the $F_{GM}\left(P\left((\overrightarrow{X}')_0^k|\overleftarrow{x}^k_{-1}\right)\middle\|P\left(\overrightarrow{X}_0^k|\overleftarrow{x}^k_{-1}\right)\right)$ is a good approximation of the information source.

There are two characteristics that makes F_{GM} a practical cost function.

1. each value of the $\sum pq$, $\sum p^2$ and $\sum q^2$ in F_{GM} can be evaluated from the statistics of the outputs efficiently;

2. with proper design of the variational circuit, its derivative can be evaluated efficiently with the help of parameter shift rule.

As the result, for a source with Markov order k, instead of evaluate the probabilities of all $2^k \times 2^k$ combinations of correlated history and future outputs with high precision, only 3 quantities needs to be measured.

Training quantum models – The quantum models are trained based on the consecutive outputs of the original source $\overleftrightarrow{X}$. Two independent set of samples are obtained from $\overleftrightarrow{X}$. The first set of data are split to M sequences with length $2k$, where k is the Markov order of the source. The values of $B_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} q^2(\overrightarrow{x}_0^k|\overleftarrow{x}_{-1}^k)$ is first evaluated by pick out the sequences whose first k outputs equals to $\overleftarrow{x}_{-1}^k$. Then by randomly pairing the sequences and count the fraction of the pairs where the two sequences are equal. This fraction is an estimation of the value $B_{\overleftarrow{x}^k}$.

At each step of the training, the values of $A_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} q^2(\overrightarrow{x}_0^k|\overleftarrow{x}_{-1}^k)$ is evaluated by running the quantum model for M times, each contains $2k$ steps. $A_{\overleftarrow{x}^k}$ can be estimated in the same method as $B_{\overleftarrow{x}^k}$. Similarly, $C_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} p(\overrightarrow{x}_0^k|\overleftarrow{x}^k) q(\overrightarrow{x}_0^k|\overleftarrow{x}^k)$ for all $\overleftarrow{x}^k$ can be evaluated by running the quantum model and count the fraction of the outputs that matches the outputs from $\overleftrightarrow{X}$ given the past $\overleftarrow{x}^k$.

Each of the partial derivative of A and C can be evaluated by repeat above process with randomized parameters. As the result, $F_{GM} = \frac{C}{AB}$ and its derivatives can be computed from the evaluated values. With the derivatives of F_{GM}, we can use ADAM method to update the parameters of the quantum model and maximize F_{GM}. This training process is summarized in Algorithm 1.

Algorithm 1 Training quantum model

Evaluate $B_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} q^2(\overrightarrow{x}_0^k|\overleftarrow{x}^k)$ for all $\overleftarrow{x}^k$;
Initialize the parameterized quantum model $U(\theta)$;
repeat
  Evaluate $A_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} p^2(\overrightarrow{x}_0^k|\overleftarrow{x}^k)$ for all $\overleftarrow{x}^k$;
  Evaluate $C_{\overleftarrow{x}^k} = \sum_{\overrightarrow{x}_0^k} p(\overrightarrow{x}_0^k|\overleftarrow{x}^k) q(\overrightarrow{x}_0^k|\overleftarrow{x}^k)$ for all $\overleftarrow{x}^k$;
  for every parameter θ_i **do**
    Evaluate $\frac{\partial}{\partial\theta_i} A_{\overleftarrow{x}^k}$ for all $\overleftarrow{x}^k$;
    Evaluate $\frac{\partial}{\partial\theta_i} C_{\overleftarrow{x}^k}$ for all $\overleftarrow{x}^k$;
    Compute F_{GM} and $\frac{\partial}{\partial\theta_i} F_{GM}$ from $A, B, C, \frac{\partial}{\partial\theta_i} A$ and $\frac{\partial}{\partial\theta_i} A$;
    Update θ_i with ADAM method;
until F_{GM} changes less than ϵ between two epochs;
return all θ_i;

4 Results

Numerical experiments – We have implemented our method in the training of a quantum model that simulates a renewal process with 3 causal states as shown in figure 3. The optimal divergence rate a 2-state ε-machine can achieve is about 0.04. But the quantum model with a 2-dimensional memory trained with out method can achieve the divergence rate of about 0.028, which is shown as the slope in figure 4.

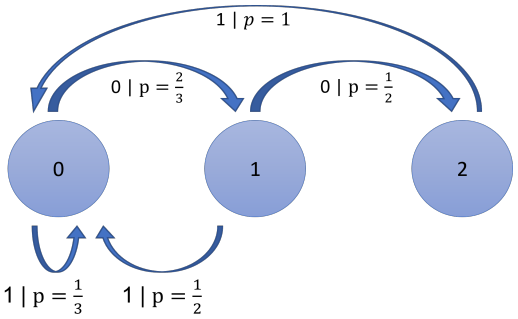


Figure 3: 3-state renewal process

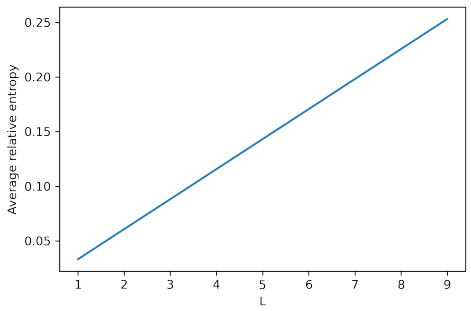


Figure 4: The averaged relative entropy plotted against the length of future. By definition, the divergence rate is the slope of the curve at $L \to \infty$.

4.1 4-state process

We also applied our algorithm to a 4 state ε-machine as shown in Fig. 5, whose Markov order is 2. For the choice of $p_{00} = 0.8$, $p_{01} = 0.1$,$p_{10} = 0.4$,$p_{11} = 0.1$, the optimal average diverge rate D_e of a 2 state classical machine is about 0.0131, while a quantum model with a 2-dimensional memory trained with our method can simulate the information source with divergence rate of 0.0009.

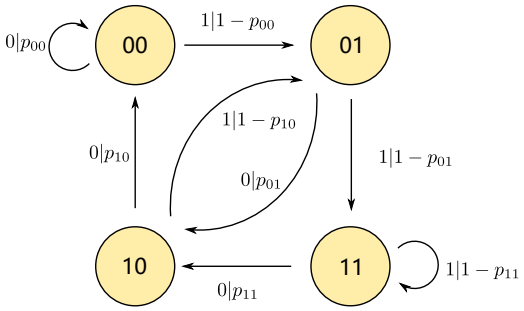


Figure 5: 4 state example

We also compare the performances of the quantum model trained with our method to a full quantum circuit that may simulate all possible distributions of a finite sequence. By varying p_{ij} several times, the results are summarized and plotted in figure 6. Although the full circuit can easily achieve a good performance on the training set, the quantum model trained with our method

can achieve a better generalization performance than the full quantum circuit.

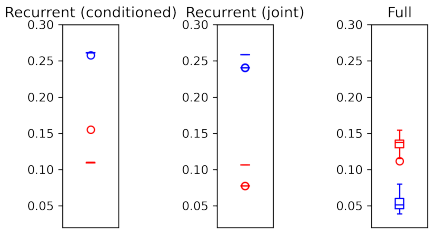


Figure 6: The distribution of the performances for i. recurrent circuit trained with the conditional distributions; ii. recurrent circuit trained with the stationary distribution; iii. the full circuit trained with the stationary distribution; are demonstrated in the three columns respectively. The red symbols represent the performances with respect to the real distribution, while the blue symbols represent the performances with respect to the distribution estimated from the samples. The performances are measured by the fidelity between the outputs of the circuit and the stationary distributions.

References

[1] R. B. Ash. *Information theory*. Courier Corporation, 2012.

[2] J. P. Crutchfield and K. Young. Inferring statistical complexity. *Physical review letters*, 63(2):105, 1989.

[3] Y.-C. Liang, Y.-H. Yeh, P. E. Mendonça, R. Y. Teh, M. D. Reid, and P. D. Drummond. Quantum fidelity measures for mixed states. *Reports on Progress in Physics*, 82(7):076001, 2019.

[4] M. Möttönen, J. J. Vartiainen, V. Bergholm, and M. M. Salomaa. Quantum circuits for general multiqubit gates. *Physical review letters*, 93(13):130502, 2004.

[5] A. Singhal et al. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.*, 24(4):35–43, 2001.

[6] X. Wang, C.-S. Yu, and X. Yi. An alternative quantum fidelity for mixed states of qudits. *Physics Letters A*, 373(1):58–60, 2008.

Fundamental limits on quantum cloning from the no-signalling principle

Yanglin Hu[1] * Marco Tomamichel[1 2] †

[1] *Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore*
[2] *Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117583, Singapore*

Abstract. The no-cloning theorem is a cornerstone of quantum cryptography. Here we generalize and rederive under weaker assumptions various upper bounds on the maximum achievable fidelity of probabilistic and deterministic cloning machines. Building on ideas by Gisin [Phys. Lett. A, 1998], our results hold even for cloning machines that do not obey the laws of quantum mechanics, as long as remote state preparation is possible and the non-signalling principle holds. We apply our general theorem to several subsets of states that are of interest in quantum cryptography.

Keywords: Remote state preparation, no-cloning, no-signaling

The no-cloning theorem states that within the framework of quantum mechanics, there does not exist any universal procedure that can replicate an unknown quantum state reliably. This fundamental principle was initially formalized in 1982 [1, 2] and later extended in various directions [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. The no-cloning theorem has rich implications in quantum cryptography, ensuring the security of primitives such as quantum money [14, 15], quantum key distribution [16] and quantum secret sharing [17, 18, 19].

However, quantum mechanics still has unsolved problems[20, 21] and may need further refinement [22, 23, 24, 25, 26, 27, 28]. It is essential to question whether the no-cloning theorem and consequentially the security of quantum cryptographic primitives still hold even if quantum mechanics is modified. The no-signalling principle states that information cannot be transmitted faster than light. It is believed to be a fundamental physical principle which will remain solid even if quantum mechanics is replaced by a more refined theory. It has been noticed that the no-cloning theorem remains true in different situations as long as the no-signalling principle is respected [29, 30, 31, 32, 33]. However, these works are restricted to specific situations.

This paper presents a general scheme to obtain bounds on the worst-case fidelity of any probabilistic or deterministic cloning machine from the no-signalling principle. Here we assume that the predictions of quantum mechanics are correct so that we can, in particular, perform a remote state preparation protocol, but we do not assume that the cloning machine itself adheres to the laws of quantum mechanics. Thus, we can restrict to cloning machines that are not necessarily linear in their input and are not necessarily positive or trace-preserving on general mixed input states.

This general scheme is versatile and can be employed to study fundamental limits for cloning machines that only attempt to clone various specific subsets of states and unitary gates. Moreover, the no-cloning bounds we derive from the no-signalling principle can sometimes reproduce the strongest no-cloning bounds that are derived assuming that the cloning machines are obeying the rules of quantum mechanics.

*yanglin.hu@u.nus.edu
†marco.tomamichel@nus.edu.sg

Probabilistic cloning machines. We consider probabilistic n-to-m cloning machines (or cloning oracles) $\mathcal{C}^{\mathcal{S}}_{n,m,p}$ for a set of pure states $\mathcal{S} = \{U|\psi_0\rangle\langle\psi_0|U^\dagger : U \in \mathcal{V}\}$ for some $|\psi_0\rangle \in \mathbb{C}^d$ and some unitary group $\mathcal{V}$ acting on $\mathbb{C}^d$. These machines map an n-fold tensor product of identical quantum states $|\psi\rangle\langle\psi| \in \mathcal{S}$ into a quantum state and a flag qubit denoting the success of the cloning machine with a non-zero success probability p, i.e.

$$\mathcal{C}^{\mathcal{S}}_{n,m,p}(|\psi\rangle\langle\psi|^{\otimes n}) = p\mathcal{R}^{\mathcal{S}}_{n,m}(|\psi\rangle\langle\psi|^{\otimes n}) \otimes |0\rangle\langle 0| + (1-p)\rho_\perp \otimes |1\rangle\langle 1|\,. \quad (1)$$

Here, $|0\rangle\langle 0|$ and $|1\rangle\langle 1|$ are flags for success and failure, respectively, $\mathcal{R}^{\mathcal{S}}_{n,m}(|\psi\rangle\langle\psi|^{\otimes n})$ is the output state in case of success and is expected to be an approximation to $|\psi\rangle\langle\psi|^{\otimes m}$, and $\rho_\perp$ is an arbitrary failure state. We do not assume that the map $\mathcal{R}^{\mathcal{S}}_{n,m}$ is linear, positive or trace-preserving, as long as it maps an n-fold quantum state from $\mathcal{S}$ to a valid quantum state on $\mathbb{C}^{d^m}$. (Our framework uses the formalism of quantum mechanics to describe the input and output spaces of the cloning machine but not the cloning machine itself.)

The performance of the cloning machine for a set $\mathcal{S}$ of states can be characterized by the worst-case global cloning fidelity, which is defined as

$$F_{\mathrm{wc}}(\mathcal{R}^{\mathcal{S}}_{n,m}) := \min_{\psi\in\mathcal{S}} F(\mathcal{R}^{\mathcal{S}}_{n,m}(|\psi\rangle\langle\psi|^{\otimes n}), |\psi\rangle\langle\psi|^{\otimes m}), \quad (2)$$

where F is the Uhlmann fidelity.

Remote identical state preparation. Let $\mathcal{W}$ be a subset of $\mathcal{S}$ with weights p_ϕ for $\phi \in \mathcal{W}$. If there exists a constant density matrix ρ_0 such that

$$(1-\epsilon)\rho_0 \le V^{\otimes n}\rho^n_{\mathcal{W}}(V^\dagger)^{\otimes n} \le (1+\epsilon)\rho_0, \forall V \in \mathcal{V}, \quad (3)$$

where $\rho^l_{\mathcal{W}} = \sum_{\phi\in\mathcal{W}} p_\phi |\phi\rangle\langle\phi|^{\otimes l}$, we can construct a remote identical state preparation protocol with ϵ accuracy in which Bob can remotely prepare identical states in $\mathcal{S}$

for Alice as follows. Initially Alice and Bob share an entangled state

$$|\Phi^n\rangle_{\mathsf{AB}} = C(\mathbb{I}\otimes L_{\mathsf{B}})|\Psi^{d,n}\rangle_{\mathsf{AB}},$$

where $|\Psi^{d,n}\rangle$ is the maximally entangled state on $\mathsf{AB}\simeq \mathbb{C}^{d^n}\otimes\mathbb{C}^{d^n}$ and C is the normalization factor. Alice and Bob measure $\{P_{\mathsf{A}},\mathbb{I}-P_{\mathsf{A}}\}$ and $\{P_{\mathsf{B}},\mathbb{I}-P_{\mathsf{B}}\}$ respectively. Their measurement outcomes are perfectly correlated. Either Alice obtains P_{A} and Bob obtains P_{B} respectively and both proceed, or Alice obtains $\mathbb{I}-P_{\mathsf{A}}$ and Bob obtains $\mathbb{I}-P_{\mathsf{B}}$ respectively and both abort. When both proceed, Bob further chooses a unitary $V\in\mathcal{V}$ according to the Haar measure on $\mathcal{V}$. Based on V, Bob measures ϕ by $p_\phi(L_{\mathsf{B}}^{-1})^\dagger\left((V|\phi\rangle\langle\phi|V^\dagger)^{\otimes n}\right)^{\intercal}L_{\mathsf{B}}^{-1}$ for $\phi\in\mathcal{W}$. Alice gets $(V|\phi\rangle\langle\phi|V^\dagger)^{\otimes n}$ if Bob gets ϕ.

Protocol 1 Remote identical state preparation

Require: Alice and Bob share $|\Phi^n\rangle_{\mathsf{AB}}$
Ensure: Alice gets $(V|\phi\rangle\langle\phi|V^\dagger)^{\otimes n}$ if Bob gets ϕ, or both abort
1: Alice measures $\{P_{\mathsf{A}},\mathbb{I}-P_{\mathsf{A}}\}$ on A
2: Bob measures $\{P_{\mathsf{B}},\mathbb{I}-P_{\mathsf{B}}\}$ on B
3: **if** Alice and Bob obtain P_{A} and P_{B} respectively **then**
4: Both know the success of the protocol
5: Bob chooses $V\in\mathcal{V}$
6: Bob measures $\phi\mapsto p_\phi(L_{\mathsf{B}}^{-1})^\dagger((V|\phi\rangle\langle\phi|V^\dagger)^{\otimes n})^{\intercal}L_{\mathsf{B}}^{-1}$
7: **else**
8: Both know the failure of the protocol
9: **end if**

No signaling principle. The no-signalling principle states that information cannot be transmitted between two separate parties without transmitting a physical system. Consider our remote identical state preparation protocol. The no-signalling principle forbids Alice to distinguish Bob's choice of measurement without communication. Thus, without knowing Bob's outcome, Alice's local state under any post-processing (including a probabilistic cloning machine conditioned on success, $\mathcal{R}^{\mathcal{S}}_{n,m}$) is independent of Bob's choice of measurement, i.e.,

$$\sum_{\phi\in\mathcal{W}} p_{\phi|V}\,\mathcal{R}\left((V|\phi\rangle\langle\phi|V^\dagger)^{\otimes n}\right)=\sigma_{\mathcal{R}}, \tag{4}$$

where $p_{\phi|V}$ is the probability of Bob measuring ϕ using V and the average output state $\sigma_{\mathcal{R}}$ is constant independent of V. Eq. 4 is the no-signalling condition we will rely on.

Main result. Our main result is a general bound on the worst-case global cloning fidelity.

Theorem 1 *Let $\mathcal{S}$ be the set of quantum states generated by some unitary group $\mathcal{V}$ and $\mathcal{W}$ be a subset of $\mathcal{S}$ with weights p_ϕ for $\phi\in\mathcal{W}$. If there exists a constant ρ_0 such that*

$$(1-\epsilon)\rho_0\le V^{\otimes n}\rho^n_{\mathcal{W}}(V^\dagger)^{\otimes n}\le(1+\epsilon)\rho_0,\forall V\in\mathcal{V}, \tag{5}$$

then $F_{\mathrm{wc}}(\mathcal{R}^{\mathcal{S}}_{n,m})$ is upper bounded by

$$F_{\mathrm{wc}}(\mathcal{R}^{\mathcal{S}}_{n,m})\le(1+\epsilon)^2F(\sigma_{\mathcal{V}},\rho^m_{\mathcal{W}}), \tag{6}$$

where $\sigma_{\mathcal{V}}=\int \mathrm{d}\mu(V)(V^\dagger)^{\otimes m}\sigma_{\mathcal{R}}V^{\otimes m}$ (the integration is over the Haar measure) and $\rho^l_{\mathcal{W}}=\sum_{\phi\in\mathcal{W}}p_\phi|\phi\rangle\langle\phi|^{\otimes l}$.

The proof, using the remote identical state preparation protocol, is given in the full text.

The intuition behind this bound is as follows. The resulting state $\sigma_{\mathcal{V}}$ has symmetry in the m-fold space because $[V^{\otimes m},\sigma_{\mathcal{V}}]=0$, while the state $\rho^l_{\mathcal{W}}$ only exhibits symmetry in the n-fold space when $l=n$ rather than in the m-fold space when $l=m$. This difference in symmetry between $\sigma_{\mathcal{V}}$ and $\rho^m_{\mathcal{W}}$ is leveraged to further upper bound the resulting fidelity, which we will explore in some examples below.

The main technical difficulties we overcome in the proof were the following. The first is constructing a general remote state preparation protocol where Bob can remotely prepare n identical quantum states for Alice. The second is reproducing the best known bounds that can be achieved assuming quantum-mechanical behaviour of the cloning machine. The latter requires a thorough understanding of the properties of $\mathcal{S}$ and $\mathcal{V}$. We achieve this for arbitrary states, multi-phase states, arbitrary Choi states (for 1-to-2 only) and multi-phase Choi states.

Example applications. Our general scheme is versatile to various subsets of quantum states $\mathcal{S}$ generated by $\mathcal{V}$ by choosing suitable L_{B}, P_{A}, P_{B}, and $\mathcal{W}$. These include arbitrary states, multi-phase states, spin coherent states, stabilizer states, Choi states of arbitrary unitaries, and Choi states of multi-phase unitaries. For illustration purposes, we present the result for arbitrary states, multi-phase states and Choi states of arbitrary unitaries.

Arbitrary states. For the set of arbitrary states $\mathcal{S}(d)$ generated by the set of arbitrary unitaries $\mathcal{U}(d)$, we can choose $(p,\mathcal{W})$ to be a weighted ϵ-approximate quantum n-design [34, 35]. A quantum n-design satisfies

$$(1-\epsilon)P^n_{\mathrm{sym}}\le V^{\otimes n}\rho^n_{\mathcal{W}}(V^\dagger)^{\otimes n}\le(1+\epsilon)P^n_{\mathrm{sym}},\forall V\in\mathcal{U}(d),$$

where P^n_{sym} is the projector to the n-fold symmetric subspace. As Eq. (3) is satisfied, we can construct the remote identical state preparation protocol as follows: let

$$P_{\mathsf{A}}=P_{\mathsf{B}}=P^n_{\mathrm{sym}},\quad L_{\mathsf{B}}=\mathbb{I}.$$

It is straight forward to check the correctness of the protocol. In later examples we will omit the remote identical state protocol. We can thus apply Theorem 1 to upper bound $F_{\mathrm{wc}}(\mathcal{R}^{\mathcal{S}(d)}_{n,m})$. Besides, the high symmetries impose that both $\sigma_{\mathcal{V}}$ and $\rho^m_{\mathcal{W}}$ in Eq. (6) are block-diagonalized, which makes an explicit no-cloning bound possible. By further optimizing $\sigma_{\mathcal{V}}$ and $\rho^m_{\mathcal{W}}$ over all possible block-diagonalized matrices under the constraint imposed by the m-fold and n-fold symmetries respectively, we can upper bound $F_{\mathrm{wc}}(\mathcal{R}^{\mathcal{S}(d)}_{n,m})$ by the size of the weighted ϵ-approximate quantum n-design divided by the dimension of the m-fold symmetric subspace.

Corollary 2 *$F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{S}(d)})$ for $\mathcal{S}(d)$ is upper bounded by the size of an ϵ-approximate quantum n-design $|\mathcal{W}|$*

$$F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{S}(d)}) \leq (1+\epsilon)^2 \frac{|\mathcal{W}|}{d[m]}. \tag{7}$$

Using the size of quantum n-design [34], for large d and fixed n and m, we obtain $F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{S}(d)}) \leq O(d^{n-m}\,\text{polylog}\,d)$ from the no-signalling principle while the optimal quantum mechanical bound is $F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{S}(d)}) \leq O(d^{n-m})$ [5].

Multi-phase states. A d-dimensional multi-phase state on $\mathbb{C}^d$ is defined as a state in the form of $|\phi_\theta\rangle = \frac{1}{\sqrt{d}}\sum_i e^{\mathrm{i}\theta_i}|i\rangle$. We consider

$$\mathcal{M}(d) = \left\{ |\phi_\theta\rangle = \frac{1}{\sqrt{d}}\sum_i e^{\mathrm{i}\theta_i}, \forall \theta \in [0,2\pi)^{\times d} \right\},$$
$$\mathcal{T}(d) = \left\{ V_\theta = \sum_i e^{\mathrm{i}\theta_i}|i\rangle\langle i|, \forall \theta \in [0,2\pi)^{\times d} \right\}.$$

Let $\alpha \in \mathbb{Z}_{n+1}^{\times d}$ with $\sum_i \alpha_i = n$. We define $|\alpha\rangle = D\sum_{\pi\in \text{Perm(n)}} \pi(|i_1\rangle...|i_n\rangle)$ on $\mathbb{C}^{d^n}$ where Perm(n) is the permutation group, $|i_1\rangle...|i_n\rangle$ is a state with α_i subsystems in state $|i\rangle$ and D is the normalization factor. Let $\beta \in \mathbb{Z}_{m+1}^{\times d}$ with $\sum_i \beta_i = m$ and similarly define $|\beta\rangle$ on $\mathbb{C}^{d^m}$. Let $\theta_0 = \frac{2\pi}{n+1}$. Readers may check that for the equally weighted set

$$\mathcal{W} = \left\{ |\phi_k\rangle = \frac{1}{\sqrt{d}}\sum_i e^{\mathrm{i}\theta_0 k_i}|i\rangle, \forall k \in \mathbb{Z}_{n+1}^{\times d} \right\},$$

we have

$$V_\theta^{\otimes}\rho_{\mathcal{W}}^n (V_\theta^\dagger)^{\otimes n} = \frac{1}{d^n}\sum_\alpha \frac{n!}{\alpha_1!...\alpha_d!}|\alpha\rangle\langle\alpha|, \forall V_\theta \in \mathcal{T}(d).$$

Therefore, Eq. (3) holds. Then we can obtain an upper bound on $F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{M}(d)})$ following Theorem 1.

Corollary 3 *$F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{M}(d)})$ for $\mathcal{M}(d)$ is upper bounded by*

$$F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{M}(d)}) \leq \sum_\nu \max_{\beta\equiv\nu \bmod (n+1)} \frac{1}{d^m}\frac{m!}{\beta_1^\nu!...\beta_d^\nu!}. \tag{8}$$

In the asymptotic limit for large n and m and fixed d, the right-hand side of Eq. (8) can be approximated by the error function, and we obtain

$$F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{M}(d)}) \leq \sqrt{d\left(1-\frac{1}{d}\right)^{d-1}}\ \text{erf}\left(\frac{dn}{2\sqrt{2(d-1)m}}\right)^{d-1}.$$

Our bound shows that multi-phase states can be super-replicated (i.e. n-to-m cloned) probabilistically only if $m = O(n^2)$ for fixed d, which reproduces the quantum mechanical upper bound [11, 12].

Choi states for arbitrary unitaries. The set of Choi states for the unitary group $\mathcal{U}(d)$ on $\mathbb{C}^d$ is

$$\mathcal{J}_{\mathcal{U}(d)} = \left\{ |J_U\rangle = (U\otimes\mathbb{I})|\Psi^{d,1}\rangle, \forall U \in \mathcal{U}(d) \right\} \tag{9}$$

where $|\Psi^{d,1}\rangle$ is the maximally entangled state on $\mathbb{C}^d\otimes\mathbb{C}^d$. The Schur-Weyl duality [36] states that

$$\mathbb{C}^{d^n} \simeq \bigoplus_\lambda \mathbb{P}_\lambda^n \otimes \mathbb{Q}_\lambda^n, \tag{10}$$

where $\mathbb{P}_\lambda^n$ are irreducible representations (irreps) of the unitary group $\mathcal{U}(d)$ and $\mathbb{Q}_\lambda^n$ are irreps of the permutation group Perm(n). We denote the Schur basis by $\{|p_\lambda^n, q_\lambda^n, \lambda\rangle, \forall\lambda, p_\lambda^n, q_\lambda^n\} \simeq \{|p_\lambda^n,\lambda\rangle|q_\lambda^n,\lambda\rangle, \forall\lambda, p_\lambda^n, q_\lambda^n\}$, where $\{|p_\lambda^n\rangle, \forall p_\lambda^n\}$ and $\{|q_\lambda^n,\lambda\rangle, \forall q_\lambda^n\}$ are basis of $\mathbb{P}_\lambda^n$ and $\mathbb{Q}_\lambda^n$ respectively. We denote the projector on $\mathbb{P}_\lambda^n$ and $\mathbb{Q}_\lambda^n$ by P_λ^n and Q_λ^n as well as the dimension of $\mathbb{P}_\lambda^n$ and $\mathbb{Q}_\lambda^n$ by $|\mathbb{P}_\lambda^n|$ and $|\mathbb{Q}_\lambda^n|$. Let

$$\Pi_\lambda^n = P_\lambda^n \otimes P_\lambda^n \otimes |\Psi_\lambda^n\rangle\langle\Psi_\lambda^n|, \tag{11}$$

be a projector on $\mathbb{C}^{d^{2n}}$ where

$$|\Psi_\lambda^n\rangle = \frac{1}{\sqrt{|\mathbb{Q}_\lambda^n|}}\sum_{q_\lambda^n} |q_\lambda^n,\lambda\rangle|q_\lambda^n,\lambda\rangle, \tag{12}$$

the maximally entangled state on $\mathbb{Q}_\lambda^n \otimes \mathbb{Q}_\lambda^n$. Readers may verify that for the weighted set of Choi states $(p,\mathcal{W})$ for the weighted relative ϵ-approximate unitary n-design $(p,\mathcal{U}_n)$ [37, 38, 39, 40], it holds

$$(1-\epsilon)\sum_\lambda \frac{|\mathbb{Q}_\lambda^n|}{d^n|\mathbb{P}_\lambda^n|}\Pi_\lambda^n \leq (V^{\otimes n}\otimes\mathbb{I})\rho_{\mathcal{W}}^n((V^\dagger)^{\otimes n}\otimes\mathbb{I}) \leq (1+\epsilon)\sum_\lambda \frac{|\mathbb{Q}_\lambda^n|}{d^n|\mathbb{P}_\lambda^n|}\Pi_\lambda^n, \forall V\in\mathcal{U}(d).$$

Eq. (3) immediately follows. We can thus construct the remote identical state preparation protocol. An upper bound on $F_{\rm wc}(\mathcal{R}_{n,m}^{\mathcal{J}_{\mathcal{U}(d)}})$ then follows from Theorem 1.

Corollary 4 *$F_{\rm wc}\left(\mathcal{R}_{n,m}^{\mathcal{J}_{\mathcal{U}(d)}}\right)$ is upper bounded by the size of an ϵ-approximate unitary n-design $|\mathcal{W}|$ via*

$$F_{\rm wc}\left(\mathcal{R}_{n,m}^{\mathcal{J}_{\mathcal{U}(d)}}\right) \leq (1+\epsilon)^2\frac{|\mathcal{W}|}{d^{2m}}\sum_\lambda |\mathbb{Q}_\lambda^m|^2. \tag{13}$$

The above upper bound is implicit, due to the hard summation. However, an explicit upper bound can be further derived for the special case $F_{\rm wc}(\mathcal{R}_{1,2}^{\mathcal{J}_{\mathcal{U}(2^r)}})$.

Corollary 5 *$F_{\rm wc}(\mathcal{R}_{1,2}^{\mathcal{J}_{\mathcal{U}(2^r)}})$ is upper bounded by*

$$F_{\rm wc}\left(\mathcal{R}_{1,2}^{\mathcal{J}(2^r)}\right) \leq \frac{2^r+\sqrt{2^{2r}-1}}{2^{3r}}. \tag{14}$$

This explicit bound exactly reproduces the quantum mechanical upper bound $F_{\rm wc}(\mathcal{R}_{1,2}^{\mathcal{J}_{\mathcal{U}(d)}}) \leq \frac{d+\sqrt{d^2-1}}{d^3}$ for $d=2^r$ [41] . When $d\neq 2^r$, numerical calculations with small d's and sub-optimal unitary 2-designs indicate that our bound approximates the quantum mechanical bound well.

References

[1] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, pp. 802–803, 1982. [Online]. Available: https://www.nature.com/articles/299802a0

[2] D. Dieks, "Communication by epr devices," *Physics Letters A*, vol. 92, no. 6, pp. 271–272, 1982. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0375960182900846

[3] N. Gisin and B. Huttner, "Quantum cloning, eavesdropping and bell's inequality," *Physics Letters A*, vol. 228, no. 1, pp. 13–21, 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0375960197000832

[4] N. Gisin and S. Massar, "Optimal quantum cloning machines," *Phys. Rev. Lett.*, vol. 79, pp. 2153–2156, Sep 1997. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.79.2153

[5] R. F. Werner, "Optimal cloning of pure states," *Phys. Rev. A*, vol. 58, pp. 1827–1832, Sep 1998. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.58.1827

[6] H. Fan, K. Matsumoto, X.-B. Wang, and M. Wadati, "Quantum cloning machines for equatorial qubits," *Phys. Rev. A*, vol. 65, p. 012304, Dec 2001. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.65.012304

[7] F. Buscemi, G. M. D'Ariano, and C. Macchiavello, "Economical phase-covariant cloning of qudits," *Phys. Rev. A*, vol. 71, p. 042327, Apr 2005. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.71.042327

[8] R. Demkowicz-Dobrzański, M. Kuś, and K. Wódkiewicz, "Cloning of spin-coherent states," *Phys. Rev. A*, vol. 69, p. 012301, Jan 2004. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.69.012301

[9] L.-M. Duan and G.-C. Guo, "Probabilistic cloning and identification of linearly independent quantum states," *Phys. Rev. Lett.*, vol. 80, pp. 4999–5002, Jun 1998. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.80.4999

[10] J. Fiurášek, "Optimal probabilistic cloning and purification of quantum states," *Phys. Rev. A*, vol. 70, p. 032308, Sep 2004. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.70.032308

[11] G. Chiribella, Y. Yang, and A. C.-C. Yao, "Quantum replication at the heisenberg limit," *Nature communications*, vol. 4, no. 1, p. 2915, 2013.

[12] W. Dür, P. Sekatski, and M. Skotiniotis, "Deterministic superreplication of one-parameter unitary transformations," *Phys. Rev. Lett.*, vol. 114, p. 120503, Mar 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.114.120503

[13] G. Chiribella, Y. Yang, and C. Huang, "Universal superreplication of unitary gates," *Phys. Rev. Lett.*, vol. 114, p. 120504, Mar 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.114.120504

[14] S. Wiesner, "Conjugate coding," *SIGACT News*, vol. 15, no. 1, p. 78–88, jan 1983. [Online]. Available: https://doi.org/10.1145/1008908.1008920

[15] S. Aaronson, E. Farhi, D. Gosset, A. Hassidim, J. Kelner, and A. Lutomirski, "Quantum money," *Commun. ACM*, vol. 55, no. 8, p. 84–92, aug 2012. [Online]. Available: https://doi.org/10.1145/2240236.2240258

[16] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, vol. 560, pp. 7–11, 2014, theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0304397514004241

[17] M. Hillery, V. Bužek, and A. Berthiaume, "Quantum secret sharing," *Phys. Rev. A*, vol. 59, pp. 1829–1834, Mar 1999. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.59.1829

[18] R. Cleve, D. Gottesman, and H.-K. Lo, "How to share a quantum secret," *Phys. Rev. Lett.*, vol. 83, pp. 648–651, Jul 1999. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.83.648

[19] D. Gottesman, "Theory of quantum secret sharing," *Phys. Rev. A*, vol. 61, p. 042311, Mar 2000. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.61.042311

[20] A. Peres and D. R. Terno, "Quantum information and relativity theory," *Rev. Mod. Phys.*, vol. 76, pp. 93–123, Jan 2004. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.76.93

[21] A. Bassi, K. Lochan, S. Satin, T. P. Singh, and H. Ulbricht, "Models of wave-function collapse, underlying theories, and experimental tests," *Rev. Mod. Phys.*, vol. 85, pp. 471–527, Apr 2013. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.85.471

[22] D. E. Kaplan and S. Rajendran, "Causal framework for nonlinear quantum mechanics," *Phys. Rev. D*, vol. 105, p. 055002, Mar 2022. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.105.055002

[23] M. Polkovnikov, A. V. Gramolin, D. E. Kaplan, S. Rajendran, and A. O. Sushkov, "Experimental limit on nonlinear state-dependent terms in quantum theory," *Phys. Rev. Lett.*, vol. 130, p. 040202,

Jan 2023. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.130.040202

[24] G. C. Ghirardi, A. Rimini, and T. Weber, "Unified dynamics for microscopic and macroscopic systems," *Phys. Rev. D*, vol. 34, pp. 470–491, Jul 1986. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevD.34.470

[25] P. Pearle, "Combining stochastic dynamical state-vector reduction with spontaneous localization," *Phys. Rev. A*, vol. 39, pp. 2277–2289, Mar 1989. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.39.2277

[26] G. C. Ghirardi, P. Pearle, and A. Rimini, "Markov processes in hilbert space and continuous spontaneous localization of systems of identical particles," *Phys. Rev. A*, vol. 42, pp. 78–89, Jul 1990. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.42.78

[27] L. Diósi, "Models for universal reduction of macroscopic quantum fluctuations," *Phys. Rev. A*, vol. 40, pp. 1165–1174, Aug 1989. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.40.1165

[28] A. Bassi, "Collapse models: analysis of the free particle dynamics," *Journal of Physics A: Mathematical and General*, vol. 38, no. 14, p. 3173, mar 2005. [Online]. Available: https://dx.doi.org/10.1088/0305-4470/38/14/008

[29] N. Gisin, "Quantum cloning without signaling," *Physics Letters A*, vol. 242, no. 1, pp. 1–3, 1998. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0375960198001704

[30] L. Hardy and D. D. Song, "No signalling and probabilistic quantum cloning," *Physics Letters A*, vol. 259, no. 5, pp. 331–333, 1999. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S037596019900448X

[31] A. K. Pati, "Probabilistic exact cloning and probabilistic no-signalling," *Physics Letters A*, vol. 270, no. 3, pp. 103–107, 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0375960100002814

[32] L. Masanes, A. Acin, and N. Gisin, "General properties of nonsignaling theories," *Phys. Rev. A*, vol. 73, p. 012112, Jan 2006. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.73.012112

[33] P. Sekatski, M. Skotiniotis, and W. Dür, "No-signaling bounds for quantum cloning and metrology," *Phys. Rev. A*, vol. 92, p. 022355, Aug 2015. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.92.022355

[34] A. Ambainis and J. Emerson, "Quantum t-designs: t-wise independence in the quantum world," in *Twenty-Second Annual IEEE Conference on Computational Complexity (CCC'07)*, 2007, pp. 129–140. [Online]. Available: https://ieeexplore.ieee.org/document/4262758

[35] A. W. Harrow, "The church of the symmetric subspace," 2013. [Online]. Available: https://arxiv.org/abs/1308.6595

[36] R. Goodman and N. R. Wallach, *Representations and invariants of the classical groups*. Cambridge University Press, 2000. [Online]. Available: https://www.cambridge.org/sg/academic/subjects/mathematics/algebra/representations-and-invariants-classical-groups?format=PB&isbn=9780521663489

[37] C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation," *Phys. Rev. A*, vol. 80, p. 012304, Jul 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.80.012304

[38] R. A. Low, "Pseudo-randomness and learning in quantum computation," 2010.

[39] Y. Nakata, D. Zhao, T. Okuda, E. Bannai, Y. Suzuki, S. Tamiya, K. Heya, Z. Yan, K. Zuo, S. Tamate, Y. Tabuchi, and Y. Nakamura, "Quantum circuits for exact unitary t-designs and applications to higher-order randomized benchmarking," *PRX Quantum*, vol. 2, p. 030339, Sep 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PRXQuantum.2.030339

[40] J. Haferkamp, "Random quantum circuits are approximate unitary t-designs in depth $O\left(nt^{5+o(1)}\right)$," *Quantum*, vol. 6, p. 795, Sep. 2022. [Online]. Available: https://doi.org/10.22331/q-2022-09-08-795

[41] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Optimal cloning of unitary transformation," *Phys. Rev. Lett.*, vol. 101, p. 180504, Oct 2008. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.101.180504

Inference-Based Quantum Sensing

C. Huerta Alderete[1 2 3 *] Max Hunter Gordon[4 5] Frédéric Sauvage[4] Akira Sone[6]
Andrew T. Sornborger[1 3] Patrick J. Coles[3 4] M. Cerezo[1 3]

[1]*Information Sciences, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2]*Materials Physics and Applications Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.*
[3]*Quantum Science Center, Oak Ridge, TN 37931, USA*
[4]*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[5]*Instituto de Física Teórica, UAM/CSIC, Universidad Autónoma de Madrid, Madrid 28049, Spain*
[6]*Aliro Technologies, Inc, Boston, MA 02135, USA*

Abstract. In a standard Quantum Sensing (QS) task one aims at estimating an unknown parameter θ, encoded into an n-qubit probe state, via measurements of the system. The success of this task relies on the ability to correlate changes in the parameter to changes in the response function, $\mathcal{R}(\theta)$. While cases have known forms of the response function, realistic scenarios lack a general closed-form expression. Our approach characterizes $\mathcal{R}(\theta)$ by measuring the response at $2n+1$ parameters, enabling inference of unknown parameters and determination of scheme sensitivity. We demonstrate that the inference error is likely smaller than δ with a number of shots scaling as $\Omega(\log^3(n)/\delta^2)$. The framework accommodates arbitrary probe states, measurement schemes, and quantum noise. We validate the method through experiments on real quantum hardware and numerical simulations.

Keywords: Quantum Sensing

1 Introduction

Quantum Sensing (QS) is a rapidly growing field of study and plays a crucial role in practical quantum technologies [1]. In QS tasks, the objective is to estimate an unknown parameter θ, which is encoded in an n-qubit probe state, by making measurements of the system. The success of this task relies on the ability to observe how changes in the parameter affect the measurement outcomes. In simpler cases, such as an idealized magnetometry experiment, the mathematical relationship, known as the response function $\mathcal{R}(\theta)$, between the system and the parameter is well-established. The commonly used probe state is the n-qubit GHZ state, combined with parity measurement [2, 3]. This leads to a response function given by $\mathcal{R}(\theta) = \cos(n\theta)$, from where one can determine the magnetic field and sensitivity [4]. However, in realistic scenarios, the explicit functional form of $\mathcal{R}(\theta)$ may not be readily accessible and would require complete characterization of the device. While recent research has attempted to address the limitations of implementing QS schemes on noisy quantum hardware [8, 6, 9, 10, 11, 12, 13], methods for recovering the true $\mathcal{R}(\theta)$ in real-time are still lacking.

In this work [14] we present a novel data-driven inference approach to recover the true response of the system in an efficient and scalable manner. We provide rigorous theoretical guarantees for the performance of our framework, which we verify with numerical simulations and experiments on IBM quantum computers.

2 Results

Leveraging tools from variational quantum algorithms, quantum machine learning and polynomial interpolation [15, 16, 17], we characterize the exact functional form of $\mathcal{R}(\theta)$ for a general class of unitary families. We focus on the case of unitary families where the parameter encoding mechanism is of the form $\mathcal{S}_\theta(\rho) = e^{-i\theta H/2}\rho\, e^{i\theta H/2} = \rho_\theta$. Here, H is a Hermitian operator such that $H = \sum_j h_j$ with $h_j^2 = 1$, and $[h_j, h_{j'}] = 0$, $\forall j, j'$ and $\rho = \mathcal{E}(\rho_{\text{in}})$ is a n-qubit probe state prepared by sending a fiduciary state ρ_{in} through a state preparation channel $\mathcal{E}$. Assumptions made over the Hamiltonian include the well-know magnetometry tasks. We allow for the possibility of sending ρ_θ through a second pre-measurement channel $\mathcal{D}$, over which we measure the expectation value of an observable O. That is, we consider that the system response is of the form $\mathcal{R}(\theta) = \text{Tr}[\mathcal{D} \circ \mathcal{S}_\theta \circ \mathcal{E}(\rho_{\text{in}})O]$. This setting encompasses cases where $\mathcal{E}$ or $\mathcal{D}$ are noisy channels, as well as cases of imperfect parameter encoding where a θ-independent noise channel acts after $\mathcal{S}_\theta$ [14].

Our first main result is to show that the system response can be exactly expressed as a trigonometric polynomial of degree n. That is,

$$\mathcal{R}(\theta) = \sum_{s=1}^{n} [a_s \cos(s\theta) + b_s \sin(s\theta)] + c \tag{1}$$

with $\{a_s, b_s\}_{s=1}^n$ and c being real valued coefficients. This *exact* functional form provides a relationship between the encoded parameter θ and the system response, which can be used to improve accuracy in tasks of parameter estimation, as well as to faithfully determine the sensitivity of the experimental set-up [18]. Since $\mathcal{R}(\theta)$ is a trigonometric polynomial of degree n, one can *fully learn* the system response by only measuring it at a set of $2n+1$ known θ parameters: $\{\mathcal{R}(\theta_k)\}_{k=1}^{2n+1}$. In practice, however, one cannot exactly evaluate each $\mathcal{R}(\theta_k)$, but rather can only estimate them up to some statistical uncertainty resulting from finite sampling. If we use the

*aldehuer@lanl.gov

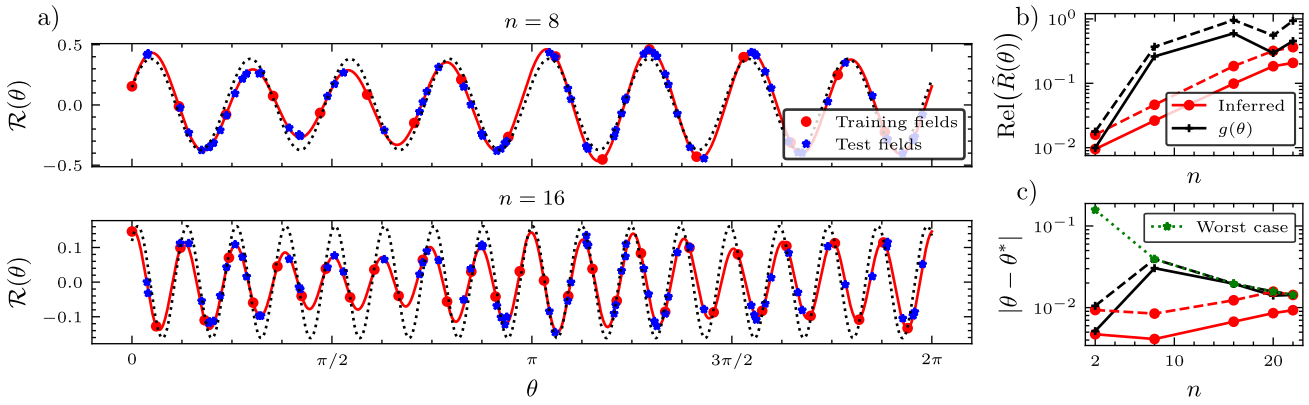


Figure 1: **Magnetometry task on IBM hardware**. a) Inferred response $\widetilde{\mathcal{R}}(\theta)$ for $n = 8, 16$ qubits. The fields used to train (red point) and test (blue star) the inference scheme, were estimated on the *IBM_Montreal* quantum computer. We depict the inferred response $\widetilde{\mathcal{R}}(\theta)$ (red solid curve) as well as the fit $g(\theta) = \alpha\cos(\beta\theta + \gamma) + \zeta$ (black dotted curve). b) Relative response error versus n. Statistics were obtained over 74 test fields and 7 experiment repetitions. The relative error is defined as the difference between the fit or inferred value and the measurement response, normalized by the average test expectation value. The red (black) points correspond to $\widetilde{\mathcal{R}}(\theta)$ ($g(\theta)$), while solid (dashed) lines represent the median (upper quartile) error. c) Parameter prediction error versus n, with green dots denoting the worst possible prediction.

N-shot estimates $\overline{\mathcal{R}}(\theta_k)$ (with $\theta_k = 2\pi(k-1)/(2n+1)$ and $k = 1, \ldots, 2n+1$) to construct an *approximated inferred response* function $\widetilde{\mathcal{R}}(\theta)$, we can rigorously show that for all θ, $|\mathcal{R}(\theta) - \widetilde{\mathcal{R}}(\theta)| \in \mathcal{O}(\varepsilon \log(n))$, where we recall that $\mathcal{R}(\theta)$ is *exact response* function, where we defined $\varepsilon = \max_{\theta_k \in P} |\mathcal{R}(\theta_k) - \overline{\mathcal{R}}(\theta_k)|$. The maximum estimation error ε is fundamentally related to the number of shots N, which enables us to derive the following theorem.

Theorem 1: *The number of shots N, necessary to ensure that with a (constant) high probability, and for all θ, the error $|\mathcal{R}(\theta) - \widetilde{\mathcal{R}}(\theta)| \leq \delta$, for an inference error δ, is in $\Omega\left(\log^3(n)/\delta^2\right)$.*

Theorem 1 implies that for fixed δ, a poly-logarithmic number of shots $N \in \Omega(\log^3(n))$ suffices to guarantee that $\widetilde{\mathcal{R}}(\theta)$ will be a good approximation for the true response $\mathcal{R}(\theta)$.

Once the inferred response $\widetilde{\mathcal{R}}(\theta)$ is obtained, it can be employed for two tasks of central importance in QS: parameter estimation and characterization of the sensitivity for a sensing apparatus. Theoretical guarantees of the performance in these two tasks is sketched below. When predicting the value of an unknown parameter θ', we assume that an estimate of the system response $\overline{\mathcal{R}}(\theta')$ is provided, and the promise that θ' is sampled from a known domain Θ. In such a case, the unknown parameter can be estimated as $\theta^* = \mathrm{argmin}_{\theta \in \Theta} |\widetilde{\mathcal{R}}(\theta) - \overline{\mathcal{R}}(\theta')|$. In many cases of interest, such as high-precision estimation of small magnetic fields, Θ will be small enough such that $\widetilde{\mathcal{R}}(\theta)$ is bijective, ensuring that the solution θ^* is unique.

Theorem 2: *Let ϵ' be the estimation error in $\overline{\mathcal{R}}(\theta')$ for some θ' in a known domain Θ where the system response is bijective. Let χ be the error introduced when estimating θ' via $\widetilde{\mathcal{R}}(\theta)$ relative to the case when the exact response $\mathcal{R}(\theta)$ is used. The number of shots, N, necessary to ensure that with a (constant) high probability χ is no greater than δ' is $\Omega(\log^3(n)/(\delta' + \varepsilon')^2)$.*

Theorem 2 certifies that $\widetilde{\mathcal{R}}(\theta)$ can be used to infer an unknown parameter from a measured system response without incurring additional uncertainties as long as enough shots are used. In fact, for fixed δ' and ε', one only needs a poly-logarithmic number of shots.

As previously mentioned, given $\overline{\mathcal{R}}(\theta)$ one can also directly compute the sensitivity of the QS scheme at a field $\theta = \theta_l$ via the error propagation formula [18]. In particular, the N-shot inferred response function leads to an approximate sensitivity $\Delta\widetilde{\theta}_l = (\Delta\widetilde{\mathcal{R}}(\theta_l))/(|\partial_\theta \widetilde{\mathcal{R}}(\theta)|_{\theta=\theta_l}|)$.

Theorem 3: *Let $\mathcal{R}(\theta)$ be the exact response function, and let $\widetilde{\mathcal{R}}(\theta)$ be its approximation obtained from the N-shot estimates $\overline{\mathcal{R}}(\theta_k)$ with θ_k. Defining the maximum estimation error $\varepsilon = \max_{\theta_k \in P} |\mathcal{R}(\theta_k) - \overline{\mathcal{R}}(\theta_k)|$, and the slope of $\mathcal{R}(\theta)$ at a field θ_k $D_l = |\partial_\theta \widetilde{\mathcal{R}}(\theta)|_{\theta=\theta_l}|$, then $|\Delta\theta - \Delta\widetilde{\theta}| \in \mathcal{O}\left(\frac{\varepsilon \log(n)}{D_l}\right)$.*

The above theorem indicates that using the inferred response to estimate the sensitivity leads to an error that scales linearly ε and $\log(n)$, but inversely proportional to D_l (which is expected as sensing schemes with flat response functions lead to high estimation errors). Which leads to the following corollary exploring how the error scales with the number of shots:

Corollary 1: *The number of shots N, necessary to ensure that with a (constant) high probability the error in the sensitivity $|\Delta\theta - \Delta\widetilde{\theta}| \leq \delta''$ at parameter $\theta = \theta_l$, for an inference sensitivity error δ'', is in $\Omega\left(\frac{\log^3(n)}{(D_l \delta'')^2}\right)$.*

Hence, we show that using an inference based approach to QS allows for good characterization of the systems performance with a resource requirement that scales poly-logarithmically in the system size. The proofs for the above results are included in [14].

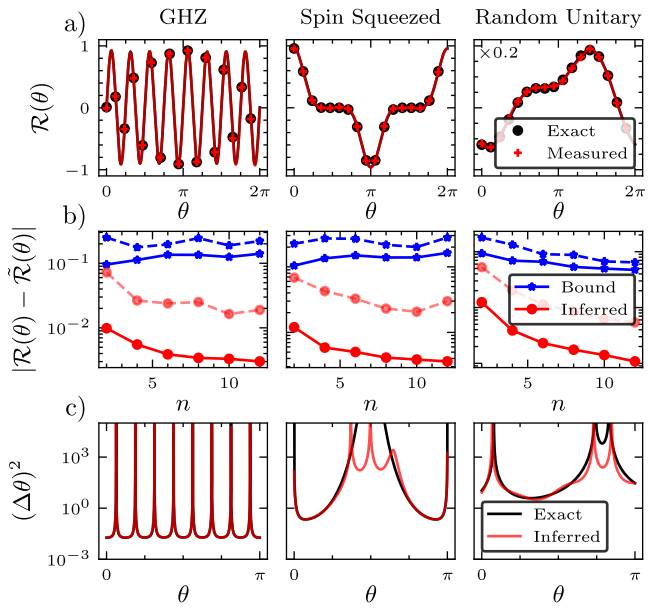


Figure 2: **Numerical results for QS tasks.** a) System response versus θ for $n = 8$ qubits in all three QS setups described in the main text. The exact response $\mathcal{R}(\theta)$ (black curve), and its value at the training fields $\mathcal{R}(\theta_k)$ (black points), were obtained with no finite sampling. In contrast, the response estimated at the training fields $\overline{\mathcal{R}}(\theta_k)$ (red crosses), and the resulting inferred function $\widetilde{\mathcal{R}}(\theta)$ (red curve), was obtained with a poly-logarithmic number of shots. b) Median (solid) and maximum (dashed) of the error $|\mathcal{R}(\theta) - \widetilde{\mathcal{R}}(\theta)|$ (red) and the bound of $\mathcal{O}(\varepsilon \log(n))$ (blue) for 10^4 test fields uniformly sampled over $(0, 2\pi)$. The statistics were obtained over 30 repetitions of the experimental setups. c) The black (red) curves depict the exact (inferred) sensitivity versus θ.

To showcase our method we implement a QS task on the *IBM_Montreal* quantum hardware for systems of up to $n = 22$ qubits, and complement our findings with numerical simulations. The experimental implementation in the *IBM* device consists in preparing the GHZ state with $H = \sum_{j=1}^{n} Z_j$, and measuring the parity operator $O = \bigotimes_{j=1}^{n} X_j$. Here, Z_j and X_j are the Pauli z and x operators acting on the j-th qubit, respectively.

The $2n+1$ training field estimates are used to infer the response $\widetilde{\mathcal{R}}(\theta)$, as well as to fit a first order approximation of a noisy response $g(\theta) = \alpha \cos(\beta\theta + \gamma) + \zeta$, where the coefficients α, β, γ and ζ account for the first order effects of hardware noise. To evaluate the ability of these two functions to recover the true response of the system, we compare their predictions against the measured system response at a set of random test fields.

In Fig. 1(a) we display inference results for $n = 8$ and $n = 16$ qubits, indicating that our method (red solid curve) is clearly able to fit the training and test fields better than the cosine response (black dotted curve). More quantitatively, in Fig. 1(b), we show the scaling of the error as a function of the system size. One can see that for all problem sizes considered our method leads to smaller response prediction error. We note that for larger n the effect of noise becomes more prominent, as the hardware noise suppresses the measured expectation values [19, 20, 21]. In this regime both methods are equally limited by finite sampling noise which becomes of the same order as the magnitude of the response. Still, even for system sizes as large as $n = 22$ qubits, the inference method reduces the relative error by a factor larger than two when compared to that of the $g(\theta)$ fit. Finally, we also use $\widetilde{\mathcal{R}}(\theta)$ and $g(\theta)$ for parameter estimation, i.e., to determine an unknown magnetic field encoded in the quantum state. As shown in Fig. 1(c), the $g(\theta)$ fit matches the worst possible prediction for $n = 8$ qubits whereas our inference method can outperform the $g(\theta)$ fit by up to one order of magnitude.

We further explore the applicability of the inference based framework with numerical results from a density matrix simulator and a realistic ion trap noise model, but where finite sampling can be included or omitted. We emulate three different sensing setups. First, once again we study the standard GHZ magnetometry setting. Second, we characterize the squeezing in a system where the probe state is a spin coherent state, $H = \sum_{j<k} X_j X_k$ is the one-axis twisting Hamiltonian [22], and $O = Z_n$. Finally, we study a scenario where the probe state is constructed by a unitary composed of 4 layers of a hardware efficient ansatz with random parameters [23, 24], $H = \sum_{j=1}^{n-1} Z_j Z_{j+1}$ and $O = \frac{1}{n}\sum_{i=1}^{n} X_i$, relevant for variational quantum metrology [8, 13, 6, 9, 12], where one wishes to prepare a probe state via some parameterized quantum circuit. As motivated in [14], $\widetilde{\mathcal{R}}(\theta)$ is inferred with $N = \lceil 5 \times 10^2 \log(n)^2 \log(2 \times 10^2 (2n+1)) \mathcal{R}eil$ shots per θ_k. Figure 2(a) shows that the inferred response (red curve) closely matches the exact one (black curve) in all three QS settings considered. In Fig. 2(b) we further show the scaling of the error $|\mathcal{R}(\theta) - \widetilde{\mathcal{R}}(\theta)|$ with respect to the system size. Indeed, we can see that allocating a number of shots N that increases poly-logarithmically with n allows the error to decrease with increasing system size. Finally, in Fig. 2(c), we use $\widetilde{\mathcal{R}}(\theta)$ to estimate the sensitivity of the three experimental setups. The sensitivity diverges in parameter regions where the experimental setup is insensitive to the field (when the response function tends to a zero) otherwise, our method (red curves) recovers the behavior of the exact sensitivity (black curves).

Conclusions. Leveraging techniques from quantum machine learning and polynomial interpolation [15, 16, 17], we have introduced a novel inference-based scheme for QS which fully characterizes the response $\mathcal{R}(\theta)$ for a general class of unitary families by only measuring the system at $2n + 1$ known parameters. Overall, this framework leads to new insights and methodology for the characterization, implementation and benchmarking of sensing protocols. One of the main advantages of our protocol is that it can be readily combined with existing sensing schemes. For instance, further research could explore the use of the inferred response function in a variational setting, involving an optimization of

the experimental setup to maximize the sensitivity and parameter prediction accuracy. This promises a new approach in data-driven quantum machine learning for QS where the optimization procedure does not require knowledge of the classical or quantum Fisher information [6, 9, 11, 10, 12, 13, 25, 26, 27, 28, 29, 30].

References

[1] C. L. Degen and F. Reinhard, and P. Cappellaro Quantum sensing, *Rev. Mod. Phys.* 89, 035002, (2017)

[2] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Bell's theorem without inequalities, *American Journal of Physics* 58, 1131–1143 (1990).

[3] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini, W. M. Itano, J. D. Jost, C. Langer, and D. J. Wineland, Toward Heisenberg-limited spectroscopy with multiparticle entangled states, Science **304**, pages 1476–1478 (2004).

[4] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone, Quantum metrology, Physical Review Letters **96**, pages 010401 (2006).

[5] Susanna F Huelga, Chiara Macchiavello, Thomas Pellizzari, Artur K Ekert, Martin B Plenio, and J Ignacio Cirac, Improvement of frequency standards with quantum entanglement, Physical Review Letters **79**, pages 3865 (1997)

[6] Bálint Koczor, Suguru Endo, Tyson Jones, Yuichiro Matsuzaki, and Simon C Benjamin, Variational-state quantum metrology, New Journal of Physics (2020).

[7] Lukas J Fiderer, Julien ME Fraïsse, and Daniel Braun, Maximal quantum Fisher information for mixed states, Physical Review Letters **123**, pages 250502 (2019)

[8] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, and Patrick J. Coles, Variational quantum algorithms, Nature Reviews Physics **3**, pages 625–644 (2021a).

[9] Jacob L Beckey, M. Cerezo, Akira Sone, and Patrick J Coles, Variational quantum algorithm for estimating the quantum Fisher information, Physical Review Research **4**, pages 013083 (2022)

[10] Akira Sone, M. Cerezo, Jacob L Beckey, and Patrick J Coles, A generalized measure of quantum Fisher information, Physical Review A **104**, pages 062602 (2021).

[11] M. Cerezo, Akira Sone, Jacob L Beckey, and Patrick J Coles, Sub-quantum Fisher information, Quantum Science and Technology (2021b).

[12] Raphael Kaubruegger, Denis V Vasilyev, Marius Schulte, Klemens Hammerer, and Peter Zoller, Quantum variational optimization of Ramsey interferometry and atomic clocks, Physical Review X **11**, pages 041045 (2021)

[13] Johannes Jakob Meyer, Johannes Borregaard, and Jens Eisert, A variational toolbox for quantum multi-parameter estimation, NPJ Quantum Information **7**, pages 1–5 (2021).

[14] Huerta Alderete, C. and Gordon, Max Hunter and Sauvage, Frédéric and Sone, Akira and Sornborger, Andrew T. and Coles, Patrick J. and Cerezo, M. Inference-Based Quantum Sensing, *Phys. Rev. Lett.*, 129, 190501,(2022).

[15] Ken M Nakanishi, Keisuke Fujii, and Synge Todo, Sequential minimal optimization for quantum-classical hybrid algorithms, Physical Review Research **2**, pages 043158 (2020).

[16] Olivia Di Matteo, Josh Izaac, Tom Bromley, Anthony Hayes, Christina Lee, Maria Schuld, Antal Száva, Chase Roberts, and Nathan Killoran, Quantum computing with differentiable quantum transforms, arXiv preprint arXiv:2202.13414 (2022).

[17] David Wierichs, Josh Izaac, Cody Wang, and Cedric Yen-Yu Lin, General parameter-shift rules for quantum gradients, Quantum (2022)

[18] Luca Pezzè, Augusto Smerzi, Markus K. Oberthaler, Roman Schmied, and Philipp Treutlein, Quantum metrology with nonclassical states of atomic ensembles, Rev. Mod. Phys. **90**, pages 035005 (2018)

[19] Samson Wang, Enrico Fontana, M. Cerezo, Kunal Sharma, Akira Sone, Lukasz Cincio, and Patrick J Coles, Noise-induced barren plateaus in variational quantum algorithms, Nature Communications **12**, pages 1–11 (2021a).

[20] Daniel Stilck França and Raul Garcia-Patron, Limitations of optimization algorithms on noisy quantum devices, Nature Physics **17**, pages 1221–1227 (2021)

[21] Samson Wang, Piotr Czarnik, Andrew Arrasmith, M. Cerezo, Lukasz Cincio, and Patrick J Coles, Can error mitigation improve trainability of noisy variational quantum algorithms? arXiv preprint arXiv:2109.01051 (2021b).

[22] Masahiro Kitagawa and Masahito Ueda, Squeezed spin states, Physical Review A **47**, pages 5138–5143 (1993).

[23] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta, Hardware-efficient

variational quantum eigensolver for small molecules and quantum magnets, Nature **549**, pages 242–246 (2017).

[24] M. Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles, Cost function dependent barren plateaus in shallow parametrized quantum circuits, Nature Communications **12**, pages 1–12 (2021c).

[25] Johannes Jakob Meyer, Fisher Information in Noisy Intermediate-Scale Quantum Applications, Quantum **5**, pages 539 (2021).

[26] Iris Cong, Soonwon Choi, and Mikhail D Lukin, Quantum convolutional neural networks, Nature Physics **15**, pages 1273–1278 (2019).

[27] Arthur Pesah, M. Cerezo, Samson Wang, Tyler Volkoff, Andrew T Sornborger, and Patrick J Coles, Absence of barren plateaus in quantum convolutional neural networks, Physical Review X **11**, pages 041011 (2021).

[28] Kunal Sharma, M. Cerezo, Lukasz Cincio, and Patrick J Coles, Trainability of dissipative perceptron-based quantum neural networks, *Physical Review Letters* 128, 180505, (2022).

[29] Jeffrey Marshall, Filip Wudarski, Stuart Hadfield, and Tad Hogg, Characterizing local noise in QAOA circuits, IOP SciNotes **1**, pages 025208 (2020).

[30] Cheng Xue, Zhao-Yun Chen, Yu-Chun Wu, and Guo-Ping Guo, Effects of quantum noise on quantum approximate optimization algorithm, *Chinese Physics Letters* **38**, pages 030302 (2021).

On the performance for the block-encoding of the matrix functions evaluated by the numerical quadrature method

Souichi Takahira[1 4 *] Asuka Ohashi[2] Tomohiro Sogabe[3] Tsuyoshi Sasaki Usuda[4]

[1] *Faculty of Information Engineering, Meijo University, 1-501 Shiogamaguchi, Tempaku, Nagoya, 468-8502, Japan*
[2] *Department of General Education, National Institute of Technology, Kagawa College, 551 Kohda, Takuma-cho, Mitoyo, Kagawa, 769-1192, Japan*
[3] *Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa, Nagoya, 464-8603, Japan*
[4] *Graduate School of Information Science and Technology, Aichi Prefectural University, 1522-3 Ibaragabasama, Nagakute, Aichi, 480-1198, Japan*

Abstract. Matrix functions such as the matrix logarithm $\log(A)$ and the matrix fractional power A^r can be represented by contour integrals. A quadrature method for this representation is proposed in [N. Hale, N.J. Higham, and L.N. Trefethen, SIAM J. Numer. Anal., 46, 5, 2505-2523, (2008)]. In this study, we apply this method to the quantum algorithm, which is described as the block-encoding framework, and evaluate its performance numerically, where the performance means the number of uses of block-encoding of A and the sub-normalization factor. We consider the matrix logarithm and matrix fractional power as examples. As a result, the order of the performance is evaluated numerically by using the maximum and minimum eigenvalues of the matrix A.

Keywords: Quantum algorithm, matrix logarithm, matrix fractional power, contour integral

1 Introduction

For a complex function $f(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ and a square matrix A, the matrix $f(A) = a_0 I + a_1 A + a_2 A^2 + \cdots$ is called a matrix function. Matrix functions appear in many fields of science and technical computations. It is known that the matrix function can be represented as the contour integral

$$f(A) = \int_\Gamma f(z)(zI - A)^{-1} \mathrm{d}z, \tag{1}$$

where Γ is a closed contour that encloses all eigenvalues of the matrix A and f is a complex function that is analytic on the inside region of Γ. Several quadratures to compute the matrix function have been proposed. For example, if $f(z)$ is an exponential function, then we can take a circle as the contour Γ and can use the trapezoidal rule. In this case, the quadrature formula $f_N(A)$ approximates $f(A)$ with exponential accuracy with respect to the number of integration points.

On the other hand, if the complex function f has a singularity, it is difficult to treat the contour integral in a naïve method. For example, $f(z) = \log(z)$ has a singularity at $z = 0$ and so we can not apply the trapezoidal rule to $\log(A)$. A method to treat such a function has been proposed in [1].

There are several quantum algorithms for matrix functions $f(A)$ based on integral representations. For example, [2, 3, 4, 5]. In [2], we have considered the case where the closed contour is a circle centered at the origin. Furthermore, in [5], we have considered the case where the closed curve can be represented as a circle or even more in general quadrature . In addition, we described the structure of the algorithm in the framework of block-encoding [6]. However, our previous study [2, 5] only treated general cases, and so does not deal with concrete matrix functions. Therefore, it is necessary to consider concrete functions.

In this study, we consider matrix logarithmic function and matrix fractional powers. Specifically, we first calculate quadrature formulas for these matrix functions using the numerical method proposed in [1]. Next, we apply the obtained quadrature formulas to the quantum algorithm described as block-encoding. Then, we evaluate the order of the subnormalization factor and the number of uses for the block-encoding of the input matrix A that is given as an oracle.

2 Preliminaries

2.1 The quadrature formula

In this study, we use method 2 in [1] to obtain a quadrature formula. Method 2 can be applied when the complex function f has no singularity on $\mathbb{C}\backslash(\infty, 0]$. Therefore, method 2 can treat logarithmic function $\log(z)$ and fractional power z^r. The quadrature formula by the method 2 is as follows: Suppose that the eigenvalue of the matrix A is positive. Let m and M be the lower and upper bounds on the eigenvalues, where $0 < m < M$. The quadrature formula $f_N(A)$ is defined as $f_N(A) = A\mathrm{Im}(T)$, where

$$T = \sum_{j=0}^{N-1} f_j (w_j I - A)^{-1}, \tag{2}$$

and

$$f_j = \frac{1}{N} \cdot \frac{-4K(mM)^{1/4}}{\pi k} \cdot \frac{f(w(t_j)^2)\mathrm{cn}(t_j)\mathrm{dn}(t_j)}{w(t_j)^2(k^{-1} - u(t_j))}, \tag{3}$$

$$w_j = w(t_j)^2. \tag{4}$$

The meaning of the symbols is the same as in [1]. For more details, see [1]. The quadrature formula $f_N(A)$ is

*takahira@meijo-u.ac.jp

exponentially accurate to the original matrix function $f(A)$. That is, $\|f_N(A) - f(A)\| = O(e^{-2\pi^2 N/(\log\kappa+6)})$ holds [1, Theorem 3.1]. where $\kappa = M/m$. In the below, we assume that $N = 2^n$ $(n \in \mathbb{N})$.

2.2 Block-encoding

We explain the block-encoding that is useful to deal with the linear algebra operations on the quantum computer.

Definition 1 ([6]) *Suppose that A is a matrix $A \in \mathbb{C}^{2^s\times 2^s}$ and $\alpha, \varepsilon_A \in \mathbb{R}_+$. Then, a unitary operation U_A is $(\alpha, a, \varepsilon_A)$-block-encoding of A if*

$$\|A - \alpha(\langle 0|^{\otimes a} \otimes I_s)U_A(|0\rangle^{\otimes a} \otimes I_s)\|_2 \le \varepsilon_A. \quad (5)$$

Here, I_s is a 2^s-by-2^s identity matrix. The number α is a *subnormalization* factor. In this study, we assume that the block-encoding U_A of the input matrix is given.

By combining some block-encodings, we obtain a block-encoding for the matrix operations of the encoded matrices on them. Let U_A be an $(\alpha, a, \varepsilon_A)$-block-encoding and U_B be a $(\beta, b, \varepsilon_B)$-block-encoding of B. Then, $(I_b \otimes U_A)(I_a \otimes U_B)$ is an $(\alpha\beta, a + b, \alpha\varepsilon_B + \beta\varepsilon_A)$-block-encoding of AB. Furthermore, the linear combination is obtained as follows: Let U_j be an $(\alpha, a, \varepsilon_A)$-block-encoding of A_j and let $\boldsymbol{c} = (c_0, c_1, \ldots, c_{2^n-1})$. We define unitary operations $P_L|0\rangle^{\otimes n} = \sum_{j=0}^{N-1}\sqrt{c'_j}|j\rangle$ and $P_R|0\rangle^{\otimes n} = \sum_{j=0}^{N-1}\sqrt{c'^*_j}|j\rangle$ where $c'_j = c_j/\|\boldsymbol{c}\|_1$. The pair of unitaries (P_L, P_R) is called state-preparation pair for $\boldsymbol{c}$. Then, $(P_L^\dagger \otimes I_{a+s})\left(\sum_{j=0}^{2^n-1}|j\rangle\langle j| \otimes U_j\right)(P_R \otimes I_{a+s})$ is an $(\alpha\|\boldsymbol{c}\|_1, a+n, \|\boldsymbol{c}\|_1\varepsilon)$-block-encoding of the linear combination $\sum_{j=0}^{2^n-1} c_j A_j$.

According to the quantum singular value transformation (QSVT), if there is an angle sequence $\Phi = (\phi_0, \ldots, \phi_d)$ such that $\langle 0|e^{\mathbf{i}\phi_j Z}\prod_j^d e^{\mathbf{i}\theta X}e^{\mathbf{i}\phi_j Z}|0\rangle = P(x)$ $(\theta = \arccos(x), x \in [-1, 1])$, then the block-encoding of the d-degree matrix polynomial $P(A)$ of U_A^Φ can be constructed by using $O(d)$ uses of U_A, and $O(d)$ primitive gates [6]. Using the polynomial that approximates the reciprocal function in a range $[1/\kappa, 1]$ $(\kappa > 1)$, we can compute the angle sequence $\Phi^{(inv)}$ that corresponds to the inverse matrix [6, 7, 8]. In other words, if we have U_A and $\Phi^{(inv)}$, we can obtain the block-encoding $U_A^{\Phi^{(inv)}}$ of A^{-1}. For more specific, suppose that the singular values of A is within $[1/\beta, \alpha]$ and satisfy $\kappa > \alpha\beta$. Then we can construct $(4\beta, a+1, 4d\sqrt{\varepsilon_A/\alpha} + \varepsilon)$-block-encoding of A^{-1}, where $d = O\left(\alpha\beta\log\left(\frac{\beta}{\varepsilon}\right)\right)$.

2.3 The constuction of the block-encoding

The purpose of this study is to implement the matrix function, Therefore, we consider to construct a block-encoding of the quadrature formula $f_N(A) = A\mathrm{Im}(T)$ defined in Equation (2). The block-encoding of this matrix is constructed as follows:

$$U_{f_N(A)} = (U_A \otimes I_{n+a+4})(U_{\mathrm{Im}T} \otimes I_a), \quad (6)$$
$$U_{\mathrm{Im}T} = |+\rangle\langle +| \otimes U_T - |-\rangle\langle -| \otimes U_T^*,$$

where

$$U_T = (I_{a+3} \otimes F_L^\dagger \otimes I_s)U_B^{\Phi^{(inv)}}(I_{a+3} \otimes F_R \otimes I_s), \quad (7)$$

and (F_L, F_R) is the state-preparation-pair for $\boldsymbol{f} = (f_0, f_1, \ldots, f_{N-1})$ and $U_B^{(inv)}$ is block-encoding of the inverse of a block-diagonal matrix $B = \sum_j |j\rangle\langle j| \otimes (w_j I - A) = \mathrm{diag}(w_0, w_1, \ldots, w_{N-1}) \otimes I - I \otimes A$. In addition, the $(w_{\max}+\alpha, a+2, 0)$-block-encoding of B is constructed as follows:

$$U_B = (\mathrm{e}^{\mathbf{i}Y\theta} \otimes I_{1+a+n+s})V_{w,A}(\mathrm{e}^{\mathbf{i}Y\theta} \otimes I_{1+a+n+s}), \quad (8)$$

where $\theta = \arccos(w_{\max}/(w_{\max} + \alpha))$ and

$$V_{w,A} = |0\rangle\langle 0| \otimes U_w \otimes I_{a+s} - |1\rangle\langle 1| \otimes I_{1+s} \otimes U_A. \quad (9)$$

Here, U_w is a $(w_{\max}, 1, 0)$-block-encoding of the diagonal matrix $\mathrm{diag}(w_0, w_1, \ldots, w_{N-1})$ constructed as follows:

$$U_w = \sum_{j=0}^{2^n-1}|j\rangle\langle j| \otimes \mathrm{e}^{\mathbf{i}\arccos(w_j/w_{\max})Y}, \quad (10)$$

where $w_{\max} = \max_j |w_j|$.

From the above, we can summarize the block-encoding of $f_N(A)$ as like the following.

Lemma 2 (Block-encoding of $f_N(A)$) *Suppose that U_A is an $(\alpha, a, 0)$-block-encoding of A. Let $1/\beta$ be the lower bound of the singular values of $w_j I - A$. Then, we can construct a $(\nu, 2a+n+4, \varepsilon)$-block-encoding $U_{f_N(A)}$ of the quadrature formula $f_N(A)$. Further $U_{f_N(A)}$ consists of $O(d)$ uses U_A and $O((1+N)d + N)$ primitive gates, where*

$$\nu = 4\alpha\beta\|\boldsymbol{f}\|_1, \quad (11)$$
$$d = O\left((\max_j |w_j| + \alpha)\beta\log\left(\frac{\alpha\beta\|\boldsymbol{f}\|_1}{\varepsilon}\right)\right). \quad (12)$$

3 Numerical result

We have described the quadrature formula and the construction of the quantum algorithm. In this section, we consider applying the quadrature formula for matrix logarithmic functions and matrix fractional powers to the quantum algorithm. Specifically, we first calculate each parameter of the quadrature formula using a numerical method. From the obtained parameters, we numerically evaluate the subnormalization factor ν given by Equation (11), which relates the performance of the block-encoding of $f_N(A)$, and the number of uses d to U_A given by Equation (12).

The settings for matrix A are as follows: First, we calculate the unitary matrix Q by using the QR decomposition for the random matrix that the each elements is generated by the uniform distribution on $[0, 1]$. Then, we set matrix A as

$$A = QDQ^\dagger \in \mathbb{R}^{20\times 20}, \quad (13)$$

where $D = \mathrm{diag}(d_0, d_1, \ldots, d_{19})$ and $d_i = m + (M - m)i/20$. In order for the quadrature formula to be sufficiently accurate, all computation were performed with $N = 64$. The MATLAB code written in [1] was used as the numerical method. The version of MATLAB is R2023a.

Table 1: $w_{\max}$

	m			
M	1	0.1	0.01	0.001
1	-	3.3	3.8	3.9
10	33.4	37.9	39.3	39.7
100	379.4	393.1	397.2	398.2
1000	3930.6	3971.5	3982.4	3983.4

Table 2: β

	m			
M	1	0.1	0.01	0.001
1	-	14.3	135.7	1339.4
10	1.4	13.6	133.9	1334.0
100	1.4	13.4	133.4	1331.9
1000	1.3	13.3	133.2	1330.0

3.1 Computation result for $w_{\max}$ and d

First, we calculate $w_{\max}$ and $1/\beta$ related with d. The computation results are shown in Tables 1 and 2. From Tables 1 and 2, we see that $w_{\max} = O(M)$ and $\beta = O(m)$, respectively. Therefore, the subnormalization factors ν and d are $\nu = O(\kappa\|\boldsymbol{f}\|_1)$ and $d = O(\kappa \log(\kappa\|\boldsymbol{f}\|_1/\varepsilon))$, respectively.

3.2 Computation result for $\|\boldsymbol{f}\|_1$

Both the subnormalization factor and the number of uses to U_A are related to f . Therefore, in the following, we compute $\|\boldsymbol{f}\|_1$ for each case of matrix logarithm function and matrix fractional power.

3.2.1 The matrix logarithm

We consider the matrix logarithm function $\log(A)$. In this case, $f(z) = \log(z)$ and f has a singularity $z = 0$. Then, method 2 of [1] can be applied. The computation results of $\|\boldsymbol{f}\|_1$ are shown in Table 3. From this table, we can see that $\|\boldsymbol{f}\|_1 = O(\log(\kappa))$. Thus, we have

$$\nu = O(\kappa \log(\kappa)), \quad d = O(\kappa \log(\kappa \log(\kappa)/\varepsilon)). \tag{14}$$

Table 3: Numerical computation result of $\sum_j |f_j|$ where f_j is defined by eq. (3), $f(z) = \log(z)$, and $N = 64$.

	m			
M	1	0.1	0.01	0.001
1	-	4.9855	10.9238	18.0700
10	4.9855	9.5357	15.2164	22.2226
100	10.9238	15.2164	20.7723	27.7191
1000	18.0700	22.2226	27.7191	34.6404

3.2.2 The matrix fractional power

We consider the matrix fractional power A^r with $r = 1/4, 2/4$, and $3/4$. In this case, $f(z) = z^r$ and f has a singularity $z = 0$. Then, method 2 of [1] can be applied. The computation results of $\|\boldsymbol{f}\|_1$ are shown in Tables 4 to 6. From this table, we can see that $\|\boldsymbol{f}\|_1 = O(M^r)$. Thus, we have

$$\nu = O(M^r \kappa), \quad d = O(\kappa \log(\kappa M^r/\varepsilon)). \tag{15}$$

Table 4: Numerical computation result of $\sum_j |f_j|$ where f_j is defined by eq. (3), $f(z) = z^{\frac{1}{4}}$, and $N = 64$.

	m			
M	1	0.1	0.01	0.001
1	-	1.7575	2.1008	2.2369
10	3.1253	3.7359	3.9778	4.0864
100	6.6434	7.0736	7.2667	7.3604
1000	12.5788	12.9222	13.0888	13.1742

Table 5: Numerical computation result of $\sum_j |f_j|$ where f_j is defined by eq. (3), $f(z) = z^{\frac{2}{4}}$, and $N = 64$.

	m			
M	1	0.1	0.01	0.001
1	-	1.6538	1.8974	1.9681
10	5.2298	6.0000	6.2238	6.2929
100	18.9737	19.6812	19.8997	19.9684
1000	62.2375	62.9285	63.1455	63.2139

Table 6: Numerical computation result of $\sum_j |f_j|$ where f_j is defined by eq. (3), $f(z) = z^{\frac{3}{4}}$, and $N = 64$.

	m			
M	1	0.1	0.01	0.001
1	-	1.7337	2.0259	2.1156
10	9.7495	11.3923	11.8968	12.0571
100	64.0636	66.9009	67.8022	68.0895
1000	376.2112	381.2800	382.8953	383.4091

4 Conclusion

In this study, we evaluate numerically the performance of quantum algorithms for matrix functions represented by contour integrals when using the quadrature method of [1]. Specifically, block-encoding of the matrix logarithm function $\log(A)$ and the matrix fractional power A^r are considered, and their subnormalization factors and the number of uses to U_A are evaluated numerically. Usually, the quantum singular value transformation (QSVT) is considered as the quantum algorithm for matrix functions. The QSVT computes a generalized matrix function, which coincides with a standard matrix function when the input matrix is a normal matrix. On the other hand, our method can handle standard matrix functions even if the input matrix is not a normal matrix. However, in this method, the eigenvalues have to be positive real numbers.

In addition, for the subnormalization factor of block-encoding obtained by our method, a factor of κ is always required caused by employing the block-encoding of the inverse. Therefore, the order of complexity will be increased compared to QSVT. On the other hand, it is not necessary to compute the angle sequence numerically for each function. That is, it is sufficient to compute the one related to the reciprocal function. Further theoretical study are left for future works.

References

[1] N. Hale, N. J. Higham, and L. N. Trefethen, "Computing $A^{\alpha}, \log(A)$, and related matrix functions by contour integrals," *SIAM Journal on Numerical Analysis*, vol. 46, no. 5, pp. 2505–2523, 2008.

[2] S. Takahira, A. Ohashi, T. Sogabe, and T. S. Usuda, "Quantum algorithm for matrix functions by Cauchy's integral formula," *Quantum Information and Computation*, vol. 20, no. 1&2, pp. 0014–0036, 2020.

[3] Y. Tong, D. An, N. Wiebe, and L. Lin, "Fast inversion, preconditioned quantum linear system solvers, and fast evaluation of matrix functions," Phys. Rev. A, 104, 032422, (2021).

[4] S. Zhang and H. Xiang, "Quantum algorithm for matrix logarithm by integral formula," Quantum Information Processing, vol.22, no.76, (2023).

[5] S. Takahira, A. Ohashi, T. Sogabe, and T. S. Usuda *Quantum Information and Computation*, vol.22, no. 11&12, pp. 0965–0979. (2022).

[6] A. Gilyén, Y. Su, G. H. Low, and N. Wiebe, "Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics," In Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC 2019), pp. 193-204, (2019).

[7] J. Haah, "Product decomposition of periodic functions in quantum signal processing," *Quantum*, vol. 3, p. 190, Oct 2019.

[8] Y. Dong, X. Meng, K.B. Whaley, and L. Lin, "Efficient phase-factor evaluation in quantum signal processing," Phys. Rev. A 103, 042419, (2021).

A Riemannian Genuine Measure of Entanglement for Pure States

R. Dharmaraj* and Radhika Vathsan†
Physics Department, BITS Pilani K K Birla Goa Campus
(Dated: May 10, 2023)

While several measures exist for entanglement of multipartite pure states, a true entanglement measure for mixed states still eludes us. A deeper study of the geometry of quantum states may be the way to address this issue, on which context we come up with a measure for pure states based on a geodesic distance on the space of quantum states. Our measure satisfies all the desirable properties of a "Genuine Measure of Entanglement" (GME), and in comparison with some of the other existing measures, shows better smoothness and discriminance.

I. INTRODUCTION

Entanglement, the property of non-separability of multipartite quantum states, has come to the forefront in recent times as an important resource in quantum computing and communication protocols. The necessity to quantify multi-party entanglement has become imperative, from the view point of both as a measure of quantum correlations and as a unit of resource consumption. This has led to the formulation of various entanglement measures from different standpoints. Some of the earliest measures such as negativity [1] and reshuffling negativity [2] are straightforward and have been proven to be computationally inexpensive. Operational measures of entanglement including entanglement cost [3] and distillable entanglement are based on the quantity of resources required to generate entanglement. Among different approaches there have been measures of entanglement based on operator algebra [4] and more recently, coherence of an entangled state in the Schmidt basis [5].

Entanglement is such a fundamental property of quantum systems that it needs to be understood from the basic notions of quantum state-space. While the geometry of quantum states has been a topic of research for a few decades [6] it was the work pioneered by Amari and Nagaoka [7] that showed that the application of differential geometry to information theory would give meaningful insights. The Riemannian structure of quantum state space was studied in great detail by Morozova and Çentsov [8]. Later along with the contribution of Petz [9], a concrete definition of a metric on quantum state space was formulated.

Using this geometric viewpoint, several authors [10–12] have attempted to provide entanglement measures using different metrics on quantum state space to find the closest separable state. However, such measures share a limitation of being computationally hard for higher dimensional state spaces, since they are based on minimisation over the infinite set of states.

In an attempt to overcome these limitations, we work with the geometry of the space of reduced density operators, which is a subspace of a hypersphere. Similar considerations for entanglement entropy have been have been studied before [13]. The spherical geometry of the Bloch sphere helps us define a Riemannian measure of two-qubit entanglement in a relatively straightforward way, avoiding minimization procedures. This can be generalised to bipartite systems of higher dimensions. We then extend this method to define a genuine measure of entanglement for multipartite systems of higher dimensions.

II. GEOMETRY OF QUANTUM STATE SPACE

A quantum state of an n-qubit system is a non-negative Hermitian operator $\hat{\rho}$ with unit trace on the d-dimensional Hilbert space $\mathcal{H}^d$, $d = 2^n$. This operator, the statistical operator or density matrix, is instrumental in extracting probabilities of all possible measurement outcomes on the system.

A unit-trace $d \times d$ Hermitian matrix can be written in terms of the identity and a set of $d^2 - 1$ traceless, Hermitian, mutually orthogonal operators σ_i [14]:

$$\hat{\rho}(x) = \frac{1}{d}\mathbb{1} + \frac{\sqrt{d-1}}{d}\sum_{i=1}^{d^2-1} x_i\sigma_i,$$

Tr $(\sigma_i\sigma_j) = \delta_{ij}d$ with real coefficients x_i. Further, a necessary (but not sufficient) condition for non-negativity of $\hat{\rho}$ is that coefficients x_i satisfy [15–17]

$$\left(\sum_{i=1}^{d^2-1} x_i^2\right)^{\frac{1}{2}} \leq 1. \tag{1}$$

The x_i thus form a $d^2 - 1$-dimensional real vector $\vec{x}$, spanning a space in the form of a hyperspherical ball of radius $r = |\vec{x}|$. We will call this the parameter space $\mathcal{P}(\mathcal{H}^d) \equiv \mathcal{P}_n$ of unit trace, Hermitian operators on $\mathcal{H}^d$. For pure states, we have the condition

$$\text{Tr } \rho^2 = 1 \implies r = 1. \tag{2}$$

For the single qubit space, Eq. (1) is also a sufficient condition for non-negativity. The parameter space is a unit ball: the Bloch ball, with the surface representing

* p20200040@goa.bits-pilani.ac.in
† radhika@goa.bits-pilani.ac.in

pure states and mixed states represented by points in the interior representing, with the center $r = 0$ being the maximally mixed state.

However, for $n > 1$, the condition Eq. (1) is not sufficient to ensure non-negativity. Further restrictions ([17]) select a subset $\mathcal{Q}_n \subset \mathcal{P}_n$ as the space of n-qubit density operators. While all n-qubit pure states lie on the boundary $\partial\mathcal{P}_n$, not all points on the surface $r = 1$ correspond to states in Q_n. Pure states are extreme points of $\mathcal{P}_n$, not just boundary points. However, for all n, the center of the hypersphere, $r = 0$ represents the maximally mixed state, from which all pure states are equidistant.

The x_i can be regarded as Cartesian coordinates of a point in $\mathcal{P}_n$, that can be transformed to hyper-spherical coordinates[18, 19]:

$$\begin{aligned} r &\in (0,1), \\ \theta_i &\in (0,\pi), \\ \phi &\in (0,2\pi), \end{aligned} \tag{3}$$

$(i = 1, ..., d^2 - 3)$, which are natural for parametrizing points in a sphere. We can then define a metric on Q_n using these coordinates.

Example: metric on single-qubit state space

The parameter space of the single qubit states is the same as the state space: the Bloch ball ($\mathcal{P}_1 \sim \mathcal{Q}_1$), and the coordinates of a state $\rho(x)$ can be determined by

$$x_i = \mathrm{Tr}(\rho\sigma_i), \tag{4}$$

where σ_i are the Pauli spin operators. The spherical polar coordinates for this point are

$$\begin{aligned} r &= ({x_1}^2 + {x_2}^2 + {x_3}^2)^{1/2}, \\ \theta &= \cos^{-1}\frac{x_3}{r}, \\ \phi &= \tan^{-1}\left(\frac{x_2}{x_1}\right). \end{aligned} \tag{5}$$

Riemannian metrics on this space are given by [20]

$$ds^2 = \frac{1}{4}\left[\frac{dr^2}{1-r^2} + \frac{1}{f\left(\frac{1-r}{1+r}\right)}\frac{r^2}{1+r}d\Omega^2\right], \tag{6}$$

where $d\Omega^2$ is the metric on the unit sphere. The function $f(.)$ is the Morozova-Čencov (MC) function [21], which is defined as any $f(t) : R_+ \to R_+$ satisfying three properties:

1. f is an operator monotone,
2. f is self-inverse,
3. $f(1) = 1$.

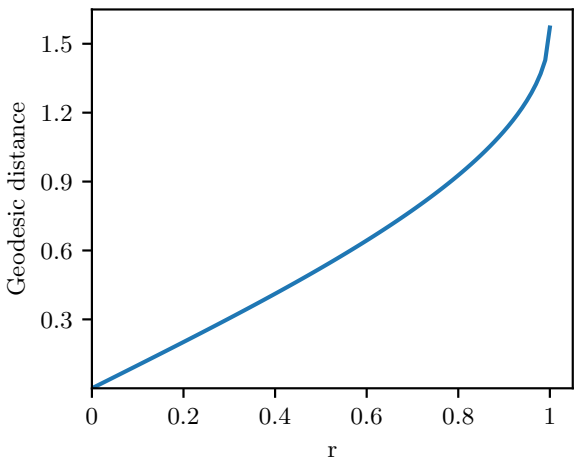


FIG. 1: Geodesic distance from the origin to a point at radial coordinate r in the Bloch Ball.

For simplicity of calculation and due to its popularity in literature, we use the MC function [22] corresponding to the Bures metric:

$$f(t) = \frac{1+t}{2}, \text{ where } t = \frac{1-r}{1+r}. \tag{7}$$

Using this in Eq. (6), a metric in the Bloch ball is defined by

$$G_{ij} = \begin{bmatrix} \frac{1}{1-r^2} & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2\sin^2\theta \end{bmatrix}. \tag{8}$$

This reduces to the Fubini-Study metric on the surface, corresponding to the space of pure single-qubit states. The geodesic distance along a curve parametrized by λ connecting two states u and v, is defined by

$$l_{uv} = \int_{\lambda_u}^{\lambda_v} \sqrt{G_{ij}\dot{u}^i\dot{v}^j}\,d\lambda. \tag{9}$$

The behaviour of l_{0r}, distance between the origin and a point at r as plotted in Fig. (1) clearly indicates the non-Euclidean nature of this metric, and brings out the curvature of the space of quantum states.

III. ENTANGLEMENT MEASURE

We wish to construct a measure of entanglement $E(\rho)$, a positive real function of the quantum state ρ, that is required to satisfy [21]: **(E1)** Monotonocity under SLOCC, **(E2)** Discriminance: $E(\rho) = 0$ iff ρ is separable and **(E3)** Convexity as primary among other desirable properties such as asymptotic continuity, normalizability and computability. Given a distance measure $d(.,.)$ on the state space, a geometric entanglement measure for a pure

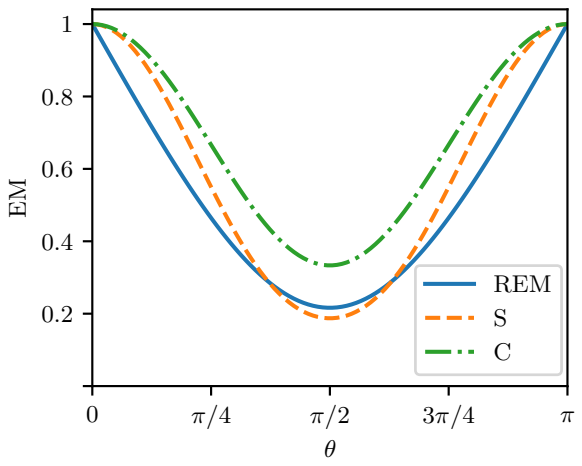


FIG. 2: Comparison of REM, von Neumann entropy (S) and concurrence (C) for $|\psi_1(\theta)\rangle$, superpositions of the 2-qubit symmetric and antisymmetric states.

state $|\psi\rangle$ is defined as the distance to the closest separable state $|\phi\rangle \in \mathcal{S}$, the set of separable states [23]:

$$E(\psi_{AB}) = \min_{\phi\in\mathcal{S}} d(\psi,\phi). \tag{10}$$

We first consider two-qubit pure states, for which a measure of entanglement is the purity of the reduced density operators. The purity of the reduced density operator ρ is related to its radial coordinate in $\mathcal{Q}_1$. Levay [24] has discussed in detail how the state space $\mathcal{Q}_1$ of single qubit density operators can be mapped to the space of two-qubit purifications (boundary of $\mathcal{Q}_2$). Therefore, distances between states on the surface of $\mathcal{Q}_2$ are related to distances in $\mathcal{Q}_1$. We can use this correspondence to construct a geometric measure of two-qubit entanglement in terms of distances in $\mathcal{Q}_1$.

The reduced density operators ρ_A and ρ_B represent pure states if and only if ψ_{AB} is separable. The partial trace operation maps the pure state ψ_{AB} to a point a in the single qubit Bloch ball corresponding to ρ_A. Since all separable states are mapped to the surface of $\mathcal{Q}_1$, the closest separable state to ρ_A will be at the radially outward point p on its surface. If l_{ap} is the geodesic distance to this point, then we can define a normalised Riemannian entanglement measure as

$$\mathrm{REM}(\psi_{AB}) = \frac{l_{ap}}{N}. \tag{11}$$

Since the metric Eq. (6) blows up at the surface $r = 1$ (Fig. 1), we need to adopt a limiting procedure to compute the integral Eq. (9) for l_{ap}. We take p very close to the surface and take the limit $r \to 1$. The normalization constant N is the distance between the center $u(0,\theta,\phi)$ and a point on the surface $v(r\to 1,\theta,\phi)$.

The REM defined in (11) overcomes the computational difficulty involved in calculating the the distance to the closest separable state by a minimization procedure. Another common measure for two qubit pure state entanglement is the largest Schmidt coefficient. Finding it involves solving the characteristic equation, which again is computationally less efficient than finding the REM.

Two qubit examples.

We test our measure by comparing with standard measures such as the von Neumann entropy $S(\psi_{AB}) = -\mathrm{Tr}(\rho_A \log \rho_A)$ and concurrence $C(\psi) = \sqrt{(1 - \mathrm{Tr}\rho_A^2}$.

Example 1: Superposition of symmetric and asymmetric states

$$\left.\begin{aligned} |\phi_0\rangle &= |00\rangle\,, \\ |\phi_1\rangle &= \frac{1}{\sqrt{2}}\big(\,|01\rangle + |10\rangle\,\big) \\ |\phi_2\rangle &= |11\rangle\,, \end{aligned}\right\} \text{ Symmetric,}$$

$$|\phi_3\rangle = \frac{1}{\sqrt{2}}\big(\,|01\rangle - |10\rangle\,\big) \text{ antisymmetric;}$$

$$|\psi_1(\theta)\rangle = \frac{\cos\theta}{\sqrt{3}}\big(\,|\phi_0\rangle + |\phi_1\rangle + |\phi_2\rangle\,\big) + \sin\theta\,|\phi_3\rangle\,.$$

Figure (2) shows a comparison of the three measures of entanglement as functions of θ.

Example 2: Superposition of a Bell state and a product of non-orthogonal states $\Big(|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}\Big)$, which is useful in information theoretic tasks[25]:

$$\begin{aligned} |e\rangle &= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\ |p\rangle &= |{+}{+}\rangle\,, \\ |\psi_2(\theta)\rangle &= \cos\theta\,|e\rangle + \sin\theta\,|p\rangle\,. \end{aligned}$$

Figure (3) compares the three entanglement measures for such states as a function of θ.

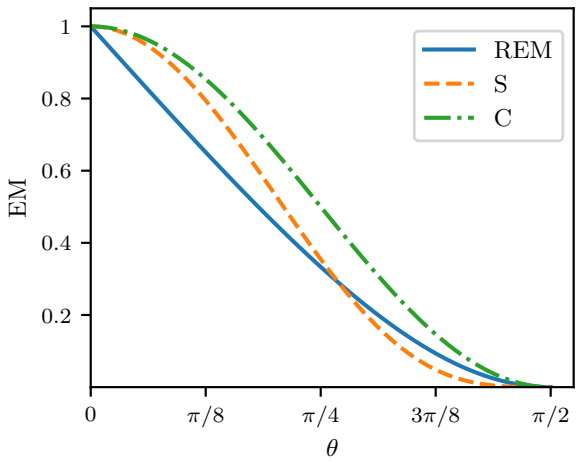


FIG. 3: Comparison of REM, von Neumann entropy (S) and concurrence (C) for $|\psi_2(\theta)\rangle$, superpositions of 2-qubit entangled states and product states in a non-orthogonal basis.

For both examples, the behaviour of REM is consistent with the behaviour of the other measures. This encourages us to move on to higher number of qubits.

IV. MULTI-QUBIT ENTANGLEMENT

For more than two qubits, the treatment of entanglement becomes more involved since there are several types of entanglement that are inequivalent[26, 27]. For instance, 3-qubit states fall under two SLOCC inequivalent classes: the GHZ-type (states equivalent under LOCC to $|000\rangle + |111\rangle$) and the W-type (superpositions of states in which only one qubit in the state $|1\rangle$). We aim to define a measure that picks genuine entanglement, i.e. not even biseparable in any way.

An n-qubit entanglement monotone is called a Genuine Measure of Entanglement (GME) if it satisfies the following properties ([28, 29]):

(P1) GME= 0 for all product states and for all biseparable states.

(P2) GME$\neq$ 0 for all non-biseparable states.

(P3) GME ranks GHZ state as more entangled than W state.

These three conditions present an interesting challenge to the existing entanglement measures. The multi-qubit entanglement measure given by Meyer *et al* [30] and interpreted by Brennen [31] satisfies **P2** and **P3**, but is non-zero for biseparable states. Later Love *et al*[32] gave a correction to this measure which qualified it as a GME. In works by Li *et al* [33], Xi *et al* [34], Sabín *et al* [35] and Markiewicz *et al* [36], bipartite measures have been utilised to construct a GME. The measures given by Pan *et al* [37] and by Carvalho *et al* [38] use addition of bipartite measures, which in general are non-zero for biseparable states and thus do not satisfy **P1**. Cai *et al* [39] have defined a GME based on von Neumann entropy. The well-known measure 3-tangle, which is an algebraic extension of concurrence given by Coffman *et al* [40] and later extended by Miyake [41], does not satisfy **P2** and is always zero for W-class entangled three qubit states. Bounds on geometric measures of entanglement using z-spectral radius for some classes of pure states were given by Xiong *et al* [42]. Haddadi *et al* [43] give an overview of different geometric measures of multi-qubit entanglement.

Now it was pointed out in by Love *et al*[32] and Yu *et al*[44] that any bipartite measure can be used to define a GME. We use this idea to define a GME, using a geometric measure for bipartite entanglement.

A. Bipartite measure of n-qubit entanglement

Along the same lines as our two qubit entanglement measure, we construct an REM for non-biseparability of tripartite states. Let us consider the bipartition $(AB)C$ of a three qubit pure state ψ_{ABC}. The bipartite measure of entanglement between the subsystems AB and C is the distance between the state ψ_{ABC} and the closest biseparable state $\phi \in \mathcal{S}_C$, the set of all biseparable states in the $(AB)C$ bipartition.

$$E\Big(\psi_{(AB)C}\Big) = \min_{\phi\in\mathcal{S}_C} d(\psi,\phi). \tag{12}$$

Calculating this by minimization is computationally hard. Using a generalization of the REM we can reduce this complexity.

The reduced density operator ρ_{AB} lives in the parameter space $\mathcal{P}_2$. The closest biseparable state ρ_p is radially outward on the surface of this hypersphere. We measure this distance using the Bures measure $D_B(\ ,\ )$ [45, 46], to define a bipartite Riemannian entanglement measure for $(AB)C$ as

$$\text{bREM}\big(\psi_{(AB)C}\big) = \frac{D_B\big(\rho_{AB},\rho_p\big)}{N_2}, \tag{13}$$

N_2 being the normalization. It is sometimes computationally easier to evaluate this distance as the complement of the distance from the state at $r = 0$ to ρ_{AB}. We use the same definition for the other two bipartitions.

Normalizing bREM

The 3-qubit GHZ state (14) possesses the maximum entanglement with respect to which we normalize our measure. While we might expect the maximum value to be 1, we find that the distance measure is typically less than 1. This is mainly because the rank of the 3-qubit state is less than the dimensionality of the AB Hilbert space. This can be seen using a generalised Schmidt decomposition to explicitly cast the many qubit state into bipartite form[47]. The AB subspace of the 3-qubit entangled ABC system is four dimensional, but the generalised Schmidt form of the GHZ state has only two vectors:

$$|GHZ\rangle_{ABC} = \frac{1}{\sqrt{2}}\left(|000\rangle + |111\rangle\right) \tag{14}$$

$$= \frac{1}{\sqrt{2}}\left(|0'0\rangle + |1'1\rangle\right) \tag{15}$$

with redefined basis vectors for AB subsystem:

$$\{|00\rangle\,,\ |11\rangle\,,\ |01\rangle\,,\ |10\rangle\}$$
$$\equiv \{|0'\rangle\,,\ |1'\rangle\,,\ |2'\rangle\,,\ |3'\rangle\}.$$

The entanglement of $(AB)C$ bipartition is then the "two-qubit" entanglement of the state expressed as Eq. (15). Since it resembles the two qubit Bell state in the generalised Schmidt basis, we expect the GHZ state to possess maximal entanglement. But there is a major difference between an actual Bell state and the state in Eq. (15). The measurement outcomes in the AB subsystem are maximally random, equivalent to those of a Bell state,

only for the projections $\{|0'\rangle\langle 0'|$ and $|1'\rangle\langle 1'|\}$. However, projections along the other two basis vectors $\{|2'\rangle\ ,\ |3'\rangle\}$ are always definite (i.e they are always zero!). Thus although the $(AB)C$ entanglement of the $|GHZ\rangle$ state is maximal, some determinacy still exists in the system. This is borne out by the value of the un-normalized bipartite measure $D_B(\rho_{AB},\rho_p) = \left(1-\sqrt{2-\sqrt{2}}\right) < 1$. We use this value of N_2 to normalize the measure bREM.

B. GME based on bREM

We define a GME by taking the geometric mean of bREMs of all possible bipartitions. We call this the GBR (Geometric mean of Bipartite Riemannian measures of entanglement):

$$\mathrm{GBR}(\psi_{ABC}) = (b_1 b_2 b_3)^{\frac{1}{3}} \tag{16}$$

where b_i is the bREM for the i^{th} bipartition. This GME is readily generalised to the n-qubit case. The number m of unique possible bipartitions is given by[33],

$$m = \begin{cases} \sum_{i=1}^{(n-1)/2} \binom{n}{i} & \text{for odd } n, \\ \sum_{i=1}^{(n-2)/2} \binom{n}{i} + \frac{1}{2}\binom{n/2}{n} & \text{for even } n. \end{cases} \tag{17}$$

For each bipartition, we calculate b_i using the Bures distance from the reduced density operator of the larger partition, to its closest pure state. We then take the geometric mean:

$$\mathrm{GBR}(\psi_n) = \prod_{i=1}^{m} (b_i)^{\frac{1}{m}}. \tag{18}$$

It is easily verified that this measure satisfies all the properties of a GME. For instance, **(P1)** and **(P2)** follow since GBR is defined using the product of bipartite measures. Calculations show that $\mathrm{GBR}(GHZ) = 1$, while $\mathrm{GBR}(W) = 0.94$, satisfying **P3**. Some useful features of this measure are discussed next.

V. DISCUSSION

Calculation of GBR avoids cumbersome minimization procedures to find the closest separable state. Since it is based on the radial distance in the space of the reduced density operators, this measure is guaranteed to be monotonic under SLOCC. One way to see this is that the Bures distance to the radially outward state is a positive function of the radial coordinate of the reduced density operator, which is monotonic under SLOCC[3].

Another interesting feature of our measure is that it takes into account the curvature of the space of states as opposed to other measures that depend on the radial (Euclidean) distance in the space of states.

We expect these features to result in properties that distinguish this measure from some of the other measures in the literature.

Comparison with recent GMEs

In recent times, several GMEs have been proposed, such as Generalised Geometric Mean (GGM)[48, 49], Geniunely Multipartite Concurrence (GMC)[50], Concurrence Fill (F)[51] and Geometric Mean of Bipartite Concurrence (GBC)[33].

Concurrence Fill is an elegant and visualisable measure for 3-qubit systems, but is hard to generalize to higher number of qubits. Though there are beautiful entanglement polygon inequalities[52], there is no reason in general to expect the GME based on bipartite concurrences to obey volume or area laws for the n-qubit case[53].

An aspect for comparison is the effectiveness of different measures in discriminating between entanglement of states that belong to LU-inequivalent classes. There are only two SLOCC inequivalent classes of three qubit states namely the GHZ-type and W-type[27]. All valid GME's including GMC and GGM can differentiate the degree of entanglement between these two classes. However, there are further six sub-classes[35, 54] based on the entanglement between different subsystems. Among these are four classes of genuinely entangled states, which form LU-equivalence classes. They can be identified by the non-zero coefficients in their Generalised Schmidt Decomposition (GSD). The GSD of a 3-qubit pure state takes the canonical form

$$|\psi\rangle = \lambda_0|000\rangle + \lambda_1|100\rangle + \lambda_2|101\rangle + \lambda_3|110\rangle + \lambda_4|111\rangle.$$

In all these classes, λ_0 and λ_4 are always non-zero.

Class 1: all of $\lambda_1, \lambda_2, \lambda_3$ zero, (GHZ-class);

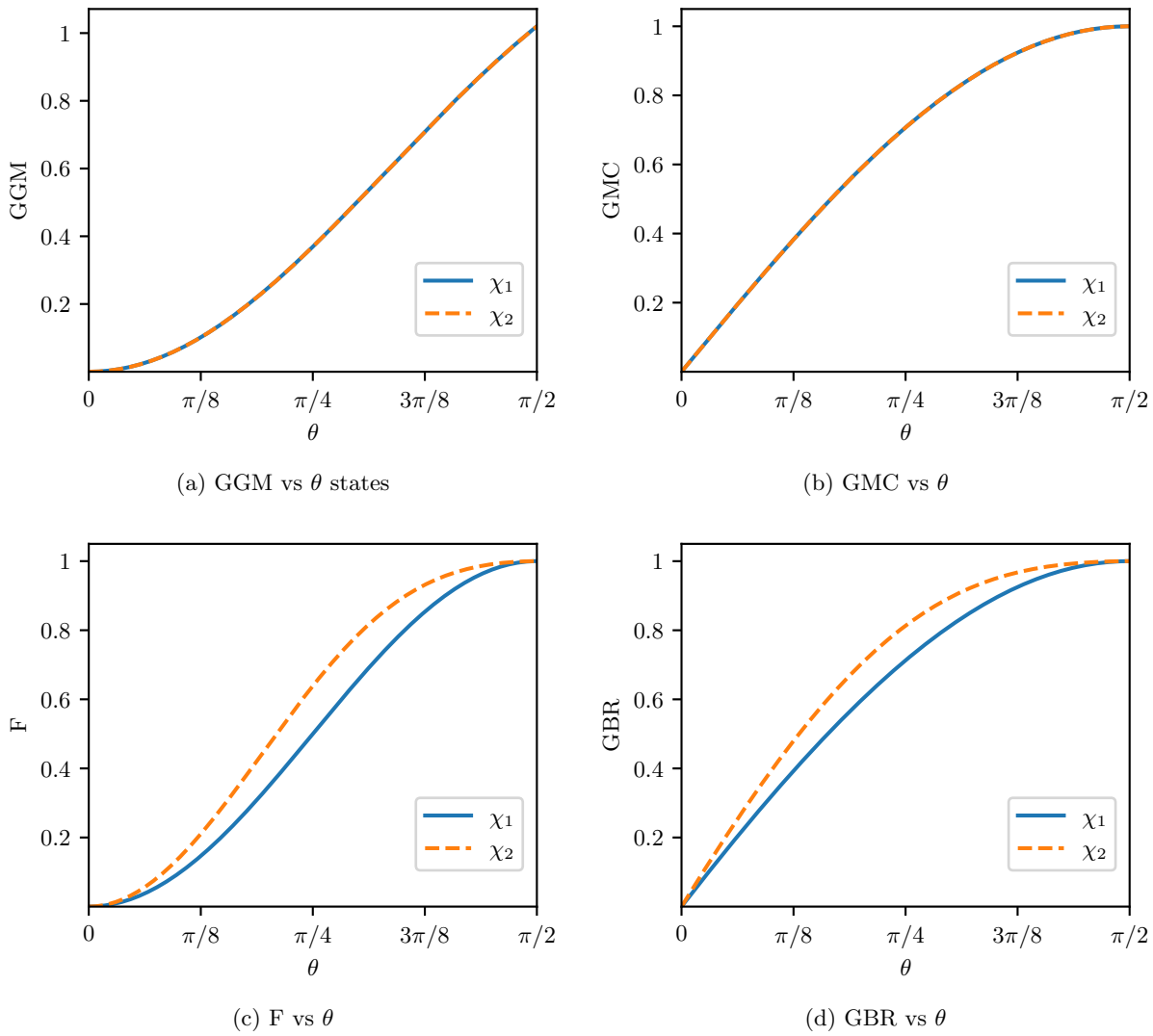
Class 2: any two of $\lambda_1, \lambda_2, \lambda_3$ zero;

Class 3: any one of $\lambda_1, \lambda_2, \lambda_3$ zero

Class 4: none of $\lambda_1, \lambda_2, \lambda_3$ zero. (W-class).

For example, GGM and GMC, both of which use the minimum entanglement among all bipartitions, are relatively poor at discriminating certain states that belong to LU-inequivalent classes. A similar argument has been put forward by Xie et al[51], where they show by example that GMC and GGM fail to differentiate the entanglement content of two states belonging to two different sub-classes, while concurrence fill successfully differentiates them. We demonstrate that GBR is also successful in discriminating these classes, using two example families of states:

$$|\chi_1(\theta)\rangle = \cos\frac{\theta}{2}|000\rangle + \sin\frac{\theta}{2}|111\rangle,$$
$$|\chi_2(\theta)\rangle = \frac{1}{\sqrt{2}}\left(\sin\theta|000\rangle + \cos\theta|110\rangle + |111\rangle\right). \tag{19}$$



(a) GGM vs θ states

(b) GMC vs θ

(c) F vs θ

(d) GBR vs θ

FIG. 4: GME for states $\chi_1(\theta)$ (Class 1) and $\chi_2(\theta)$ (Class 2) of Eq (19): (a) GGM and (b) GMC fail to distinguish these two classes, while the difference shows up in (c) Concurrence Fill and (d) GBR.

$|\chi_1(\theta)\rangle$ and $|\chi_2(\theta)\rangle$ belong to **Class 1** and **Class 2** respectively for $\theta \in (0, \pi/2]$. Fig. 4(a) and (b) show that GMC and GGM put them on equal footing, whereas concurrence fill in Fig. 4(c) and GBR in Fig. 4(d) distinguish their entanglement contents.

Entanglement being an intrinsic property of a quantum system, it is expected to vary smoothly with any single parameter of a family of states. For instance, consider the state:

$$|\chi_3(\theta)\rangle = \cos\theta\,|001\rangle + \frac{1}{\sqrt{2}}\sin\theta\big(\,|010\rangle + |100\rangle\,\big). \quad (20)$$

We plot GGM, GMC, GBC and GBR for this class in Fig. 5. GGM and GMC contain undesirable sharp peaks, a limitation of measures involving non-analytic functions such as the minimum. In contrast, we see that GBC and GBR are smooth. Fig. 5(b) is a zoomed-in view to highlight the difference between GBC and GBR, though they appear to overlap in Fig. 5(a).

VI. CONCLUSION

In this work we give a measure for genuine pure state entanglement using the Riemannian structure of quantum state space. The two-qubit measure REM is a computationally efficient method of calculating the distance to the closest separable state. This inspired an extension to multi-party systems through bipartitions. We construct a function of the n-qubit state (GBR) that satisfies all the requisite properties of a Genuine Measure of Entanglement.

It is noteworthy that this measure uses entanglement information from all bipartitions. Therefore it is better at discriminating different LU-invariant sub-classes of three qubit states than some of the existing GME's in liter-

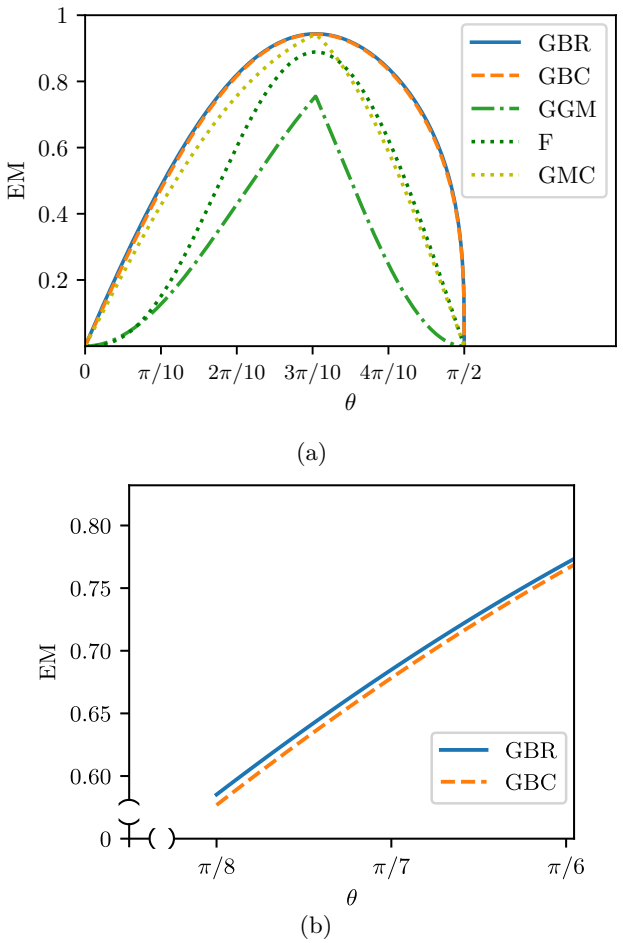


FIG. 5: (a) Comparison of GBR with other GME's for the state of class $\chi_3(\theta)$. In this scale, GBR and GBC appear to overlap, but their formulations are completely different. (b) highlights the difference between the curves in higher resolution.

ature that are based on minimum entanglement among bipartitions.

GBR explicitly uses the Riemannian structure of state space, and consequently varies smoothly with state parameters. Entropic measures for instance, are functions of the Euclidean distance, and do not pick up the curvature of the state space.

While measures like concurrence fill do not easily generalise to n-qubits, GBR by its definition can be constructed for all finite dimensional pure states.

As for mixed states, this measure can be readily extended using convex roof construction. A detailed analysis of computational cost reduction is work in progress.

ACKNOWLEDGEMENTS

We acknowledge support from the Department of Science and Technology, Government of India, through the project DST/ICPS/QuST/Theme-3/2019/Q109. We would like to thank Kinjal Banerjee for useful discussions.

[1] J. Eisert and M. B. Plenio, Journal of Modern Optics **46**, 145 (1999).
[2] K. Chen and L.-A. Wu, Physics Letters A **306**, 14 (2002).
[3] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Physical Review A **54**, 3824 (1996).
[4] A. P. Balachandran, T. R. Govindarajan, A. R. de Queiroz, and A. F. Reyes-Lega, Phys. Rev. A **88**, 022301 (2013).
[5] N. Pathania and T. Qureshi, International Journal of Theoretical Physics **61** (2022), 10.1007/s10773-022-05030-z.
[6] J. P. Provost and G. Vallee, Communications in Mathematical Physics **76**, 289 (1980).
[7] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Translations of mathematical monographs (American Mathematical Society, 2000).
[8] E. A. Morozova and N. N. Chentsov, Journal of Soviet Mathematics **56**, 2648 (1991).
[9] D. Petz and C. Sudár, Journal of Mathematical Physics **37**, 2662 (1996).
[10] V. Vedral and M. B. Plenio, Physical Review A **57**, 1619 (1998).
[11] J. Eisert, K. Audenaert, and M. B. Plenio, Journal of Physics A: Mathematical and General **36**, 5605 (2003).
[12] C. Witte and M. Trucks, Physics Letters A **257**, 14 (1999).
[13] P. Calabrese, J. Cardy, and B. Doyon, Journal of Physics A: Mathematical and Theoretical **42**, 500301 (2009).
[14] J. Avron and O. Kenneth, Reviews in Mathematical Physics **32**, 2030001 (2019).
[15] O. Gamel, Phys. Rev. A **93**, 062320 (2016).
[16] G. Kimura, Physics Letters A **314**, 339 (2003).
[17] M. S. Byrd and N. Khaneja, Physical Review A **68** (2003), 10.1103/physreva.68.062322.
[18] S. R. Hedemann, *Hyperspherical Bloch Vectors with Applications to Entanglement and Quantum State Tomography*, Ph.D. thesis, Stevens Institute of Technology, New

Jersey (2014).
[19] L. E. Blumenson, The American Mathematical Monthly **67**, 63 (1960).
[20] D. Petz and C. Sudár, Journal of Mathematical Physics **37**, 2662 (1996), https://doi.org/10.1063/1.531535.
[21] K. Zyczkowski and I. Bengtsson, Cambridge University Press , 363–414 (2006).
[22] F. Kubo and T. Ando, Mathematische Annalen **246**, 205 (1980).
[23] A. Shimony, Annals of the New York Academy of Sciences **755**, 675 (1995).
[24] P. Levay, J. Phys. A **37**, 1821 (2004), arXiv:quant-ph/0306115.
[25] S. Halder and U. Sen, "Separability and entanglement in superpositions of quantum states," (2021).
[26] J. Schlienz and G. Mahler, Phys. Rev. A **52**, 4396 (1995).
[27] W. Dür, G. Vidal, and J. I. Cirac, Physical Review A **62** (2000), 10.1103/physreva.62.062314.
[28] Z.-H. Ma, Z.-H. Chen, J.-L. Chen, C. Spengler, A. Gabriel, and M. Huber, Physical Review A **83** (2011), 10.1103/physreva.83.062325.
[29] S. Xie and J. H. Eberly, Phys. Rev. Lett. **127**, 040403 (2021).
[30] D. A. Meyer and N. R. Wallach, Journal of Mathematical Physics **43**, 4273 (2002), https://doi.org/10.1063/1.1497700.
[31] G. K. Brennen, Quantum Info. Comput. **3**, 619–626 (2003).
[32] P. J. Love, A. M. van den Brink, A. Y. Smirnov, M. H. S. Amin, M. Grajcar, E. Il'ichev, A. Izmalkov, and A. M. Zagoskin, Quantum Information Processing **6**, 187 (2007).
[33] Y. Li and J. Shang, Phys. Rev. Research **4**, 023059 (2022).
[34] S. Xie and J. H. Eberly, "Managing the three-party entanglement challenge," (2022).
[35] C. Sabín and G. García-Alcaine, The European Physical Journal D **48**, 435 (2008).
[36] M. Markiewicz, W. Laskowski, T. Paterek, and M. Żukowski, Phys. Rev. A **87**, 034301 (2013).
[37] F. Pan, D. Liu, G. Lu, and J. P. Draayer, International Journal of Theoretical Physics **43**, 1241 (2004).
[38] A. R. R. Carvalho, F. Mintert, and A. Buchleitner, Phys. Rev. Lett. **93**, 230501 (2004).
[39] J.-M. Cai, Z.-W. Zhou, X.-X. Zhou, and G.-C. Guo, Phys. Rev. A **74**, 042338 (2006).
[40] V. Coffman, J. Kundu, and W. K. Wootters, Phys. Rev. A **61**, 052306 (2000).
[41] A. Miyake, Physical Review A **67** (2003), 10.1103/physreva.67.012108.
[42] L. Xiong, J. Liu, and Q. Qin, Quantum Information Processing **21**, 102 (2022).
[43] S. Haddadi and M. Bohloul, International Journal of Theoretical Physics **57**, 3912 (2018).
[44] C. Yu and H. Song, Physics Letters A **330**, 377 (2004).
[45] D. Bures, Transactions of the American Mathematical Society **135**, 199 (1969).
[46] A. Uhlmann, Reports on Mathematical Physics **9**, 273 (1976).
[47] S. Xie and J. H. Eberly, Contemporary Physics **62**, 189 (2021).
[48] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).
[49] A. Sen(De) and U. Sen, Phys. Rev. A **81**, 012308 (2010).
[50] S. M. Hashemi Rafsanjani, M. Huber, C. J. Broadbent, and J. H. Eberly, Phys. Rev. A **86**, 062303 (2012).
[51] S. Xie and J. H. Eberly, Phys. Rev. Lett. **127**, 040403 (2021).
[52] X. Yang, Y.-H. Yang, and M.-X. Luo, Phys. Rev. A **105**, 062402 (2022).
[53] Y. Guo, Y. Jia, X. Li, and L. Huang, Journal of Physics A: Mathematical and Theoretical **55**, 145303 (2022).
[54] A. Acín, A. Andrianov, L. Costa, E. Jané, J. I. Latorre, and R. Tarrach, Phys. Rev. Lett. **85**, 1560 (2000).

Catalytic and asymptotic equivalence for quantum entanglement

Ray Ganardi[1*] Tulja Varun Kondra[1†] Alexander Streltsov[1‡]

[1] *Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland*

Abstract. Many quantum information processing tasks rely on singlets, but obtaining them is challenging due to the presence of noise. Typically, procedures involving asymptotically many copies of a state are considered to distill singlets from noisy states. When it comes to manipulating entangled systems on a single copy level, using entangled states as catalysts can significantly broaden the range of achievable transformations. Our results demonstrate a full equivalence between the asymptotic and catalytic settings for all distillable states. As an important consequence, we demonstrate that using an entangled catalyst cannot enhance the asymptotic singlet distillation rate of a distillable quantum state.

Keywords: Entanglement theory, distillable entanglement, entanglement catalysis

Entanglement is a key feature of quantum mechanics, and it plays a vital role in many areas of quantum information science. Being a strong form of correlations between quantum systems, entanglement enables a wide range of applications and protocols that have the potential to revolutionize information processing and communication [BBC+93, Eke91]. The study of entanglement and its properties has led to significant advancements in our understanding of quantum mechanics, and it has provided insights into how to manipulate and harness its power for practical applications [HHHH09].

To understand the pivotal role of entanglement as a resource in quantum information processing, we can consider the distant lab paradigm [BBP+96, BDSW96]. This scenario involves two parties, Alice and Bob, who are located in different quantum laboratories and can exchange classical messages to communicate with each other. In this setting, entangled states shared between Alice and Bob become a valuable resource, allowing them to perform tasks that would otherwise be impossible [HHHH09].

One of the most significant applications of entanglement is in the field of quantum communication, including quantum teleportation [BBC+93] and quantum cryptography [Eke91]. These tasks typically rely on singlets, which are pure highly entangled states of two qubits. However, in practice, Alice and Bob may only have access to noisy states. In order to use noisy states for singlet-based protocols, they can employ entanglement distillation [BBP+96, BBPS96], which is a special case of asymptotic state transformations. In this process, n copies of an initial state are transformed approximately into rn copies of the final state, where r is the transformation rate. Quantum states which can be distilled into singlets at a nonzero rate are called distillable. There exist noisy entangled states which cannot be distilled into singlets, a phenomenon known as bound entanglement [HHH98].

Another way how Alice and Bob can gain access to singlets from noisy states is to use entanglement catalysis. In this process, an auxiliary entangled state, known as a catalyst, is employed to aid in the transformation of one entangled state to another without altering the catalyst itself [JP99]. Recent work [KDS21, LBS21, RT22, DKMS22b] extended this idea to approximate catalysis, where the transformation can be achieved with a certain degree of inaccuracy. This concept has proven to be instrumental in advancing our understanding of catalytic entanglement manipulation and its potential applications [DKMS22a].

At first glance, catalytic and asymptotic transformations may seem like distinct concepts, but recent research has uncovered a strong connection between them. Initial evidence for a connection between these concepts was presented in [DFY05, DFLY05], and subsequent work has made significant progress in this direction, particularly through the use of approximate catalysis [KDS21, RT22]. Furthermore, it has been shown that in quantum thermodynamics, catalysis and many-copy transformations with a unit rate are fully equivalent [SS21, Wil21]. Given the shared features between quantum entanglement and thermodynamics [HOH02, PSW06, BP08, LR23], it is plausible that a similar equivalence may exist in entanglement theory.

In this article, we resolve this question by considering catalytic and asymptotic protocols which can establish a non-vanishing amount of correlations. This provides a more flexible and practical approach for studying catalysis and asymptotic transformations and their applications in quantum information processing. In this setting, we prove that for distillable states, catalysis and asymptotic transformations with unit rate are fully equivalent notions of entangled state manipulation. We discuss several applications of our results, including the crucial finding that the addition of a catalyst cannot increase the distillable entanglement of a noisy distillable state.

Asymptotic entanglement manipulations and catalysis

As previously discussed, asymptotic transformations are a powerful tool for understanding the structure and manipulation of quantum entanglement. For instance, consider two bipartite pure states $|\psi\rangle$ and $|\phi\rangle$. The objective is to use local operations and classical communication (LOCC) to transform n copies of $|\psi\rangle$ into m copies of $|\phi\rangle$, allowing for an error margin that vanishes in the limit of large n. The maximal ratio m/n defines the transformation rate. This

*r.ganardi@cent.uw.edu.pl
†t.ko@cent.uw.edu.pl
‡a.streltsov@cent.uw.edu.pl

framework is particularly useful if the target is the singlet state $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$, in which case the optimal rate is known as distillable entanglement [BBP+96, BBPS96, PV07]. It coincides with the entanglement entropy $E(|\psi\rangle) = S(\psi^A)$ of the initial state, and $S(\rho) = -\mathrm{Tr}[\rho \log_2 \rho]$ is the von Neumann entropy [BBPS96]. In a similar way, it is possible to define transformation rates for noisy states, we refer to the Methods section for more details. A state is called asymptotically reducible onto another state if the transformation can be achieved with a rate at least one [BPR+00]. This reflects the intuition that if $|\psi\rangle$ is reducible onto $|\phi\rangle$, then $|\psi\rangle$ is at least as valuable as $|\phi\rangle$ for any application that allows for asymptotic transformations.

Entanglement catalysis is a phenomenon where an entangled catalyst is used to facilitate the single-copy transformation of an entangled state into another without changing the state of the catalyst [JP99, EW00]. Given an entangled state $|\psi\rangle$ and a target state $|\phi\rangle$, the aim is to find a catalytic state $|\eta\rangle$ such that the transformation $|\psi\rangle\otimes|\eta\rangle \to |\phi\rangle\otimes|\eta\rangle$ is possible by LOCC. The catalyst is particularly useful if it enables the transformation of $|\psi\rangle$ into $|\phi\rangle$ which is not possible without the catalyst. Recently, the notion of catalysis has been extended to approximate catalysis in [KDS21, LBS21, RT22, DKMS22b], which allows for some degree of inaccuracy in the catalytic transformation. The notion of approximate catalysis provides a more realistic model for practical implementations of catalytic entanglement manipulation and enables a broader range of applications [DKMS22a]. It has been demonstrated in [KDS21] that transformations between bipartite pure states in this scenario are fully determined by the entanglement entropy of the corresponding states. Therefore, for bipartite pure states approximate catalysis is fully equivalent to reducibility. Catalytic phenomena have been extensively studied not only in the context of entanglement, but also in other areas of quantum physics, such as quantum thermodynamics [BHN+15, MÏ8, SS21, BEG+19, Wil21], where they are essential for understanding and manipulating quantum systems subject to constraints imposed by energy conservation.

As has been shown in [KDS21], there is a close connection between asymptotic state transformations and catalysis. More precisely, if a state ρ is asymptotically reducible to another state σ, then a transformation from ρ into σ can also be achieved on the single-copy level with approximate catalysis [KDS21]. However, it has remained a crucial open question whether the converse is also true, i.e., whether catalysis and asymptotic reducibility are fully equivalent notions for entangled state transformations. In this article, we introduce the frameworks of marginal asymptotic transformations and correlated catalysis, which allows us to resolve this question and establish the equivalence between catalysis and reducibility for all distillable quantum states.

Correlated catalysis and marginal reducibility

In the context of entanglement catalysis, an important generalization is to consider *correlated catalysis*, where the catalyst is allowed to have non-vanishing correlations with the system throughout the transformation process. This means that the system and the catalyst can remain correlated in the final state. More precisely, we say that ρ can be converted into σ via correlated catalysis if for any error margin $\varepsilon > 0$ there is an LOCC protocol Λ and a catalyst state τ such that

$$\begin{aligned} \mu^{SC} &= \Lambda(\rho^S \otimes \tau^C), \\ \|\mu^S - \sigma^S\|_1 &< \varepsilon, \quad \mu^C = \tau^C. \end{aligned} \tag{1}$$

Here, S denotes a possibly multipartite system, C denotes the catalyst, and $\|M\|_1 = \mathrm{Tr}\sqrt{M^\dagger M}$ is the trace norm. In other words, the state μ^{SC} is obtained by applying an LOCC protocol Λ to the state $\rho^S \otimes \tau^C$, such that the marginal on C is preserved and the resulting state on S can be made arbitrarily close to the target state σ. Previous studies in quantum thermodynamics have explored the significance of correlations for catalytic state transformations, revealing that the presence of correlations between the system and catalyst can increase the transformation power of the procedure [WGE17, BEG+19, Wil21, SS21].

We now introduce the notion of *marginal reducibility*. We say that ρ can be reduced into σ in the marginals if for any arbitrarily small error margin, there exists an LOCC protocol which can transform n copies of ρ into a state with approximately m marginals, and each marginal being close to the desired state σ. Specifically, we require for any $\varepsilon, \delta > 0$ there exists an LOCC protocol Λ and integers $m \leq n$ such that the following conditions hold for all $i \leq m$:

$$\begin{aligned} \Lambda\left(\rho^{\otimes n}\right) &= \mu_m, \\ \left\|\mu_m^{(i)} - \sigma\right\|_1 &< \varepsilon, \\ \frac{m}{n} + \delta &> 1. \end{aligned} \tag{2}$$

Here, μ_m is a state on m subsystems, each shared by Alice and Bob, and $\mu_m^{(i)}$ is the reduced state of μ_m on i-th subsystem. Marginal asymptotic transformations have been previously studied in continuous variable systems in [FLTP23].

It is worth to discuss the difference between marginal reducibility and the notion of reducibility introduced in [BPR+00]. The latter is more stringent as it requires that the final state μ_m is close to m copies of σ as a whole. However, for many quantum information processing tasks that rely on pure states $|\phi\rangle$, such as singlets in the bipartite case or GHZ states in the multipartite case, small perturbations of the state do not significantly affect its usefulness. For marginal reducibility, it suffices that $\rho^{\otimes n}$ can be approximately converted into $\mu_\varepsilon^{\otimes n}$ for any $\varepsilon > 0$, which as we have argued above is enough for many tasks based on pure states. Therefore, we suggest that the framework of marginal reducibility is particularly suitable when one aims to produce pure entangled states of high quality that are intended to be used independently.

In the remainder of this article we will focus on the relationship between correlated catalysis and marginal reducibility. Unless otherwise specified, we will refer to these concepts simply as catalysis and reducibility, respectively.

Catalysis-reducibility equivalence

As previously noted, there have been indications that catalysis and reducibility are interchangeable concepts for entangled state transformations [DFY05, KDS21, RT22,

DFLY05]. The key contribution of this article is to establish this equivalence for any pair of distillable states, using the notions of catalysis and reducibility which include correlations, as outlined above in this article. We recall that a distillable state is a quantum state which can be converted into singlets at nonzero rate in the asymptotic limit.

Theorem 1. *For any pair of distillable states ρ and σ reducibility and catalysis are fully equivalent.*

We present a brief overview of the techniques employed for proving the theorem, more details can be found in the Methods section and the Supplementary information. Firstly, we demonstrate that if the state ρ can be reduced to the state σ, then it is possible to achieve a catalytic transformation from ρ to σ, using techniques similar to those presented in prior works [SS21, KDS21, LBS21, RT22]. Subsequently, we establish the converse by explicitly constructing a reduction protocol that utilizes a catalytic conversion protocol from a distillable state ρ into σ. This involves several technical steps that are described in detail in the Supplementary information. By combining these two results, we conclusively demonstrate the full equivalence of reducibility and catalysis for any pair of distillable states ρ and σ.

Since all entangled two-qubit states are distillable [HHH97], Theorem 1 implies that catalysis and reducibility are fully equivalent for all two-qubit states. For states beyond two qubits, Theorem 1 also applies if the target state σ is not distillable. Moreover, catalysis is generally at least as powerful as reducibility. With this in mind, Theorem 1 leaves open the possibility that there exist bound entangled states ρ that cannot be reduced to some state σ, yet a catalytic conversion from ρ to σ is possible. Thus, if catalysis and reducibility are not equally powerful on all quantum states, catalysis must show an advantage on some bound entangled initial states. This underscores the importance of investigating the relationship between these concepts in the general case, as it can provide insights into the nature of bound entanglement and the power of entanglement catalysis. Additionally, our findings can have practical implications for quantum information processing tasks where bound entangled states are known to play a significant role [HHH99, SST01, VB14].

Going one step further, we investigate the role of catalysis for asymptotic transformation rates. Our findings reveal that the addition of a catalyst does not alter the asymptotic rate of transformation from a distillable state ρ into another state σ, again under the assumption that correlations can be established in the procedure. An important application of this result pertains to the scenario where the target state is a singlet $|\psi\rangle^-$. In this context, our analysis reveals that the correlations, which are typically established in the catalytic and asymptotic procedures considered earlier, vanish. This property allows us to explore the features of distillable entanglement when a catalyst is incorporated into the transformation, bringing us to the second main result of this article.

Theorem 2. *Catalysis cannot increase the distillable entanglement of a distillable state.*

The proof of the theorem combines the previously mentioned results on asymptotic transformation rates with the additional finding that correlations usually established in the involved procedures disappear if the target state is pure. We refer to the Methods section for the proof and more details. Recalling that all entangled two-qubit states can be distilled into singlets [HHH97], it follows that Theorem 2 applies to all two-qubit states. In general, our results leave open the possibility that bound entangled states could be activated into singlets through catalysis.

Our results have implications also beyond the scope of bipartite systems. It is worth noting that Theorem 1 can be generalized to the multipartite scenario. To this end, we consider multipartite distillable states, which are those multipartite states that can be distilled into singlets between each pair of parties with some nonzero rate in the asymptotic limit This includes all pure states which are entangled across any bipartition [SVW05, HOW05]. With this in mind, we can extend Theorem 1 to state that for any pair of multipartite distillable states, reducibility and catalysis are fully equivalent. Furthermore, Theorem 2 is also applicable to this scenario, indicating that the addition of a catalyst cannot enhance the multipartite distillable entanglement of any multipartite distillable state, we refer to the Methods section for more details. The results obtained in the multipartite setting are in line with those in the bipartite setting and imply that if catalysis offers any benefit over reducibility, it can only be observed when the initial state is not distillable.

These findings offer a better understanding of the relationship between entanglement catalysis and many-copy transformations, and can have practical implications for the exploitation of entanglement in quantum information processing tasks.

References

[BBC+93] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70:1895–1899, Mar 1993. URL: https://link.aps.org/doi/10.1103/PhysRevLett.70.1895, doi:10.1103/PhysRevLett.70.1895.

[BBP+96] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. *Phys. Rev. Lett.*, 76:722–725, Jan 1996. URL: https://link.aps.org/doi/10.1103/PhysRevLett.76.722, doi:10.1103/PhysRevLett.76.722.

[BBPS96] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher. Concentrating partial entanglement by local operations. *Phys. Rev. A*, 53:2046–2052, Apr 1996. URL: https://link.aps.org/doi/10.1103/PhysRevA.53.2046, doi:10.1103/PhysRevA.53.2046.

[BDSW96] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996. URL: https://link.aps.org/doi/10.1103/PhysRevA.54.3824, doi:10.1103/PhysRevA.54.3824.

[BEG+19] Paul Boes, Jens Eisert, Rodrigo Gallego, Markus P. Müller, and Henrik Wilming. Von Neumann Entropy from Unitarity. *Phys. Rev. Lett.*, 122:210402, May 2019. URL: https://link.aps.org/doi/10.1103/PhysRevLett.122.210402, doi:10.1103/PhysRevLett.122.210402.

[BHN+15] Fernando Brandão, Michał Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. The second laws of quantum thermodynamics. *Proc. Natl. Acad. Sci. U.S.A.*, 112(11):3275–3279, 2015. doi:10.1073/pnas.1411728112.

[BP08] Fernando G. S. L. Brandão and Martin B. Plenio. Entanglement theory and the second law of thermodynamics. *Nat. Phys.*, 4(11):873–877, Nov 2008. doi:10.1038/nphys1100.

[BPR+00] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63:012307, Dec 2000. URL: https://link.aps.org/doi/10.1103/PhysRevA.63.012307, doi:10.1103/PhysRevA.63.012307.

[DFLY05] Runyao Duan, Yuan Feng, Xin Li, and Mingsheng Ying. Multiple-copy entanglement transformation and entanglement catalysis. *Phys. Rev. A*, 71:042319, Apr 2005. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.042319, doi:10.1103/PhysRevA.71.042319.

[DFY05] Runyao Duan, Yuan Feng, and Mingsheng Ying. Entanglement-assisted transformation is asymptotically equivalent to multiple-copy transformation. *Phys. Rev. A*, 72:024306, Aug 2005. URL: https://link.aps.org/doi/10.1103/PhysRevA.72.024306, doi:10.1103/PhysRevA.72.024306.

[DKMS22a] Chandan Datta, Tulja Varun Kondra, Marek Miller, and Alexander Streltsov. Catalysis of entanglement and other quantum resources. *arXiv:2207.05694*, 2022. URL: https://arxiv.org/abs/2207.05694.

[DKMS22b] Chandan Datta, Tulja Varun Kondra, Marek Miller, and Alexander Streltsov. Entanglement catalysis for quantum states and noisy channels. *arXiv:2202.05228*, 2022. URL: https://arxiv.org/abs/2202.05228.

[Eke91] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991. URL: https://link.aps.org/doi/10.1103/PhysRevLett.67.661, doi:10.1103/PhysRevLett.67.661.

[EW00] Jens Eisert and Martin Wilkens. Catalysis of Entanglement Manipulation for Mixed States. *Phys. Rev. Lett.*, 85:437–440, Jul 2000. URL: https://link.aps.org/doi/10.1103/PhysRevLett.85.437, doi:10.1103/PhysRevLett.85.437.

[FLTP23] Giovanni Ferrari, Ludovico Lami, Thomas Theurer, and Martin B. Plenio. Asymptotic State Transformations of Continuous Variable Resources. *Commun. Math. Phys.*, 398(1):291–351, Feb 2023. doi:10.1007/s00220-022-04523-6.

[HHH97] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Inseparable Two Spin-$\frac{1}{2}$ Density Matrices Can Be Distilled to a Singlet Form. *Phys. Rev. Lett.*, 78:574–577, Jan 1997. URL: https://link.aps.org/doi/10.1103/PhysRevLett.78.574, doi:10.1103/PhysRevLett.78.574.

[HHH98] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature? *Phys. Rev. Lett.*, 80:5239–5242, Jun 1998. URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.5239, doi:10.1103/PhysRevLett.80.5239.

[HHH99] Paweł Horodecki, Michał Horodecki, and Ryszard Horodecki. Bound Entanglement Can Be Activated. *Phys. Rev. Lett.*, 82:1056–1059, Feb 1999. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.1056, doi:10.1103/PhysRevLett.82.1056.

[HHHH09] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009. URL: https://link.aps.org/doi/10.1103/RevModPhys.81.865, doi:10.1103/RevModPhys.81.865.

[HOH02] Michał Horodecki, Jonathan Oppenheim, and Ryszard Horodecki. Are the Laws of Entanglement Theory Thermodynamical? *Phys. Rev. Lett.*, 89:240403, Nov 2002. URL: https://link.aps.org/doi/10.1103/PhysRevLett.89.240403, doi:10.1103/PhysRevLett.89.240403.

[HOW05] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Partial quantum information. *Nature*, 436:673–676, Aug 2005. doi:10.1038/nature03909.

[JP99] Daniel Jonathan and Martin B. Plenio. Entanglement-Assisted Local Manipulation of Pure Quantum States. *Phys. Rev. Lett.*, 83:3566–3569, Oct 1999. URL: https://link.aps.org/doi/10.1103/PhysRevLett.83.3566, doi:10.1103/PhysRevLett.83.3566.

[KDS21] Tulja Varun Kondra, Chandan Datta, and Alexander Streltsov. Catalytic Transformations of Pure Entangled States. *Phys. Rev. Lett.*, 127:150503, Oct 2021. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.150503, doi:10.1103/PhysRevLett.127.150503.

[LBS21] Patryk Lipka-Bartosik and Paul Skrzypczyk. Catalytic Quantum Teleportation. *Phys. Rev. Lett.*, 127:080502, Aug 2021. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.080502, doi:10.1103/PhysRevLett.127.080502.

[LR23] Ludovico Lami and Bartosz Regula. No second law of entanglement manipulation after all. *Nat. Phys.*, 19(2):184–189, Feb 2023. doi:10.1038/s41567-022-01873-9.

[MÏ8] Markus P. Müller. Correlating Thermal Machines and the Second Law at the Nanoscale. *Phys. Rev. X*, 8:041051, Dec 2018. URL: https://link.aps.org/doi/10.1103/PhysRevX.8.041051, doi:10.1103/PhysRevX.8.041051.

[PSW06] Sandu Popescu, Anthony J. Short, and Andreas Winter. Entanglement and the foundations of statistical mechanics. *Nat. Phys.*, 2(11):754–758, Nov 2006. doi:10.1038/nphys444.

[PV07] Martin B. Plenio and S. Virmani. An introduction to entanglement measures. *Quant. Inf. Comput.*, 7:1, 2007. URL: https://arxiv.org/abs/quant-ph/0504163.

[RT22] Roberto Rubboli and Marco Tomamichel. Fundamental Limits on Correlated Catalytic State Transformations. *Phys. Rev. Lett.*, 129:120506, Sep 2022. URL: https://link.aps.org/doi/10.1103/PhysRevLett.129.120506, doi:10.1103/PhysRevLett.129.120506.

[SS21] Naoto Shiraishi and Takahiro Sagawa. Quantum Thermodynamics of Correlated-Catalytic State Conversion at Small Scale. *Phys. Rev. Lett.*, 126:150502, Apr 2021. URL: https://link.aps.org/doi/10.1103/PhysRevLett.126.150502, doi:10.1103/PhysRevLett.126.150502.

[SST01] Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Nonadditivity of Bipartite Distillable Entanglement Follows from a Conjecture on Bound Entangled Werner States. *Phys. Rev. Lett.*, 86:2681–2684, Mar 2001. URL: https://link.aps.org/doi/10.1103/PhysRevLett.86.2681, doi:10.1103/PhysRevLett.86.2681.

[SVW05] John A. Smolin, Frank Verstraete, and Andreas Winter. Entanglement of assistance and multipartite state distillation. *Phys. Rev. A*, 72:052317, Nov 2005. URL: https://link.aps.org/doi/10.1103/PhysRevA.72.052317, doi:10.1103/PhysRevA.72.052317.

[VB14] Tamás Vértesi and Nicolas Brunner. Disproving the Peres conjecture by showing Bell nonlocality from bound entanglement. *Nat. Commun.*, 5(1):5297, Nov 2014. doi:10.1038/ncomms6297.

[WGE17] Henrik Wilming, Rodrigo Gallego, and Jens Eisert. Axiomatic Characterization of the Quantum Relative Entropy and Free Energy. *Entropy*, 19(6):241, 2017. doi:10.3390/e19060241.

[Wil21] H. Wilming. Entropy and Reversible Catalysis. *Phys. Rev. Lett.*, 127:260402, Dec 2021. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.260402, doi:10.1103/PhysRevLett.127.260402.

Valid and efficient entanglement verification with finite copies of a quantum state

Paweł Cieśliński,[1] Jan Dziewior,[2,3,4] Lukas Knips,[2,3,4] Waldemar Kłobus,[1] Jasmin Meinecke,[2,3,4] Tomasz Paterek,[1,5] Harald Weinfurter,[2,3,4] and Wiesław Laskowski[1,6]

[1] *Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics, and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland*
[2] *Max Planck Institute for Quantum Optics, 85748 Garching, Germany*
[3] *Faculty of Physics, Ludwig Maximilian University, 80799 Munich, Germany*
[4] *Munich Center for Quantum Science and Technology, 80799 Munich, Germany*
[5] *School of Mathematics and Physics, Xiamen University Malaysia, 43900 Sepang, Malaysia*
[6] *International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-308 Gdańsk, Poland*

Detecting entanglement of multipartite quantum states is an inherently probabilistic process, typically with a small number of measured samples. The level of confidence in entanglement detection can be used to quantify the validity of the detection scheme via the probability that the measured signal is coming from a separable state, and provides a meaningful figure of merit for big data sets. Yet, for limited sample sizes, to avoid serious misinterpretations of the experimental results, one should not only consider the probability that a separable state gave rise to the measured signal but should also include information about the probability that the signal came from an entangled state, i.e. the efficiency of the detection scheme. We demonstrate this explicitly and propose a general method to optimise both the validity and the efficiency in small data sets. The method is applicable to arbitrary entanglement witnesses and is based on an analytical model of finite statistics effects on correlation functions. As an example, we derive the optimal number of measurement settings and distribution of clicks per setting that guarantee high validity and efficiency of entanglement verification with only 20 copies of a state. The analysis takes into account both a Frequentist as well as Bayesian approach.

Quantum entanglement is long recognised as an important prerequisite of modern quantum technologies. Its detection is accordingly a well studied topic with a plethora of different methods available. The field has evolved towards strategies directly applicable to experimental data which inevitably is limited to a finite number of detection

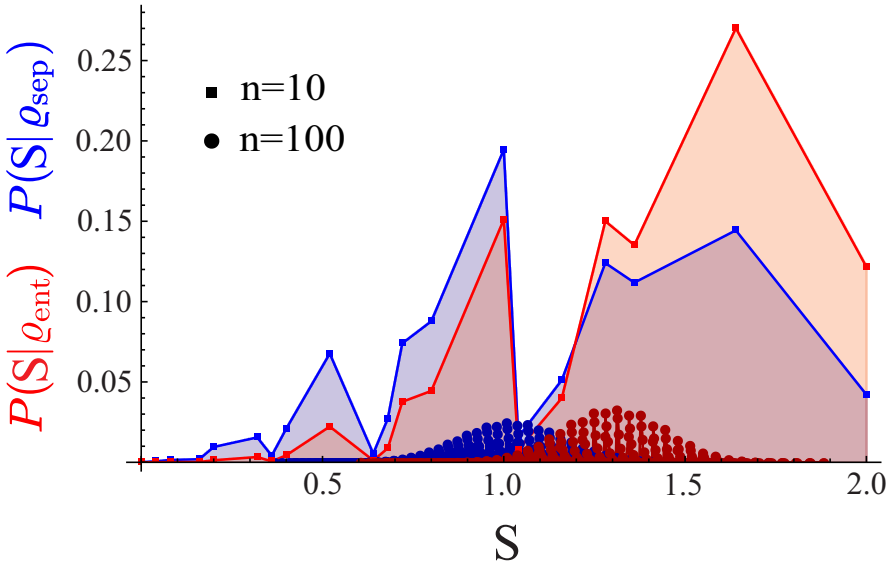


FIG. 1. **Illustration of finite statistics effects on entanglement verification**. The plot shows the probability distributions to obtain a specific value of a certain nonlinear entanglement witness S. The witness uses only two correlation measurements, each estimated either with $n = 10$ (squares) or $n = 100$ (circles) copies of the states. The probability distribution given an entangled state is marked in red. The probability in blue is computed for a separable state that saturates the value of the witness in the case of infinite data sets, i.e. a separable state that tends to yield large values of S. For larger statistics (see circles), the two distributions have small overlap making it easy to verify that entangled state was prepared. For smaller statistics (see squares) the overlap is very significant and care has to taken even if big values of the witness are observed. Because of this issue, we have developed universal tools, applicable to any entanglement witness, that extend it to the domain of finite statistics and can be employed to efficiently use every single copy of a state.

events. If this number is large, various forms of entanglement witnesses [1, 2] provide practically deterministic entanglement verification [3]. Interestingly, also the analysis of smaller data sets allows to detect entanglement [4], most recently with quantum state verification methods [5, 6] and tailored game-like protocols in which particles are measured one by one [7–10]. These methods are of high practical relevance in all cases where only a limited amount or only partial data is accessible. The finiteness of data sets used to derive a conclusion about entanglement inevitably leads to a probabilistic nature of this conclusion. The finiteness of data sets used to derive a conclusion about entanglement inevitably leads to a probabilistic nature of this conclusion. By far the two most prominent ways to quantify the validity of such a probabilistic statement are the confidence and the credibility, respectively related to the Frequentist and Bayesian approaches. Both of these measures capture different important aspects of statistical validity.

We showed explicitly that for small data sets it is necessary to consider not only $P(Q|\text{sep})$, i.e. the probability of a separable state to yield particular values of the witness Q, but also $P(Q|\text{ent})$ of entangled states. This is illustrated in Fig. 1, which shows that for smaller statistics one is required to go to more and more strict criteria to reliably distinguish between results compatible with entangled states and results compatible with separable states. In consequence, due to the increased strictness it becomes also more and more improbable that an entangled state will pass the test. Thus, especially in the case of small statistics, confidence or credibility is not the only figure of merit and we argue that one also has to ensure the *efficiency* of the method quantifying what fraction of entangled systems is expected to pass the criterion. A discussion of this parameter has to be an integral part of any useful measurement scheme. Only with these two quantifiers it becomes possible to optimize the usage of the available experimental resources. We have introduced a universal procedure that allows the extension of any entanglement witness to the domain of finite statistics. It is based on an analytical calculation of probability distributions over possible outcomes when estimating correlation functions experimentally with finite resources. Based on these probability distributions we defined the measures of validity and efficiency, applicable to any entanglement detection scheme, from the points of view of both Frequentist and Bayesian interpretation. We provide illustrative examples based on linear as well as nonlinear witnesses and broad families of states. The methods introduced are directly applicable to raw data and should be especially helpful for a resource-efficient estimation of the performance of a known apparatus subject to variations of external parameters or in multipartite experiments with rare detection events, e.g., multi-photon setups based on coincidence clicks. As an example, our scheme could be employed to quickly certify the quality of a large quantum processor before a time-consuming computation task.

[1] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[3] J. M. Arrazola, O. Gittsovich, J. M. Donohue, J. Lavoie, K. J. Resch, and N. Lütkenhaus, Reliable entanglement verification, Phys. Rev. A **87**, 062331 (2013).

[4] R. Blume-Kohout, J. O. S. Yin, and S. J. van Enk, Entanglement verification with finite data, Phys. Rev. Lett. **105**, 170501 (2010).

[5] K. Wang and M. Hayashi, Optimal verification of two-qubit pure states, Phys. Rev. A **100**, 032315 (2019).

[6] W.-H. Zhang, C. Zhang, Z. Chen, X.-X. Peng, X.-Y. Xu, P. Yin, S. Yu, X.-J. Ye, Y.-J. Han, J.-S. Xu, G. Chen, C.-F. Li, and G.-C. Guo, Experimental optimal verification of entangled states using local measurements, Phys. Rev. Lett. **125**, 030506 (2020).

[7] A. Dimić and B. Dakić, Single-copy entanglement detection, npj Quant. Inf. **4**, 11 (2018).

[8] V. Saggio, A. Dimić, C. Greganti, L. A. Rozema, P. Walther, and B. Dakić, Experimental few-copy multipartite entanglement detection, Nat. Phys. **15**, 935 (2019).

[9] J. Morris, V. Saggio, A. Gočanin, and B. Dakić, Quantum verification and estimation with few copies, Adv. Quant. Tech. , 2100118 (2022).

[10] V. Saggio and P. Walther, Few-copy entanglement detection in the presence of noise, Annalen der Physik **534**, 2100597 (2022).

Indefinite causal key distribution

Hector Spencer-Wood[1] *

[1] *School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, Scotland*

Abstract. We propose a quantum key distribution (QKD) protocol that is carried out in an indefinite causal order (ICO). In QKD, one considers a setup in which two parties, Alice and Bob, share a key with one another in such a way that they can detect whether an eavesdropper, Eve, has learnt anything about the key. To our knowledge, in all QKD protocols proposed until now, Eve is detected by publicly comparing a subset of Alice and Bob's key and checking for errors. If operations can be applied in an indefinite causal order, we show that, interestingly, the presence of Eve can be detected by Alice alone, without publicly comparing any information about the key with Bob. Indeed, we prove that both correlated and uncorrelated eavesdroppers cannot extract any useful information about the shared key without inducing a nonzero probability of being detected.

Keywords: Indefinite causal order, quantum cryptography, quantum information

Introduction

In our everyday, classical world, we are used to events occurring in a well defined order: A happens before B or vice versa. Remarkably, it appears that, in the quantum world, events can happen in a superposition of orders [1, 2, 3]. This phenomenon has been termed indefinite causal order (ICO) and, aside from the foundational interest in this topic, a number of applications have been proposed that show differences when compared to their definite causal counterparts [1, 4, 5, 6, 7]. Here, we propose another such application, this time in the well established field of quantum key distribution (QKD).

QKD is concerned with the scenario in which two parties, conventionally named Alice and Bob, would like to share a private key (a string of 0s and 1s) in such a way that they are confident an eavesdropping third party, called Eve, has not been listening in. There have been a number of protocols proposed [8, 9, 10, 11, 12, 13, 14], the first by Charles Bennett and Giles Brassard in 1984 (BB84) [8]. The security of these protocols comes from the fact that Eve can be detected. This is possible because, when Eve is present, due to the quantum phenomenon of measurement disturbance, a non-zero probability of error in Bob's key, with respect to Alice's, is induced. So, if one could somehow detect these errors induced by Eve, assuming noiseless and lossless transmission, it could be concluded that an eavesdropper had been listening in. These errors are normally detected by Alice and Bob publicly comparing a subset of their respective keys. Now public information, this subset is subsequently discarded regardless of whether they conclude Eve is there or not.

To our knowledge, this public comparison is a feature of all QKD protocols so far proposed. In this work, presented fully in [15], we show that using ICO, it is possible to determine whether eavesdroppers are there or not *without* having to publicly compare, and subsequently discard, a subset of the distributed key. This work therefore highlights a new area in quantum information science that appears to exhibit different features in a world that allows ICO compared to one that does not.

Quantum key distribution

Suppose two parties, Alice and Bob, would like to share a private key to use for some cryptographic task. This is often done, as in BB84, by having Alice prepare qubits in states that correspond to the 0s and 1s of the private key and sending them to Bob to be measured. Indeed, in BB84, Alice and Bob respectively prepare and measure, independently and randomly, in one of two non-orthogonal bases. In this work, we will use the Pauli x and z-bases: $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$ respectively, where $|\pm\rangle := (|0\rangle \pm |1\rangle)/\sqrt{2}$. If Alice (Bob) prepared (measured) the qubit to be in the state $|0\rangle$ or $|+\rangle$, she (he) will have a corresponding key bit of 0. Likewise, if $|1\rangle, |-\rangle$ the corresponding key bit will be 1. Once Bob has measured the qubit Alice sent him, the two parties publicly discuss which bases they chose. If they chose different bases, there is only a 50% chance of them agreeing on the key bit value, so they discard the corresponding key bit. If, however, they chose the same basis, when no eavesdroppers are present, Bob's measurement result is guaranteed to correspond to the state that Alice prepared, assuming noiseless and lossless transmission, as we will do throughout. Therefore, Alice and Bob can use the corresponding ordered set of key bit values as their shared key.

To make this protocol secure, notice that when an eavesdropper, Eve, intercepts the transmission from Alice to Bob and tries to learn the key bit value being shared, she disturbs the quantum state being sent with non-zero probability. This means that, even if Alice and Bob agree on the basis chosen, there is a non-zero probability that they disagree on the state of the qubit, meaning that there is a chance of an error in Bob's key with respect to Alice's. To detect these errors, Alice and Bob take a subset of their final keys and compare them *publicly*. Let us now see how this protocol can be adapted to an indefinite causal ordered setting.

*h.spencer-wood.1@research.gla.ac.uk

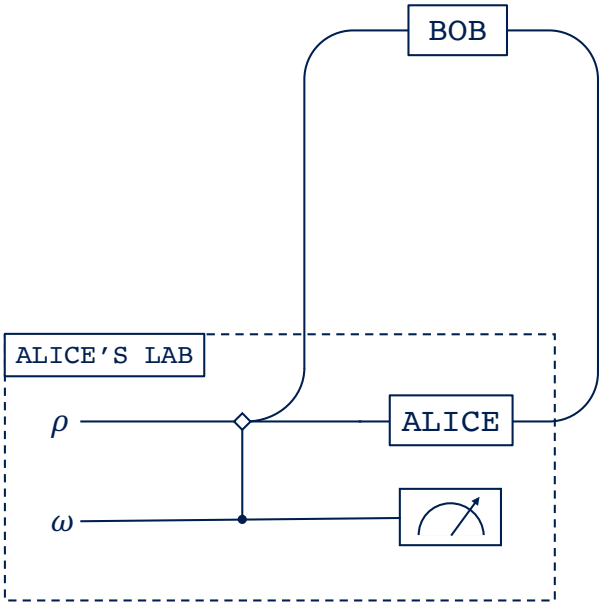


Figure 1: Indefinite causal key distribution with no eavesdroppers. A key is shared between Alice and Bob by sending a qubit ρ to them in a superposition of orders controlled by the qubit ω. Alice and Bob perform projective measurements randomly in either the Pauli x or z-basis. After discarding cases in which Alice and Bob measured in different bases, they are left with identical keys. Regardless of the superposition of orderings, ω never changes when there are no eavesdroppers.

Quantum key distribution in an indefinite causal order

In BB84, Alice would prepare the qubits to be sent to Bob in a certain state. When considering an indefinite causal ordered scheme, Alice is simultaneously sending and receiving the qubit from Bob, so having one party prepare the state makes no sense. To avoid this, *both Alice and Bob* measure the qubit being used, which, because of how states are updated following projective measurements, allows them to both be the preparer and measurer of the shared qubit. This method has similarities to how the key is generated in protocols like E91 [9]. Taking this approach, the key would be made up of the results of a projective measurement on some qubit ρ, but only when Alice and Bob agree they had performed the *same* projective measurement.

Thinking of the key generation in this way, we can consider a scheme in which a key is distributed in an indefinite causal order. Here, we send a state ρ to two parties, Alice and Bob, in a superposition of two orders: Alice before Bob and Bob before Alice. As shown in FIG. 1, this superposition is controlled by the qubit state ω: if $\omega = |0\rangle\langle 0|$, ρ travels around the loop in one direction, if $\omega = |1\rangle\langle 1|$, ρ travels around the loop in the opposite direction, and if $\omega = |\varphi\rangle\langle\varphi|$ is in some superposition of $|0\rangle$ and $|1\rangle$, ρ travels around the loop in a superposition of both directions. Alice and Bob then both make a random choice to measure either in the Pauli z-basis $\{|0\rangle, |1\rangle\}$ or x-basis $\{|+\rangle, |-\rangle\}$. We can therefore think of Alice and Bob as acting on the state by putting it through a quantum channel $\mathcal{A}$, defined by the Kraus operators

$$\begin{aligned} A_{0/1} &= \frac{1}{\sqrt{2}}|0/1\rangle\langle 0/1|, \\ A_{+/-} &= \frac{1}{\sqrt{2}}|+/-\rangle\langle +/-|, \end{aligned} \tag{1}$$

where the factors of $1/\sqrt{2}$ arise because we are assuming Alice and Bob are both equally likely to measure in the x or z-basis. For convenience, define the set containing the Kraus operator indices by $I := \{0, 1, +, -\}$.

Following their measurements, Alice and Bob then publicly discuss the basis they chose for each measurement and only keep the measurement outcomes in which they measured ρ in the same basis. Assuming no errors occur between Alice and Bob's measurements, their keys, made up of the measurement outcomes they kept, should be identical. In what follows, similarly to what we discussed earlier, a measurement outcome of 0 and $+$ will correspond to a 0 in the key. Likewise, 1 and $-$ correspond to a 1 in the key.

Let's see in more detail out what happens to the state ρ when it is put through the setup in FIG. 1. Following [7], the channel that ρ goes through, corresponding to a superposition of being measured by Alice first then Bob, and vice versa, is given by

$$\mathcal{S}_\omega(\mathcal{A}, \mathcal{A})(\rho) = \sum_{i,j\in I} S_{ij}\rho \otimes \omega S_{ij}^\dagger, \tag{2}$$

where

$$S_{ij} = A_i A_j \otimes |0\rangle\langle 0| + A_j A_i \otimes |1\rangle\langle 1|. \tag{3}$$

After some algebra and index relabelling, it can be shown that, following the public discussion of basis used,

$$\begin{aligned} \mathcal{S}_\omega(\mathcal{A}, \mathcal{A})(\rho) \to \frac{1}{2}\sum_{S\in B}\sum_{i,j\in S} \big(\{A_i, A_j\}\rho\{A_i, A_j\}^\dagger \otimes \omega \\ + [A_i, A_j]\rho[A_i, A_j]^\dagger \otimes \sigma_z\omega\sigma_z\big) \end{aligned} \tag{4}$$

where the prefactor is found by requiring normalisation, σ_z is the z Pauli operator and $B = \{\{0, 1\}, \{+, -\}\}$. Noting the form of A_k given in Eq. (1), the terms in these sums have the following properties

$$\begin{aligned} \{A_i, A_j\} &= \sqrt{2}A_i\delta_{ij}, \\ [A_i, A_j] &= 0, \end{aligned} \tag{5}$$

for all i, j, where δ_{ij} is the Kronecker delta. This confirms that Alice and Bob must agree in their measurement outcomes. Overall, we have that

$$\mathcal{S}_\omega(\mathcal{A}, \mathcal{A})(\rho) \to \sum_{i\in I} A_i\rho A_i^\dagger \otimes \omega. \tag{6}$$

So, when there are no eavesdroppers present, the control qubit ω stays in its original state and this situation is ultimately no different from that when the causal order is definite. Let us introduce an eavesdropper to see what changes.

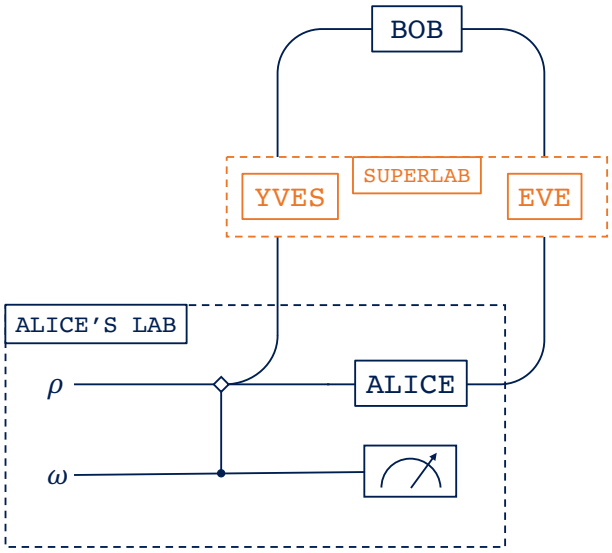


Figure 2: In this protocol, there are two places eavesdroppers can reside, indicated by Eve and Yves. Indeed, these eavesdroppers can work together, depicted here as some hypothetical "superlab".

Introducing eavesdroppers

Notice that, unlike in BB84, there are two places an eavesdropper can reside (see FIG. 2). Having said this, to obtain some intuition as to how eavesdroppers change things, let us first consider introducing just a single eavesdropper, Eve, between Alice and Bob. Denote the channel corresponding to Eve's measurement by $\mathcal{E}$, defined by the Kraus operators $\{E_i\}$. As before, allowing a qubit ρ to be acted on by Alice, Eve and Bob in an indefinite causal order controlled by ω, the channel ρ passes through is given by

$$\mathcal{S}_\omega(\mathcal{A},\mathcal{E},\mathcal{A})(\rho) = \sum_{i,j,k} S_{ijk}\rho \otimes \omega S_{ijk}^\dagger, \tag{7}$$

where

$$S_{ijk} := A_i E_j A_k \otimes |0\rangle\langle 0| + A_k E_j A_i \otimes |1\rangle\langle 1|. \tag{8}$$

Note that E_j is always in the middle since Eve is in between Alice and Bob. Again, after some algebra and index relabelling, and after basis comparison,

$$\mathcal{S}_\omega(\mathcal{A},\mathcal{E},\mathcal{A})(\rho) \to \frac{1}{2}\sum_{\substack{S\in B\\ j\\ i,k\in S}} \big[(A_{\{i}E_jA_{k\}})\rho(A_{\{i}E_jA_{k\}})^\dagger \otimes \omega + (A_{[i}E_jA_{k]})\rho(A_{[i}E_jA_{k]})^\dagger \otimes \sigma_z\omega\sigma_z\big]. \tag{9}$$

Here,

$$A_{\{i}E_jA_{k\}} := A_iE_jA_k + A_kE_jA_i, \tag{10}$$

and analogously for the commutator brackets.

From this, we can see that, like before, the ω terms survive. But more interestingly, notice that the $\sigma_z\omega\sigma_z$ terms can survive too. For example, suppose Alice and Bob measure in the z-basis and Eve measures in the x-basis, then it is possible for Alice to obtain an outcome of 0, and Bob an outcome of 1. This combination allows for $[A_0, E_\pm, A_1] \neq 0$.

We may therefore hypothesise that if Eve attempts to extract information about the state when in between Alice and Bob, she induces a nonzero $\sigma_z\omega\sigma_z$ term. So, if we were to let $\omega = |+\rangle\langle+|$ (and therefore $\sigma_z\omega\sigma_z = |-\rangle\langle-|$), if someone were to perform the measurement $\{|+\rangle\langle+|, |-\rangle\langle-|\}$ on the control qubit ω, and obtain an outcome of $-$, they could conclude that there was an eavesdropper in between Alice and Bob.

Main result

It may be shown that both correlated and uncorrelated attacks by the two possible eavesdroppers, Eve and Yves, can be detected. By taking the probability of detection P_{detect} to be the probability of the control qubit, initially in the state $\omega = |+\rangle\langle+|$, being measured to be in the state $|-\rangle\langle-|$, we can state the main result of this work.

Theorem 1. *For both correlated and uncorrelated attacks of Eve and Yves, $P_{detect} = 0$ if and only if Eve and Yves gain no information about the key shared between Alice and Bob.*

Conclusion and discussion

In the work, presented fully in [15], we have shown that, with the use of indefinite causal order, it is possible to detect eavesdroppers during a QKD task *without* publicly comparing any subset of a shared private key between the two parties involved, Alice and Bob. As far as we are aware, this differs from all other QKD protocols which require a public comparison to detect eavesdroppers. In contrast to some of these other protocols, however, there are two locations eavesdroppers can reside, allowing for correlated and uncorrelated attacks. These have both been considered and it was shown that the eavesdroppers can be detected if they extract any useful information about the shared key.

It is natural to ask whether this protocol is physically realisable, let alone practical. The difficulties lie in that ρ must go through (projective) measurement apparatuses and carry on around the loop while simultaneously doing the same in the opposite direction along the same loop. When it comes to practicality, consider using a Sagnac interferometer or something similar to create an indefinite causal ordering of operations [16]. In order for the ICO to be legitimate, the coherence length of the light used must be considerably larger than the path length of the interferometer [3], perhaps indicating a limit to how practical such a protocol would be. Another thing to notice is that, in practical QKD, in order for privacy amplification to be performed, error rates are required, which don't seem to be accessible without public comparison.

Having mentioned these limitations, the purpose of this work is not necessarily to propose a new practical protocol, rather, it is to explore an interesting new connection between indefinite causal structures and QKD. Indeed, it is a connection that yields a unique feature not observed in other QKD protocols.

References

[1] G. Chiribella et al. "Quantum computations without definite causal structure". *Phys. Rev. A* 88 (2 2013), p. 022318.

[2] O. Oreshkov, F. Costa, and Č. Brukner. "Quantum correlations with no causal order". *Nature communications* 3.1 (2012), pp. 1–8.

[3] K. Goswami et al. "Indefinite causal order in a quantum switch". *Phys. Rev. Lett.* 121.9 (2018), p. 090503.

[4] M. Araújo, F. Costa, and Č. Brukner. "Computational Advantage from Quantum-Controlled Ordering of Gates". *Phys. Rev. Lett.* 113 (25 2014), p. 250402.

[5] P. A. Guérin et al. "Exponential Communication Complexity Advantage from Quantum Superposition of the Direction of Communication". *Phys. Rev. Lett.* 117 (10 2016), p. 100502.

[6] X. Zhao, Y. Yang, and G. Chiribella. "Quantum Metrology with Indefinite Causal Order". *Phys. Rev. Lett.* 124 (19 2020), p. 190503.

[7] G. Chiribella et al. "Indefinite causal order enables perfect quantum communication with zero capacity channels". *New J. Phys.* 23.3 (2021), p. 033039.

[8] C. H. Bennett and G. Brassard. "Quantum cryptography: Public key distribution and coin tossing". *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing* (1984).

[9] A. K. Ekert. "Quantum cryptography based on Bell's theorem". *Phys. Rev. Lett.* 67 (6 1991), pp. 661–663.

[10] W-Y. Hwang. "Quantum key distribution with high loss: toward global secure communication". *Phys. Rev. Lett.* 91.5 (2003), p. 057901.

[11] V. Scarani et al. "Quantum cryptography protocols robust against photon number splitting attacks for weak laser pulse implementations". *Phys. Rev. Lett.* 92.5 (2004), p. 057901.

[12] V. Scarani et al. "The security of practical quantum key distribution". *Rev. Mod. Phys.* 81.3 (2009), p. 1301.

[13] M. Koashi and N. Imoto. "Quantum cryptography based on split transmission of one-bit information in two steps". *Phys. Rev. Lett.* 79.12 (1997), p. 2383.

[14] M. Lucamarini and S. Mancini. "Secure deterministic communication without entanglement". *Phys. Rev. Lett.* 94.14 (2005), p. 140501.

[15] H. Spencer-Wood. "Indefinite causal key distribution". *arXiv:2303.03893* (2023).

[16] T. Strömberg et al. "Demonstration of a quantum SWITCH in a Sagnac configuration". *arXiv:2211.12540* (2022).

Accurate Harmonic Vibrational Frequencies for Diatomic Molecules via Quantum Computing

Shih-Kai Chou[1] Jyh-Pin Chou[2 7] Alice Hu[3 4] Yuan-Chung Cheng[5 6 7]
Hsi-Sheng Goan[1 6 7 *]

[1] *Department of Physics and Center for Theoretical Physics, National Taiwan University, Taipei 10617, Taiwan*
[2] *Department of Physics, National Changhua University of Education, Changhua 50007, Taiwan*
[3] *Department of Mechanical Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR 999077, China*
[4] *Department of Materials Science and Engineering, City University of Hong Kong, Kowloon, Hong Kong SAR 999077, China*
[5] *Department of Chemistry, National Taiwan University, Taipei 10617, Taiwan*
[6] *Center for Quantum Science and Engineering, National Taiwan University, Taipei 10617, Taiwan*
[7] *Physics Division, National Center for Theoretical Sciences, Taipei, 10617, Taiwan*

Abstract. During the noisy intermediate-scale quantum (NISQ) era, quantum computational approaches refined to overcome the challenge of limited quantum resources are highly valuable. A comprehensive benchmark for a quantum computational approach in this spirit could provide insights toward further improvements. In this work, we present such an investigation by benchmarking harmonic vibrational frequencies for 43 diatomic molecules via the variational quantum eigensolver algorithm. Using the accurate Hamiltonian constructed from Kohn-Sham orbitals expanded in the Daubechies wavelet basis set, we show that the results from the exact diagonalization of the corresponding qubit Hamiltonian in great agreement with the experimental data, and the chemistry-inspired UCCSD ansatz could achieve the same accuracy except for systems whose Mayer bond order indices are larger than 2. For those systems, we then demonstrate that the heuristic hardware-efficient RealAmplitudes ansatz can provide significant improvement over the UCCSD ansatz, verifying that the harmonic vibrational frequencies could be calculated accurately by quantum computation in the NISQ era.

Keywords: quantum computing, quantum chemistry, VQE, Daubechies wavelet, vibrational frequency

1 Introduction

Quantum computation of quantum chemistry has been considered as a promising application of quantum computing [1, 2]. With the quantum nature of wavefunctions, quantum computing makes use of superposition, entanglement and interference to prepare and manipulate quantum states, offering the potential for exponential speedup over classical computing. The method of the unitary coupled cluster with single and double excitations (UCCSD) can be efficiently implemented on a quantum computer [3, 4], making UCCSD a powerful tool for quantum computing on chemical systems.

In the noisy intermediate-scale quantum (NISQ) era [5], quantum computers with a limited number of qubits are noisy without error correction. As a result, the number of consecutive quantum gates that can be reliably run on the NISQ machines is also restricted. Therefore, reducing the requirement on the number of qubits and thus the depth of the quantum circuit is one of the major strategies in the NISQ era. To address this challenge, a hybrid quantum-classical algorithm called the variational quantum eigensolver (VQE) [3] has been proposed and widely used. However, the circuit depths of VQE UCCSD are still too deep for current NISQ devices. Therefore, a heuristic hardware-efficient ansatz is proposed to take advantage of its shorter circuit depth than that of the chemistry-inspired UCCSD ansatz on NISQ devices [6].

Besides the variational ansatz encoded in the trial wavefunction, high-quality representation of the molecular Hamiltonian is essential for accurate predictions of chemical properties, which was often overlooked in the quantum computing community. Recently, Hong et al. [7] demonstrated that a minimal basis (MB) set constructed from Daubechies wavelet (DW) molecule orbitals (MOs) basis calculated from BigDFT [8] can yield accurate results in harmonic vibrational frequencies for H_2, LiH, and BeH_2 on quantum simulator with noisy model implemented from the real devices. That is, VQE quantum computations with accuracy comparable with that of the full configuration interaction calculation using the cc-pVDZ basis set, whereas the computational cost the same as that of a STO-3G calculation, has been achieved for this small set of molecules. It is necessary at this moment to carry out a benchmark study in order to evaluate the true performances of quantum algorithms and to point out the possible directions of improvements. Such benchmark should be comprehensive and able to compare with the experimental data so that the results may be extended to more general situations. Here, we propose to represent the Hamiltonian using DW MOs with different exchange-correlation (XC) functionals, where an active space based on the MP2 energy criterion is selected to further reduce the required number of qubits. We perform the VQE benchmark on harmonic vibrational frequencies for 43 diatomic molecules with results in great agreement with the experimental data. We find that the approach of using the MB set of DW MOs [7] does not provide adequate results of vibrational fre-

*goan@phys.ntu.edu.tw

quencies by always giving smooth potential energy curves for all the 43 molecules, and also has significantly larger root-mean-square deviation (RMSD) values, even though the number of active MOs used is considerably higher than our approach. We attribute the great performance of our approach to three factors: (i) a better description of the Hamiltonian by introducing the DW MOs, (ii) incorporating the electron correlation effect into the MOs via the XC functional of PBE0, (iii) a suitable selection of active space by MP2.

2 Results

2.1 Accuracy of the represented Hamiltonian

The performance of the proposed approach is benchmarked against the classical CASCI and CCSD(T) methods in the cc-pVDZ basis set on 43 diatomic harmonic vibrational frequencies, where RMSDs are presented in Figure 1. The dataset considered consists of diatomic molecules that are neutral, closed-shell and formed by row 1 to row 4 atoms, excluding the transition metal elements, K atom, some molecules containing Na atom (no smooth curve calculated from BigDFT), and those whose experimental data are unavailable. Besides, C_2 and F_2 owing to severe static correlation are also excluded. Our notation reads "Method[active space selection method]-XC/Basis Set", and EDQC denotes the exact diagonalization method of quantum computing (after the qubit Hamiltonian is constructed) whose results are regarded as the best results achievable by quantum computation. As clearly seen from Figure 1, our proposed approach

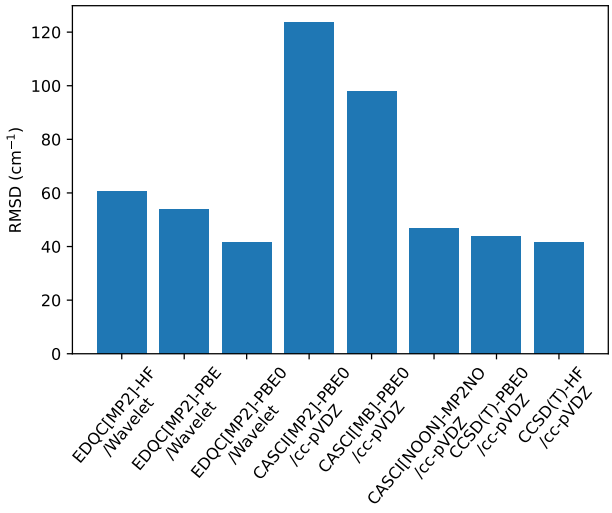


Figure 1: RMSDs of the harmonic vibrational frequencies (in cm^{-1}) obtained by comparing the results to their corresponding experimental values for a variety of methods, where [MB] stands for imitating the size of minimal basis set, NO and NOON for nature orbital and occupation number, and MP2NO sources from MP2 wavefunction.

of EDQC[MP2]-PBE0/Wavelet has the best performance among all the methods.

2.2 VQE(UCCSD) Benchmark

After equipped with the [MP2]-PBE0/Wavelt Hamiltonian, we present the UCCSD calculations using VQE, denoted as VQE(UCCSD), in Figure 2 (see the blue points). The results show that VQE(UCCSD) can be as accurate as the exact diagonalization except for the BeO family, the CO family, and some of the N_2 family. Previous study [9] showed that for systems with strongly correlated electrons, UCCSD would not give results achieving chemical accuracy even in the region near the equilibrium point. In strongly correlated systems, the states resulting from the action of the UCCSD cluster operators that include only single and double excitations might not encompass all those important configurations. We choose the Mayer bond order [10], a good electron correlation descriptor applicable to strongly correlated systems, calculated by DFT-PBE0/cc-pVDZ to present the relation with the harmonic vibrational frequencies calculated by VQE(UCCSD) in Figure 2. The results show that systems for which UCCSD does not yield accurate results have the Mayer bond order indices > 2.

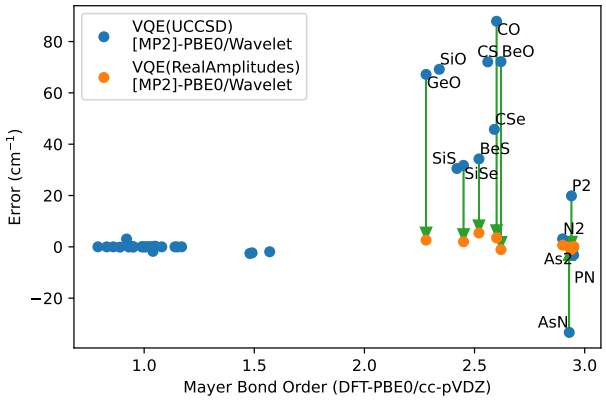


Figure 2: Mayer bond order indices calculated by DFT-PBE0/cc-pVDZ versus the error (difference) in the harmonic vibrational frequencies (in cm^{-1}) calculated by VQE(UCCSD)[MP2]-PBE0/Wavelet with respect to those by EDQC[MP2]-PBE0/Wavelet for diatomics in the benchmark dataset. The orange dots denote the vibrational frequency results for the specified molecules calculated by VQE(RealAmplitudes)[MP2]-PBE0/Wavelet. The relation between blue and orange dots is from Table 1, and the green arrows point toward the directions of improvement from UCCSD to RealAmplitudes.

2.3 VQE(UCCSD) VS VQE(RealAmplitudes)

For those systems whose Mayer bond order indices are larger than 2, we then consider a heuristic hardware-efficient ansatz, the RealAmplitudes ansatz implemented in Qiskit [11], since it can go beyond the restriction of the accessible Hilbert space of the chemistry-inspired UCCSD ansatz. In Table 1, the results of the harmonic vibrational frequencies between UCCSD and RealAmplitudes with linear entanglement using SLSQP and L-BFGS-B optimizers of SciPy [12] are compared. Despite having shallower circuit depths, RealAmplitudes could

Table 1: Comparisons of the harmonic vibrational frequencies and the relevant circuit information between the VQE(UCCSD) and VQE(RealAmplitudes) calculations using the Hamiltonian in the [MP2]-PBE0/Wavelet approach for the systems whose Mayer bond order indices are larger than 2 with the required number of qubits up to 10. Here the indices $[n;m]$ denote n electrons and m MOs (2m spin orbitals) in the active space for the molecule, and with parity encoding the number of qubits used is $2m-2$. The value inside the parenthesis in the harmonic vibrational frequency denotes the difference between VQE and EDQC.

	VQE(UCCSD)			State	VQE(RealAmplitudes)	Dep[a]		State
Mol.	[MP2]-PBE0/Wavelet	Dep[a]	Nθ[b]	Fidelity[c]	[MP2]-PBE0/Wavelet	(Rep[d])	Nθ[b]	Fidelity[c]
BeO	[6,6]1,508.87 (72.15)	8,460	92	0.98634	[6,6]1,435.63 (-1.09)	99(30)	310	0.9990823
BeS	[6,6] 983.82 (34.27)	8,460	92	0.99086	[6,6] 955.98 (5.39)	99(30)	310	0.9981542
CO	[8,6]2,336.76 (87.90)	8,460	92	0.99729	[8,6]2,249.18 (3.49)	84(25)	260	0.9999526
SiSe	[6,6] 613.98 (31.75)	8,460	92	0.99515	[6,6] 584.19 (1.96)	99(30)	310	0.9996153
GeO	[6,6]1,066.79 (67.18)	8,460	92	0.99068	[6,6]1,012.25 (2.59)	84(25)	260	0.9988251
N_2	[4,4]2,402.52 (3.10)	1,480	26	0.99991	[4,4]2,400.07 (0.65)	35(10)	66	0.9999982
PN	[4,4]1,372.77 (-3.21)	1,480	26	0.99775	[4,4]1,376.12 (0.14)	35(10)	66	0.9999997
P_2	[4,4] 798.40 (19.87)	1,480	26	0.99988	[4,4] 777.54 (-0.99)	35(10)	66	0.9999974
AsN	[4,4]1,055.19 (-33.35)	1,480	26	0.99785	[4,4]1,088.34 (-0.20)	35(10)	66	0.9999982
As_2	[4,4] 414.10 (-0.18)	1,480	26	0.99971	[4,4] 414.31 (0.03)	35(10)	66	0.9999996

[a] Dep denotes the circuit depths.
[b] Nθ denotes the number of tunable circuit parameters.
[c] State fidelity denotes the average state fidelity over points employed to calculate the vibrational frequency.
[d] Rep denotes the number of repetitions of the unit pattern circuit.

still achieve higher state fidelities than UCCSD, and for the cases with a smaller number of qubits outstanding performance can be achieved more easily. This is a clear indication that a heuristic hardware-efficient quantum circuit can span a state space larger than that spanned by UCCSD. To the best of our knowledge, our investigation is the first systematical benchmark study to demonstrate that a heuristic hardware-efficient ansatz could outperform a chemistry-inspired UCCSD ansatz in predicting accurate molecular properties by quantum computation.

3 Conclusion

We propose a quantum computational approach that combines DW MOs with the XC functional and an optimal active space determined by MP2 energy criterion, resulting in a significantly reduced qubit number requirement while maintaining excellent accuracy. Our calculations show that a quantum computer capable of carrying out calculations on 10 qubits with circuit depth < 100 can accurately predict the vibrational frequencies of neutral closed-shell diatomic molecules, and these quantum resource requirements should be able to be achieved on near-term NISQ devices (e.g., IBM 100×100 Challenge in 2024). Our benchmark investigation here provides a critical assessment on the power of quantum computation of molecular properties and insights on further improvements.

References

[1] Y. Cao, et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.*, 119(19):10856–10915, 2019.

[2] S. McArdle, et al. Quantum computational chemistry. *Rev. Mod. Phys.*, 92(1):015003, 2020.

[3] A. Peruzzo, et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.*, 5(1):4213, 2014.

[4] M.-H. Yung, et al. From transistor to trapped-ion computers for quantum chemistry. *Sci. Reports*, 4(1):1–7, 2014.

[5] J. Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[6] A. Kandala, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, 2017.

[7] C.-L. Hong, et al. Accurate and efficient quantum computations of molecular properties using daubechies wavelet molecular orbitals: A benchmark study against experimental data. *PRX Quantum*, 3(2):020360, 2022.

[8] L. Genovese, et al. Daubechies wavelets for high performance electronic structure calculations: The bigdft project. *Comptes Rendus Mécanique*, 339(2-3):149–164, 2011.

[9] J. Lee, et al. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.*, 15(1):311–324, 2018.

[10] I. Mayer. Charge, bond order and valence in the ab initio scf theory. *Chem. Phys. Lett.*, 97(3):270–274, 1983.

[11] M. S. ANIS, et al. Qiskit: An open-source framework for quantum computing, 2021.

[12] P. Virtanen, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods*, 17:261–272, 2020.

A Protein-Folding Entangler for the Variational Quantum Eigensolver

Samanvay Sharma, AQUA Group, Master's Program, Keio University Shonan Fujisawa Campus
Mentors: Rodney Van Meter, Takahiko Satoh, Hiroyuki Kusumoto, Atsushi Matsuo

Problem Definition / Abstract

The Variational Quantum Eigensolver (VQE) algorithms are one of the most popular solution approaches for quantum optimization problems in the era of Noisy-Intermediate State Quantum (NISQ) devices, but there are certain limitations with the toolset and software libraries used for creating initial circuit states that rarely consider constraints specific to the problem itself. Our objective is to create circuit libraries for an open-source quantum software development tool called Qiskit that will take into consideration the different types of constraints for different problem cases when creating an initial circuit guess for our VQE algorithm, and improve upon existing performance benchmarks. This project will eventually allow users to improve performance for their own optimization problems ranging from simulating complex molecules to solving graph and network problems using VQE. For this research undertaking, we will solve a simple model of the protein folding problem, which is a highly complex computational task and has major applications in biochemical fields such as material discovery and designing targeted molecules. By utilizing these tools, we will be able to accurately calculate important information about a desired protein system such as accurate energy values and reduced computation time.

Details / Background

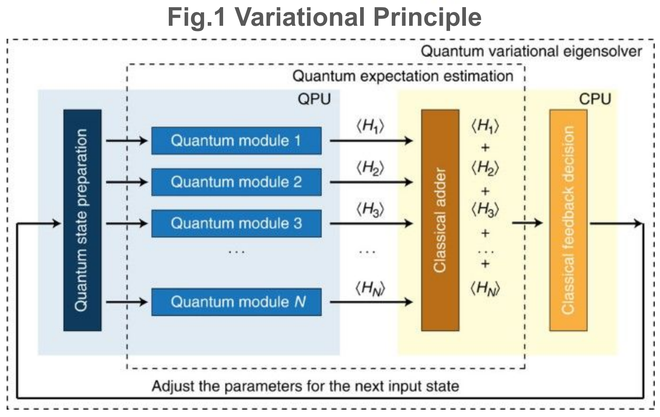
Variational Quantum Eigensolver (VQE):

Variational principle/method states that the expectation value (probability) of any wave function (our initial guess of the solution) will always greater or equal to the minimum eigenvalue (solutions of system) associated with its Hamiltonian, which is a matrix that describes the possible energies of a physical system. To give a brief example of how we use this, we can select a random circuit which represents a state for our selected system, described as an eigenstate "psi", and we write down a Hamiltonian matrix which describes our selected system, and calculating the probability values of such a system will get us closer and closer to the actual value of the system.

The Variational Quantum Eigensolver algorithm is a quantum computing algorithm based on this principle that helps to estimate the ground state energy of a given quantum mechanical system. This is well suited for solving certain classes of optimization problems.
The basic idea is simply a feedback loop that starts with an initial guess for the eigenstate, measures the differences in the resulting expectation values, and adjusts the initial guess until it reaches a saturation point or minima.

$$\lambda_{min} \leq \langle H \rangle_\psi = \langle \psi | H | \psi \rangle = \sum_{i=1}^{N} \lambda_i |\langle \psi_i | \psi \rangle|^2$$

Fig.1 Variational Principle



[1] Fig.2 VQE Architecture

Entanglers / Variational Forms / Parametrized Quantum Circuits (PQCs):

We optimize our initial state or "ansatz" by using a parametrized circuit with a fixed form, interchangeably called entanglers or variational forms, which is applied to our initial state and generate an output state which on further iteration will give an expectation value close to the minimum eigenvalue.

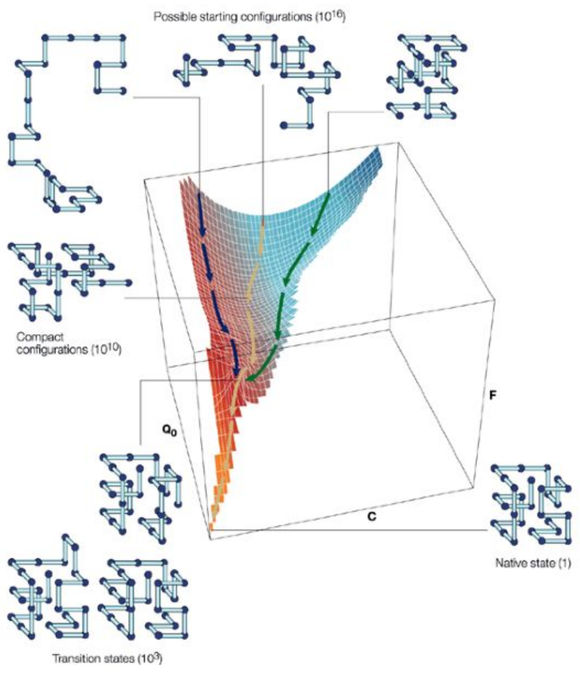
For this project we use Qiskit, which is an open-source quantum development platform designed by IBM Quantum primarily utilizing the Python language and libraries to run code and algorithms on real quantum hardware and simulators. We will be collaborating with IBM Quantum services and the Qiskit toolkit to make additions to the Qiskit Applications modules. [2]

Limitations:

(1) Very few types of entangler circuits are known or available. Even the state-of-the-art library for quantum computers has only four types of entangler circuits such as Ry, RyRz, SwapRz and UCCSD. They are all general entangler circuits with static structures and are used generically for most problems.
(2) Existing entangler circuits do not take into account the feasibility of output answers, and they often output infeasible answers. Results must be feasible answers of corresponding optimization problems when using the VQE algorithm for optimization problems.

Protein-Folding:

Proteins "fold" in nature to their native 3D conformation which enables them to become biologically functional. The structure and functions of many proteins are still not well understood, which means that processes that make use of the knowledge of proteins such as material discovery and designing targeted molecules for applications such as vaccinations, is also heavily underdeveloped.
This is a challenge because unfolded proteins have a very high degree of freedom and thus an enormous number of potential configurations in 3D space which grows exponentially upto the order of 10^{47}, which makes them a computationally intensive task for classical devices alone.
Using quantum computers however, we can solve this problem linearly as N^4 in the number of our Hamiltonian terms, where the number of protein molecules is N, thus providing with an exponential speed-up. [3] The number of qubits required to map these conformations scales quadratically.



[4] Fig.3 Energy landscape for protein folding

Key Idea

We initiate by writing an entangler library for problem-specific VQE circuits as described in [5] to improve upon the limitations of current systems defined earlier. The overall vision is to create multiple libraries for different kinds of constraints and for different problems, and combine them into a single circuit library.

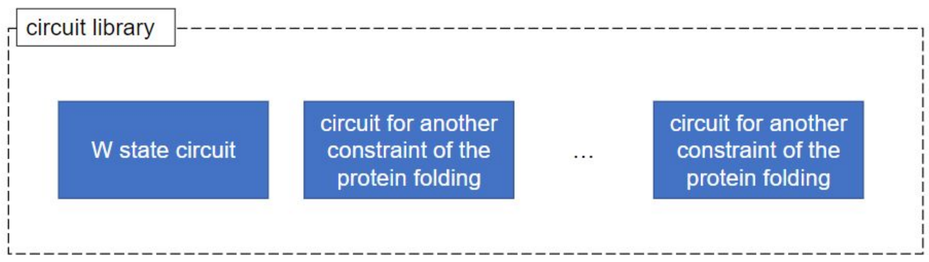
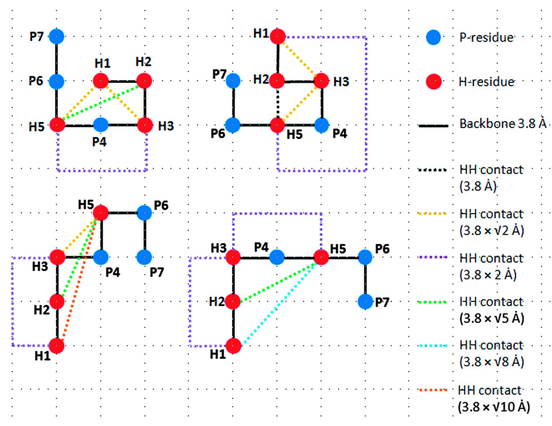


Fig.4 Circuit library schematic for protein folding class

For our project we want to solve the protein folding problem for a very simple protein chain. We start by encoding the different constraints for our sample Hamiltonian. Two of the constraints that we identify are configuration or geometrical constraints, which describe the relative positions of each molecule in the chain and governs its growth, and interaction constraints, which represents energy values that emerge due to electronic interactions between neighbouring molecules in a chain. We will also want to encode penalty terms for representing physical constraints such as local overlap between molecules in a chain, avoiding redundant rotations over an axis, and chirality terms. See Fig. 5 for an example of how these constraints might look like in 2D space.



[6] Fig.5 Scheme representation of 2D hydrophobic models

We then start preparing our PQC class for this protein folding problem by incorporating multi-qubit entanglement schemes to represent our constraints. One such method is to use the W-state circuits which is a superposition of states where exactly one qubit is in state |1> while all others are in state |0>, i.e. hamming weight is 1. Finally we create multiple qubit registers to accommodate for different Hamiltonian constraints and deploying multi-qubit entanglement for each set of identified constraints.
Once the ProblemSpecificVQE class is prepared, we pass a QuadraticProgram instance to this class. It will make an ansatz for our VQE algorithm based on constraints defined above, using circuits in the circuit library and checking for each constraint. If it matches to the specific pattern, it can then call for corresponding circuits in the circuit library.

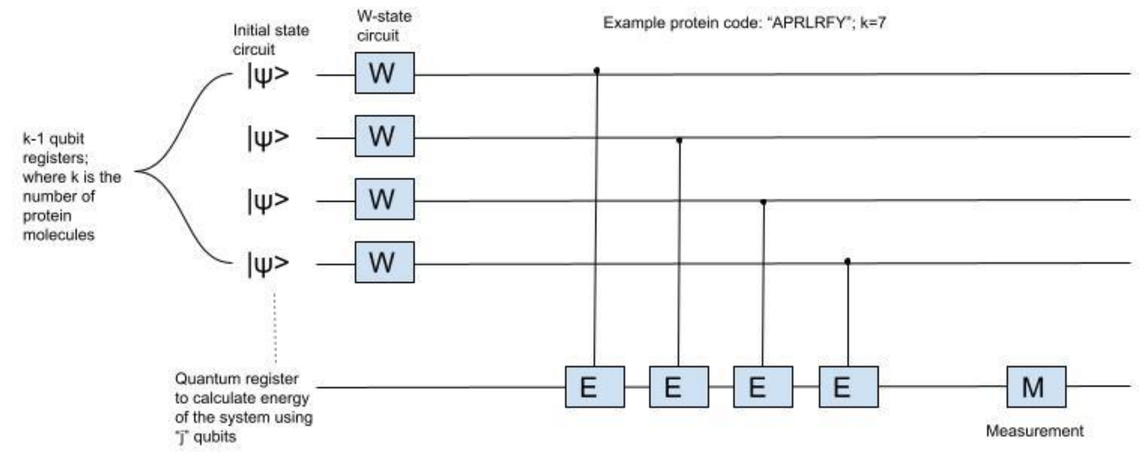


Fig.6 Initial circuit schematic for ProblemSpecificVQE_class using parameterized W states

Expected Results

Once we complete all steps involving designing code, merging libraries with source, and running simulations over different hardware to benchmark our results against existing solutions, we expect to achieve two milestones. One is to have initiated a circuit library that creates dynamic entanglers to reflect the constraints of a simple protein folding model, which in the future can allow additions to the overall library itself to provide other users the flexibility to create improved solutions for their own optimization problems using similar approaches. The other one would be in the simulation results themselves where we expect lower conformational energy values of the system, indicating improved accuracy of the algorithm, and lower optimization time, indicating overall computational speed-up.

[1] Peruzzo, A., McClean, J., Shadbolt, P. et al. A variational eigenvalue solver on a photonic quantum processor. Nat. Commun. 5, 4213 (2014).
[2] Qiskit. Qiskit: An open-source framework for quantum computing. https://www.qiskit.org/.
[3] Robert, A., Barkoutsos, P.K., Woerner, S. et al. Resource-efficient quantum algorithm for protein folding. npj Quantum Inf 7, 38 (2021).
[4] Dobson, C. Protein Folding and Disease: a view from the first Horizon Symposium. Nat Rev Drug Discov 2, 154–160 (2003).
[5] Matsuo, A., Yudai S. and Shigeru Y., Problem-specific parameterized quantum circuits of the VQE algorithm for optimization problems. arXiv preprint arXiv:2006.05643 (2020).
[6] Onofrio, Angelo, et al. Distance-dependent hydrophobic–hydrophobic contacts in protein folding simulations. Physical Chemistry Chemical Physics 16.35 (2014): 18907-18917.

Locally unidentifiable set of quantum states as resource for secret password distribution

Pratik Ghosal[1] Arkaprabha Ghosal[2] Subhendu B. Ghosh[3] Amit Mukherjee[4]

[1] *Department of Physics, Bose Institute, Unified Academic Campus, EN 80, Sector V, Bidhan Nagar, Kolkata 700 091, India*
[2] *Optics and Quantum Information Group, The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai 600113, India*
[3] *Physics and Applied Mathematics Unit, Indian Statistical Institute, Kolkata, 203 B. T. Road, Kolkata 700108, India*
[4] *Department of Physics, Bose Institute, Unified Academic Campus, EN 80, Sector V, Bidhan Nagar, Kolkata 700 091, India*

Abstract. Nonlocality makes quantum theory nontrivially sacred and useful in the information processing paradigm. Aside from Bell nonlocality, there is a different kind of quantum nonlocality which is associated with indistinguishability of orthogonal multipartite quantum states by local operations and classical communication (LOCC). Based on such nonlocality, we propose a distributed task, namely, *local subset identification* that calls for identification of subsets of a known set of orthogonal multipartite states by spatially separated agents using LOCC. Failure to accomplish this task results in a previously unexplored form of quantum nonlocality, called *local subset unidentifiability*. This is a stronger quantum nonlocality than its predecessors, such as local indistinguishability and local unmarkability. Further, we present a multipartite version of this nonlocality, namely, *genuine unidenifiability* - where a set of states remain unidentifiable unless all the agents come together and perform global measurement. We also demonstrate an intriguing application of this nonlocality in a certain *secret password distribution protocol*, where this form of nonlocality serves as a more useful resource than its predecessors.

Keywords: Quantum communication, Quantum protocols, Nonlocality, Quantum information processing

1 Introduction

Recently, stronger forms of quantum nonlocality, such as local state irreducibility and local unmarkability, have been developed in addition to local indistinguishability [1, 3, 2]. These nonlocalities arise from the impossibility of locally realizing certain distributed tasks. Two paradigms exist for these tasks: (i) local state discrimination and (ii) local state elimination. The former includes examples like local indistinguishability and unmarkability [7], while the latter is represented by local state irreducibility [2]. This work focuses on the first paradigm and introduces an even stronger form of quantum nonlocality found in sets of mutually orthogonal entangled states. To demonstrate this, a distributed task called Local Subset Identification (LSI) is introduced. In LSI, multiple spatially separated agents share more than one state chosen from a known set of mutually orthogonal states. The objective is to identify these shared states perfectly using LOCC (Local Operations and Classical Communication). This work shows that the incapability of accomplishing this task exhibits a stronger nonlocality compared to existing literature. For instance, considering a set of bipartite orthogonal states, if any two states from the set are shared between spatially separated agents, LSI requires the agents to locally recognize the identity of the given states perfectly, i.e., identify which two states they were given using LOCC. The inability to accomplish this task perfectly demonstrates a form of quantum nonlocality known as Local Subset Unidentifiability. This proposed nonlocality arises from the inability to perfectly distinguish certain sets of subspaces of rank more than one using LOCC. Since addressing the (in)distinguishability of sets of subspaces is inherently more complex than that of sets of vectors, this proposed nonlocality stands out compared to previous nonlocalities. Furthermore, this work also presents an information processing application of the proposed nonlocality.

Any nonlocal feature of quantum systems gets more intricate when multipartite scenario comes into the picture. In case of LSI, we also explore scenarios involving more than two spatially separated agents. Interestingly, we come up with a further stronger version of the nonlocality we are introducing here. In particular, we present sets of multipartite states that shows local subset unidentifiability (or *locally unidentifiable*) when all the agents are spatially separated. We show that these sets retain local unidentifiablity in all possible bi-partitions. Therefore, to perfectly accomplish the LSI task, *all* the agents need to come together or must resort to additional quantum resources. We term it as *Genuine Unidentifiablity*. We also illustrate that any set that is genuinely locally unidentifible must also show *genuine* unmarkability which is hitherto a uncharted notion.

Definition 1 ($(n, \mathcal{S}')$-Local Subset Identification) *Consider a set $\mathcal{S} = \{|\psi_1\rangle, |\psi_2\rangle, \ldots, |\psi_n\rangle\} \subset \bigotimes_{k=1}^{N} \mathbb{C}^{d_k}$ of n N-party orthogonal quantum states. A subset $\mathcal{S}' \subset \mathcal{S}$ containing $1 \leq \mathcal{S}' < n$ quantum states is randomly chosen and distributed among spatially separated agents keeping its identity hidden. The task of $(n, \mathcal{S}')$-local identification is to perfectly determine the elements of the set $\mathcal{S}'$.*

Observation 1 *For a given set $\mathcal{S}$ the task of $(n, 1)$-LSI cor-*

responds to the well-known task of Local State Discrimination (LSD) [?].

Observation 2 *If a set $\mathcal{S}$ is (n, m)-markable, then it readily follows that it is also (n, m)-identifiable.*

2 Theorems

Theorem 2 *If a set $\mathcal{S}$ is (n, m)-identifiable, then it is not necessarily (n, m)-markable.*

Theorem 3 *The set of four two-qubit maximally entangled states is (4,2)-unidentifiable.*

Theorem 4 *Consider a complete basis set $\mathcal{S}$ of maximally entangled states (MES) in $\mathbb{C}^d \otimes \mathbb{C}^d$. There are $\binom{d^2}{k}$ possible subsets, each containing k distinct states from $\mathcal{S}$. The set $\mathcal{S}$ is (d^2, k)-unidentifiable, if $\binom{d^2}{k} > d^k$. Moreover, the set of D maximally entangled states ($D < d^2$) will also be (D, k)-unidentifiable, provided $\binom{D}{k} > d^k$.*

Genuine tripartite unidentifiability: Consider a set of eight 3-qubit GHZ states, $\mathcal{S} := \{\frac{1}{\sqrt{2}}(|000\rangle \pm |111\rangle), \frac{1}{\sqrt{2}}(|001\rangle \pm |110\rangle), \frac{1}{\sqrt{2}}(|010\rangle \pm |101\rangle), \frac{1}{\sqrt{2}}(|100\rangle \pm |011\rangle)\}$.

Theorem 5 *The set $\mathcal{S}$ is locally (D, k)-unidentifiable if $\binom{D}{k} > 2^{2k}$ (where $D \in (2, 8]$ and $k \in (1, D)$) even when two out of three parties come together in a same lab (collaborate).*

Genuine four party unidentifiability: Consider the set $\mathcal{S} := \{|\Omega_\alpha\rangle\}_{\alpha=1}^{16}$, where

$$
\begin{aligned}
|\Omega_{1,2}\rangle &= \tfrac{1}{2}(|0000\rangle \pm |0111\rangle + |1010\rangle \pm |1101\rangle),\\
|\Omega_{3,4}\rangle &= \tfrac{1}{2}(|0000\rangle \pm |0111\rangle - |1010\rangle \mp |1101\rangle),\\
|\Omega_{5,6}\rangle &= \tfrac{1}{2}(|0001\rangle \pm |0110\rangle + |1011\rangle \pm |1100\rangle),\\
|\Omega_{7,8}\rangle &= \tfrac{1}{2}(|0001\rangle \pm |0110\rangle - |1011\rangle \mp |1100\rangle),\\
|\Omega_{9,10}\rangle &= \tfrac{1}{2}(|0010\rangle \pm |0101\rangle + |1000\rangle \pm |1111\rangle),\\
|\Omega_{11,12}\rangle &= \tfrac{1}{2}(|0010\rangle \pm |0101\rangle - |1000\rangle \mp |1111\rangle),\\
|\Omega_{13,14}\rangle &= \tfrac{1}{2}(|0011\rangle \pm |0100\rangle + |1001\rangle \pm |1110\rangle),\\
|\Omega_{15,16}\rangle &= \tfrac{1}{2}(|0011\rangle \pm |0100\rangle - |1001\rangle \mp |1110\rangle).
\end{aligned}
$$

Theorem 6 *The set $\mathcal{S}$ is locally (D, k)-unidentifiable if $\binom{D}{k} > 2^{3k}$ (where $D \in (2, 16]$ and $k \in (1, D)$) in all $1:3$ as well $2:2$ bi-partitions.*

Remark 1 *Theorem 5 and Theorem 6 implies necessary and sufficient condition genuine tripartite and four party local unidentifiability for the given set $\mathcal{S}$.*

3 Application in secret password distribution

Suppose a sender wants to distribute some hidden information (*a locked password, for example*) among several spatially separated receivers, who can send classical information among themselves. The sender desires that the receivers cannot know the identity of the password

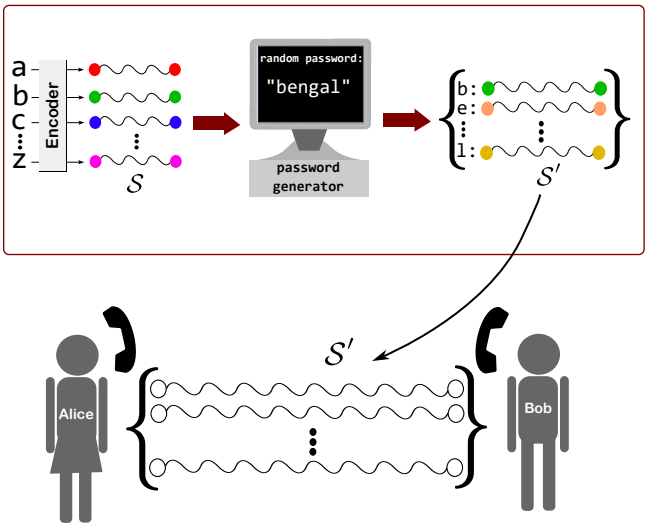


Figure 1: A random password-string of m letters is generated from a known alphabet of $n > m$ letters, to be distributed among spatially separated parties. The sender wishes to keep the identity of the password hidden unless all the parties physically come together in a common location. If the letters of the alphabet are encoded into a set of multipartite orthogonal states such that the set is locally (n, m)-unidentifiable, then the spatially separated parties cannot locally reveal the hidden password (not even the letters) with certainty. This provides a tighter security condition than encoding the letters into a locally unmarkable set, where the receivers may locally predict the identity of each letter in the password with certainty.

as long as they are spatially separated. Precisely, consider that the sender wish to share a password – a string $\mathcal{X} := x_1 x_2 \cdots x_m$ of m letters, x_i being the ith letter in the string – among the receivers. Each letter in the string is to be chosen without repetition from an alphabet $\mathcal{A} = \{a_k\}_{k=1}^n$ of $n > m$ letters which is known to the receivers as well. Now, the sender and the receivers agree upon an encoding scheme: the letters of the alphabet $\mathcal{A}$ are encoded in a set $\mathcal{S} := \{|\psi_k\rangle\}_{k=1}^n$ of pairwise orthogonal pure multipartite quantum states. Accordingly, the sender encodes their password $\mathcal{X}$ into a string of quantum states: $\mathcal{X} \mapsto |\xi_1\rangle \otimes |\xi_2\rangle \otimes \cdots \otimes |\xi_m\rangle$, where the state $|\xi_i\rangle$ can be any state from $\mathcal{S}$ with the only restriction that $|\xi_i\rangle \neq |\xi_j\rangle, \forall i, j$. Subsequently, the sender shares this composite state among the receivers (see Fig. 1). If $\mathcal{S}$ is locally (n, m)-unmarkable, then the spatially separated parties will not be able to perfectly discriminate the received string from the ${}^nP_m = \frac{n!}{(n-m)!}$ possible strings of quantum states by LOCC. However, it may so happen that the receivers can perfectly predict the identity of the m individual states $\{|\xi_i\rangle\}$ if the encoding is done in a locally $(n, m)-$unmarkable, but locally $(n, m)-$identifiable set of quantum states. Then, they will be able to guess the correct permutation of the letters with success probability $\frac{1}{m!}$. Another alternative which

they may opt for is to imperfectly discriminate (*i.e.*, with a nonzero probability $P_{imp} < 1$, bounded by an upper limit discussed in [**?**]) the received string. The encoding set $\mathcal{S}$ determines which of the success probabilities is higher.

Nevertheless, if the sender encodes the password in a locally (n, m)-unidentifiable set of states, then the receivers will not be able to even identify the individual letters perfectly by LOCC, and hence they will not follow the former strategy. Furthermore, if, for such sets, $P_{imp} < \frac{1}{m!}$, then the security of the password is enhanced significantly. We have found examples of $(n, m)-$unidentifiable sets, $\mathcal{S}_d \subset \mathbb{C}^d \otimes \mathbb{C}^d$ containing maximally entangled states for which $P_{imp} < \frac{1}{m!}$ (follows from Theorem 4) as long as $|\mathcal{S}_d| \geq d + 1$.

4 Discussion

We come up with a new distributed task – LSI. We show that impossibility of accomplishing this task gives birth to a unique version of quantum nonlocality – local subset unidentifiability. We show that this is by-far the strongest quantum nonlocality in the state discrimination paradigm, that arises from the impossibility of discriminating certain mutually orthogonal subspaces of rank more than one. In the multipartite framework, we introduce the notion of genuine unidentifiability which says that a set of quantum states may remain locally unidentifiable even in all possible bi-partitions. Along this line, we also introduce the notion of genuine unmarkability in multipartite scenarios. Interestingly, we also propose a cryptographic application of this proposed nonlocality. In secret password distribution scheme, we demonstrate that local unidentifiability provides a strictly better encoding for protecting password secrecy than its predecessors. While we explore local unidentifiability only in entangled states, we believe it is not necessarily an exclusive characteristic of entanglement. Exploring the same feature in orthogonal product states would be quite intriguing.

This paper is available in *arXiv:2209.10954v2* [*quant-ph*].

References

[1] A. Peres. Optimal detection of quantum information. In *Phys. Rev. Lett. 66, 1119–1122 (1991).*

[2] Saronath Halder, Manik Banik, Sristy Agrawal, and Somshubhro Bandyopadhyay, "Strong quantum nonlo- cality without entanglement," *Physical review letters 122, 040403 (2019).*

[3] Charles H Bennett, David P DiVincenzo, Christopher A Fuchs, Tal Mor, Eric Rains, Peter W Shor, John A Smolin, and William K Wootters, "Quantum nonlocality without entanglement," *Physical Review A 59, 1070 (1999).*

[4] Jonathan Walgate, Anthony J. Short, Lucien Hardy, and Vlatko Vedral, "Local distinguishability of multipartite orthogonal quantum states," *Phys. Rev. Lett. 85, 4972– 4975 (2000).*

[5] Michał Horodecki, Aditi Sen(De), Ujjwal Sen, and Karol Horodecki, "Local indistinguishability: More nonlocal- ity with less entanglement," *Phys. Rev. Lett. 90, 047902 (2003).*

[6] Samrat Sen, Edwin Peter Lobo, Sahil Gopalkrishna Naik, Ram Krishna Patra, Tathagata Gupta, Subhendu B Ghosh, Sutapa Saha, Mir Alimuddin, Tamal Guha, Some Sankar Bhattacharya, et al., "Local quantum state marking," *Physical Review A 105, 032407 (2022).*

[7] H. J. Kimble, "The quantum internet," *Nature 453, 1023– 1030 (2008).*

Fiber-based NIR entanglement distribution for short quantum communication networks

Anindya Banerji[1] * Rui Wang[1] Alexander Ling[1 2]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore*
[2] *Department of Physics, National University of Singapore, 2 Science Drive 3, 117551, Singapore*

Abstract. For deployable quantum communication in urban environments, the quantum channel in the form of telecommunication optical fiber (confirming to ITU G.652D standards) are available, but practical semiconductor-based detectors in this range typically have low efficiency. We demonstrate polarization entanglement distribution using non-degenerate entangled photon pairs of wavelength in the near infrared through standard telecommunication fiber. This technique benefits from the high efficiency of the NIR single photon detectors and the mature design of setups around these wavelengths. Here we obtain high quality entanglement after an overall distance of 12 km, corresponding to about -36 dB of fiber induced loss.

Keywords: Entanglement, near infrared, quantum communication, quantum network

1 Introduction

Quantum entanglement has emerged as the pivotal concept in a variety of applications, such as quantum teleportation[1], quantum key distribution (QKD)[2], quantum metrology[3], distributed quantum sensing and quantum computation. It is also going to play an important role in the emerging field of quantum networks which would involve entanglement distribution across the network and the teleportation of quantum states between nodes[4, 5]. In all these applications of entanglement, the quality of entanglement being distributed over a network will have a significant impact.

The choice of wavelength is a significant consideration when integrating high efficiency entangled photon sources with existing optical fiber communication infrastructure that is designed to benefit from the low-loss transmission window around 1550 nm. However, existing semiconductor detectors like InGaAs single photon detectors designed for these wavelengths suffer from lower efficiency, a higher dark count rate and longer effective dead time.

It might be possible to perform entanglement distribution with photons generated in the near infrared region when the use-case is for relatively short distances, e.g. within a campus or data-centre. This is due to the fact that although the shorter wavelengths experience higher levels of attenuation in these fibers (-3 dB/km compared to -0.22 dB/km), the greater efficiency of silicon-based single photon detectors can be used to offset the effect of fiber attenuation for short distances. In fact, it has been shown[6, 7] that this translates to a lower system loss suffered by 800 nm photons compared to 1550 nm photons over a transmission distance of about 2.4 km.

The use of NIR wavelengths in fibers optimised for telecom wavelength introduces additional complication. These fibers are no longer single-mode for the shorter wavelengths. This leads to excitation of higher order modes giving rise to mode dispersion effects. This can affect the quality of the quantum signal resulting in reduced entanglement visibility.

*cqtab@nus.edu.sg

The following results demonstrate that polarization entanglement could be robustly distributed through telecommunication fiber using near-infrared non-degenerate entangled photons. We show that, only with temporal filtering applied in the form of narrow coincidence window of 1 ns, high quality entanglement after a transmission distance of 12 km, corresponding to fiber losses of -36 dB, could be obtained.

2 Methods

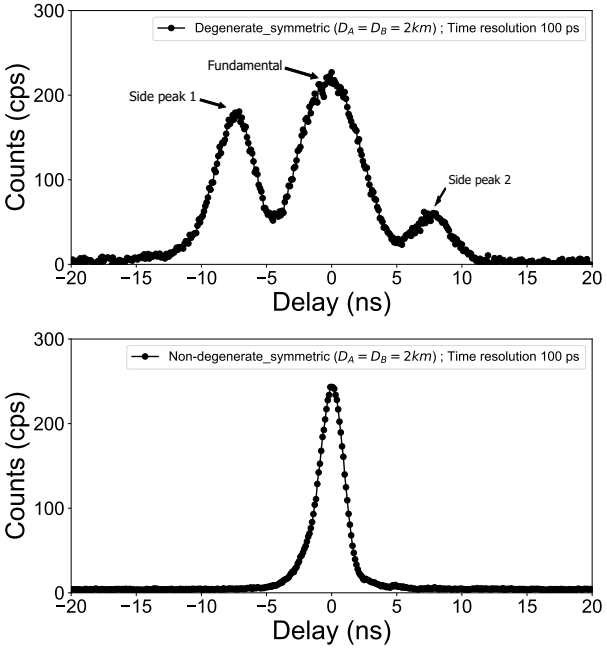


Figure 1: The cross-correlation measurement with both signal and idler photons transmitted through equal distance of fiber. (a) Degenerate situation. The distance between the fundamental mode and higher order mode is proportional to the fiber length. (b) Non-degenerate situation. The side peaks are absent and the main peak has a reduced full-width at half-maximum.

A type-0 periodically poled potassium titanyl phosphate (PPKTP) crystal under collinear phase-matching condition is used for generating polarization-entangled photon pairs[8] in the following Bell state:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|H_A H_B\rangle \pm e^{i\phi}|V_A V_B\rangle\right) \quad (1)$$

The spectrum of the emitted pair of photons can be controlled via temperature tuning of the crystal. Non-degenerate phase matching is achieved at a temperature of 34 °C, generating signal photon at 774 nm and idler photon at 850 nm while degenerate phase matching occurs at a crystal temperature of 26.5 °C. A wavelength division multiplexer (WDM) is used to separate the signal and idler photons into two arms which are then sent through equal lengths of SMF-28 fiber in each arm.

3 Results

The fiber in the experiment conforms to the ITU G.652D standard [9] and behaves as a few-mode fiber for near-infrared photons[10]. The total number of modes that can exist in the fiber for the NIR photons can be obtained from the normalized frequency V_{norm} of the fiber. The calculated V_{norm} number is 4.3. This means multiple modes could exist inside the fiber [11] with the most likely ones being the LP_{01} and LP_{11} modes. Results

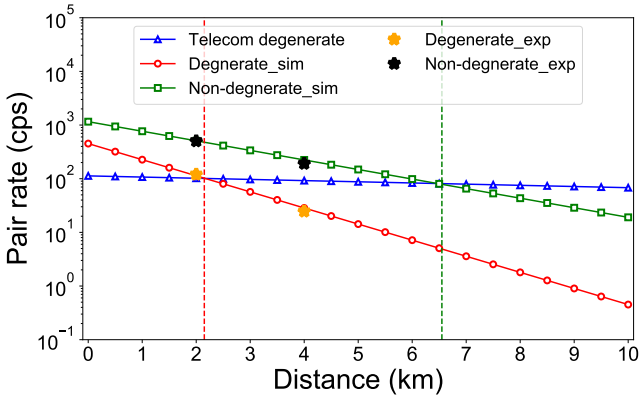


Figure 2: Comparison of the relationship between the detected pair rate for a given transmission distance using different wavelengths for the entangled photons. Telecom degenerate refers to an entangled photon source operating in the degenerate situation with similar parameters as the NIR entangled photon source used in this work. The vertical green dashed line marks the distance at which the non-degenerate NIR system would no longer outperform a telecom system purely from a loss perspective. The red vertical line mark the same for degenerate NIR entangled photons. The values corresponding to the black and yellow stars are obtained experimentally and are in excellent agreement with the predicted values.

of the cross correlation measurement between the signal and idler arms are shown in Fig. 1 for the degenerate and non-degenerate cases. Multiple peaks signal the presence of higher order modes in case of degenerate phase

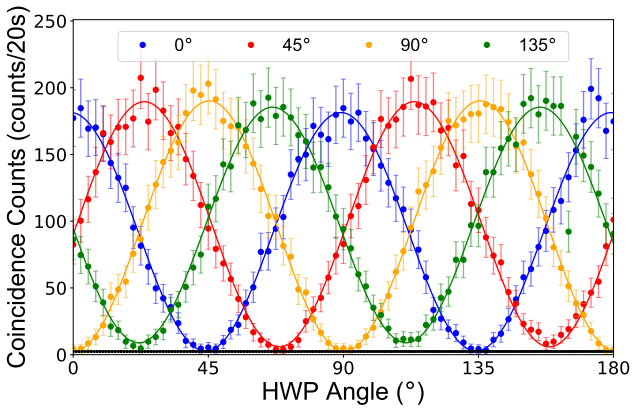


Figure 3: Polarization correlation in both H/V and D/A bases measured after 6 km telecom fiber in each arm corresponding to -36 dB of fiber loss and an estimated detector efficiency of 50%. The slight reduction in observed visibility is due to polarization mode dispersion. This effect can be further minimised if the linewidth of the NIR photons were further reduced.

matching, while no such modes were excited in the case of non-degenerate phase matching.

Typically, only when the photon pair is in the fundamental mode can their entanglement be readily measured and used. The presence of higher order modes is therefore detrimental to detection rates, because photons are removed from the fundamental mode by intermodal dispersion.

Next, we model the performance of NIR and telecom systems (including source, channel losses and detector efficiencies) and present the results in Fig. 2. We assume an entangled photon source produces polarization entangled photons in the telecom wavelength of 1550 nm with the same source brightness as in the NIR regime. Additionally, we assume Si GM-APD efficiency of about 50% for the NIR signal and commercial state of the art InGaAs GM-APD efficiency of 25% [12] for telecom wavelength. This model selects only the fraction of photons that remain within the fundamental mode.

The transmission losses can be grouped into three categories: intrinsic loss due to fiber attenuation, losses arising from intermodal dispersion and, detector inefficiency losses. Our model shows non-degenerate NIR entangled photon systems compare favorably with a telecom system for up to -18 dB of fiber loss.

Finally, the quality of entanglement was verified by measuring the entanglement visibility[13]. Raw coincidences were recorded by using the coincidence window as a temporal filter. Following our model, 6 km fiber spools were installed in each arm of the experiment. This resulted in a separation of 12 km or a total of -36 dB of fiber loss between the signal and idler measurement stations. The pump power and the coincidence window were set at 0.5 mw and 1 ns respectively. The visibility measurement is shown in Fig. 3. In the H/V basis, the raw visibilities were 97.1% and 97.2% while in the D/A basis, they were 94.0% and 90.8%. This leads to an average visibility of

94.6%, without any spatial filtering. For comparison, the average entanglement visibility was recorded to be 98.1% at source and 96.1% after 4 km separation.

4 Conclusion

These results highlight that high quality entanglement can be preserved over an effective distance of 12 km using non-degenerate NIR photons through telecom fiber. This can be useful in campus-type implementations or short metropolitan networks, for example in the deployment of entanglement based QKD.

References

[1] Bouwmeester, D. et al. Experimental quantum teleportation. Nature 390, 575–579, DOI: 10.1038/37539 (1997).

[2] Gisin, N., Ribordy, G., Tittel, W. & Zbinden, H. Quantum cryptography. Rev. Mod. Phys. 74, 145–195, DOI: 10.1103/ RevModPhys.74.145 (2002).

[3] Fink, M. et al. Entanglement-enhanced optical gyroscope. New J. Phys. 21, 053010, DOI: 10.1088/1367-2630/ab1bb2 (2019).

[4] Gisin, N. & Thew, R. Quantum communication. Nat. Photonics 1, 165–171, DOI: 10.1038/nphoton.2007.22 (2007).

[5] Kimble, H. J. The quantum internet. Nature 453, 1023–1030, DOI: 10.1038/nature07127 (2008).

[6] Meyer-Scott, E. et al. Quantum entanglement distribution with 810 nm photons through telecom fibers. Appl. Phys. Lett. 97, DOI: 10.1063/1.3460920 (2010).

[7] Holloway, C., Meyer-Scott, E., Erven, C. & Jennewein, T. Quantum entanglement distribution with 810 nm photons through active telecommunication fibers. Opt. Express 19, 20597, DOI: 10.1364/OE.19.020597 (2011).

[8] Lohrmann, A., Perumangatt, C., Villar, A. & Ling, A. Broadband pumped polarization entangled photon-pair source in a linear beam displacement interferometer. Appl. Phys. Lett. 116, 021101, DOI: 10.1063/1.5124416 (2020).

[9] Itu-t characteristics of a single-mode optical fibre and cable (2016).

[10] Gloge, D. Weakly guiding fibers. Appl. Opt. 10, 2252, DOI: 10.1364/AO.10.002252 (1971)

[11] C. Marand, K. J. B., S. J. D. Phoenix, Barnett, S. M. & Townsend, P. D. Secure optical communications systems using quantum cryptography. Philos. Transactions Royal Soc. London. Ser. A: Math. Phys. Eng. Sci. 354, 805–817, DOI: 10.1098/rsta.1996.0033 (1996).

[12] ID Quantique SA. Id230 infrared single-photon detector. Datasheet (2022).

[13] Hübel, H. et al. High-fidelity transmission of polarization encoded qubits from an entangled source over 100 km of fiber, Opt. Express 15, 7853, DOI: 10.1364/OE.15.007853 (2007).

Exact distributed quantum algorithm for Simon's problem*

Hao Li[1 2] Daowen Qiu[1 2 †] Le Luo[3] Paulo Mateus[4]

[1] *Institute of Quantum Computing and Computer Theory, School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China;*
[2] *The Guangdong Key Laboratory of Information Security Technology, Sun Yat-sen University, 510006, China; and QUDOOR Technologies Inc., Zhuhai, China*
[3] *School of Physics and Astronomy, Sun Yat-sen University, 519082 Zhuhai, China; and QUDOOR Technologies Inc., Zhuhai, China*
[4]*Instituto de Telecomunicações, Departamento de Matemática, Instituto Superior Técnico, Avenida Rovisco Pais, 1049-001 Lisbon, Portugal*

Abstract. Simon's problem is one of the most important problems demonstrating the power of quantum algorithm. In this paper, we propose an exact quantum algorithm for solving Simon's problem in distributed scenario. Our algorithm combines distributed Simon's quantum algorithm with the quantum amplitude amplification technique to ensure its determinism. Compared with the current best distributed quantum algorithm for solving Simon's problem, our algorithm has the advantage of exactness.

Keywords: Simon's problem, distributed quantum algorithms, amplitude amplification, circuit depth

1 INTRODUCTION

Quantum computing [1] has been proved to have great potential in factorizing large numbers [2], searching unordered database [3] and solving linear systems of equations [4]. At present, quantum technology has been entered to the noisy intermediate-scale quantum (NISQ) era [5], which makes it possible to implement quantum algorithms on middle-scale circuits.

Distributed quantum computing is a novel computing architecture, which combines quantum computing with distributed computing [6, 7, 8, 9, 10, 11, 12]. In distributed quantum computing architecture, multiple quantum computing nodes communicate with each other and cooperate to complete computing tasks. Compared with centralized quantum computing, the size and depth of circuit can be reduced by using distributed quantum computing, which is beneficial to improve the performance of circuit against noise.

Simon's problem is one of the most important problems in quantum computing [13]. For solving Simon's problem, quantum algorithms have the advantage of exponential acceleration over the best classical algorithms [14]. Simon's algorithm greatly inspired the proposal of Shor's algorithm [2].

Tan, Xiao and Qiu et al. [10] proposed a distributed quantum algorithm to solve Simon's problem. Their algorithm has the advantage of exponential acceleration compared with the distributed classical algorithm for Simon's problem. However, their algorithm is not exact.

We apply quantum amplitude amplification technique to design an exact distributed quantum algorithm for solving Simon's problem [15], which has the advantage of exactness compared with the current best distributed quantum algorithm for Simon's problem [10]. See the Appendix for the proof of the correctness of our algorithm.

*This work is supported in part by NSFC (No. 61876195).
†Corresponding author: issqdw@mail.sysu.edu.cn

2 PRELIMINARIES

Simon's problem is a special kind of hidden subgroup problem [16], which can be described as follows. Consider a function $f : \{0,1\}^n \to \{0,1\}^m$, where we promise that there is a string $s \in \{0,1\}^n$ with $s \neq 0^n$, such that $f(x) = f(y)$ if and only if $x = y$ or $x \oplus y = s$. We have an oracle that can query the value of function f. In classical computing, for any $x \in \{0,1\}^n$ and any $y \in \{0,1\}^m$, if we input (x, y) into the oracle, then $(x, y \oplus f(x))$ is outputted. In quantum computing, for any $x \in \{0,1\}^n$ and any $y \in \{0,1\}^m$, if $|x\rangle|y\rangle$ is inputted into the oracle, we will get $|x\rangle|y \oplus f(x)\rangle$. Our goal is to find s by performing the minimum number of queries (using oracle) to f.

3 Exact distributed quantum algorithm for Simon's problem

Simon's problem in distributed scenario is described as follows. Suppose there are 2^t people, each of whom has an oracle O_{f_w} ($w \in \{0,1\}^t$ is each person's unique identifier) that can query $f_w(u) = f(uw)$ for any $u \in \{0,1\}^{n-t}$. Each person can access 2^{n-t} values of f. They need to find the hidden string s by querying their own oracle and exchanging information as few times as possible.

Let s be the target string to be found in Simon's problem, and denote $s = s_1 s_2$, where the length of s_1 is $n-t$, the length of s_2 is t. Since s_1 may be 0^{n-t} or not, we discuss it in two cases. Firstly, assume $s_1 = 0^{n-t}$, and then we apply Algorithm 2 to find s_2. If $s_2 \neq 0^t$, then $s = 0^{n-t} s_2$. Otherwise, $s_1 \neq 0^{n-t}$, then we use Algorithm 4 to find s_1 and Algorithm 2 to find s_2.

See the Appendix for definitions of function $S(u)$ and operator U_{sort} in Algorithm 1, and a related theorem for $S(u)$. To make Algorithm 1 exact, we can do this by making sure $(Y \setminus \{0^{n-t}\}) \cup \{z\}$ is always linearly independent when we get the measured result z of the first register. Let $K = \{0^{n-t}, s_1\}$, making sure the measured result of the first register is in $K^{\perp}$ but not in $\langle Y \rangle$. First, we

Algorithm 1 Distributed quantum algorithm for finding s_1 (2^t distributed computing nodes)

1: **procedure** DISTRIBUTEDSIMON(integer n, integer t, integer m, operator O_{f_w})
2: $Y \leftarrow \{0^{n-t}\}$;
3: **repeat**
4: $|\psi_0\rangle = |0^{n-t}\rangle \left(\bigotimes_{w\in\{0,1\}^t} |0^m\rangle \right) \left|0^{2^t m}\right\rangle$;
5: $|\psi_1\rangle = \left(H^{\otimes n-t} \otimes I^{\otimes 2^{t+1}m}\right)|\psi_0\rangle$
$= \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} |u\rangle \left(\bigotimes_{w\in\{0,1\}^t} |0^m\rangle \right) \left|0^{2^t m}\right\rangle$;
6: Each computing node queries its own oracle under the control of the first register:
$|\psi_2\rangle = \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} |u\rangle \left(\bigotimes_{w\in\{0,1\}^t} |f_w(u)\rangle \right) \left|0^{2^t m}\right\rangle$
7: The (2^t+2)-th register performs its own U_{Sort} under the control of the middle 2^t registers:
$|\psi_3\rangle = \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} |u\rangle \left(\bigotimes_{w\in\{0,1\}^t} |f_w(u)\rangle \right) |S(u)\rangle$;
8: Each computing node queries its own oracle under the control of the first register:
$|\psi_4\rangle = \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} |u\rangle \left(\bigotimes_{w\in\{0,1\}^t} |0^m\rangle \right) |S(u)\rangle$;
9: $|\psi_5\rangle = \left(H^{\otimes n-t} \otimes I^{\otimes 2^{t+1}m}\right)|\psi_4\rangle$;
10: Measure the first register, get the result z;
11: **if** $z \notin \langle Y\rangle$ **then**
12: $Y \leftarrow Y \cup \{z\}$;
13: **end if**
14: **until** $|Y| = n - t$;
15: Solve the system of exclusive-or equations, get $s_1 \in \langle Y\rangle^{\perp} \setminus \{0^{n-t}\}$;
16: **return** s_1.
17: **end procedure**

can use Algorithm 1 to ensure that the measured result of the first register is in $K^{\perp}$. Then, we can utilize quantum amplitude amplification technique to ensure that the measured result of the first register is not in $\langle Y\rangle$.

Note that $\{0,1\}^{n-t}$ can be partitioned into 2^{n-t-1} pairs of strings of the form $\{x, x\oplus s_1\}$ ($x \in \{0,1\}^{n-t}$). Let T be a subset of $\{0,1\}^{n-t}$ consisting of exactly one representative from each of these pairs. Let $\left|K^{\perp}, 0^{2^t m}, S(T)\right\rangle$ be the registered state after line 9 of Algorithm 1. Let $\mathcal{A}$ denote the combined unitary operators from line 5 to line 9 in Algorithm 1. We define operators $\mathcal{R}_0(\phi) : \{0,1\}^{n-t+2^{t+1}m} \to \{0,1\}^{n-t+2^{t+1}m}$ and $\mathcal{R}_{\mathcal{A}}(\varphi, Y) : \{0,1\}^{n-t} \to \{0,1\}^{n-t}$ as follows.

$$\mathcal{R}_0(\phi)\,|x,b\rangle = \begin{cases} |x,b\rangle\,, & x \neq 0^{n-t} \text{ or } b \neq 0^{2^{t+1}m}; \\ e^{i\phi}\,|x,b\rangle\,, & x = 0^{n-t} \text{ and } b = 0^{2^{t+1}m}. \end{cases} \tag{1a}$$

$$\mathcal{R}_{\mathcal{A}}(\varphi, Y)\,|x\rangle = \begin{cases} e^{i\varphi}\,|x\rangle\,, & x \notin \langle Y\rangle; \\ |x\rangle\,, & x \in \langle Y\rangle. \end{cases} \tag{1b}$$

With the definitions of $\mathcal{R}_0(\phi)$ and $\mathcal{R}_{\mathcal{A}}(\varphi, Y)$, we define the quantum amplitude amplification operator as follows.

$$\mathcal{Q} = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^{\dagger}\left(\mathcal{R}_{\mathcal{A}}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right). \tag{2}$$

Let

$$X = K^{\perp} \setminus \langle Y\rangle. \tag{3}$$

Definition 1 *Let $|\Psi_X\rangle$ and $|\Psi_Y\rangle$ denote the projections onto the good and bad state subspaces, respectively, i.e., the subspaces spanned by $\left\{\, |x,b\rangle \mid x \in X, b \in \{0,1\}^{2^{t+1}m}\right\}$ and $\left\{\, |y,b\rangle \mid y \in \langle Y\rangle, b \in \{0,1\}^{2^{t+1}m}\right\}$.*

Denote by

$$\left|K^{\perp}, 0^{2^t m}, S(T)\right\rangle = |\Psi_X\rangle + |\Psi_Y\rangle\,. \tag{4}$$

For a related lemma and a theorem on the quantum amplitude amplification operator $\mathcal{Q}$, see the Appendix.

Algorithm 2 Distributed quantum algorithm for finding s_2 (2^t distributed computing nodes)

1: Query each oracle O_{f_w} once in parallel to get $f(0^{n-t}w)$ ($w \in \{0,1\}^t$);
2: Query oracle $O_{f_{0^t}}$ once to get $f(s_1 0^t)$;
3: Find a $v \in \{0,1\}^t$ such that $f(0^{n-t}v) = f(s_1 0^t)$;
4: Obtain $s_2 = v$;

Combining Algorithm 1 with Algorithm 3, we obtain Algorithm 4, whose proof of correctness see the Appendix.

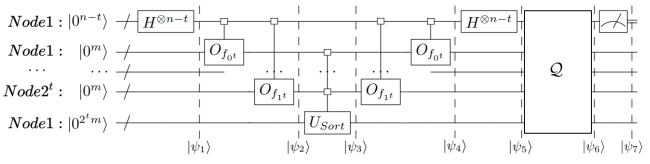


Figure 1: The circuit for the quantum part of exact distributed quantum algorithm for finding s_1 (2^t computing nodes) (Algorithm 4).

Algorithm 3 Quantum amplitude amplification for measuring good states

Require: Input parameters satisfy associated definitions

Ensure: $z \in X$

1: **procedure** QAMPAMP(registers $\left|K^{\perp}, 0^{2^t m}, S(T)\right\rangle$, integer n, integer t, integer m, operator $\mathcal{A}$, set Y)
2: $l \leftarrow |Y|$;
3: $\phi \leftarrow 2\arctan\left(\sqrt{\frac{2^{n-t-l}}{3\cdot 2^{n-t-l}-4}}\right)$;
4: $\varphi \leftarrow \arccos\left(\frac{2^{n-t-l-1}-1}{2^{n-t-l}-1}\right)$;
5: Apply $\mathcal{Q}$ to $\left|K^{\perp}, 0^{2^t m}, S(T)\right\rangle$ where $\mathcal{Q} = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^{\dagger}\left(\mathcal{R}_{\mathcal{A}}(\varphi, Y)\otimes I^{\otimes 2^{t+1}m}\right)$;
6: Measure the first register, get the result z;
7: **return** z.
8: **end procedure**

Algorithm 4 Exact distributed quantum algorithm for finding s_1 (2^t computing nodes)

1: **procedure** EXACTDISTRIBUTEDSIMON(integer n, integer t, integer m, operator O_{f_w})
2: $Y \leftarrow \{0^{n-t}\}$;
3: **repeat**
4: Prepare $|0^{n-t}\rangle\left(\bigotimes_{w\in\{0,1\}^t}|0^m\rangle\right)\left|0^{2^t m}\right\rangle$;
5: Apply $\mathcal{A}$ to the registers where $\mathcal{A}$ denote the combined unitary operators from line 5 to line 9 in Algorithm 1;
6: $z \leftarrow$ QAMPAMP($\left|K^{\perp}, 0^{2^t m}, S(T)\right\rangle$, n, t, m, $\mathcal{A}$, Y);
7: $Y \leftarrow Y \cup \{z\}$;
8: **until** $|Y| = n - t$;
9: Solve the system of exclusive-or equations, get $s_1 \in \langle Y\rangle^{\perp} \setminus \{0^{n-t}\}$;
10: **return** s_1.
11: **end procedure**

4 Comparison with other algorithms

The comparison of our algorithm with the distributed classical algorithm for Simon's problem, the distributed quantum algorithm for Simon's problem[17] and the Simon's algorithm is the same situation as the comparison of the algorithm in [10].

We compare our algorithm with the current best quantum algorithm for solving Simon's problem in distributed scenario [10]. The algorithm in [10] is not exact. Our algorithm combines the algorithm in [10] with quantum amplitude amplification technique, which can solve Simon's problem in distributed scenario exactly.

5 Conclusion

In this paper, we have designed an exact distributed quantum algorithm to solve Simon's problem, which combines the current best quantum algorithm for solving Simon's problem in distributed scenario with quantum amplitude amplification technique. By means of multiple quantum computing nodes processing in parallel, each node needs to query their own oracle with fewer times. This reduces the depth of query complexity for each oracle, and therefore it reduces circuit noise and likely makes it easier to be implemented in the current NISQ era.

References

[1] M. A. Nielsen and I. L. Chuang, *Quantum computation and quantum information*, 10th ed. (Cambridge University Press, Cambridge, 2010).

[2] P. W. Shor, SIAM J. Comput. **26**, 1484 (1997).

[3] L. K. Grover, in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing* (ACM Press, Philadelphia,1996), pp. 212-219.

[4] A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).

[5] J. Preskill, Quantum **2**, 79 (2018).

[6] H. Buhrman and H. Röhrig, in *Mathematical Foundations of Computer Science 2003*, edited by G. Goos, J. Hartmanis, J. van Leeuwen, B. Rovan, and P. Vojtáš, Lecture Notes in Computer Science Vol. 2747 (Springer, Berlin, 2003), pp. 1–20.

[7] R. Beals, S. Brierley, O. Gray, A. W. Harrow, S. Kutin, N. Linden, D. Shepherd, and M. Stather, Proc. R. Soc. A. **469**, 20120686 (2013).

[8] K. Li, D. W. Qiu, L. Z. Li, S. G. Zheng, and Z. B. Rong, Inf. Process. Lett. **120**, 23 (2017).

[9] D. W. Qiu, L. Luo, and L.G. Xiao, arXiv:2204.10487v3.

[10] J. W. Tan, L. G. Xiao, D. W. Qiu, L. Luo, and P. Mateus, Phys. Rev. A **106**, 032417 (2022).

[11] L. G. Xiao, D. W. Qiu, L. Luo, and P. Mateus, Quantum Inf. Comput. **23**, 1-2 (2023).

[12] L. G. Xiao, D. W. Qiu, L. Luo, and P. Mateus, arXiv:2304.12100.

[13] D. R. Simon, SIAM J. Comput. **26**, 1474 (1997).

[14] G. Y. Cai and D.W. Qiu, J. Comput. Syst. Sci. **97**, 83 (2018).

[15] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Contemp. Math. **305** (2002).

[16] P. Kaye, R. Laflamme, and M. Mosca, *An introduction to quantum computing* (Oxford University Press, Oxford, 2007).

[17] J. Avron, O. Casper, and I. Rozen, Phys. Rev. A **104**, 052404 (2021).

6 Appendix

In the following, we introduce the function $S(u)$ in Algorithm 1 and its associated theorem [10].

Definition 2 *For any $u \in \{0,1\}^{n-t}$, let $S(u)$ represent a string of length $2^t m$ by concatenating all strings $f_w(u)$ ($w \in \{0,1\}^t$) according to lexicographical order, that is,*

$$S(u) = f_{w_0}(u) f_{w_1}(u) \cdots f_{w_{2^t-1}}(u), \tag{5}$$

where $f_{w_0}(u) \le f_{w_1}(u) \le \ldots \le f_{w_{2^t-1}}(u) \in \{0,1\}^m$ with $w_i \in \{0,1\}^t (i = 0, 1, \ldots, 2^t - 1)$ where $w_i \neq w_j$ for any $i \neq j$.

The following theorem concerning $S(u)$ is useful and important, which is proved in [10].

Theorem 3 *Suppose function $f : \{0,1\}^n \to \{0,1\}^m$, satisfies that there is a string $s \in \{0,1\}^n$ with $s \neq 0^n$, such that $f(x) = f(y)$ if and only if $x = y$ or $x \oplus y = s$. Then $\forall u, v \in \{0,1\}^{n-t}, S(u) = S(v)$ if and only if $u \oplus v = 0^{n-t}$ or $u \oplus v = s_1$, where $s = s_1 s_2$.*

In the following, we describe the operator U_{Sort} in Algorithm 1. The effect of U_{Sort} in Algorithm 1 is to sort the values in the 2^t control registers by lexicographical order, and XOR to the target register. When $t = 1$, the effect of operator U_{Sort} is:

$$\begin{aligned} U_{Sort}|a\rangle|b\rangle|c\rangle =& |a\rangle|b\rangle|c \oplus (\min(a,b)\max(a,b))\rangle \\ =& \begin{cases} |a\rangle|b\rangle|c \oplus (ab)\rangle, & a \le b, \\ |a\rangle|b\rangle|c \oplus (ba)\rangle, & a > b, \end{cases} \end{aligned} \tag{6}$$

where $a, b \in \{0,1\}^m$ and $c \in \{0,1\}^{2m}$.

In order to make Algorithm 1 work, the crucial step is to eliminate all states in $\langle Y\rangle$ from the first register. In quantum amplitude amplification process, one can accomplish this by choosing appropriate $\phi, \varphi \in \mathbb{R}$ such that after applying $\mathcal{Q}$ on $\left|K^\perp, 0^{2^t m}, S(T)\right\rangle$, the amplitudes of all states in $\langle Y\rangle$ of the first register become zero.

In the following, we introduce a related lemma and a theorem on the quantum amplitude amplification operator $\mathcal{Q}$ [14].

Lemma 4 *Let $\mathcal{A}$ denote the combined unitary operators from line 5 to line 9 in Algorithm 1. Let $\mathcal{Q} = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^\dagger \left(\mathcal{R}_\mathcal{A}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right)$. Then*

$$\begin{aligned} \mathcal{Q}\,|\Psi_X\rangle =& e^{i\varphi}\left((1-e^{i\phi})(1-2^{l-n+t}) - 1\right)|\Psi_X\rangle + \\ & e^{i\varphi}(1-e^{i\phi})(1-2^{l-n+t})\,|\Psi_Y\rangle, \\ \mathcal{Q}\,|\Psi_Y\rangle =& (1-e^{i\phi})2^{l-n+t}\,|\Psi_X\rangle - \\ & \left((1-e^{i\phi})(1-2^{l-n+t}) + e^{i\phi}\right)|\Psi_Y\rangle, \end{aligned}$$

where $l = |Y|$.

Proof. From Eq. (1a), we can write $\mathcal{R}_0(\phi)$ as follows:

$$\begin{aligned} \mathcal{R}_0(\phi) =& I^{\otimes n-t+2^{t+1}m} \\ & - \left(1-e^{i\phi}\right)\left|0^{n-t}, 0^{2^{t+1}m}\right\rangle\left\langle 0^{n-t}, 0^{2^{t+1}m}\right|. \end{aligned} \tag{7}$$

From the definitions of $\mathcal{R}_\mathcal{A}(\varphi, Y)$, $|\Psi_X\rangle$ and $|\Psi_Y\rangle$, we have:

$$\left(\mathcal{R}_\mathcal{A}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right)|\Psi_X\rangle = e^{i\varphi}\,|\Psi_X\rangle. \tag{8}$$

$$\left(\mathcal{R}_\mathcal{A}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right)|\Psi_Y\rangle = |\Psi_Y\rangle. \tag{9}$$

Let $\mathcal{U}(\mathcal{A}, \phi) = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^\dagger$, then according to Eq. (2), $\mathcal{Q}$ can be written as:

$$\mathcal{Q} = \mathcal{U}(\mathcal{A}, \phi)\left(\mathcal{R}_\mathcal{A}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right). \tag{10}$$

For $\mathcal{U}(\mathcal{A}, \phi)$, we have:

$$\begin{aligned} \mathcal{U}(\mathcal{A}, \phi) =& -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^\dagger \\ =& -\mathcal{A}\left(I^{\otimes n-t+2^{t+1}m} - \right. \\ & \left.\left(1-e^{i\phi}\right)\left|0^{n-t}, 0^{2^{t+1}m}\right\rangle\left\langle 0^{n-t}, 0^{2^{t+1}m}\right|\right)\mathcal{A}^\dagger \\ =& \left(1-e^{i\phi}\right)\left(\mathcal{A}\left|0^{n-t}, 0^{2^{t+1}m}\right\rangle\left\langle 0^{n-t}, 0^{2^{t+1}m}\right|\mathcal{A}^\dagger\right) \\ & - I^{\otimes n-t+2^{t+1}m} \\ =& \left(1-e^{i\phi}\right)\left|K^\perp, 0^{2^t m}, S(T)\right\rangle\left\langle K^\perp, 0^{2^t m}, S(T)\right| \\ & - I^{\otimes n-t+2^{t+1}m} \\ =& \left(1-e^{i\phi}\right)\left(|\Psi_X\rangle + |\Psi_Y\rangle\right)\left(\langle\Psi_X| + \langle\Psi_Y|\right) \\ & - I^{\otimes n-t+2^{t+1}m}. \end{aligned} \tag{11}$$

Because $|\langle Y\rangle| = 2^{l-1}$and $|K^\perp| = 2^{n-t-1}$, according to the definition of $|\Psi_X\rangle$ and $|\Psi_Y\rangle$, we have:

$$\begin{aligned} &\langle\Psi_X\,|\,\Psi_X\rangle = 1 - 2^{l-n+t}. \\ &\langle\Psi_Y\,|\,\Psi_Y\rangle = 2^{l-n+t}. \\ &\langle\Psi_X\,|\,\Psi_Y\rangle = 0. \end{aligned} \tag{12}$$

Then we can get:

$$\begin{aligned} \mathcal{Q}\,|\Psi_X\rangle =& \mathcal{U}(\mathcal{A}, \phi)\left(\mathcal{R}_\mathcal{A}(\varphi, Y) \otimes I^{\otimes 2^{t+1}m}\right)|\Psi_X\rangle \\ =& e^{i\varphi}\mathcal{U}(\mathcal{A}, \phi)\,|\Psi_X\rangle \\ =& e^{i\varphi}\Big(\left(1-e^{i\phi}\right)\left(|\Psi_X\rangle + |\Psi_Y\rangle\right)\left(\langle\Psi_X| + \langle\Psi_Y|\right) \\ & - I^{\otimes n-t+2^{t+1}m}\Big)|\Psi_X\rangle \\ =& e^{i\varphi}\left(1-e^{i\phi}\right)\left(|\Psi_X\rangle + |\Psi_Y\rangle\right)\left(\langle\Psi_X| + \langle\Psi_Y|\right) \\ & |\Psi_X\rangle - e^{i\varphi}\,|\Psi_X\rangle \\ =& e^{i\varphi}\left(1-e^{i\phi}\right)\langle\Psi_X\,|\,\Psi_X\rangle\left(|\Psi_X\rangle + |\Psi_Y\rangle\right) \\ & - e^{i\varphi}\,|\Psi_X\rangle \\ =& e^{i\varphi}\left(1-e^{i\phi}\right)\left(1-2^{l-n+t}\right)\left(|\Psi_X\rangle + |\Psi_Y\rangle\right) \\ & - e^{i\varphi}\,|\Psi_X\rangle \\ =& e^{i\varphi}\left((1-e^{i\phi})(1-2^{l-n+t}) - 1\right)|\Psi_X\rangle \\ & + e^{i\varphi}(1-e^{i\phi})(1-2^{l-n+t})\,|\Psi_Y\rangle. \end{aligned} \tag{13}$$

$$
\begin{aligned}
\mathcal{Q}\,|\Psi_Y\rangle =&\mathcal{U}(\mathcal{A},\phi)\left(\mathcal{R}_{\mathcal{A}}(\varphi,Y)\otimes I^{\otimes 2^{t+1}m}\right)|\Psi_Y\rangle\\
=&\mathcal{U}(\mathcal{A},\phi)\,|\Psi_Y\rangle\\
=&\Big(\left(1-e^{i\phi}\right)\left(|\Psi_X\rangle+|\Psi_Y\rangle\right)\left(\langle\Psi_X|+\langle\Psi_Y|\right)\\
&-I^{\otimes n-t+2^{t+1}m}\Big)|\Psi_Y\rangle\\
=&\left(1-e^{i\phi}\right)\left(|\Psi_X\rangle+|\Psi_Y\rangle\right)\left(\langle\Psi_X|+\langle\Psi_Y|\right)|\Psi_Y\rangle\\
&-|\Psi_Y\rangle\\
=&\left(1-e^{i\phi}\right)\langle\Psi_Y\,|\,\Psi_Y\rangle\left(|\Psi_X\rangle+|\Psi_Y\rangle\right)-|\Psi_Y\rangle\\
=&\left(1-e^{i\phi}\right)2^{l-n+t}\left(|\Psi_X\rangle+|\Psi_Y\rangle\right)-|\Psi_Y\rangle\\
=&(1-e^{i\phi})2^{l-n+t}\,|\Psi_X\rangle\\
&-\left((1-e^{i\phi})(1-2^{l-n+t})+e^{i\phi}\right)|\Psi_Y\rangle\,.
\end{aligned}
\tag{14}
$$

□

Theorem 5 *Let $\mathcal{A}$ denote the combined unitary operators from line 5 to line 9 in Algorithm 1. Let $\mathcal{Q} = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^\dagger\left(\mathcal{R}_{\mathcal{A}}(\varphi,Y)\otimes I^{\otimes 2^{t+1}m}\right)$, $\phi = 2\arctan\left(\sqrt{\frac{2^{n-t-l}}{3\cdot 2^{n-t-l}-4}}\right)$, $\varphi=\arccos\left(\frac{2^{n-t-l-1}-1}{2^{n-t-l}-1}\right)$. Then*

$$
\mathcal{Q}\left|K^{\perp},0^{2^t m},S(T)\right\rangle=\mathcal{Q}(|\Psi_X\rangle+|\Psi_Y\rangle)=|\Psi_X\rangle\,.
$$

Proof. By making sure the resulting superposition $\mathcal{Q}(|\Psi_X\rangle+|\Psi_Y\rangle)$ has inner product zero with $|\Psi_Y\rangle$, according to Eq. (13) and Eq. (14) of Lemma 4, we can obtain the following equation.

$$
e^{i\varphi}(1-e^{i\phi})(1-2^{l-n+t})=(1-e^{i\phi})(1-2^{l-n+t})+e^{i\phi}. \tag{15}
$$

We denote $b=(1-2^{l-n+t})$, then according to Eq. (15), we have:

$$
\begin{aligned}
b&=e^{-i\varphi}\left(b+\frac{1}{e^{-i\phi}-1}\right)\\
&=e^{-i\varphi}\left(b+\frac{1}{\cos\phi-1-i\sin\phi}\right)\\
&=e^{-i\varphi}\left(b+\frac{\cos\phi-1}{(\cos\phi-1)^2+\sin^2\phi}\right.\\
&\qquad\left.+i\frac{\sin\phi}{(\cos\phi-1)^2+\sin^2\phi}\right)\\
&=e^{-i\varphi}\left(b-\frac{1}{2}+i\frac{\sin\phi}{2-2\cos\phi}\right).
\end{aligned}
\tag{16}
$$

Taking the square of b, we can further have:

$$
b^2=\left(b-\frac{1}{2}\right)^2+\frac{\sin^2\phi}{4\,(1-\cos\phi)^2}. \tag{17}
$$

Arrange Eq. (17) to get:

$$
4b-1=\frac{\sin^2\phi}{(1-\cos\phi)^2}=\cot^2\frac{\phi}{2}. \tag{18}
$$

From Eq. (18), we can obtain:

$$
\begin{aligned}
\phi&=2\arctan\left(\sqrt{\frac{1}{4b-1}}\right)\\
&=2\arctan\left(\sqrt{\frac{2^{n-t-l}}{3\cdot 2^{n-t-l}-4}}\right).
\end{aligned}
\tag{19}
$$

Since b is a real number, then the Eq. (16) also needs to be a real number, so we can obtain:

$$
\varphi=\arccos\left(\frac{b-\frac{1}{2}}{b}\right)=\arccos\left(\frac{2^{n-t-l-1}-1}{2^{n-t-l}-1}\right). \tag{20}
$$

□

Since $1\le l\le n-t-1$, we can always obtain $\phi,\varphi\in\mathbb{R}$. Thus we get Algorithm 3. Actually, it is easy to see that for all $1\le l\le n-t-1$, $\langle\Psi_X\,|\,\Psi_X\rangle\ge 1/2$. That is, the success probability for measuring good states is already greater than or equal to $1/2$ without amplitude amplification. Therefore, we only need a constant number of iterations of amplitude amplification to make the success probability exactly 1 in our cases, and our analysis above further shows that this "constant" is 1. That is the reason why only one iteration of amplitude amplification is needed in our case.

In the following, we prove the correctness of Algorithm 4. First, we write out the state after the first step of Algorithm 4 in FIG. 1.

$$
\begin{aligned}
|\psi_1\rangle&=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\left(\bigotimes_{w\in\{0,1\}^t}|0^m\rangle\right)\left|0^{2^t m}\right\rangle\\
&=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\underbrace{|0^m\rangle\ldots|0^m\rangle}_{2^t}\left|0^{2^t m}\right\rangle.
\end{aligned}
\tag{21}
$$

Then Algorithm 4 queries each oracle to get the following state:

$$
|\psi_2\rangle=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\underbrace{|f(u0^t)\rangle\ldots|f(u1^t)\rangle}_{2^t}\left|0^{2^t m}\right\rangle. \tag{22}
$$

After sorting by using U_{Sort}, we have the following state:

$$
|\psi_3\rangle=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\underbrace{|f(u0^t)\rangle\ldots|f(u1^t)\rangle}_{2^t}|S(u)\rangle\,. \tag{23}
$$

After that, we query each oracle again and restore the state of the 2^t m-bit registers to $|0^m\rangle$. Then we obtain the following state:

$$
\begin{aligned}
|\psi_4\rangle&=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\underbrace{|0^m\rangle\ldots|0^m\rangle}_{2^t}|S(u)\rangle\\
&=\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle\left|0^{2^t m}\right\rangle|S(u)\rangle\,.
\end{aligned}
\tag{24}
$$

From Theorem 3, we know that the structure of Simon's problem exists in function S.

After Hadamard transform on the first register, we get the following state:

$$\begin{aligned}
|\psi_5\rangle =& \left(H^{\otimes n-t} \otimes I^{\otimes 2^{t+1}m}\right)|\psi_4\rangle \\
=& \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \left(\sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \right. \\
& \left. + \sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \right) \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \left(\sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \right. \\
& \left. + \sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\oplus s_1\rangle\right)\left|0^{2^t m}\right\rangle |S(u\oplus s_1)\rangle \right) \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \left(\sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \right. \\
& \left. + \sum_{u\in\{0,1\}^{n-t}} \left(H^{\otimes n-t}|u\oplus s_1\rangle\right)\left|0^{2^t m}\right\rangle |S(u)\rangle \right) \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \sum_{u\in\{0,1\}^{n-t}} \left[H^{\otimes n-t}\left(|u\rangle + |u\oplus s_1\rangle\right)\right] \\
& \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \sum_{u\in\{0,1\}^{n-t}} \left(\frac{1}{\sqrt{2^{n-t}}} \sum_{z\in\{0,1\}^{n-t}} \right. \\
& \left. \times \left[(-1)^{u\cdot z} + (-1)^{(u\oplus s_1)\cdot z}\right]|z\rangle \right) \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{\sqrt{2^{n-t+2}}} \sum_{u\in\{0,1\}^{n-t}} \left(\frac{1}{\sqrt{2^{n-t}}} \sum_{z\in\{0,1\}^{n-t}} \right. \\
& \left. (-1)^{u\cdot z}\left[1+(-1)^{s_1\cdot z}\right]|z\rangle \right) \left|0^{2^t m}\right\rangle |S(u)\rangle .
\end{aligned} \tag{25}$$

Note that if $s_1 \cdot z = 1$ we have $1+(-1)^{s_1\cdot z} = 0$ and the basis state $|z\rangle$ vanishes in the above state. If $s_1 \cdot z = 0$, we have $1 + (-1)^{s_1\cdot z} = 2$, so we have:

$$\begin{aligned}
|\psi_5\rangle =& \frac{1}{\sqrt{2^{n-t+2}}} \sum_{u\in\{0,1\}^{n-t}} \left(\frac{1}{\sqrt{2^{n-t}}} \sum_{z\in\{0,1\}^{n-t}} \right. \\
& \left. (-1)^{u\cdot z}\left[1+(-1)^{s_1\cdot z}\right]|z\rangle \right) \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{\sqrt{2^{n-t}}} \sum_{u\in\{0,1\}^{n-t}} \left(\frac{1}{\sqrt{2^{n-t}}} \sum_{z\in s_1^\perp} (-1)^{u\cdot z}|z\rangle \right) \left|0^{2^t m}\right\rangle \\
& |S(u)\rangle \\
=& \frac{1}{2^{n-t}} \sum_{u\in\{0,1\}^{n-t}} \sum_{z\in s_1^\perp} (-1)^{u\cdot z}|z\rangle \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{2^{n-t}} \sum_{z\in s_1^\perp} |z\rangle \sum_{u\in\{0,1\}^{n-t}} (-1)^{u\cdot z} \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \frac{1}{2^{n-t-1}} \sum_{z\in s_1^\perp} |z\rangle \sum_{u\in T} (-1)^{u\cdot z} \left|0^{2^t m}\right\rangle |S(u)\rangle \\
=& \left|K^\perp, 0^{2^t m}, S(T)\right\rangle \\
=& |\Psi_X\rangle + |\Psi_Y\rangle .
\end{aligned} \tag{26}$$

Apply $\mathcal{Q}$ to $|\psi_5\rangle$, then we have:

$$|\psi_6\rangle = \mathcal{Q}\left|K^\perp, 0^{2^t m}, S(T)\right\rangle = \mathcal{Q}(|\Psi_X\rangle + |\Psi_Y\rangle), \tag{27}$$

where $\mathcal{Q} = -\mathcal{A}\mathcal{R}_0(\phi)\mathcal{A}^\dagger\left(\mathcal{R}_{\mathcal{A}}(\varphi, Y)\otimes I^{\otimes 2^{t+1}m}\right)$, $\phi = 2\arctan\left(\sqrt{\frac{2^{n-t-l}}{3\cdot 2^{n-t-l}-4}}\right)$, $\varphi = \arccos\left(\frac{2^{n-t-l-1}-1}{2^{n-t-l}-1}\right)$.

According to Theorem 5, we have:

$$|\psi_6\rangle = |\Psi_X\rangle . \tag{28}$$

After measurement on the first register, we can get a string that is in $K^\perp \setminus \langle Y\rangle$. After $n-t-1$ repetitions of the Algorithm 4, we can obtain $n-t$ elements in the $K^\perp \setminus \langle Y\rangle$. Then, using the classical Gaussian elimination method, we can obtain s_1.

If we have already found s_1, we can use Algorithm 2 to find out s_2. Since $f(s_1 0^t) = f((s_1 0^t)\oplus s)$, we have $f(s_1 0^t) = f(0^{n-t}s_2)$. So we can find a v such that $f(s_1 0^t) = f(0^{n-t}v)$. Then we can obtain $s_2 = v$. At last, we can obtain $s = s_1 s_2$.

Remark 1 *Note that the oracle query of each quantum computing node in Algorithm 1 and Algorithm 4 can actually be completed in parallel. With the help of auxiliary $(2^t-1)(n-t)$ qubits, they can change the state of the control register after the first Hadamard transform $\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}|u\rangle$ to $\frac{1}{\sqrt{2^{n-t}}}\sum_{u\in\{0,1\}^{n-t}}\underbrace{|u\rangle|u\rangle\ldots|u\rangle}_{2^t}$. In fact, after the first Hadamard transform, they can teleport each group of $n-t$ control bits to every quantum computing node and use this to control the oracle of the computing node.*

Commutation simulator for open quantum dynamics

Jaewoo Joo[1] * Timothy P. Spiller[2]

[1] *School of Mathematics and Physics, University of Portsmouth, Portsmouth PO1 3QL, UK*
[2] *York Centre for Quantum Technologies, Department of Physics, University of York, York, YO10 5DD, U.K*

Abstract. Recent progress in quantum simulation and algorithms has demonstrated a rapid expansion in capabilities. The search continues for new techniques and applications to exploit quantum advantage. Here we propose an innovative method to investigate directly the properties of a time-dependent density operator $\hat{\rho}(t)$. Using generalised quantum commutation simulators, we can directly compute the expectation value of the commutation relation and thus of the rate of change of $\hat{\rho}(t)$. The approach can be utilised as a quantum eigen-vector solver for the von Neumann equation and a decoherence investigator for the Lindblad equation, by using just the statistics of single-qubit measurements. A simple but important example is demonstrated in the single-qubit case and we discuss extension of the method for practical quantum simulation with many qubits, towards investigation of more realistic quantum systems.

Keywords: Quantum simulation, Open quantum dynamics

1 Introduction

A century ago, very early in the development of quantum mechanics, commutation relations emerged in various crucial roles [1, 2]. Pairs of non-commuting operators (e.g., the position and momentum operators for a particle) describe the complementary nature of their corresponding physical properties, leading to uncertainty relations between these quantities for quantum systems. Commutators also underpin the time evolution of quantum systems, whether this be of general operators in the Heisenberg picture (or the relevant part of the Interaction picture), or the system density operator in the Schrödinger picture, where the time dependence resides in the quantum state or density operator. In this latter picture, the quantum state of a system given by $|\psi(t)\rangle$ evolves according to the Schrödinger equation (with units where $\hbar = 1$) [3], given by

$$\hat{\mathcal{H}}|\psi(t)\rangle = i\frac{\partial}{\partial t}|\psi(t)\rangle \,, \tag{1}$$

where $\hat{\mathcal{H}}$ is the system Hamiltonian. For an initial state defined as $|\psi(0)\rangle = |\psi_0\rangle$ at $t = 0$ and a time-independent $\hat{\mathcal{H}}$, the evolution from 0 to t is determined by the unitary operator $\hat{U}(t)$, such that $|\psi(t)\rangle = \hat{U}(t)|\psi_0\rangle = e^{-i\hat{\mathcal{H}}t}|\psi_0\rangle$.

An equivalent alternative description is via the density operator, so defining this as $\hat{\rho}(t) = \hat{U}(t)|\psi_0\rangle\langle\psi_0|(\hat{U}(t))^\dagger$ the evolution is given by the von Neumann equation [4] as

$$\frac{d}{dt}\hat{\rho}(t) = i[\hat{\rho}(t), \hat{\mathcal{H}}] \,, \tag{2}$$

where the commutation relation between $\hat{x}$ and $\hat{y}$ is given by $[\hat{x}, \hat{y}] = \hat{x}\hat{y} - \hat{y}\hat{x}$.

The form of the von Neumann equation is very interesting because the time-dependence of the system is expressed directly in terms of the commutation relation between the density operator and the Hamiltonian. The density operator approach provides a direct statistical representation because the diagonal parts of $\frac{d}{dt}\hat{\rho}(t)$ give the rate of change of the system probability density. These always correspond to real numbers, which can be measured for the actual physical system, either through repeated measurements on an identically prepared and evolved single pure system, or through measurements on an ensemble of identical systems all equivalently prepared and evolved. We refer to these equivalent approaches as an "ensemble measurement". The density matrix approach can also be used to incorporate classical uncertainty (lack of knowledge), in addition to quantum superposition, via (finite-entropy) mixtures of pure quantum states. In this work we will use the density operator approach $\hat{\rho}(t)$, both from the perspective of the reversible von Neumann equation (2) but also to provide scope for the inclusion of classical uncertainty and irreversible evolution.

In quantum theory, the irreversibility inherent in open systems—those coupled to additional environment degrees of freedom—can be modelled by modification and addition of noise terms to either the Heisenberg equation for system operators or the Schrödinger equation for system states [5]. However, the density matrix approach forms a very important method for investigating the dynamics of open quantum systems, beyond just the Schrödinger equation. The Lindblad master equation is a very widely used and applicable example. This commonly describes an open system interacting weakly with its environment, describing the effects of the environment on the system (generally, decoherence mechanisms) using Lindblad operators $\hat{\mathcal{L}}_j$. These operators modify the von Neumann equation (2) to

$$\frac{d}{dt}\hat{\rho}(t) = i\left[\hat{\rho}(t), \hat{\mathcal{H}}\right] + \sum_j \left(\hat{\mathcal{L}}_j \hat{\rho}(t)\, \hat{\mathcal{L}}_j^\dagger - \frac{1}{2}\left\{\hat{\rho}(t)\,, \hat{\mathcal{L}}_j^\dagger \hat{\mathcal{L}}_j\right\}\right), \tag{3}$$

where the anti-commutation relation between $\hat{x}$ and $\hat{y}$ is given by $\{\hat{x}, \hat{y}\} = \hat{x}\hat{y} + \hat{y}\hat{x}$ [6, 7]. In general the Lindblad operators are not Hermitian and act to introduce decoherence to the system, changing its entropy. The particular case of Hermitian Lindblad operators can be

*jaewoo.joo@port.ac.uk

used to model quantum measurements, or noisy external source terms in the system Hamiltonian. For $\hat{\mathcal{L}}_j = \hat{1}$ ($\hat{1}$: identity operator), the Lindblad terms disappear and only the unitary term $i\left[\hat{\rho}(t), \hat{\mathcal{H}}\right]$ survives, thus returning to the von Neumann equation and the unitary evolution of a closed quantum system in time.

In this work, we propose a novel method to directly compute, or simulate, matrix elements of $\frac{d}{dt}\hat{\rho}(t)$, by measuring expectation values of the commutation relation in the von Neumann equation (2) and the more general Lindblad equation (3). Consider the case where the system of interest comprises L qubits, so the density operator $\hat{\rho}(t)$ can be represented by a $2^L \times 2^L$ matrix. Our approach provides the (diagonal and off-diagonal) matrix elements $\frac{d}{dt}\rho_{n,m}(t) = \langle n|\frac{d}{dt}\hat{\rho}(t)|m\rangle$ with $n, m = 0, ..., 2^L - 1$ ranging over a suitable basis of the system. So, for example, if we seek the expectation of the rate of state change in time, given by $\langle\Phi|\frac{d}{dt}\hat{\rho}(t)|\Phi\rangle$ for some chosen reference state $|\Phi\rangle$, we can perform quantum processing to determine this by measuring the expectation value of the commutator in the von Neumann equation (2), given by $i\langle\Phi|[\hat{\rho}(t), \hat{\mathcal{H}}]|\Phi\rangle$ in the case of closed quantum systems. For the off-diagonal terms, we can compute $i\langle\Phi|[\hat{\rho}(t), \hat{\mathcal{H}}]|\Phi'\rangle$ by a sum of expectation values given by another controlled-operator gate, with operator $\hat{A}$ for $|\Phi'\rangle = \hat{A}|\Phi\rangle$. For the case of open quantum systems, it is required to perform additional quantum processing to compute the extra Lindblad terms that depend on the $\hat{\mathcal{L}}_j$.

2 Algorithm for quantum commutation simulation

We first provide the protocol describing the algorithm, followed by a detailed explanation of the quantum commutation simulator. The simulation is built upon the following resources, as employed in Fig. 1: The system S, assumed to be of dimension 2^L, or $2^L \times 2^L$ in density matrix form; a separate reference system M of the same size as the system; a separate control qubit C. The state of the total system is denoted by $|\Psi\rangle$. With reference to the full quantum circuit shown in Fig. 1, the protocol for the simulation runs as follows.

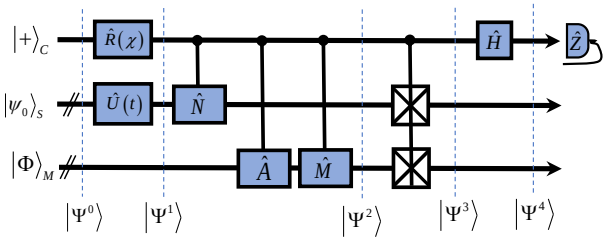


Figure 1: Schematic of the generalised quantum circuit to simulate the expectation value of the commutation relation. Three controlled-operators are given by operators $\hat{N}$, $\hat{A}$ and $\hat{M}$ and the total number of qubits required in the simulator is $2L + 1$ to describe an L-qubit system. The detailed protocol is described in Section 2.1.

2.1 Protocol

1. Initialise the total system state $|\Psi^0\rangle$ as a product of system state $|\psi_0\rangle_S$, reference $|\Phi\rangle_M$ and control qubit $|+\rangle_C$.

2. Perform a single-qubit gate $\hat{R}(\chi)$ on $|+\rangle_C$ and system unitary evolution operator $\hat{U}(t)$ on $|\psi_0\rangle_S$ to produce $|\Psi^1\rangle$.

3. Apply controlled-operator gate $\hat{N}$ between control C and system S as well as two controlled-operator gates $\hat{A}$ and $\hat{M}$ between C and reference M to produce $|\Psi^2\rangle$.

4. Apply a block controlled-SWAP gate from control C between system S and reference M to produce $|\Psi^3\rangle$ [8].

5. Apply a Hadamard gate $\hat{H}$ to control C to produce $|\Psi^4\rangle$.

6. Measure a single qubit in C in the Pauli-Z gate (the computational) basis, to obtain the expectation value $\langle\hat{Z}\rangle$.

2.2 Quantum commutation simulator

Using qubit terminology, we first explain the operation of the generalised quantum commutation simulator, shown in Fig. 1. Generally, by default, qubits are assumed to be initialised in $|0\rangle$ for a quantum circuit, but here we assume some additional preparation. The control qubit C is prepared in the state $|+\rangle_C = \hat{H}|0\rangle_C$ using a Hadamard gate $\hat{H}$. As shown in Fig. 1, there are two L-qubit states, for the system S and the reference M. For the system S, the initial state $|\psi_0\rangle$ is assumed to be created by a suitable prior quantum circuit, specified by the chosen initial conditions of the target problem to be simulated at $t = 0$. For the reference M, we can simply utilise one of the computational basis states (e.g., $|0\rangle^{\otimes L}$), or any other interesting reference state $|\Phi\rangle$ to be evolved dynamically. Note that only the control qubit C is measured at the end of the process and thus the outcome provides us with the expectation value of a quantum operator for the other degrees of freedom, effectively given by a $2^L \times 2^L$ matrix for each of system S and reference M.

At the end, ensemble measurement of the control qubit C in the computational basis will generate the probabilities of the outcomes $|0\rangle_C$ and $|1\rangle_C$, defined respectively as P_0 and P_1. The expectation value of Pauli operator $\hat{Z}$ is equal to $\left\langle\Psi^4\middle|\hat{Z}\middle|\Psi^4\right\rangle = P_0 - P_1$. Since the results are given as a difference of scalar values (details in arXiv:2206.00591), we can interchange these to reformulate the expectation value of $\hat{Z}$ for qubit C as

$$\left\langle\Psi^4\middle|\hat{Z}\middle|\Psi^4\right\rangle \equiv \langle\Phi|\hat{Z}_A^{\chi}|\Phi\rangle, \tag{4}$$

defining a new quantum operator as

$$\hat{Z}_A^{\chi} = \frac{1}{2}\left(e^{i\chi}\hat{N}\hat{\rho}(t)\hat{M}\hat{A} + e^{-i\chi}\hat{A}^{\dagger}\hat{M}^{\dagger}\,\hat{\rho}(t)\,\hat{N}^{\dagger}\right). \tag{5}$$

Note that this new operator contains actions of the controlled gates $\hat{A}$ and $\hat{M}$, in a manner that depends on the chosen rotation angle χ.

As an example, for an identity $\hat{N} = \hat{A} = \hat{1}$ and $\hat{M}$ being a Hermitian operator $\hat{M}^\dagger = \hat{M}$, the value of χ can then determine whether the result delivers the expectation value of the commutation or anti-commutation relation between the time-dependent density matrix $\hat{\rho}(t)$ and the operator $\hat{M}$. These follow from the statistics of single-qubit measurements through

$$\left\langle \Phi \left| \left\{ \hat{\rho}(t), \hat{M} \right\} \right| \Phi \right\rangle = 2 \langle \Phi | \hat{Z}_{\hat{1}}^{0} | \Phi \rangle, \quad (6)$$

$$i \left\langle \Phi \left| \left[\hat{\rho}(t), \hat{M} \right] \right| \Phi \right\rangle = 2 \langle \Phi | \hat{Z}_{\hat{1}}^{\pi/2} | \Phi \rangle \, . \quad (7)$$

For $\hat{N} \neq \hat{1}$ and $\hat{A} \neq \hat{1}$, we are further able to utilise the outcome of the expectation value to evaluate both $\Re\left(\left\langle \Phi \left| \hat{N}\hat{\rho}(t)\hat{M} \right| \Phi' \right\rangle\right) = \langle \Phi | \hat{Z}_{A}^{0} | \Phi \rangle$ and $\Im\left(\left\langle \Phi \left| \hat{N}\hat{\rho}(t)\hat{M} \right| \Phi' \right\rangle\right) = -\langle \Phi | \hat{Z}_{A}^{\pi/2} | \Phi \rangle$, for $|\Phi'\rangle = \hat{A}|\Phi\rangle$, where $\Re()$ and $\Im()$ represent real and imaginary parts respectively.

3 Summary and Remarks

In summary, we have proposed a new quantum algorithm to simulate the dynamics of open and closed quantum systems in quantum circuits. Two interesting applications of this approach are investigations of: (i) steady states in a closed quantum system, via the von Neumann equation; and (ii) decoherence mechanisms in an open quantum system, via the Lindblad equation. For a large quantum system, the von Neumann method is beneficial for computing a transition rate between two specific quantum states and this result can be also used for the study of its open quantum system. For example, although the sizes of the system qubits state $|\psi_0\rangle$ and the reference qubits state $|\Phi\rangle$ is each given by a $2^L \times 1$ column vector for L qubits, the probability transition rate between two specific states is always given by the expectation value of the off-diagonal elements in a 2×2 matrix form. Correspondingly, the changes of each state probability follow from the diagonal elements.

Clearly in general Lindblad evolution is irreversible, with a change in mixture (entropy) of the density matrix. In the simulation approach this changing mixture is introduced because the outputs of different simulations have to be combined, and each of these simulations involve measurements, to compute the expectation values that comprise the various matrix elements of the (rate of change of the) density matrix.

Extensions of the Lindblad equation simulation method have the potential to generate innovative approaches for general purpose master equations. This is because the simulation approach preserves a probabilistic interpretation for the system even for open systems, generating the evolution of the system probabilities from the diagonal elements of a density. The investigation of possible quantum advantage in such open system applications will be an interesting topic for future study, for example to investigate the quantum speed limit of simulating an open quantum system [9, 10] and the study of quantum channels [11].

References

[1] Born M and Jordan P 1925 The 1925 Born and Jordan paper "On quantum mechanics" *Z. Phys* **34** 858

[2] Dirac P A M 1925 The fundamental equations of quantum mechanics *Proc. R. Soc. London, Ser. A* **109** 642

[3] Schrödinger E 1926 An undulatory theory of the mechanics of atoms and molecules *Phys. Rev.* **28** 1049

[4] von Neumann J 1927 *Göttinger Nachrichten* **245**

[5] Plenio M B and Knight P L 1998 *Rev. Mod. Phys.* **70** 101

[6] Lindblad G 1976 On the generators of quantum dynamical semigroups *Commun. Math. Phys.* **48** 119

[7] Gorini V, Kossakowski A and Sudarshan E C G 1976 Completely positive dynamical semigroups of N-level systems *Jour. of Math. Phys.* **17** 821

[8] Joo J and Moon H 2021 Quantum variational PDE solver with machine learning arXiv:2109.09216

[9] Funo K, Shiraishi N and Saito K 2019 Speed limit for open quantum systems *New J. Phys.* **21** 013006

[10] Van Vu T and Hasegawa Y 2021 Lower bound on irreversibility in thermal relaxation of open quantum systems *Phys. Rev. Lett.* **127** 190601

[11] Kai Wang K and Dong-Sheng Wang D-S 2023 Quantum circuit simulation of superchannels *New J. Phys.* **25** 043013

A Variational Approach to Unique Determinedness in Pure-state Tomography

Chao Zhang[1] * Xuanran Zhu[1] † Bei Zeng[1] ‡

[1] *Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China*

Abstract. In this study, we present a variational approach to address the unique determinedness (UD) problem in pure-state tomography. Our algorithm effectively differentiates between UD and non-UD measurement schemes by minimizing a specialized loss function, resulting in the discovery of optimal pure-state Pauli measurement schemes across various dimensions. We also observe an alignment between unique determinedness among pure states (UDP) and all states (UDA) in qubit systems. This work advances our understanding of UD in quantum state tomography, providing valuable insights for experimental applications while emphasizing the importance of balancing mathematical optimality with practical considerations.

Keywords: Pure-state tomography, Unique determinedness, Variational approach

[1]Quantum state tomography (QST) is a pivotal technique in quantum information science since it enables the accurate reconstruction and characterization of quantum states [1, 2, 3, 4]. As an essential tool in quantum devices and protocols, QST has far-reaching implications in various domains, including quantum computing [5, 6, 7], quantum communication [8, 9, 10], and quantum cryptography [11, 12, 13].

General QST necessitates d^2 measurement outcomes to recover an arbitrary d-dimensional state. Several positive operator-valued measure (POVM) schemes, such as symmetric informationally complete POVM [14] and mutually unbiased bases POVM [15], offer satisfactory state recovery. As for many-body systems, it has been demonstrated that a minimum of 3^n separable projective measurement settings are required for n-qubit systems [16]. This number can be reduced to $2^n + 1$ by allowing nonseparable measurements [17]. However, these measurement schemes for general QST can be prohibitively costly for experimental implementation due to the exponential complexity.

In the realm of quantum information tasks, there is a strong emphasis on pure states, stemming from both theoretical and experimental interests. The former can be traced back to the famous Pauli's problem [18], which questions whether the position and the momentum distributions can uniquely determine the wave function, while the latter leads to various experimental realizations [19, 20, 21, 22]. The presence of prior information can significantly reduce the number of measurements needed to achieve state recovery, giving rise to a new concept known as pure-state tomography, which can also be extended to rank-r states [23, 24, 25, 26] or matrix product states [27].

In this study, we primarily focus on the problem about unique determinedness (UD) of pure states, given the specific measurement scheme $\mathbf{A}$ consists of observables $\{A_0 = \mathbb{I}, A_1, A_2, ..., A_m\}$. A measurement scheme $\mathbf{A}$ is classified as UD if any pure state $|\psi\rangle$ is uniquely determined among pure states (UDP) or among all states (UDA) by measuring the given observables, i.e., any other pure state or mixed state cannot have the same measurement results as those of $|\psi\rangle$. These definitions are consistent with the notions of (strictly) informationally complete measurements [28, 14, 29].

Necessary and sufficient conditions for UD measurement have been established by investigating the eigenvalue structure of the orthogonal space with respect to $\mathbf{A}$. It is known that one condition for a UDP measurement scheme is that any nonzero Hermitian operator belonging to the orthogonal space has at least 2 nonzero eigenvalues. In the case of UDA, the condition requires the operator to have at least 2 negative and 2 positive eigenvalues. This analysis led to the discovery of a gap between UDA and UDP concerning the number of required observables [23, 30]. A similar gap is also present for projective measurements [31, 32]. Furthermore, experimental aspects have been examined, such as the stability of state recovery comparing UDP and UDA [21, 33].

However, systematically constructing UD measurement schemes proves to be quite challenging, primarily due to the complexity of verifying UD's conditions about the eigenvalue structure. Here, we propose an effective algorithm to determine whether a given measurement scheme is UD by minimizing a suitably defined loss function. Specifically, we construct a variational Hermitian matrix that lacks the aforementioned eigenvalue structure to search for counterexamples that violate those UD conditions. The loss function is then defined associating with the orthogonality between our variational matrix and measurement space of $\mathbf{A}$. If a counterexample is found, the minimum loss approaches a value close to zero, subject to calculational precision. Conversely, if no counterexample is present, the loss function yields a distinct nonzero value.

In n-qubit systems, the numerical results show a clear

*czhangdh@connect.ust.hk
†xzhube@connect.ust.hk
‡zengb@ust.hk

[1]The full paper is available on arXiv https://arxiv.org/abs/2305.10811

gap in minimized loss between UD and non-UD measurement schemes using Pauli measurement. The former exhibits a discernibly non-zero value, while the latter approaches zero within the machine precision. Consequently, we can set a threshold δ based on the non-UD's minimum loss to determine whether a minimized loss is effectively zero or non-zero; if the value is above δ, we can regard the scheme as UD.

Furthermore, with the assistance of random sampling techniques, we successfully identify numerous optimal pure-state Pauli measurement schemes across various dimensions, including previous minimum operator sets for $2, 3$-qubit UDA Pauli measurement [20] (see Table 1). Here, "optimal" means the size of the operator set is a local (global) minimum through our search algorithm. Intriguingly, our findings reveal that in qubit systems, UDP invariably aligns with UDA when employing Pauli measurements, a phenomenon not commonly observed in other contexts. This insight implies that UD Pauli measurement schemes intrinsically possess a convex property owned by UDA, shedding light on the underlying mechanism of robustness in quantum compressed sensing [33].

Table 1: UDP/UDA scheme with Pauli measurements. The column $m \times n$ denotes we find m different UDP/UDA Pauli measurement schemes with n Pauli operators (including the identity), which could be Clifford equivalent. The minimized loss function $\mathcal{L}$ is evaluated for the scheme with the least operators.

#qubits	UDA $\mathcal{L}$	UDP $\mathcal{L}$	$m \times n$
2	1	2	$6 \times 11, 19 \times 13$
3	0.519	2	$176 \times 31, 258 \times 32$
4	0.280	1.788	$3 \times 106, 14 \times 107$
5	0.202	1.951	$1 \times 393, 1 \times 395$

Nevertheless, our numerical results also indicate that for most types of optimal UD measurement schemes, the non-zero minimum loss tends to decay toward zero as the dimension increases. This trend is also observed in Pauli measurements, albeit not significantly. In other words, the gap between non-UD and UD schemes in terms of their minimum losses becomes increasingly less distinct, which poses the challenge of finding optimal schemes in higher dimensions. It inspires us to explore what the decay of the minimum loss signifies and what impact it may have.

We successfully establish a connection between our loss function and the stability of UD measurement schemes. Firstly, we prove that the systematic noise error allowed for a specific UD scheme is bounded by the square root of minimum loss. When the noise error exceeds that threshold, the measurement scheme can no longer qualify as UD. Secondly, we define a stability coefficient related to the fidelity of the reconstructed state, which exhibits an inverse relationship with the square root of the minimum loss. A larger stability coefficient corresponds to lower constructed fidelity in the presence of noise. These two observations suggest that an optimal UD measurement scheme with a lower minimum loss is more vulnerable to noise and results in suboptimal state recovery. revealing a clear trade-off between mathematical optimality and experimental pragmatism in higher dimensions.

Taking the UD scheme constructed by polynomial bases [31, 32] as an example, we have verified that this d-dimensional optimal scheme displays minimum losses that decay towards zero considerably as the dimension increases, leading to instability in state recovery. We also identify the most vulnerable state during the recovery process. Although for random states, the state recovery remains stable against noise, in the worst case, the fidelity deteriorates considerably as the dimension increases, e.g., the fidelity drops to around 0.5 for $d = 9$. For comparison, we also investigate the stability of the optimal UD schemes constructed by Pauli operators (in Table 1), which exhibits a less pronounced decay in the minimum loss. This kind of UD scheme yields perfect state recovery, even in the worst case.

In contrast to the unextendible mathematical techniques commonly employed in other works, we introduce a variational approach to investigate the specific eigenvalue structure of a certain matrix space. This approach applies to various measurement schemes and can even be extended to measurement settings [34], determining whether a given measurement scheme fulfills the conditions for pure-state tomography. In lower dimensions, this method can be utilized to construct optimal UD measurement schemes with relatively large minimum losses. However, in higher dimensions, the minimum loss for optimal measurement schemes inevitably decays toward zero, posing a potential risk of instability. Experimentalists can establish a reasonable threshold δ based on the instruments' noise levels to identify suitable measurement schemes for pure-state tomography, even if they may not be mathematically optimal.

Our study not only propels the understanding of UD in quantum state tomography forward but also delivers valuable practical insights for experimental applications, highlighting the need for a balanced approach between mathematical optimality and experimental pragmatism. The applicability of the proposed algorithm transcends pure-state tomography, extending to other quantum information tasks, such as entanglement detection [35] and super-activation [36]. This method holds the potential to develop into a comprehensive framework for examining the unique structure of a space, which could consist of desired elements like operators, matrices, and states.

References

[1] Matthias Christandl and Renato Renner. "Reliable quantum state tomography". In: *Physical Review Letters* 109.12 (2012), p. 120403.

[2] Alexander I Lvovsky and Michael G Raymer. "Continuous-variable optical quantum-state tomography". In: *Reviews of modern physics* 81.1 (2009), p. 299.

[3] Rob T Thew et al. "Qudit quantum-state tomography". In: *Physical Review A* 66.1 (2002), p. 012303.

[4] Bo Qi et al. "Quantum state tomography via linear regression estimation". In: *Scientific reports* 3.1 (2013), pp. 1–6.

[5] Andrew Steane. "Quantum computing". In: *Reports on Progress in Physics* 61.2 (1998), p. 117.

[6] Jeremy L O'brien. "Optical quantum computing". In: *Science* 318.5856 (2007), pp. 1567–1570.

[7] John Preskill. "Quantum computing in the NISQ era and beyond". In: *Quantum* 2 (2018), p. 79.

[8] Nicolas Gisin and Rob Thew. "Quantum communication". In: *Nature photonics* 1.3 (2007), pp. 165–171.

[9] Daniele Cozzolino et al. "High-dimensional quantum communication: benefits, progress, and future challenges". In: *Advanced Quantum Technologies* 2.12 (2019), p. 1900038.

[10] Gilles Brassard. "Quantum communication complexity". In: *Foundations of Physics* 33 (2003), pp. 1593–1616.

[11] Nicolas Gisin et al. "Quantum cryptography". In: *Reviews of modern physics* 74.1 (2002), p. 145.

[12] Stefano Pirandola et al. "Advances in quantum cryptography". In: *Advances in optics and photonics* 12.4 (2020), pp. 1012–1236.

[13] Charles H Bennett, Gilles Brassard, and Artur K Ekert. "Quantum cryptography". In: *Scientific American* 267.4 (1992), pp. 50–57.

[14] Joseph M Renes et al. "Symmetric informationally complete quantum measurements". In: *Journal of Mathematical Physics* 45.6 (2004), pp. 2171–2180.

[15] William K Wootters and Brian D Fields. "Optimal state-determination by mutually unbiased measurements". In: *Annals of Physics* 191.2 (1989), pp. 363–381.

[16] Mark D De Burgh et al. "Choice of measurement sets in qubit tomography". In: *Physical Review A* 78.5 (2008), p. 052122.

[17] RBA Adamson and Aephraim M Steinberg. "Improving quantum state estimation with mutually unbiased bases". In: *Physical review letters* 105.3 (2010), p. 030406.

[18] Wolfgang Pauli. *Die allgemeinen prinzipien der wellenmechanik*. Springer, 1933.

[19] Tao Xin et al. "Quantum pure state tomography via variational hybrid quantum-classical method". In: *Physical Review Applied* 13.2 (2020), p. 024013.

[20] Xian Ma et al. "Pure-state tomography with the expectation value of Pauli operators". In: *Physical Review A* 93.3 (2016), p. 032140.

[21] H Sosa-Martinez et al. "Experimental study of optimal measurements for quantum state tomography". In: *Physical Review Letters* 119.15 (2017), p. 150401.

[22] Wei-Tao Liu et al. "Experimental quantum state tomography via compressed sampling". In: *Physical review letters* 108.17 (2012), p. 170403.

[23] Teiko Heinosaari, Luca Mazzarella, and Michael M Wolf. "Quantum tomography under prior information". In: *Communications in Mathematical Physics* 318.2 (2013), pp. 355–374.

[24] Claudio Carmeli et al. "Tasks and premises in quantum state determination". In: *Journal of Physics A: Mathematical and Theoretical* 47.7 (2014), p. 075302.

[25] Michael Kech and Michael M Wolf. "Constrained quantum tomography of semi-algebraic sets with applications to low-rank matrix recovery". In: *Information and Inference: A Journal of the IMA* 6.2 (2017), pp. 171–195.

[26] Charles H Baldwin, Ivan H Deutsch, and Amir Kalev. "Strictly-complete measurements for bounded-rank quantum-state tomography". In: *Physical Review A* 93.5 (2016), p. 052105.

[27] Marcus Cramer et al. "Efficient quantum state tomography". In: *Nature communications* 1.1 (2010), p. 149.

[28] Steven T Flammia, Andrew Silberfarb, and Carlton M Caves. "Minimal informationally complete measurements for pure states". In: *Foundations of Physics* 35 (2005), pp. 1985–2006.

[29] Andrew J Scott. "Tight informationally complete quantum measurements". In: *Journal of Physics A: Mathematical and General* 39.43 (2006), p. 13507.

[30] Jianxin Chen et al. "Uniqueness of quantum states compatible with given measurement results". In: *Phys. Rev. A* 88 (1 July 2013), p. 012109. DOI: 10.1103/PhysRevA.88.012109. URL: https://link.aps.org/doi/10.1103/PhysRevA.88.012109.

[31] Claudio Carmeli et al. "How many orthonormal bases are needed to distinguish all pure quantum states?" In: *The European Physical Journal D* 69 (2015), pp. 1–11.

[32] Claudio Carmeli et al. "Stable pure state quantum tomography from five orthonormal bases". In: *Europhysics Letters* 115.3 (2016), p. 30001.

[33] David Gross et al. "Quantum state tomography via compressed sensing". In: *Physical review letters* 105.15 (2010), p. 150401.

[34] Jun Li et al. "Optimal design of measurement settings for quantum-state-tomography experiments". In: *Physical Review A* 96.3 (2017), p. 032307.

[35] Otfried Gühne and Géza Tóth. "Entanglement detection". In: *Physics Reports* 474.1-6 (Apr. 2009), pp. 1–75. DOI: 10.1016/j.physrep.2009.02.004. URL: https://doi.org/10.1016%2Fj.physrep.2009.02.004.

[36] Runyao Duan. "Super-activation of zero-error capacity of noisy quantum channels". In: *arXiv preprint arXiv:0906.2527* (2009).

Optimization of Grover's Search Algorithm using ZX-calculus

Natchapol Patamawisut[1] * Wanchai Pijitrojana[2] † Ruchipas Bavontaweepanya[3] ‡

[1] *Quantum Computing and Information Research Centre, King Mongkut's University of Technology Thonburi, Thailand*

[2] *Department of Electrical and Computer Engineering, Thammasat School of Engineering, Thailand*

[3] *Division of Physics, Faculty of Science and Technology, Thammasat University Research Unit in Quantum Technology, Thammasat University, Thailand*

Abstract. The implementation of Grover's Search, on real quantum devices is impeded by significant challenges arising from a high gate count. Grover's algorithm is particularly susceptible to such issues, which can lead to computation inaccuracies due to quantum decoherence. In this paper, we address this problem by leveraging the mathematical framework of the ZX-calculus to optimize Grover's Search algorithm. The ZX-calculus offers an effective and graphical language for reasoning about quantum circuits, enabling the simplification and optimization of quantum algorithms. Our study focuses on the application of ZX-calculus for optimizing the quantum gate count of the diffuser part of Grover's Search algorithm.

Keywords: Quantum computing, Quantum algorithm, ZX-calculus

1 Introduction

Grover's algorithm is an important quantum search algorithm that guarantees quadratic speed-up in unstructured searches over classical methods. This quantum algorithm works on amplitude amplification which allows finding a targeted element in $\sqrt{N}$ iterations. In principle, Grover's algorithm requires an oracle to mark the target element and a diffusion operator to amplify the amplitude of the marked element.

There are many applications that use Grover's algorithm to solve problems such as finding the minimum of an unsorted list [1], image pattern matching [2] or string matching [3]. However, when the problems become more complex, Grover's algorithm needs more qubits and quantum gates to operate the amplitude amplification, leading to expensive implementation.

This study aims to study single-qubit gate and two-qubit gate reduction in the diffusion part of Grover's algorithm by using ZX calculus. We use the PyZX Python library to perform ZX calculus. We also compare the results of circuit reduction using ZX calculus and those using the transpilers from QISKIT.

2 Background

Grover's search algorithm: The algorithm, invented by Lov Grover in 1996 [4, 5], is a quantum algorithm that significantly improves the efficiency of unstructured searches. It locates a specific item in a list of N items in roughly $\sqrt{N}$ queries, compared to N queries in a classical setting. It achieves this by using a quantum operator to amplify the probability of the target item, resulting in a high likelihood of locating the item after approximately $\sqrt{N}$ iterations. This algorithm runs on amplitude amplification, which marks the target element and amplifies the marked element.

ZX-calculus: The ZX-calculus is a graphical language [6] developed for simplifying and automating the reasoning about quantum circuits and protocols. In ZX calculus language, the typical primitive of the quantum circuit is consist of CNOT gate, Hadamard gate, Z-phase gate, and X-phase gate. The qubits in the primitive circuit will be mapped to the wires, and the quantum gates will be decomposed to the boxes in ZX language. With ZX rules, we can connect the diagram to reduce the quantum gates in the primitive circuit.

3 Methodology

In this research, we reduced the quantum gates in the diffusion part of Grover's algorithm by using ZX calculus. The diffusion operator contains the Toffoli gate, which can be decomposed into multiple single-qubit gates and two-qubit gates (CNOT gate, Hadamard gate, T gate, and RZ gate). The decomposed circuit will be converted to ZX language by using The PyZX package. We examine the algorithm's complexity and performance, particularly in the count of single-qubit gates and two-qubit gates with different numbers of qubits, ranging from 3 to 10 qubits. We compared the results from ZX calculus with those from the transpiler at level 1 to level 3 in QISKIT package.

4 Result and Discussion

The results indicate that ZX-calculus was more effective in reducing the number of single-qubit gates in the diffusion operator than Qiskit's transpiler at all optimization levels. However, when it comes to the optimization of two-qubit gates, Qiskit's transpiler showed better performance at every level, consistently optimizing the number of two-qubit gates, than ZX calculus. This suggests that ZX calculus is effective in reducing single-qubit gates in a diffusion operator, but it has low performance in two-qubit gates reduction.

*natchapol.pat@gmail.com
†pwanchai@engr.tu.ac.th
‡ruchipas@tu.ac.th

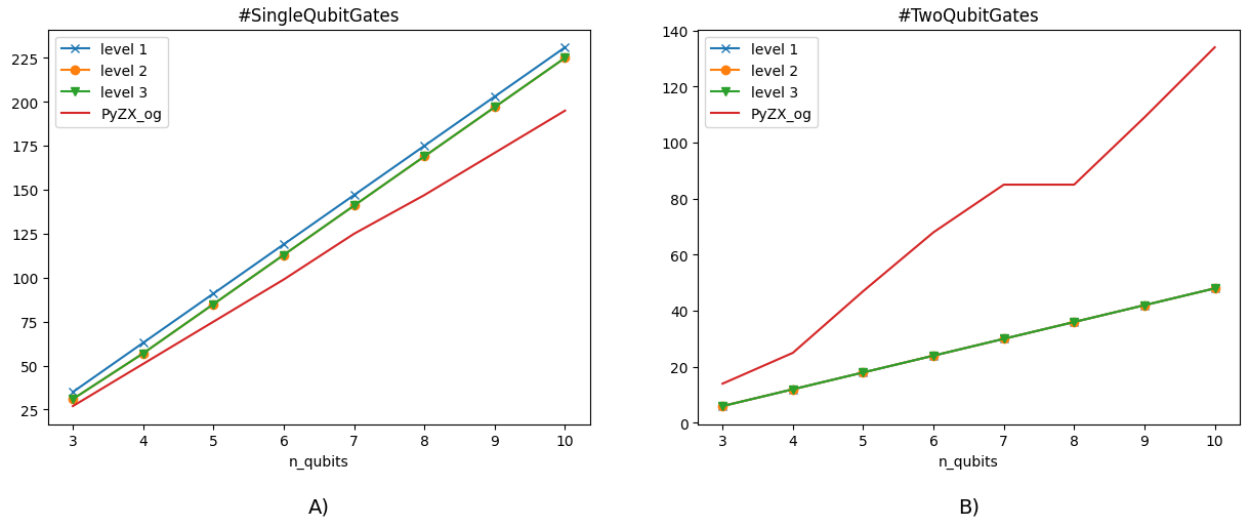


Figure 1: Comparison between Qiskit's transpiler at levels 1-3 and PyZX optimized circuits in terms of A) the number of single-qubit gates and B) the number of two-qubit gates

References

[1] Dürr, C. & Høyer, P. (1996). A quantum algorithm for finding the minimum. Preprint at quant-ph/9607014.

[2] Tezuka, Hiroyuki and Nakaji, Kouhei and Satoh, Takahiko and Yamamoto, Naoki (2022). Grover search revisited: Application to image pattern matching. Phys. Rev. A 105(3), 032440. https://link.aps.org/doi/10.1103/PhysRevA.105.032440

[3] Niroula, P., Nam, Y. (2021). A quantum algorithm for string matching. npj Quantum Inf 7, 37. https://doi.org/10.1038/s41534-021-00369-3

[4] Grover, Lov K. (1996). A Fast Quantum Mechanical Algorithm for Database Search. Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing, Philadelphia, Pennsylvania, USA, 212–219. https://doi.org/10.1145/237814.237866

[5] Nielsen, M. A., Chuang, I. L. (2010). Quantum Computation and Quantum Information: 10th Anniversary Edition. Cambridge University Press.

[6] Coecke, B., Kissinger, A. (2017). Picturing Quantum Processes. A First Course in Quantum Theory and Diagrammatic Reasoning. Cambridge University Press.

Detection of Beyond-Quantum Non-locality based on Standard Local Quantum Observables

(The full paper is available as [53].)

Hayato Arai[1] * Baichu Yu[2 3] † Masahito Hayashi[4 3 1] ‡

[1] *Graduate School of Mathematics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8602, Japan*
[2] *Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Nanshan District, Shenzhen, 518055, China*
[3] *International Quantum Academy (SIQA), Shenzhen 518048, China*
[4] *School of Data Science, The Chinese University of Hong Kong, Shenzhen, Longgang District, Shenzhen, 518172, China*

Abstract. Device independent detections of quantum non-locality like Bell-CHSH inequality are important methods to detect quantum non-locality because the whole protocol can be implemented by uncertified local observables. However, this detection is not sufficient for the justification of standard quantum theory, because there are theoretically many types of beyond-quantum non-local states in General Probabilistic Theories. One important class is Entanglement Structures (ESs), which contain beyond-quantum non-local states even though their local systems are completely equivalent to standard quantum systems. This paper shows that any device independent detection cannot distinguish beyond-quantum non-local states from standard quantum states. To overcome this problem, this paper gives a device dependent detection based on local observables to distinguish any beyond-quantum non-local state from all standard quantum states. Especially, we give a way to detect any beyond-quantum non-local state in two-qubit ESs by observing only spin observables on local systems.

Keywords: beyond-quantum state, general probabilistic theories, Bell's inequality, device-independent detection

Introduction—Bell's inequality [1] (or CHSH inequality [2]) is one of the important ways to detect quantum non-locality in our physical systems. Bell-CHSH inequality (hereinafter, CHSH inequality) consists of bipartite players and their local operations. It is especially important that the protocol of CHSH inequality can be implemented by local observables. In other words, by implementing the protocol of CHSH inequality as a bipartite communication task, we can experimentally detect quantum non-locality of our physical systems when Bell-CHSH inequality is violated. Actually, the violation of CHSH inequality is confirmed in physical experiments [3, 4, 5, 6, 7, 8]. Moreover, CHSH inequality can be implemented without certification of measurement devices. Such detection without certification of measurement devices is called *device independent* (DI) detection [10, 11, 12, 13, 14, 15, 16, 17, 18]. These remarkable results played an important role in the early studies of quantum physics and quantum information theory to ensure that our physical systems truly possess quantum non-locality.

However, it is not sufficient for the strict verification of quantum theory to detect standard quantum non-locality because there are many other theoretical models with non-locality than quantum systems. Such models can be described as General Probabilistic Theories (GPTs) [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 42, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 45]. GPT is a framework for general theoretical models with states and measurements, including classical and quantum theories. Important models are called *Entanglement Structures* (ESs) with local quantum subsystems [32, 37, 38, 39, 40, 45], including not only the Standard Entanglement Structure (SES), i.e., the standard quantum model defined by the tensor product but also many other models. Some ESs have fewer non-local states than the SES [37, 39], and also, some ESs has beyond-quantum non-local states, i.e., non-local states that do not belong to the SES [42, 38, 40, 45]. In order to ensure that our physical systems obey truly standard quantum theory, it is also necessary to verify whether beyond-quantum non-local states exist or not. However, preceding studies [43, 44, 45] have revealed that all ESs satisfy Tirelson's bound, i.e., CHSH inequality cannot distinguish the SES from any beyond-quantum non-local state in ESs.

Furthermore, as we show in this paper, not only CHSH inequality, but also any DI detection cannot distinguish any beyond-quantum non-local state from the SES. Although a similar statement was shown by the reference [18], this paper also shows a corresponding statement in our setting with GPTs as Theorem 2. Therefore, this paper deals with a device dependent detection of an arbitrary beyond-quantum non-local state in ESs by an experimental protocol. First, we give a device dependent detection separating an arbitrary given beyond-quantum state from all standard quantum states as an inequality defined by local observables (Theorem 3). Next, we give a bipartite protocol to implement the above detection. Our protocol consists of local operations by bipartite players Alice and Bob and classical communication by them. In the protocol, Alice and Bob detect whether a target

*m18003b@math.nagoya-u.ac.jp
†yubc@sustech.edu.cn
‡hmasahito@cuhk.edu.cn

state is beyond-quantum or not. If the target state is truly beyond-quantum, Alice and Bob conclude that the target state is beyond-quantum with high probability.

Our criterion and protocol are implemented by a complicated sequence of local observables in general. However, in the 2-qubits case, we give a simple detection of a beyond-quantum non-local state by observing Pauli's spin observables in a specific order. Moreover, like Bell's scenario, any beyond-quantum non-local "pure" state can be detected by sequential local observables biased in the same way as $\sigma_x, \sigma_y, \sigma_z$ (Theorem 4). As a result, we give a convenient detection for beyond-quantum non-local pure states like Bell's inequality in the 2-qubits case.

Settings and Definition of Beyond-Quantum States—As a preliminary, we denote the set of Hermitian matrices and the set of positive semi-definite matrices on a finite dimensional Hilbert space $\mathcal{H}$ as $\mathcal{L}_{\rm H}(\mathcal{H})$ and $\mathcal{L}^+_{\rm H}(\mathcal{H})$, respectively.

This paper deals with bipartite composite models of GPTs whose local systems are equal to standard quantum systems. In the setting of GPTs, there exist infinitely many such composite models, and they called *Entanglement Structures* (ESs). In ESs, we can regard certain non-positive Hermitian matrices as states in the setting of GPTs.

Definition 1 (Beyond-Quantum States) *We say that a Hermitian matrix $\rho \in \mathcal{L}_{\rm H}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is a beyond-quantum state if $\rho \in {\rm SEP}^*(A;B) \setminus {\rm SES}(A;B)$ and ${\rm Tr}\,\rho = 1$, where the sets ${\rm SES}(A;B)$ and ${\rm SEP}^*(A;B)$ are defined as*

$$ {\rm SES}(A;B) := \mathcal{L}^+_{\rm H}(\mathcal{H}_A \otimes \mathcal{H}_B), \tag{1} $$

$$ {\rm SEP}^*(A;B) := \Big\{ x \in \mathcal{L}_{\rm H}(\mathcal{H}_A \otimes \mathcal{H}_B) \Big| \, {\rm Tr}\, xy \geq 0 \ \forall y \in \mathcal{L}^+_{\rm H}(\mathcal{H}_A) \otimes \mathcal{L}^+_{\rm H}(\mathcal{H}_B) \Big\}. \tag{2} $$

We denote the set of all beyond-quantum states and all standard quantum states as $\mathcal{S}({\rm SEP}^*(A;B))$ and $\mathcal{S}({\rm SES}(A;B))$, respectively.

Here, we simply give the definition of beyond-quantum states, but the definition is derived from the standard setting of GPTs and local quantum structures [21, 32, 38, 42]. You can check the detailed setting in our preprint [53].

Except for beyond-quantum states, this paper deals with standard objects in local standard quantum systems, for example, local Positive Operator Valued Measures (POVMs) and local standard quantum observables. Then, our interest is how we detect beyond-quantum states by local operations if they exist.

Impossibility of Device-Independent Detection for Beyond-Quantum States—First, we consider the possibility of the *Device-Independent* (DI) detection of a beyond-quantum state (Figure 1). In the device-independent detection, we have no certificate of measurement devices. Therefore, it is natural to consider that a beyond-quantum state ρ_0 is distinguished device-independently by local measurements $M^A_a := \{M^A_{a;i}\}_{i\in I}$ and $M^B_b := \{M^B_{b;j}\}_{j\in J}$ from all standard quantum states when no pair of a standard quantum state and local POVMs simulates the pair of the state ρ_0 and local POVMs $M^A_a := \{M^A_{a;i}\}_{i\in I}$ and $M^B_b := \{M^B_{b;j}\}_{j\in J}$, i.e., there does not exist a pair of a standard quantum state $\rho_1 \in \mathcal{S}({\rm SES}(A;B))$ and local POVMs $M^A_a := \{M^{A\prime}_{a;i}\}_{i\in I}$ and $M^B_b := \{M^{B\prime}_{b;j}\}_{j\in J}$ such that the relation

$$ {\rm Tr}\, \rho_0 M^A_{a;i} \otimes M^B_{b;j} = {\rm Tr}\, \rho_1 M^{A\prime}_{a;i} \otimes M^{B\prime}_{b;j} \tag{3} $$

holds for any a, b, i, j. In other words, a beyond-quantum state ρ_0 is distinguished device-independently from all standard quantum states when there exist local measurements $M^A_a := \{M^A_{a;i}\}_{i\in I}$ and $M^B_b := \{M^B_{b;j}\}_{j\in J}$ to satisfy the above condition. Therefore, the above device-independent detectability is equivalent to the impossibility of the simulation by a pair of a standard quantum state and local POVMs.

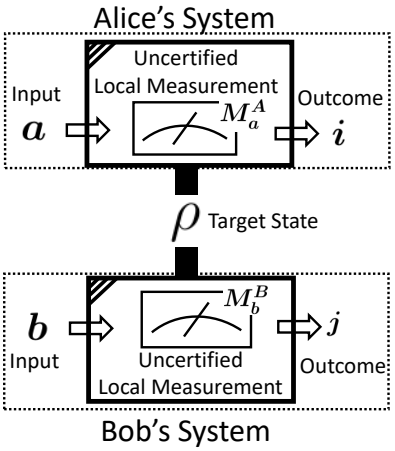


Figure 1: In DI detection, Alice and Bob apply uncertified local measurements $M^A_a := \{M^A_{a;i}\}_{i\in I}$ and $M^B_b := \{M^B_{b;j}\}_{j\in J}$ to a given non-local state ρ. Then, Alice and Bob determine whether ρ is beyond-quantum by the probability ${\rm Tr}\,\rho M^A_{a;i} \otimes M^B_{b;j}$.

However, previous studies [43, 44, 45] showed that CHSH inequality cannot detect any beyond-quantum states by noticing steering condition. Furthermore, the following theorem holds.

Theorem 2 *For any pair of a beyond-quantum state ρ_0 and local POVMs $M^A_a := \{M^A_{a;i}\}_{i\in I}$ and $M^B_b := \{M^B_{b;j}\}_{j\in J}$, there exists a pair of a standard quantum state $\rho_1 \in {\rm SES}(A;B)$ and local POVMs $M^A_a := \{M^{A\prime}_{a;i}\}_{i\in I}$ and $M^B_b := \{M^{B\prime}_{b;j}\}_{j\in J}$ to satisfy the condition (3).*

Although the reference [18] proved a similar statement, it does not formulate the problem with GPTs. Further, while the proof in [18] has a problem caused by an inverse of a key operator, our proof does not have such a problem because our proof is straightforward and different from that of the reference [18], as shown in our preprint [53, Appendix]. Due to Theorem 2, it is impossible to distinguish a beyond-quantum state from all

standard quantum states. To resolve this problem, instead of measurement devices without certification, we need to employ measurement devices that are identified with certifications. This problem setting is called device-dependent (DD) detection.

Device Dependent Detection of Beyond-Quantum State and Its Implementation—

Now, we discuss a DD detection of an arbitrary given beyond-quantum state in ESs. In the following analysis, instead of the joint distribution, as a simple indicator, we focus on the sum of an expectation of a function $f(a,i,b,j)$, i.e., $\sum_{a,i,b,j} f(a,i,b,j) \operatorname{Tr} \rho M^A_{a;i} \otimes M^B_{b;j}$ so that the magnitude relationship of this indicator makes the required discrimination. For our simple analysis, we assume $f(a,i,b,j) = f(a,i) f(b,j)$. Then, this value can be rewritten as

$$\begin{aligned} &\sum_{a,i,b,j} f(a,i,b,j) \operatorname{Tr} \rho M^A_{a;i} \otimes M^B_{b;j} \\ =& \sum_{a,b} \operatorname{Tr} \rho \mathcal{O}^A_a \otimes \mathcal{O}^B_b, \end{aligned} \tag{4}$$

where $\mathcal{O}^A_a := \sum_i f(a,i) M^A_{a;i}$, $\mathcal{O}^B_b := \sum_j f(b,j) M^B_{b;j}$. The Hermitian matrices $\mathcal{O}^A_a$ and $\mathcal{O}^B_b$ can be regarded as standard quantum observables with the POVMs $M^A_{a;i}$, $M^B_{b;j}$ and outcomes $f(a,i)$, $f(b,j)$, respectively. Therefore, the value $\operatorname{Tr} \rho \mathcal{O}^A_a \otimes \mathcal{O}^B_b$ corresponds to the expectation value of the standard quantum observable $\mathcal{O}^A_a \otimes \mathcal{O}^B_b$ with the state ρ. Hereinafter, we abbreviate the pair of POVMs and outcomes in the left-hand side of (4) to the right-hand side of (4) by using observables, according to this correspondence.

Based on the sum of the expectation of standard quantum local observables, the following theorem gives a DD detection of any beyond-quantum state from all standard quantum states.

Theorem 3 *Given an arbitrary state ρ_0, there exist families of local observables $\{\mathcal{O}^A_k\}_{k=1}^m$ and $\{\mathcal{O}^B_k\}_{k=1}^m$ and a real number α satisfying the following two properties:*

1. $\operatorname{Tr} \rho_0 \sum_{k=1}^m \mathcal{O}^A_k \otimes \mathcal{O}^B_k > \alpha$.
2. $\sup_{\rho_1 \in \mathrm{SES}(A;B)} \operatorname{Tr} \rho_1 \sum_{k=1}^m \mathcal{O}^A_k \otimes \mathcal{O}^B_k \le \alpha$.

The proof of Theorem 3 can be available in our preprint [53, Appendix], but we remark that we can find $\{\mathcal{O}^A_k \otimes \mathcal{O}^B_k\}_{k=1}^m$ and α by a deterministic way. Theorem 3 guarantees that the joint distribution with ρ_0 cannot be simulated by the joint distribution with any standard quantum state ρ_1 under the common local measurements. The above discussion can be understand in terms of the Semi-Definite Programing (SDP) with the target function $\operatorname{Tr} \rho_1 \sum_{k=1}^m \mathcal{O}^A_k \otimes \mathcal{O}^B_k$ and the trace 1 condition. The second relation in Theorem 3 shows that the solution of the SDP is upper bounded by α. The first relation in Theorem 3 states that ρ_0 attains a strictly larger value than the solution, and therefore, ρ_0 is not positive semi-definite, i.e., beyond-quantum.

In our preprint [53], we give a protocol to implement the detection given as Theorem 3 based on local standard observables. Therefore, any beyond-quantum state can be detected by a finite number of certified local quantum observables with large probability. In general cases, this detection require large costs because we need to certify a number of local quantum observables dependent on a target state. However, in the 2×2 dimensional case, we give a detection of any beyond-quantum pure state with the certification of only three observables of Pauli's spin observables.

Let us consider an arbitrary ES of two local quantum systems with dimension 2, i.e., we consider the case $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = 2$. First, we define the following function $\mathbb{A}_{\mathrm{Pauli}}(\rho; U_A, U_B)$ using Pauli's spin matrices:

$$\begin{aligned} &\mathbb{A}_{\mathrm{Pauli}}(\rho; U_A, U_B) \\ :=& \sum_{c=x,y,z} \operatorname{Tr} (U_A \otimes U_B) \rho \left(U_A^\dagger \otimes U_B^\dagger \right) \sigma_c \otimes \sigma_c, \end{aligned} \tag{5}$$

where U_A, U_B and $\sigma_x, \sigma_y, \sigma_z$ denote unitary matrices on $\mathcal{H}_A, \mathcal{H}_B$, Pauli's spin observables defined as

$$\begin{aligned} \sigma_x :=& \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_z :=& \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \end{aligned} \tag{6}$$

respectively.

The value $\mathbb{A}_{\mathrm{Pauli}}(\rho; U_A, U_B)$ detects all beyond-quantum pure states, i.e., all beyond-quantum states that cannot be written as any convex mixtures of other beyond-quantum states.

Theorem 4 *The following two properties hold:*

1. *For any beyond-quantum pure state ρ_0, there exist unitary matrices U_A, U_B such that $\mathbb{A}_{\mathrm{Pauli}}(\rho_0; U_A, U_B) > 1$.*
2. $\sup_{U_A, U_B : \text{unitary}} \mathbb{A}_{\mathrm{Pauli}}(\rho_1; U_A, U_B) \le 1$.
   $\rho_1 \in$ (SES$(A;B)$)

The proof of Theorem 4 can be available in our preprint [53, Appendix]. Theorem 4 implies that in 2×2 dimensional case, if a target state ρ is beyond-quantum pure, there exists a pair of unitary matrices U_A and U_B such that $\mathbb{A}_{\mathrm{Pauli}}(\rho; U_A, U_B)$ detects the beyond-quantum pure state ρ_0 from all standard quantum states ρ_1. If we can apply the unitary operations in the whole protocol, it is not necessary to certify the description of unitary matrices. In other words, we only need to certify the observables $\sigma_x, \sigma_y, \sigma_z$ for all detections of any beyond-quantum pure state.

acknowledgments

HA is supported by a JSPS Grant-in-Aids for JSPS Research Fellows No. JP22J14947. MH is supported in part by the National Natural Science Foundation of China (Grant No. 62171212).

References

[1] J. S. Bell, "On the Einstein Podolsky Rosen paradox." *Phys. Phys. Fiz.* **1**, 195 (1964).

[2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories." *Phys. Rev. Lett.* **23**, 880 (1970).

[3] S. J. Freedman and J. F. Clauser, "Experimental Test of Local Hidden-Variable Theories." *Phys. Rev. Lett.* **28**, 938 (1972).

[4] A. Aspect, J. Dalibard, and G. Roger, "Experimental Test of Bell's Inequalities Using Time-Varying Analyzers." *Phys. Rev. Lett.* **49**, 1804 (1982).

[5] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's Inequality under Strict Einstein Locality Conditions." *Phys. Rev. Lett.* **81**, 5039 (1998).

[6] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, ""Event-ready-detectors" Bell experiment via entanglement swapping." *Phys. Rev. Lett.* **71**, 4287 (1993).

[7] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe and D. J. Wineland, "Experimental violation of a Bell's inequality with efficient detection." *Nature* **409**, 791–794 (2001).

[8] T. Scheidl, R. Ursin, J. Kofler, and A. Zeilinger, "Violation of local realism with freedom of choice." *PNAS* **107**, 19708 (2010).

[9] B. S. Cirel'son, "Quantum generalizations of Bell's inequality." *Lett. Math. Phys.* **4**, 93–100 (1980).

[10] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner "Bell Nonlocality." *Rev. Mod. Phys.* **86**, 419 (2014).

[11] M. Navascués, S. Pironio, and A. Acín, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations." *New. J. Phys.* **10**, 073013 (2008).

[12] V. Scarani, "The device-independent outlook on quantum physics." *Acta Phys. Slovaca* **62**(4), 347-409 (2012).

[13] K. T. Goh, J. Kaniewski, E. Wolfe, et.al., "Geometry of the set of quantum correlations." *Phys. Rev. A* **97**(2), 022104 (2018).

[14] I. Šupić and J. Bowles, "Self-testing of quantum systems: a review." *Quantum* **4**, 337 (2020).

[15] A. Acín, N. Brunner, N. Gisin, et al., "Device-independent security of quantum cryptography against collective attacks." *Phys. Rev. Lett.* **98**(23), 230501 (2007).

[16] S. Pironio, A. Acín, N. Brunner, et al, "Device-independent quantum key distribution secure against collective attacks." *New J. Phys.* **11**(4), 045021 (2009).

[17] U. Vazirani and T. Vidick. "Fully device independent quantum key distribution." *Commun. ACM* **62**(4), 133-133 (2019).

[18] H. Barnum, S. Beigi, S. Boixo, et.al., "Local quantum measurement and no-signaling imply quantum correlations." *Phys. Rev. Lett.*, **104**(14) 140401 (2010).

[19] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom." *Found. Phys.* **24**, 379 (1994).

[20] M. Plávala and M. Ziman, "Popescu-Rohrlich box implementation in general probabilistic theory of processes." *Phys. Rett. A* **384**, 126323 (2020).

[21] M. Plavala, "General probabilistic theories: An introduction." *arXiv*:2103.07469, (2021).

[22] M. Pawłowski., T. Patere., D. Kaszlikowski, et al., "Information causality as a physical principle." *Nature* **461**, 1101–1104 (2009).

[23] A. J. Short and S. Wehner, "Entropy in general physical theories." *New J. Phys.* **12**, 033023 (2010).

[24] H. Barnum, J. Barrett, L. O. Clark, et.al., "Entropy and Information Causality in General Probabilistic Theories." *New J. Phys.* **14**, 129401 (2012).

[25] J. Barrett, "Information processing in generalized probabilistic theories." *Phis. Rev. A* **75**, 032304 (2007).

[26] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Probabilistic theories with purification." *Phys. Rev. A* **81**, 062348 (2010).

[27] G. Chiribella, C. M. Scandolo, "Operational axioms for diagonalizing states." *EPTCS* **195**, 96-115 (2015).

[28] G. Chiribella, C. M. Scandolo, "Entanglement as an axiomatic foundation for statistical mechanics." *arXiv*:1608.04459 (2016).

[29] M. P. Müller and C. Ududec, "Structure of Reversible Computation Determines the Self-Duality of Quantum Theory." *PRL* **108**, 130401 (2012).

[30] H Barnum, C. M. Lee, C. M. Scandolo, J. H. Selby, "Ruling out Higher-Order Interference from Purity Principles." *Entropy* **19**, 253 (2017).

[31] H. Barnum and J. Hilgert, "Strongly symmetric spectral convex bodies are Jordan algebra state spaces." *arXiv*:1904.03753 (2019).

[32] P. Janotta and H. Hinrichsen, "Generalized probability theories: what determines the structure of quantum theory?" *J. Phys. A: Math. Theor.* **47**, 323001 (2014).

[33] M. Krumm, H. Barnum, J. Barrett, and M. P. Müller, "Thermodynamics and the structure of quantum theory." *New J. Phys.* **19**, 043025 (2017).

[34] K. Matsumoto and G. Kimura, "Information storing yields a point-asymmetry of state space in general probabilistic theories." *arXiv*:1802.01162 (2018).

[35] R. Takagi and B. Regula, "General Resource Theories in Quantum Mechanics and Beyond: Operational Characterization via Discrimination Tasks." *Phys. Rev. X* **9**, 031053 (2019).

[36] R. Takakura, K. morisue, I. Watanabe, and G. Kimura, "Trade-off relations between measurement dependence and hiddenness for separable hidden variable models." *arXiv*:2208.13634 [quant-ph] (2022).

[37] H. Arai, Y. Yoshida, and M. Hayashi, "Perfect discrimination of non-orthogonal separable pure states on bipartite system in general probabilistic theory." *J. Phys. A* **52**, 465304 (2019).

[38] G. Aubrun, L. Lami, C. Palazuelos, et al., "Entangleability of cones." *Geom. Funct. Anal.* **31**, 181-205 (2021).

[39] Y. Yoshida, H. Arai, and M. Hayashi, "Perfect Discrimination in Approximate Quantum Theory of General Probabilistic Theories." *PRL*, **125**, 150402 (2020).

[40] G. Aubrun, L. Lami, C. Palazuelos, et al., "Entanglement and superposition are equivalent concepts in any physical theory." *arXiv*:2109.04446 (2021).

[41] S. Minagawa and H. Arai and F. Buscemi, "von Neumann's information engine without the spectral theorem," *Physical Review Research* **4**, 033091 (2022).

[42] H.Arai and M. Hayashi, "Pseudo standard entanglement structure cannot be distinguished from standard entanglement structure." *New. J. Phys.* **25**, 023009 (2023).

[43] M. Banik, MD. R. Gazi, S. Ghosh, and G. Kar, "Degree of Complementarity Determines the Nonlocality in Quantum Mechanics." *Phys. Rev. A.* **87**, 052125 (2013).

[44] N. Stevens and P. Busch, "Steering, incompatibility, and Bell inequality violations in a class of probabilistic theories." *Phys. Rev. A.* **89**, 022123 (2014).

[45] H. Barnum, C. Philipp, and A. Wilce, "Ensemble Steering, Weak Self-Duality, and the Structure of Probabilistic Theories" *Found. Phys.* **43**, 1411–1427 (2013).

[46] D. Chruściński and G. Sarbicki, "Entanglement witnesses: construction, analysis and classification." *J. Phys. A* **47**, 483001 (2014).

[47] M. Marciniak, "On extremal positive maps acting between type I factors Noncommutative Harmonic Analysis with Application to Probability II." *Polish Academy of Sciences Mathematics* **89** 201 (2010).

[48] A. Peres, "Separability Criterion for Density Matrices." *Phys. Rev. Lett.* **77**, 1413 (1996).

[49] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators." *Rep. Math. phys.* **3** 4, 275-278 (1972).

[50] M. Hayashi, K. Matsumoto, and Y. Tsuda, "A study of LOCC-detection of a maximally entangled state using hypothesis testing." *J. Phys. A: Math. Gen.* **39** 14427 (2006).

[51] H. Zhu and M. Hayashi "Optimal verification and fidelity estimation of maximally entangled states." *Phys. Rev. A* **99**, 052346 (2019).

[52] S. Boyd and L. Vandenberge, "Convex Optimization." *Cambridge University Press* (2004).

[53] H. Arai, B. Yu, and M. Hayashi, "Detection of Beyond-Quantum Non-locality based on Standard Local Quantum Observables." arXiv:2301.04196 [quant-ph] (2023).

Quantum Ridgelet Transform: Winning Lottery Ticket of Neural Networks with Quantum Computation

Hayata Yamasaki[1] * Sathyawageeswar Subramanian[2] Satoshi Hayakawa[3]
Sho Sonoda[4]

[1] *The University of Tokyo*
[2] *University of Warwick, University of Cambridge*
[3] *University of Oxford*
[4] *RIKEN AIP*

Abstract. Ridgelet transform has been a fundamental mathematical tool in the theoretical studies of neural networks, but the practical applicability of ridgelet transform to conducting learning tasks was limited since its numerical implementation by conventional classical computation requires an exponential runtime $\exp(O(D))$ as data dimension D increases. To address this problem, we develop a quantum ridgelet transform (QRT), which implements the ridgelet transform of a quantum state within a linear runtime $O(D)$ of quantum computation. As an application, we also show that one can use QRT as a fundamental subroutine for QML to efficiently find a sparse trainable subnetwork of large shallow wide neural networks without conducting large-scale optimization of the original network. This application discovers an efficient way in this regime to demonstrate the lottery ticket hypothesis on finding such a sparse trainable neural network. These results open an avenue of QML for accelerating learning tasks with commonly used classical neural networks.

Technicial version [1] is to appear in the proceeding of ICML2023.

Keywords: quantum machine learning, lottery ticket hypothesis, neural network, quantum algorithm, ridgelet transform

Quantum machine learning (QML) is an emerging field of research to take advantage of quantum computation for accelerating tasks in machine learning [2]. Quantum computation can achieve significant speedups compared to the best existing algorithms with conventional classical computation in solving various computational tasks, but it is still challenging to establish its application to common tasks in machine learning. QML indeed has advantages in learning data obtained from quantum states [3], yet machine learning commonly deals with classical data rather than quantum states. For a classical dataset constructed carefully so that its classification reduces to a variant of Shor's algorithm, QML is known to achieve the classification superpolynomially faster than classical algorithms [4]; however, the applicability of such QML to practical datasets has been unknown. Meanwhile, motivated by the success of neural networks, various attempts have been made to apply quantum computation to more practical tasks for neural networks. For example, one widely studied approach in QML is to use parameterized quantum circuits, often called "quantum neural networks", as a potential substitute for conventional classical neural networks; however, problematically, the parameterized quantum circuits do not successfully emulate essential components of the neural networks, e.g., perceptrons and nonlinear activation functions, due to linearity of the transformation implemented by the quantum circuits [2]. Thus, a significant challenge in QML has been to develop a novel technique to bridge the gap between quantum computation and classical neural networks, so as to clarify what advantage QML could offer on top of the empirically proven merit of the classical neural networks.

To address this challenge, we here develop a fundamental quantum algorithm for making the tasks for classical neural networks more efficient, based on ridgelet transform. Ridgelet transform, one of the well-studied integral transforms in signal processing, is a fundamental mathematical tool for studying neural networks [5]. Let $f : \mathbb{R}^D \to \mathbb{R}$ denote a function with D-dimensional input, to be learned with a neural network. For an activation function $g : \mathbb{R} \to \mathbb{R}$ such as the rectified linear unit (ReLU), a shallow feed-forward neural network with a single hidden layer is represented by $f(\boldsymbol{x}) \approx \sum_{n=1}^{N} w_n g(\boldsymbol{a}_n^\top \boldsymbol{x} - b_n)$, where N is the number of nodes in the hidden layer, and w_n is the weight of the map $g(\boldsymbol{a}_n^\top \boldsymbol{x} - b_n)$ parameterized by $(\boldsymbol{a}_n, b_n)$ at node $n \in \{1, \dots, N\}$. In the over-parameterized (continuous) limit $N \to \infty$, the representation simplifies into an integral representation of the neural network, i.e.,

$$f(\boldsymbol{x}) = S[w](\boldsymbol{x}) \coloneqq \int_{\mathbb{R}^D \times \mathbb{R}} d\boldsymbol{a}\, db\, w(\boldsymbol{a}, b) g(\boldsymbol{a}^\top \boldsymbol{x} - b), \quad (1)$$

where $(\boldsymbol{a}, b)$ runs over all possible parameters in the continuous space, and $w : \mathbb{R}^D \times \mathbb{R} \to \mathbb{R}$ at each $(\boldsymbol{a}, b)$ corresponds to the weight w_n at the node n with parameter $(\boldsymbol{a}_n, b_n) = (\boldsymbol{a}, b)$. With a ridgelet function $r : \mathbb{R}^D \to \mathbb{R}$ that we appropriately choose corresponding to g, the D-dimensional **ridgelet transform** $R[f]$ is defined as an inverse transform of $S[w]$ given by [6]

$$w(\boldsymbol{a}, b) = R[f](\boldsymbol{a}, b) \coloneqq \int_{\mathbb{R}^D} d\boldsymbol{x}\, f(\boldsymbol{x}) r(\boldsymbol{a}^\top \boldsymbol{x} - b), \quad (2)$$

which characterizes a weight w to represent f. A wide

*hayata.yamasaki@phys.s.u-tokyo.ac.jp

class of function f is representable in terms of the integral representation; moreover, as long as g and r satisfy a certain admissibility condition, we can reconstruct f from the ridgelet transform of f, i.e., $f \propto S[R[f]]$, up to a normalization factor [6]. For theoretical analysis, an essential benefit of the integral representation is to simplify the analysis by the linearity; that is, we can regard the right-hand side of the above definition of $S[w](\boldsymbol{x})$ as the linear combination of an non-orthogonal over-complete basis of functions, i.e., $\{g(\boldsymbol{a}^\top \boldsymbol{x} - b) : (\boldsymbol{a}, b) \in \mathbb{R}^D \times \mathbb{R}\}$. The meaning of the ridgelet transform of f is the weight $w(\boldsymbol{a}, b)$ of the nodes in the hidden layer of this linearized large neural network $S[w](\boldsymbol{x})$ to represent $f(\boldsymbol{x})$.

Progressing beyond using the ridgelet transform for theoretical analysis, our key idea is to study its use for conducting tasks for neural networks. However, D-dimensional ridgelet transform has been computationally hard to use in practice since the existing classical algorithms for ridgelet transform require $\exp(O(D))$ runtime as D increases [7]. After all, the D-dimensional ridgelet transform is a transform of D-dimensional functions in an $\exp(O(D))$-size space, and classical algorithms for such transforms conventionally need $\exp(O(D))$ runtime; e.g., fast Fourier transform may be a more established transform algorithm than that for the ridgelet transform but still needs $O(n \log(n)) = \exp(O(D))$ runtime for the space of size $n = \exp(O(D))$. To solve these problems, we discover that we can employ quantum computation. Our results are as follows.

1. To make exact implementation of ridgelet transform possible for computer with a finite number of bits and qubits, we formulate a new discretized version of ridgelet transform, which we call **discrete ridgelet transform** (Sec. 2.1 of Technical Version [1]). We identify the conditions for the discrete ridgelet transform to be an isometry transform. We prove that our discrete ridgelet transform can be used for exactly representing any function on the discretized domain (Sec. 2.2 of Technical Version [1]).

2. We develop a quantum algorithm to apply the D-dimensional discrete ridgelet transform to a quantum state of $O(D)$ qubits, i.e., a state in an $\exp(O(D))$-dimensional space, only within linear runtime $O(D)$ (Algorithm 1 in Sec. 3 of Technical Version [1]). We call this quantum algorithm **quantum ridgelet transform (QRT)**. QRT is exponentially faster in D than the $\exp(O(D))$ runtime of the best existing classical algorithm for ridgelet transform in the $\exp(O(D))$-size space, in the same spirit as quantum Fourier transform (QFT) being exponentially faster than the corresponding classical algorithm of fast Fourier transform.

3. As an application, we demonstrate that we can use QRT to learn a sparse representation of an unknown function f by sampling a subnetwork of a shallow wide neural network to approximate f well. We analytically show the advantageous cases of our algorithm and also conduct a numerical simulation to support the advantage (Sec. 4 of Technical Version [1]). This application is important as a demonstration of the **lottery ticket hypothesis**, as explained below.

Technical contributions in developing QRT.— For the purpose of QML, we formulate the **discrete ridgelet transform**. Our key development is a **Fourier slice theorem** that characterizes our discrete ridgelet transform using Fourier transform. Although multiple definitions of discrete versions of ridgelet transform have been proposed, none of them has such Fourier expression [7]. By contrast, the significance of the Fourier slice theorem is that it makes the ridgelet analysis tractable with the well-established techniques for the Fourier transform, which we will use for constructing the quantum algorithm as well.

Using the discrete ridgelet transform, we show that any function f on the discretized domain has an **exact representation** in terms of a shallow neural network with a finite number of parameters in the discretized space. In the continuous case, any square-integrable function f is represented as the shallow neural network in the over-parameterized (continuous) limit, i.e., $f = S[w]$, with the weight given by the ridgelet transform $w \propto R[f]$, as explained in (1) and (2). With discretization, it is non-trivial to show such an exact representation due to finite precision in discretizing the real number. Nevertheless, we prove that any function $f(\boldsymbol{x})$ on a discretized domain can be exactly represented as $f(\boldsymbol{x}) = \mathcal{S}[w](\boldsymbol{x})$ using our discretized version $\mathcal{S}$ of the integral representation S as well, with the weight given by our formulation of the discrete ridgelet transform $w \propto \mathcal{R}[f]$. We call this exact representation a **discretized neural network**.

Furthermore, we introduce **QRT**, an efficient quantum algorithm for applying the discrete ridgelet transform to a given quantum state. In various quantum algorithms, we may use QFT as a fundamental subroutine. In addition to QFT, various discrete transforms are known to be implemented efficiently with quantum computation, such as wavelet transform, Radon transform, fractional Walsh transform, Hartley transform, and curvelet transform. However, the existing discrete versions of ridgelet transform [7] were lacking implementation by quantum computation, due to the lack of the Fourier slice theorem. In contrast, our QRT opens a way to use the discrete ridgelet transform as a fundamental subroutine for QML to deal with tasks for classical neural networks.

The advantage of QRT is its linear runtime $O(D)$ in the data dimension D, which is exponentially faster than the best existing classical algorithm for ridgelet transform in the $\exp(O(D))$-size space requiring $\exp(O(D))$ runtime. This advantage is in the same spirit as QFT being exponentially faster than the corresponding classical algorithm of the fast Fourier transform. On top of our development of QRT, we further clarify that QRT has an application to accelerate the task of finding the winning ticket of neural networks, as explained below.

Impact on QML with neural networks.— State-of-the-art neural networks have billions of parameters to attain high learning accuracy, but such large-scale networks may be problematic for practical use, e.g., with mobile devices and embedded systems. Pruning techniques for obtaining neural networks with a smaller number of parameters are gaining growing importance in learning with neural networks. The **lottery ticket hypothesis** claims that, given a large-scale neural network, one can find a sparse trainable subnetwork, which has been nominated as the best paper in one of the top conferences of machine learning (ICLR2019) and attracts significant attention in the community of machine learning [8].

Learning a general class of function $f : \mathbb{R}^D \to \mathbb{R}$ would be inevitably demanding, since a neural network to represent f would require $\exp(O(D))$ parameters to specify the values $f(\boldsymbol{x})$ for exponentially many points $\boldsymbol{x}$ in the worst case as the data dimension D increases. Indeed, as explained above, any function f can be represented by our discretized neural network $f(\boldsymbol{x}) = \mathcal{S}[w](\boldsymbol{x})$, but this network is exponentially wide in D. By contrast, the goal of our demonstration of the lottery ticket hypothesis is to achieve an approximation of this exponentially large original neural network $\mathcal{S}[w]$ feasibly with much fewer parameters, using a subnetwork of the original network. To this goal, it is conventional in statistical learning theory to consider a reasonably restricted class of functions, e.g., those with bounded norms; correspondingly, we work on a setting where the norm of the weights in the original network for representing f should be bounded.

However, still in such a setting, it is computationally demanding to search for the appropriate subnetwork hidden in the large-scale neural network. One existing way to find a trainable subnetwork to approximate the original network is to train the overall large original network and then perform masking to eliminate the low-weight nodes while keeping those with higher weights [8]. This approach is inefficient since one needs large-scale optimization to train the large original network before the pruning. Then, more recent studies have suggested that one should be able to find the subnetwork only by pruning the initial network directly, even without the optimization for training. Still, to perform this pruning appropriately, the existing classical methods need to store the parameters of the large-scale network in the classical memory, to perform a large-scale search for the subnetwork within the parameter space of the large original neural network. Thus, as D increases, it would become infeasible to deal with the exponentially large original network in D for training or searching, as long as we use the existing methods based on classical computation.

To apply QML to this pruning problem, our idea is to use QRT for preparing a quantum state so that, by measuring the state, we can sample the parameters of the important nodes for the subnetwork with high probability. This quantum algorithm is presented in Sec. 4.2 and Appendix D of Technical Version [1] (Algorithm 2). To see how this quantum algorithm works, recall that the weights of the nodes of the neural network to represent f can be given by the ridgelet transform of f. Given M examples of input-output pairs $(\boldsymbol{x}_1, f(\boldsymbol{x}_1)), \ldots, (\boldsymbol{x}_M, f(\boldsymbol{x}_M))$, as the input model, our algorithm uses preparation of $O(D)$-qubit quantum states that represent $f(\boldsymbol{x})$ and the empirical distribution $\hat{p}_{\mathrm{data}}(\boldsymbol{x})$ by their amplitude. Then, roughly speaking, our quantum algorithm performs QRT of this input state and measure it; in this way, we can sample parameters of the high-weight nodes with high probability since the amplitude of the state to be measured are the weights given by the QRT of f. By repeating this sampling, we can collect the set of high-weight nodes (i.e., important nodes for learning f) efficiently in our setting.

Using this quantum algorithm, we further clarify the overall algorithm for finding the sparse subnetwork to approximate the original network with accuracy ϵ, as described in Sec. 4.3 and Appendix E in Technical Version [1] (Algorithm 3). The overall runtime for finding this sparse trainable subnetwork is $\widetilde{O}(D \times \mathrm{poly}(1/\epsilon))$, avoiding $\exp(O(D))$ runtime. The advantage of our quantum algorithm over a conventional classical algorithm is also verified with numerical simulation.

These quantum algorithms are designed to avoid the overhead of input and output throughout achieving the learning task. Regarding the input model, we explicitly show how to prepare the required input state by quantum circuit within $O(D\,\mathrm{polylog}(M))$ depth (see Sec. 4.2 and Appendix D of Technical Version [1] for our assumption). Note that M dependency does not matter for our quantum speedup in D. To make this QML algorithm efficient, we never store all parameters of the large original neural network in the classical memory but represent them by the amplitude of the quantum state prepared directly from given data. Also for the output, our algorithm finds a high-weight node per single state preparation and measurement. If one were using some algorithms that estimate expectation values of observables or classical description of the prepared quantum state, their overheads could cancel out the speedup in QML, but our algorithm is designed to avoid these overheads.

Consequently, these results show that QRT can be used as a fundamental subroutine for QML to accelerate the tasks for the classical neural networks. Remarkably, our algorithms have the theoretical guarantee in a conventional learning setting, progressing beyond heuristic QML for neural networks. Quantum computation often uses QFT to achieve large quantum speedups for various search problems, such as period finding and more recent "verifiable quantum advantage without structure". By contrast, we make quantum computation applicable to searching in the parameter space of neural networks, by developing QRT to be used in place of QFT. A potential drawback may be that our current technique is designed simply for the shallow neural networks with a single hidden layer; however, studies of shallow networks capture various essential features of neural networks. We leave the generalization to deep neural networks for future research, but our development opens a promising route in this direction.

References

[1] Technical version attached to this submission, to appear in the proceeding of Fortieth International Conference on Machine Learning (ICML2023), text overlap with an early version arXiv:2301.11936, 2023.

[2] M. Schuld and F. Petruccione, *Machine learning with quantum computers*, Springer, 2021.

[3] R. Sweke, J.-P. Seifert, D. Hangleiter, and J. Eisert, Quantum **5**, 417, 2021; H.-Y. Huang, R. Kueng, and J. Preskill, Phys. Rev. Lett. **126**, 190505, 2021; S. Chen, J. Cotler, H. Huang, and J. Li, 2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS2021), 2021.

[4] Y. Liu, S. Arunachalam, and K. Temme, Nat. Phys., **17**, 1013, 2021.

[5] E. J. Candes, PhD thesis, Stanford University, 1998.

[6] S. Sonoda and N. Murata, Applied and Computational Harmonic Analysis, **43**, 233, 2017.

[7] M. N. Do and M. Vetteri, IEEE Trans. Image Representation, **12**, 16, 2003; P. Carre and E. Andes, Signal Processing, **84**, 2165, 2004; D. Helbert, P. Carre, and E. Andes, IEEE Trans. Image Processing, **15**, 3701, 2006.

[8] J. Frankle and M. Carbin, the proceedings of International Conference on Learning Representations (ICLR2019), 2019.

On the universality of S_n-equivariant k-body gates

Martín Larocca[1] * Sujay Kazi[1 2] M. Cerezo[1] †

[1] *Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2] *New York University, New York, New York 10012, USA*

Abstract. The importance of symmetries has recently been recognized in quantum machine learning from the simple motto: if a task exhibits a symmetry (given by a group $\mathfrak{G}$), the learning model should respect said symmetry. In this work we study how the interplay between symmetry and locality in the generators of a parametrized quantum circuit affect its expressiveness. We focus on the case of $\mathfrak{G} = S_n$, the symmetric group acting on an n-qubit system by qubit permutation. We find that if the Quantum Neural Network (QNN) is generated by arbitrary one- and two-body S_n-equivariant gates, the it is semi-universal but not universal. Universality can only be achieved by employing up-to-n-body generators.

Keywords: Quantum Circuits, Universality, Symmetry, Representation Theory

1 Introduction

Numerous endeavors have been undertaken to create learning models that are tailored specifically to a given task. Among these, Geometric Quantum Machine Learning (GQML) has emerged as one of the most promising approaches [1, 2, 3, 4, 5, 6, 7]. The fundamental idea behind GQML is to leverage the symmetries present in the task to develop sharp inductive biases for the learning models. The GQML program consists of several steps. First, one needs to identify the group $\mathfrak{G}$ of transformations preserving some important property of the data (e.g., a symmetry that preserves the labels in supervised learning). While the theoretical foundations for GQML have been established, it is still unclear what the true expressive power of group-invariant QNNs is. In this work, we will focus on $\mathfrak{G} = S_n$, the symmetric group of permutations, with its action on n qubits. This group is of special interest as it is the relevant symmetry group for a wide range of learning tasks related to problems defined on sets, graphs and grids, molecular systems, multipartite entanglement, and distributed quantum sensing [8, 9, 10, 11, 12, 13, 14, 15, 16, 17].

Consider n-qubit systems with $d = 2^n$-dimensional state space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$. A QNN is a parametrized quantum circuit

$$U(\boldsymbol{\theta}) = \prod_{m=1}^{M} e^{-i\theta_m H_m}\,, \tag{1}$$

where H_m are Hermitian operators taken from a given *set of generators* $\mathcal{G}$, and $\boldsymbol{\theta} = (\theta_1, \dots \theta_M) \in \mathbb{R}^M$ are trainable parameters. As shown in Fig. 1, we consider the case where the QNN contains only up-to-k-body gates, i.e., gates acting non-trivially on at most k qubits. Such restriction is usually physically motivated and arises when working with gates that are native to some specific hardware [18]. At this point, we find it convenient to recall that in the absence of symmetries, 2-local gates are sufficient to generate any d-dimensional unitary. However, the same is generally not true when the operators in $\mathcal{G}$ are local, but also chosen to respect certain symmetry group $\mathfrak{G}$ [19, 20]. Hence, given these constraints, it is critical to quantify the QNN's expressiveness, i.e., the breadth of unitaries that $U(\boldsymbol{\theta})$ can generate when varying the parameters $\boldsymbol{\theta}$ (see Fig. 1(b)). While several measures of expressiveness exist [21, 22, 23], here we will focus on the so-called Dynamical Lie Algebra (DLA) [24], which captures the potential expressiveness of the QNN. Given a set of Hermitian generators $\mathcal{G}$, the DLA is the subspace of operator space $\mathfrak{u}(d)$ spanned by the repeated nested commutators of the elements in $\mathcal{G}$. That is, $\mathfrak{g} = \langle i\mathcal{G}\rangle_{\text{Lie}}$, where $\langle\,\cdot\,\rangle_{\text{Lie}}$ denotes the Lie closure. Notably, the DLA fully determines the ultimate expressiveness of the QNN, as we have $U(\boldsymbol{\theta}) \in \mathbb{G} = e^{\mathfrak{g}} \subseteq \mathbb{U}(d)$.

The main ingredient to build symmetry respecting circuits is through the concept of equivariance.

Definition 1 (Equivariance) *Given a compact group $\mathfrak{G}$, an operator H is $\mathfrak{G}$-equivariant if it commutes with every group element.*

Note that if all the generators in $\mathcal{G}$ are equivariant, the QNN can be readily shown to be equivariant itself [4, 2]. We also highlight the fact that Defini-

*larocca@lanl.gov
†cerezo@lanl.gov

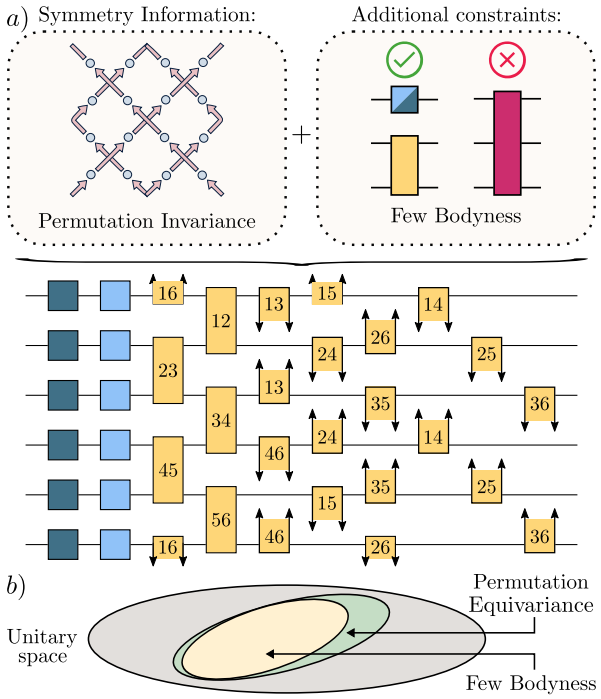


Figure 1: a) We study the impact of both symmetry and locality constraints on parametrized circuit generators. We focus on the case of permutation invariance and one and two qubit gates (same color share a common parameter). b) Imposing S_n-equivariance can appropriately reduce the QNN's expressiveness to a region of unitaries respecting the task's symmetry. Imposing additional restrictions, such as few-bodyness, further restricts its expressiveness.

tion 1 implies $H \in \mathfrak{comm}(\mathfrak{G})$, where $\mathfrak{comm}(\mathfrak{G})$ denotes the *commutant algebra* of the (representation R of the) group $\mathfrak{G}$, i.e., the associative matrix algebra of linear operators that commute with every element in $\mathfrak{G}$:

$$\mathfrak{comm}(\mathfrak{G}) = \{A \in \mathcal{B}(\mathcal{H}) \mid [A, R(g)] = 0\,,\ \forall g \in \mathfrak{G}\}\,. \quad (2)$$

Here, $\mathcal{B}(\mathcal{H})$ denotes the space of bounded linear operators in $\mathcal{H}$.

Next, it is fundamental to recall that the representation R admits an isotypic decomposition

$$R(g \in \mathfrak{G}) \cong \bigoplus_{\lambda=1}^{L} r_\lambda(g) \otimes \mathbb{1}_{m_\lambda}\,, \quad (3)$$

where r_λ is a $d_\lambda = \dim(r_\lambda)$-dimensional irreducible representation (irrep) of $\mathfrak{G}$, and $\mathbb{1}_{m_\lambda}$ is an identity of dimension m_λ. Using $\mathbb{P}^\nu_\lambda$ to denote the projector onto the subspace associated with each irrep, we can focus on the part of the DLA that acts non-trivially on each $\mathcal{H}^\nu_\lambda$, given by $\mathfrak{g}^\nu_\lambda = \{\mathbb{P}^\nu_\lambda\, iH\,,\ iH \in \mathfrak{g}\}$.

The previous motivates us to define three important Lie subalgebras of $\mathfrak{u}(d)$. First, we define the *maximal* $\mathfrak{G}$-symmetric subalgebra $\mathfrak{u}^{\mathfrak{G}}(d) = \mathfrak{comm}(\mathfrak{G}) \cap \mathfrak{u}(d)$ as

$$\mathfrak{u}^{\mathfrak{G}}(d) = Q\left(\bigoplus_{\lambda=1}^{L} \mathfrak{u}(m_\lambda)\right) = \bigoplus_{\lambda=1}^{L} \mathbb{1}_{d_\lambda} \otimes \mathfrak{u}(m_\lambda)\,, \quad (4)$$

where Q is a representation defined by the right-hand-side of Eq. (4). We then define the *maximal special* subalgebra $\mathfrak{su}^{\mathfrak{G}}(d) = \mathfrak{comm}(\mathfrak{G}) \cap \mathfrak{su}(d)$, as

$$\mathfrak{su}^{\mathfrak{G}}(d) = s\left[\bigoplus_{\lambda=1}^{L} \mathbb{1}_{d_\lambda} \otimes \mathfrak{u}(m_\lambda)\right]\,, \quad (5)$$

where $s[\cdot]$ denotes keeping the operators with vanishing trace. Finally, we also define the *maximal centerless* $\mathfrak{G}$-symmetric subalgebra

$$\mathfrak{su}^{\mathfrak{G}}_{\text{cless}}(d) = Q\left(\bigoplus_{\lambda=1}^{L} \mathfrak{su}(m_\lambda)\right) = \bigoplus_{\lambda=1}^{L} \mathbb{1}_{d_\lambda} \otimes \mathfrak{su}(m_\lambda)\,. \quad (6)$$

We now introduce three key definitions that will allow us to study controllability and degrees of universality when there are symmetries in play.

Definition 2 *The QNN, or its associated DLA, is said to be: i) Subspace controllable, if for all* λ, $\mathfrak{su}(m_\lambda) \subseteq \mathfrak{g}_\lambda \subseteq \mathfrak{u}(m_\lambda)$, *ii) Semi-universal, if* $\mathfrak{su}^{\mathfrak{G}}_{\text{cless}}(d) \subseteq \mathfrak{g} \subseteq \mathfrak{u}^{\mathfrak{G}}(d)$, *and iii) Universal, or completely controllable, if* $\mathfrak{su}^{\mathfrak{G}}(d) \subseteq \mathfrak{g} \subseteq \mathfrak{u}^{\mathfrak{G}}(d)$.

Given a Lie algebra $\mathfrak{h} \subseteq \mathfrak{u}(d)$, its center $\mathfrak{z}(\mathfrak{h})$ is composed of all the elements in $\mathfrak{h}$ that commute with every element in $\mathfrak{h}$, i.e.,

$$\mathfrak{z}(\mathfrak{h}) = \{iH \in \mathfrak{h} \mid [H, H'] = 0\,,\ \forall iH' \in \mathfrak{h}\}\,. \quad (7)$$

In particular, $\mathfrak{z}(\mathfrak{u}^{\mathfrak{G}}(d)) = Q\left(\bigoplus_{\lambda=1}^{L} \mathfrak{u}(1)\right)$ and together with the centerless *su* they form the maximal $\mathfrak{G}$-symmetric unitary subalgebra

$$\mathfrak{u}^{\mathfrak{G}}(d) = \mathfrak{su}^{\mathfrak{G}}_{\text{cless}}(d) \cup \mathfrak{z}(\mathfrak{u}^{\mathfrak{G}}(d))\,. \quad (8)$$

2 Results

We consider $\mathfrak{G}$ to be the symmetric group S_n and R the qubit-permuting representation of S_n on n qubits, $R(\pi \in S_n) \bigotimes_{i=1}^n |\psi_i\rangle = \bigotimes_{i=1}^n |\psi_{\pi^{-1}(i)}\rangle$. Given a set of equivariant k-local generators, it is natural to ask: *can we achieve subspace controllability, or even (semi-)universality?* In particular, consider the following set of one and two body S_n-equivariant generators (see Fig. 1).

$$\mathcal{G}_2 = \left\{\sum_{j=1}^{n} X_j, \sum_{j=1}^{n} Y_j, \sum_{j_1<j_2}^{n} Z_{j_1} Z_{j_2}\right\}\,. \quad (9)$$

The expressiveness of $\mathcal{G}_2$ is captured by the following theorem.

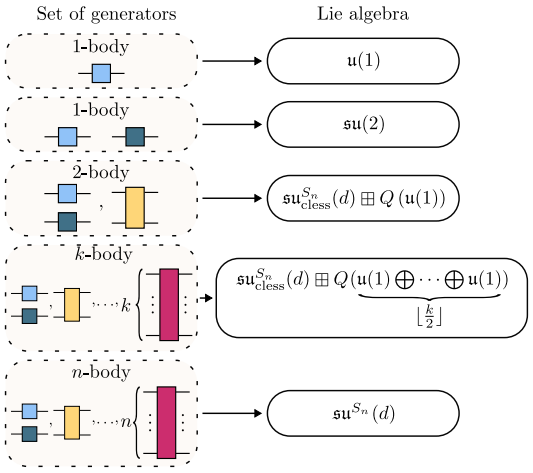


Figure 2: Here we review the main results of our work.

Theorem 1 *$\mathcal{G}_2$ The DLA for $\mathcal{G}_2$ is*

$$\mathfrak{g}_2 = \mathfrak{su}^{S_n}_{\text{cless}}(d) \boxplus Q\left(\mathfrak{u}(1)\right) , \tag{10}$$

where $\boxplus$ denotes the Minkowski sum[1] *and where $Q(\mathfrak{u}(1)) \subset \mathfrak{z}(\mathfrak{u}^{S_n}(d))$.*

Theorem 1 suggests the set $\mathcal{G}_2$ is semi-universal (and thus subspace controllable) according to Definitions 2. We now present the main steps in its derivation. In addition, we can infer from this theorem that the dimension of $\mathfrak{g}_2$ is $\dim(\mathfrak{g}_2) = \binom{n+3}{3} - \lfloor \frac{n}{2} \rfloor$.

Our previous results proved that sets of generators with S_n-equivariant 1-body and 2-body operators are not sufficient to generate $\mathfrak{su}^{S_n}(d)$. A natural question then is whether this can be fixed by including in the set of k-body S_n-equivariant generators (for $k \geq 3$). Defining as $\mathfrak{g}_k$ the DLA associated with a set of generators containing all S_n-equivariant k-body gates, we have

Theorem 2 *The DLA for the set $\mathcal{G}_k$ is*

$$\mathfrak{g}_k = \mathfrak{su}^{S_n}_{\text{cless}}(d) \boxplus Q\left(\underbrace{\mathfrak{u}(1) \oplus \cdots \oplus \mathfrak{u}(1)}_{\lfloor k/2 \rfloor}\right) , \tag{11}$$

where $Q\left(\mathfrak{u}(1) \oplus \cdots \oplus \mathfrak{u}(1)\right)$ is a $\lfloor k/2 \rfloor$-dimensional subalgebra of $\mathfrak{z}(\mathfrak{u}^{S_n}(d))$.

Theorem 2 shows that one element in the center is generated in the DLA every time one adds a generator with *even* bodyness.

[1] Given two sets A and B, their Minkowski sum is defined as $A \boxplus B = \{a + b \,|\, a \in A, b \in B\}$.

Corollary 1 *Any set consisting of at-most-k-body S_n-equivariant operators will always be insufficient to generate $\mathfrak{su}^{S_n}(d)$ unless $k = n$ for n even or $k \geq n - 1$ for n odd.*

Theorem 2 and Corollary 1 show that in order for a QNN with S_n-equivariant k-body gates to be universal, one needs to include in the set of generators up-to-n-body gates (for n even) or up-to-$(n-1)$-body gates (for n odd). Hence, this corollary imposes a fundamental limitation of the universality of QNNs with S_n-equivariant k-local gates.

3 Discussion

S_n-equivariant QNNs hold unique properties such as the absence of barren plateaus, generalization with few training points, and the capacity for efficient overparameterization [5]. However, the influence of locality constraints on the gate generators could limit their power. Our work shows that S_n-equivariant QNNs constructed with elementary k-body gates exhibit semi-universality and subspace controllability, but not universality.

This might appear to be a restriction; however, the fact that S_n-equivariant QNNs with only 2-body gates achieve semi-universality is a significant result. Importantly, our findings dictate that for achieving universality, one must include up-to-n-body interactions for even n, and up-to-$(n-1)$-body interactions for odd n. This outcome carries several implications: it not only reveals a fundamental limitation in reaching universality with local permutation-invariant gates, but it also corrects a previously held belief that universality can be attained with one- and two-body S_n-equivariant gates.

Furthermore, our work establishes an intrinsic connection with the study in Ref. [25] through the Schur-Weyl duality. This relationship offers an insightful lens for exploring how central projections condition impacts generators on each "side" of the duality. Moreover, it highlights the potential of similar results for dual reductive pairs.

As a final note, both our work and those in Refs. [19, 25, 26, 20] focus primarily on obstructions to universality coming from the central projections condition. We thus believe that a more general investigation into the failure modes for gate sets with additional constraints (beyond symmetry-equivariance) to generate the full symmetry-invariant algebra will be a fruitful area of study.

Link to the manuscript

https://arxiv.org/abs/2303.00728

References

[1] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, PRX Quantum **3**, 030341 (2022).

[2] J. J. Meyer, M. Mularski, E. Gil-Fuster, A. A. Mele, F. Arzani, A. Wilms, and J. Eisert, arXiv preprint arXiv:2205.06217 (2022).

[3] A. Skolik, M. Cattelan, S. Yarkoni, T. Bäck, and V. Dunjko, arXiv preprint arXiv:2205.06109 (2022).

[4] Q. T. Nguyen, L. Schatzki, P. Braccia, M. Ragone, M. Larocca, F. Sauvage, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2210.08566 (2022).

[5] L. Schatzki, M. Larocca, F. Sauvage, and M. Cerezo, arXiv preprint arXiv:2210.09974 (2022).

[6] M. Ragone, Q. T. Nguyen, L. Schatzki, P. Braccia, M. Larocca, F. Sauvage, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2210.07980 (2022).

[7] F. Sauvage, M. Larocca, P. J. Coles, and M. Cerezo, arXiv preprint arXiv:2207.14413 https://doi.org/10.48550/arXiv.2207.14413 (2022).

[8] H. Maron, O. Litany, G. Chechik, and E. Fetaya, in *Proceedings of the 37th International Conference on Machine Learning*, Proceedings of Machine Learning Research, Vol. 119, edited by H. D. III and A. Singh (PMLR, 2020) pp. 6734–6744.

[9] H. Maron, H. Ben-Hamu, N. Shamir, and Y. Lipman, in *International Conference on Learning Representations* (2019).

[10] N. Keriven and G. Peyré, in *Advances in Neural Information Processing Systems*, Vol. 32, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., 2019).

[11] H. Maron, H. Ben-Hamu, H. Serviansky, and Y. Lipman, Provably powerful graph networks (Curran Associates Inc., Red Hook, NY, USA, 2019).

[12] G. Verdon, T. McCourt, E. Luzhnica, V. Singh, S. Leichenauer, and J. Hidary, arXiv preprint arXiv:1909.12264 (2019).

[13] I. Cong, S. Choi, and M. D. Lukin, Nature Physics **15**, 1273 (2019).

[14] J. L. Beckey, N. Gigena, P. J. Coles, and M. Cerezo, Phys. Rev. Lett. **127**, 140501 (2021).

[15] X. Guo, C. R. Breum, J. Borregaard, S. Izumi, M. V. Larsen, T. Gehring, M. Christandl, J. S. Neergaard-Nielsen, and U. L. Andersen, Nature Physics **16**, 281 (2020).

[16] C. Huerta Alderete, M. H. Gordon, F. Sauvage, A. Sone, A. T. Sornborger, P. J. Coles, and M. Cerezo, Phys. Rev. Lett. **129**, 190501 (2022).

[17] E. R. Anschuetz, A. Bauer, B. T. Kiani, and S. Lloyd, arXiv preprint arXiv:2211.16998 (2022).

[18] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[19] I. Marvian, Nature Physics **18**, 283 (2022).

[20] I. Marvian, arXiv preprint arXiv:2302.12466 (2023).

[21] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, Advanced Quantum Technologies **2**, 1900070 (2019).

[22] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, PRX Quantum **3**, 010313 (2022).

[23] Z. Zimborás, R. Zeier, T. Schulte-Herbrüggen, and D. Burgarth, Physical Review A **92**, 042309 (2015).

[24] R. Zeier and T. Schulte-Herbrüggen, Journal of mathematical physics **52**, 113510 (2011).

[25] I. Marvian, H. Liu, and A. Hulse, arXiv preprint arXiv:2202.01963 (2022).

[26] I. Marvian, H. Liu, and A. Hulse, arXiv preprint arXiv:2105.12877 (2021).

No-Regret Learning in Quantum Games: Equilibration, Correlation and Entanglement

Wayne Lin[1] * Georgios Piliouras[1] † Ryann Sim[1] ‡ Antonios Varvitsiotis[1] §

[1] *Singapore University of Technology and Design, Singapore*

Abstract. Recent works have studied quantum versions of Nash and correlated equilibria, but the ability of distributed quantum agents to reach such equilibria is poorly understood. We study no-regret learning in quantum games through the notion of quantum coarse correlated equilibria (QCCE), which we introduce in this work. We show that the time-average behavior of no-regret learning dynamics converges in time-average to the set of separable QCCE. In the special case of quantum zero-sum games, no-regret learning implies time-average convergence of the players' marginal distributions to quantum Nash equilibria. Finally, we show that QCCE computation is reducible to a semidefinite programming formulation.

Keywords: Quantum game theory, no-regret learning, coarse correlated equilibrium

1 Introduction and Preliminaries

In this work, we study the following questions: What new solution concepts emerge at the intersection of online learning, game theory and quantum information theory? How do agents that combine online optimization and regret minimization with quantum information behave when interacting in multi-agent systems?

Our results. We form a connection between no-regret learning and quantum games, introducing the notion of quantum coarse correlated equilibrium (QCCE). We provide structural and computational characterizations for QCCEs, providing a semidefinite programming formulation thereof. Moreover, we study in more detail the convergence properties of algorithms that satisfy the no-external-regret property. Specifically, for two-player quantum zero-sum games, we show that the time-average behaviour of both players using no-external-regret algorithms converges to the set of quantum Nash equilibria (QNE). On the other hand, for general quantum games we show that the time-average behaviour if all players use no-external-regret algorithms converges to the set of separable QCCEs.

Classical game theory and equilibrium notions. A powerful mathematical framework for modeling and studying the interactions among strategic agents is normal-form game theory. A normal-form game consists of a set of players $\mathcal{N} = \{1, ..., k\}$ where player i may select from a finite set of pure strategies S_i. Each player has a payoff function $u_i : S \equiv \prod_i S_i \to \mathbb{R}$ assigned over pure strategy profiles $s = (s_1, \ldots, s_k)$. The expected payoff of the i-th player using joint strategy (i.e. probability distribution over pure strategies) $p \in \Delta(S_1 \times \ldots S_k)$ is given by $u_i(p) = \sum_{s_1 \in S_1, \ldots, s_k \in S_k} p(s_1, \ldots, s_k) u_i(s_1, \ldots, s_k)$, where $p(s_1, \ldots, s_k)$ is the probability the system is in the pure state $(s_1, \ldots, s_k)$.

The analysis of normal-form games typically involves studying if the players converge to an equilibrium. The most famous form of equilibrium is the Nash equilibrium [23], defined as a product distribution wherein players cannot increase their utilities by unilateral deviations. However, in recent years doubt has been cast on the practicality of the Nash equilibrium as a solution concept for strategic interaction. This is because even though a Nash equilibrium (in mixed strategies) always exists, computing them (even approximately) is intractable [8]. Several alternative equilibrium concepts have been proposed. The first is the notion of a correlated equilibrium (CE) [1], defined as a mixed strategy $p \in \Delta(S_1 \times \ldots \times S_k)$ so that for all players i and $s_i, s_i' \in S_i$:

$$\sum_{s_{-i}} p(s_{-i}|s_i) u_i(s_{-i}, s_i) \geq \sum_{s_{-i}} p(s_{-i}|s_i) u_i(s_{-i}, s_i'), \quad \text{(CE)}$$

where $p(s_{-i}|s_i)$ denotes the conditional distribution player i computes on $\times_{-i} S_i$ given that they received recommendation s_i. A second important alternative is the coarse correlated equilibrium (CCE) [22], defined as a joint strategy $p \in \Delta(S_1 \times \ldots \times S_k)$ satisfying:

$$u_i(p) \geq \sum_{s_i} x_{is_i} \sum_{s_{-i}, s_i'} p(s_i', s_{-i}) u_i(s_i, s_{-i}), \qquad \text{(CCE)}$$

for all players i and $x_i = \{x_{is_i}\} \in \Delta(S_i)$. Unlike Nash equilibria, optimization over the sets of CE and CCE is tractable. Additionally, many algorithms have been proposed to compute them efficiently in a variety of game-theoretic settings (see e.g. [10, 17, 12, 6, 3, 30, 31, 7, 27]).

Learning in games via online optimization. Modern machine learning applications call for efficient algorithms that learn from data as they arrive. A mathematical framework that has emerged in recent years for studying real-time decision making is *online optimization* [29], where at each epoch $t = 1, \ldots, T$ the optimizing agent commits to a decision x_t in a (typically compact convex) set $\mathcal{X}$ without prior knowledge of the loss function and receives feedback with which they use to improve the quality of their decisions over time.

*wayne_lin@mymail.sutd.edu.sg
†georgios@sutd.edu.sg
‡ryann_sim@mymail.sutd.edu.sg
§antonios@sutd.edu.sg

A well-studied measure of the performance of an online algorithm $\mathbf{A}$ is the notion of (external) regret [16], defined as the difference between the expected cumulative loss of the algorithm after T iterations and the loss incurred by the best fixed decision in hindsight. In the learning in games setting where players obtain utilities at each timestep, $\text{regret}_i^T(\mathbf{A}) = \max_{p_i} \sum_{t=1}^T u_i^t(p_i) - \sum_{t=1}^T u_i^t(p_i^t)$.

An algorithm is called no-external-regret if it performs on average as well as the best fixed strategy in hindsight, i.e., $\frac{1}{T}\text{regret}_T(\mathbf{A}) = o(T)$. If all players in a normal-form game employ external regret-minimizing algorithms to select their strategies, the players' time-average behavior converges to a coarse correlated equilibrium [28, 17]. In the setting of two-player zero-sum games, the product of the marginals of the players' individual time-averaged strategies converges to the set of Nash equilibria [11, 24].

The no-external-regret benchmark considers what would have happened if at every time $t = 1, \ldots, T$ the agent replaces x_t with $\phi(x_t)$ where ϕ is any admissible deviation mapping. In the same spirit, for any set Φ of deviation mappings, [14] defines a corresponding notion of Φ-regret, which for the game-theoretic setting gives $\text{regret}_i^{\Phi,T}(\mathbf{A}) = \max_{\phi\in\Phi} \sum_{t=1}^T u_i^t(\phi(p_i)) - \sum_{t=1}^T u_i^t(p_i^t)$, and show that for any family Φ of linear deviation maps, no-Φ-regret algorithms exist. Moreover, the time-average behavior of players using a no-Φ-regret algorithm converges to the corresponding notion of Φ-equilibria in general normal-form games. This line of work has helped us better understand the landscape of regret minimization in games and beyond [13, 30, 2, 25].

Learning in quantum games. Recently, the concept of *quantum games*, where players process and exchange quantum information, has garnered much interest in the literature [9, 15, 5, 34, 33]. Recent works such as [19], [5], [34] and [32] have formulated and studied quantum versions of Nash equilibria and correlated equilibria, exploring the advantages players can derive by playing quantum strategies. However, the ability of players to reach such equilibria in a distributed, online fashion is not well understood. Previously, [19] studied Matrix Multiplicative Weights Update (MMWU) in two-player zero-sum quantum games and showed time-average convergence to the quantum Nash equilibrium. Recently, [18] study a continuous time variant of MMWU, showing that the dynamics exhibit a cyclical behaviour known as Poincaré recurrence. This mirrors results classical results which show that regret minimization alone does not suffice for convergence to equilibria [26, 4, 21]. Beyond the zero-sum setting, [20] studies a version of MMWU in quantum potential games. Other than that, little is known about the behaviour of learning algorithms in quantum games.

Quantum preliminaries. Given a finite-dimensional Hilbert space $\mathcal{H} = \mathbb{C}^n$, we denote by $\mathrm{L}(\mathcal{H})$ the set of linear operators acting on $\mathcal{H}$. When two quantum registers with associated spaces $\mathcal{A}$ and $\mathcal{B}$ of dimension n and m respectively are considered as a joint quantum register, the associated state space is given by the density operators on the tensor product space, i.e., $D(\mathcal{A}\otimes\mathcal{B})$. A linear operator that maps matrices to matrices, i.e., a mapping $\Theta : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$, is called a *super-operator*. The adjoint super-operator $\Theta^\dagger : \mathrm{L}(\mathcal{A}) \to \mathrm{L}(\mathcal{B})$ is uniquely determined by the equation $\langle A, \Theta(B)\rangle = \langle \Theta^\dagger(A), B\rangle$. A super-operator $\Theta : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ is *positive* if it maps PSD matrices to PSD matrices. There exists a linear bijection between matrices $R \in \mathrm{L}(\mathcal{A}\otimes\mathcal{B})$ and super-operators $\Theta : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ known as the *Choi-Jamiołkowski isomorphism*. For a super-operator Θ its *Choi matrix* is: $C_\Theta = \sum_{1\le i,j\le m} \Theta(E_{i,j}) \otimes E_{i,j} \in \mathrm{L}(\mathcal{A}\otimes\mathcal{B})$, where $\{E_{i,j}\}_{i,j=1}^m$ is the standard orthonormal basis of $\mathrm{L}(\mathcal{B}) = \mathbb{C}^{m\times m}$. Conversely, given an operator $R = \sum_{1\le i,j\le m} A_{i,j} \otimes E_{i,j} \in \mathrm{L}(\mathcal{A}\otimes\mathcal{B})$, we can define $\Theta_R : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ by setting $\Theta_R(E_{i,j}) = A_{i,j}$ from which it easily follows that $C_{\Theta_R} = R$. Explicitly, we have $\Theta_R(B) = \mathrm{Tr}_{\mathcal{B}}(R(\mathbb{1}_{\mathcal{A}} \otimes B^\top))$, where the partial trace $\mathrm{Tr}_{\mathcal{B}} : \mathrm{L}(\mathcal{A}\otimes\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ is the *unique* function that satisfies: $\mathrm{Tr}_{\mathcal{B}}(A\otimes B) = A\,\mathrm{Tr}(B),\ \forall A, B$. Moreover, the adjoint map is $\mathrm{Tr}_{\mathcal{B}}^\dagger(A) = A \otimes \mathbb{1}_{\mathcal{B}}$. Lastly, a superoperator Θ is completely positive (i.e., $\mathbb{1}_m \otimes \Theta$ is positive for all $m \in \mathbb{N}$) iff the Choi matrix of Θ is positive semidefinite.

Quantum games. In a quantum game (QG), there are k players and each player i has register $\mathcal{H}_i$ and selects a density matrix $\rho_i \in \mathrm{D}(\mathcal{H}_i)$. A joint strategy is given by a joint state $\rho \in \mathrm{D}(\bigotimes_i \mathcal{H}_i)$ and the utility function for each player is the (multilinear) expected value of an observable R_i on the joint state, i.e.,

$$u_i(\rho) = \mathrm{Tr}(\rho R_i) \tag{QG}$$

for some Hermitian $R_i \in \mathrm{L}(\otimes_i \mathcal{H}_i)$. We henceforth refer to R_i as player i's utility tensor. If $\rho = \rho_i \otimes \rho_{-i}$ for some $i \in [k]$, $\rho_{-i} \in \mathrm{D}(\bigotimes_{i'\neq i} \mathcal{H}_{i'})$, we can write the utility (QG) in the alternative form $u_i(\rho) = \mathrm{Tr}(\rho R_i) = \left\langle \rho_i, \Theta_i(\rho_{-i}^\top)\right\rangle$ where $\Theta_i : \mathrm{L}(\bigotimes_{i'\neq i} \mathcal{H}_{i'}) \to \mathrm{L}(\mathcal{H}_i)$ and $R_i \in \mathrm{L}(\bigotimes_{i'} \mathcal{H}_{i'}) = \mathrm{L}(\mathcal{H}_i \otimes (\bigotimes_{i'\neq i} \mathcal{H}_{i'}))$ are related via the Choi-Jamiołkowski isomorphism. Finally, if a state $\rho \in \mathrm{D}(\bigotimes_i \mathcal{H}_i)$ can be written as a convex combination of product states i.e., $\rho = \sum_j \lambda_j \bigotimes_i \rho_{i,j}$, then it is called *separable*, and otherwise it is *entangled*.

Definition 1 (QNE) *We call a product state $\rho = \rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_k$ a quantum Nash equilibrium (QNE) if $\forall i \in [k], \rho_i' \in \mathrm{D}(\mathcal{H}_i)$,*

$$u_i(\rho) \ge u_i\left(\rho_i' \otimes \left(\bigotimes_{i'\neq i} \rho_{i'}\right)\right) \tag{QNE}$$

Moreover ρ is called an ϵ-approximate quantum Nash equilibrium (ϵ-QNE) if the inequality is satisfied up to an additive error of ϵ.

2 Quantum Coarse Correlated Equilibria

By taking the non-commutative extension of the Φ-regret framework to the space of density strategies and setting Φ to be constant CPTP maps on this space, we have that ρ is a quantum coarse correlated equilibrium if

$$u_i(\rho) \ge u_i((\phi_i \otimes \mathbb{I}_{-i})(\rho)) \tag{dev-QCCE}$$

for all replacement channels $\phi_i : \mathrm{L}(\mathcal{H}_i) \to \mathrm{L}(\mathcal{H}_i),\ X \mapsto \mathrm{Tr}(X)\rho_i'$ for some $\rho_i' \in \mathrm{L}(\mathcal{H}_i)$. We can equivalently express dev-QCCE using partial traces.

Definition 2 (QCCE) *A state ρ is called a quantum coarse correlated equilibrium (QCCE) if for each player $i \in [k]$ and $\rho_i' \in \mathrm{D}(\mathcal{H}_i)$, we have:*

$$u_i(\rho) \geq u_i(\rho_i' \otimes \mathrm{Tr}_i\, \rho) \qquad \text{(QCCE)}$$

where $\mathrm{Tr}_i : \mathrm{L}(\bigotimes_{i'} \mathcal{H}_{i'}) \to \mathrm{L}(\bigotimes_{i' \neq i} \mathcal{H}_{i'})$ is the partial trace with respect to player i's subsystem. Moreover ρ is an ϵ-approximate quantum coarse correlated equilibrium (ϵ-QCCE) if the inequality is satisfied up to an additive error of ϵ. Finally, if a QCCE ρ is a separable state (i.e., it can be expressed as a convex combination of product states), we call it a separable QCCE.

Spectrahedral characterization of QCCEs. Analogous to the classical setting, the set of QCCEs of a game can be described as the feasible set of a semidefinite program (SDP). For each $i \in [k]$, $\Theta_i : \mathrm{L}(\bigotimes_{i' \neq i} \mathcal{H}_{i'}) \to \mathrm{L}(\mathcal{H}_i)$ and $R_i \in \mathrm{L}(\bigotimes_{i'} \mathcal{H}_{i'}) = \mathrm{L}(\mathcal{H}_i \otimes (\bigotimes_{i' \neq i} \mathcal{H}_{i'}))$ are related via the Choi-Jamiołkowski isomorphism which allows us to write the following characterization:

$$\mathrm{QCCEs} = \{\rho^* \in \mathcal{H} : \mathrm{Tr}(R_i \rho^*) I - \Theta_i((\mathrm{Tr}_i\, \rho^*)^\top) \succeq 0 \quad \forall i \in [k],\ \mathrm{Tr}\, \rho^* = 1,\ \rho^* \succeq 0\}.$$

These conic inequality constraints can be combined into a single, block-diagonal LMI in terms of the entries of ρ^*, giving us an SDP characterization of the set of QCCEs of a given game. Hence, the set of QCCEs is a spectrahedron, which mirrors the classical result that the set of CCEs is a polyhedron.

3 Quantum Zero-Sum Games

In this section, we consider the special case of quantum zero-sum games, where the utility is defined such that the sum of all players' utilities is always zero. Moreover, let us restrict ourselves to the two-player case in order to present an analogue of a standard classical result - no-regret algorithms converge to Nash in two-player zero-sum games. In our setting, two players Alice and Bob play density matrices $\rho \in \mathrm{D}(\mathcal{A})$ and $\sigma \in \mathrm{D}(\mathcal{B})$ respectively. For notational simplicity, Alice's payoff is $u_A(\rho, \sigma) = \langle \rho, \Theta(\sigma) \rangle = \mathrm{Tr}\big(R(\rho \otimes \sigma^\top)\big)$, where $\Theta : \mathrm{L}(\mathcal{B}) \to \mathrm{L}(\mathcal{A})$ and $R \in \mathrm{L}(\mathcal{A} \otimes \mathcal{B})$ are related by the Choi-Jamiołkowski isomorphism. The main result in this section shows that no-regret algorithms converge to the set of QNE in two-player quantum zero-sum games.

Theorem 3 *If both players in a two-player quantum zero-sum game update their strategies with a no-regret algorithm that guarantees a time-averaged regret of $\leq \epsilon$ after T timesteps, then the product of the time-averages of their individual sequences of play after T timesteps is a 2ϵ-QNE. In particular, the set of limit points of the product of these time-averaged sequences with infinite time horizon is equal to the the set of QNEs, and the utility attained by the time-averaged strategies converges to the value of the game.*

Corollary 4 *Suppose a two-player quantum zero-sum game satisfies*

$$-1 \leq \min_{\substack{\rho \in \mathrm{D}(\mathcal{A}),\\ \sigma \in \mathrm{D}(\mathcal{B})}} \langle \rho, \Theta(\sigma) \rangle \leq \max_{\substack{\rho \in \mathrm{D}(\mathcal{A}),\\ \sigma \in \mathrm{D}(\mathcal{B})}} \langle \rho, \Theta(\sigma) \rangle \leq 1,$$

and let n be the larger of the dimensions of their registers. For any $\epsilon \leq 4$, the players are guaranteed to have the product of their time-averaged play after $T = \frac{16 \ln n}{\epsilon^2}$ timesteps be an ϵ-QNE if they use MMWU with fixed stepsize $\eta = \frac{\epsilon}{4}$ to select their strategies.

4 Beyond Quantum Zero-Sum Games

Main Theorem. *For any quantum game (QG), the set of separable QCCEs is equal to the limit points of the time-averaged history of joint play of players using no-regret algorithms.*

Theorem 5 *If all players in a quantum game update their strategies with a no-regret algorithm that guarantees a time-averaged regret of $\leq \epsilon$ after T timesteps, then the time-averaged joint history of play after T timesteps is a separable ϵ-QCCE. In particular, the limit points of the time-averaged joint history of play generated by no-regret algorithms are separable QCCEs.*

Corollary 6 *Suppose a k-player quantum game satisfies*

$$|\,\mathrm{Tr}\big(R(\otimes_{i=1}^k \rho_i)\big)| \leq 1$$

for all $\rho_1 \in \mathrm{D}(\mathcal{H}_1),\ \rho_2 \in \mathrm{D}(\mathcal{H}_2),\ \ldots,\ \rho_k \in \mathrm{D}(\mathcal{H}_k)$, and let n be the largest dimension of the players' registers. For any $\epsilon \leq 2$, the players are guaranteed to have the their time-averaged joint history of play after $T = \frac{4 \ln n}{\epsilon^2}$ timesteps be a separable ϵ-QCCE if they each use MMWU with fixed stepsize $\eta = \frac{\epsilon}{2}$ to select their strategies.

Theorem 7 *For any separable QCCE ρ^*, there exist no-regret algorithms for each player such that, if all players follow them, their time-averaged joint history of play $\left(\frac{1}{T}\sum_{t=1}^T \bigotimes_i \rho_i^t\right)_T$ converges to ρ^* as $T \to \infty$.*

Combining Theorems 5 and 7 suffices to complete the proof for the Main Theorem.

5 Conclusion

In this work we provide a general class of quantum games that fits with and subsumes prior formulations. We introduce and focus on quantum coarse correlated equilibria and show that for general quantum games, the set of separable QCCEs is actually the set of limits points of the time-averaged distribution of joint play when players use no-regret algorithms. Moreover, in the two-player zero-sum case, no-regret algorithms result in convergence to quantum Nash equilibria. Overall, this indicates a rich connection between the worlds of online optimization, learning in games, and quantum information theory.

Full version of paper attached as appendix.

References

[1] R. J. Aumann. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*, 1(1):67–96, 1974.

[2] Y. Bai, C. Jin, S. Mei, Z. Song, and T. Yu. Efficient ϕ-regret minimization in extensive-form games via online mirror descent. *arXiv preprint arXiv:2205.15294*, 2022.

[3] A. Blum and Y. Mansour. From external to internal regret. *Journal of Machine Learning Research*, 8(6), 2007.

[4] V. Boone and G. Piliouras. From Darwin to Poincaré and von Neumann: Recurrence and cycles in evolutionary and algorithmic game theory. In *Web and Internet Economics: 15th International Conference, WINE 2019, New York, NY, USA, December 10–12, 2019, Proceedings 15*, pages 85–99. Springer, 2019.

[5] J. Bostanci and J. Watrous. Quantum game theory and the complexity of approximating quantum Nash equilibria. *Quantum*, 6:882, 2022.

[6] N. Cesa-Bianchi and G. Lugosi. *Prediction, learning, and games*. Cambridge university press, 2006.

[7] C. Daskalakis, M. Fishelson, and N. Golowich. Near-optimal no-regret learning in general games. *Advances in Neural Information Processing Systems*, 34:27604–27616, 2021.

[8] C. Daskalakis, P. W. Goldberg, and C. H. Papadimitriou. The complexity of computing a Nash equilibrium. *Communications of the ACM*, 52(2):89–97, 2009.

[9] J. Eisert and M. Wilkens. Quantum games. *Journal of Modern Optics*, 47(14-15):2543–2556, 2000.

[10] D. P. Foster and R. V. Vohra. Calibrated learning and correlated equilibrium. *Games and Economic Behavior*, 21(1-2):40, 1997.

[11] Y. Freund and R. E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(1-2):79–103, 1999.

[12] D. Fudenberg, F. Drew, D. K. Levine, and D. K. Levine. *The theory of learning in games*, volume 2. MIT press, 1998.

[13] G. J. Gordon, A. Greenwald, and C. Marks. No-regret learning in convex games. In *Proceedings of the 25th international conference on Machine learning*, pages 360–367, 2008.

[14] A. Greenwald and A. Jafari. A general class of no-regret learning algorithms and game-theoretic equilibria. In *COLT*, volume 3, pages 2–12, 2003.

[15] G. Gutoski and J. Watrous. Toward a general theory of quantum games. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 565–574, 2007.

[16] J. Hannan. Approximation to Bayes risk in repeated play. *Contributions to the Theory of Games*, 3:97–139, 1957.

[17] S. Hart and A. Mas-Colell. A simple adaptive procedure leading to correlated equilibrium. *Econometrica*, 68(5):1127–1150, 2000.

[18] R. Jain, G. Piliouras, and R. Sim. Matrix multiplicative weights updates in quantum zero-sum games: Conservation laws & recurrence. *arXiv preprint arXiv:2211.01681*, 2022.

[19] R. Jain and J. Watrous. Parallel approximation of non-interactive zero-sum quantum games. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 243–253. IEEE, 2009.

[20] W. Lin, G. Piliouras, R. Sim, and A. Varvitsiotis. Quantum potential games, replicator dynamics, and the separability problem. *arXiv preprint arXiv:2302.04789*, 2023.

[21] P. Mertikopoulos, C. Papadimitriou, and G. Piliouras. Cycles in adversarial regularized learning. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2703–2717. SIAM, 2018.

[22] H. Moulin and J. P. Vial. Strategically zero-sum games: the class of games whose completely mixed equilibria cannot be improved upon. *International Journal of Game Theory*, 7:201–221, 1978.

[23] J. Nash. Non-cooperative games. *Annals of Mathematics*, 54(2):286–295, 1951.

[24] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani. *Algorithmic game theory*. Cambridge university press, 2007.

[25] G. Piliouras, M. Rowland, S. Omidshafiei, R. Elie, D. Hennes, J. Connor, and K. Tuyls. Evolutionary dynamics and phi-regret minimization in games. *Journal of Artificial Intelligence Research*, 74:1125–1158, 2022.

[26] G. Piliouras and J. S. Shamma. Optimization despite chaos: Convex relaxations to complex limit sets via Poincaré recurrence. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 861–873. SIAM, 2014.

[27] G. Piliouras, R. Sim, and S. Skoulakis. Beyond time-average convergence: Near-optimal uncoupled online learning via clairvoyant multiplicative weights update. *arXiv preprint arXiv:2111.14737*, 2021.

[28] T. Roughgarden. *Twenty lectures on algorithmic game theory*. Cambridge University Press, 2016.

[29] S. Shalev-Shwartz et al. Online learning and online convex optimization. *Foundations and Trends in Machine Learning*, 4(2):107–194, 2012.

[30] G. Stoltz and G. Lugosi. Learning correlated equilibria in games with compact sets of strategies. *Games and Economic Behavior*, 59(1):187–208, 2007.

[31] V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire. Fast convergence of regularized learning in games. *Advances in Neural Information Processing Systems*, 28, 2015.

[32] Z. Wei and S. Zhang. Full characterization of quantum correlated equilibria. *Quantum Inf. Comput.*, 13(9-10):846–860, 2013.

[33] J. Wu. A new mathematical representation of game theory, i. *arXiv preprint quant-ph/0404159*, 2004.

[34] S. Zhang. Quantum strategic game theory. In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, pages 39–59, 2012.

Robust and efficient verification of measurement-based quantum computation

Zihao Li[1] Huangjun Zhu[1] * Masahito Hayashi[2] †

[1]*Department of Physics and State Key Laboratory of Surface Physics, Institute for Nanoelectronic Devices and Quantum Computing, Center for Field Theory and Particle Physics, Fudan University*
[2]*School of Data Science, The Chinese University of Hong Kong, Shenzhen, International Quantum Academy (SIQA), and Graduate School of Mathematics, Nagoya University*

Abstract. To achieve reliable measurement-based quantum computation, it is crucial to verify whether the resource graph states are accurately prepared in the adversarial scenario. Previous verification protocols for this task are resource consuming or noise susceptible. Here, we propose a robust and efficient protocol for verifying arbitrary graph states with any prime local dimension in the adversarial scenario, which can be applied immediately to verifying measurement-based quantum computation. Our protocol requires only local Pauli measurements and is easy to realize with current technologies. It achieves the optimal scaling behaviors with respect to the system size and the target precision, and exponentially enhances the scaling behavior with the significance level.

Keywords: robust verification, MBQC, graph state, adversarial scenario

The technical version of this work is available on arXiv:2305.10742.

1 Introduction

Quantum computation offers the promise of exponential speedups over classical computation on certain important problems [1–3]. The very power of quantum computation raises the challenging problem of verifying the correctness of computation results. This problem lies at the heart of the active research field of quantum characterization, verification, and validation (QCVV) [4–9]. However, it is extremely difficult to construct robust and efficient verification protocols that apply to noisy, intermediate-scale quantum (NISQ) devices [3, 10, 11].

Measurement-based quantum computation (MBQC) [12–16] is a powerful and flexible model of quantum computation, where graph states are used as resources and local projective measurements on qudits are used to drive the computation. Compared with the preparation of multipartite entangled states, it is in general much easier to perform local projective measurements accurately. So the main challenge in the verification of MBQC lies in the verification of the underlying resource states.

In this paper, we consider the problem of verifying the resource graph states in the following adversarial scenario [17–23], which is pertinent to blind and cloud quantum computing [19, 24–26]: Alice is a client who can only perform single-qudit projective measurements, and Bob is a server who can prepare arbitrary quantum states. To perform MBQC, Alice delegates the preparation of the resource graph state $|G\rangle \in \mathcal{H}$ to Bob, who then prepares a state ρ on the whole system $\mathcal{H}^{\otimes M}$ and sends it to Alice qudit by qudit. If Bob is honest, then he is supposed to generate M copies of $|G\rangle$; while if he is malicious, then he can mess up the computation of Alice by generating an arbitrary correlated or even entangled state. To obtain reliable computation results, Alice needs to verify the resource state prepared by Bob with suitable tests on some systems of ρ. If the test results satisfy certain conditions, then she can guarantee that the reduced state on the remaining system is close enough to $|G\rangle$, and can safely use it for MBQC; otherwise, she rejects Bob's state. Since there is no communication from Alice to Bob after the preparation of ρ, the client's privacy is kept against the server by the no-signaling principle [24]. Hence, the procedure above is also suitable to verifying blind MBQC.

According to the above discussion, the problem of verifying MBQC reduces to the problem of verifying the resource graph state in the adversarial scenario [17, 19, 27]. However, it is highly nontrivial to construct robust and efficient verification protocols in the adversarial scenario. Although various protocols have been proposed [17–23], most protocols known so far are too resource consuming. Moreover, most protocols are not robust to experimental noise: the state prepared by Bob will be rejected with a high probability even if it has a very small deviation from the ideal state $|G\rangle$. However, in practice, it is unrealistic to ask honest Bob to generate the perfect resource state. In addition, if the state deviation from $|G\rangle$ is small enough, then it is still useful for MBQC [21, 22]. Therefore, a practical protocol should accept nearly ideal quantum states with a sufficiently high probability. Otherwise, Alice needs to repeat the verification protocol many times to accept such states, which may substantially increase the sampling complexity. Unfortunately, no protocol known in the literature can achieve this goal. A fault-tolerant verification protocol [26] that accepts noisy but still error-correctable states has been proposed, but it is robust only to certain correctable errors, and is difficult to realize in the current era of NISQ devices [3, 10, 11].

In this work, we propose a robust and efficient protocol for verifying blind MBQC. To achieve this goal we propose a robust and efficient protocol for verifying general qudit graph states with a prime local dimension in the

*zhuhuangjun@fudan.edu.cn
†hmasahito@cuhk.edu.cn

adversarial scenario. Our protocol is appealing to practical applications because it only requires stabilizer tests based on local Pauli measurements, which are easy to implement with current technologies. It is robust against arbitrary types of noise in state preparation, as long as the fidelity is sufficiently high. Moreover, our protocol can achieve optimal scaling behaviors with respect to the system size and target precision, and the sampling cost is comparable to the counterpart in the nonadversarial scenario. As far as we know, such a high efficiency has never been achieved before when robustness is taken into account. In addition to verifying MBQC, our approach can also be applied to verifying many other important quantum states in the adversarial scenario.

2 Verification of MBQC

Recently, a homogeneous strategy [17, 27] for testing qudit stabilizer states was proposed [17, 28]. Here we use a variant strategy for testing qudit graph states (d is a prime), which serves as an important subroutine of our verification protocol. The strategy is based on stabilizer tests with local Pauli measurements (see the technical version for details), and can be characterized by a two-outcome measurement $\{\Omega, \mathbb{I}-\Omega\}$, where the outcome Ω corresponds to passing the test, and the outcome $\mathbb{I}-\Omega$ corresponds to the failure. The operator Ω is called a strategy, and can be constructed as the form

$$\Omega = |G\rangle\langle G| + \lambda(\mathbb{I} - |G\rangle\langle G|) \tag{1}$$

for any $1/d \leq \lambda < 1$. We denote by $\nu := 1-\lambda$ the spectral gap of Ω from the largest eigenvalue.

Suppose Alice intends to perform MBQC on the graph state $|G\rangle$ prepared by Bob. Our verification protocol runs as follows. First, Bob produces a state ρ on the whole space $\mathcal{H}^{\otimes(N+1)}$ with $N \geq 1$ and sends it to Alice. After receiving the state, Alice randomly permutes the $N+1$ systems of ρ (due to this procedure, we can assume that ρ is permutation invariant without loss of generality) and applies the strategy Ω in Eq. (1) to the first N systems. If at most k failures are observed among the N tests, Alice accepts the reduced state σ_{N+1} on the remaining system and uses it for MBQC; otherwise, she rejects it. Here the integer k is called the number of allowed failures, which is chosen by Alice before performing the tests.

With this verification protocol, Alice aims to achieve three goals: completeness, soundness, and robustness. The completeness means Alice does not reject the correct state. Since $|G\rangle$ can always pass each test, this goal is automatically guaranteed. The soundness means the following: once accepting, Alice needs to ensure with a small significance level δ that her state σ_{N+1} for MBQC has a sufficiently high fidelity (at least $1-\epsilon$) with $|G\rangle$. The threshold ϵ is called the target infidelity, which together with δ characterize the target verification precision. Among all known verification protocols, only the protocol of Refs. [17, 27] achieves the optimal sampling complexity with respect to all δ, ϵ, and the qudit number n of $|G\rangle$, even without considering the robustness. Although the condition of soundness looks simple, it is highly nontrivial to determine the degree of soundness. Even in the special case $k=0$, this problem was resolved only very recently after quite a lengthy analysis [17, 27].

To characterize the robustness of a protocol, we need to consider the case in which honest Bob prepares independent and identically distributed (i.i.d.) quantum states, that is, $\rho = \tau^{\otimes(N+1)}$ with $\tau \in \mathcal{D}(\mathcal{H})$. Due to inevitable noise, τ may not equal the ideal state $|G\rangle\langle G|$. Nevertheless, if the infidelity $\epsilon_\tau := 1 - \langle G|\tau|G\rangle$ is smaller than ϵ, then τ is still useful for MBQC. For a robust verification protocol, such a state should be accepted with a high probability. On the other hand, the probability that Alice accepts τ reads

$$p^{\text{iid}}_{N,k}(\tau) = B_{N,k}\big(1 - \text{tr}(\Omega\tau)\big) = B_{N,k}(\nu\epsilon_\tau), \tag{2}$$

where N is the number of tests, k is the number of allowed failures, and $B_{N,k}(p) := \sum_{j=0}^{k} \binom{N}{j} p^j (1-p)^{N-j}$ is the binomial cumulative distribution function.

To construct a robust verification protocol, k should be sufficiently large, so that $p^{\text{iid}}_{N,k}(\tau)$ is sufficiently high. However, most previous protocols can only reach a meaningful conclusion when $k=0$ [17–20, 27], in which case the probability $p^{\text{iid}}_{N,k=0}(\tau) = (1-\nu\epsilon_\tau)^N$ decreases exponentially with N, which is not satisfactory. Consequently, many repetitions are necessary to ensure that Alice accepts the state τ at least once. When $\epsilon_\tau = \epsilon/2$ for example, the number of repetitions is at least $\Theta(\exp[1/(4\delta)])$ for the protocol in [19] and $\Theta(\delta^{-1/2})$ for the protocol in [17, 27] as shown in the technical version, which substantially increases the actual sampling cost. Therefore, although some protocols known in the literature are reasonably efficient in achieving the soundness, they are not useful in verifying blind MBQC in a realistic scenario.

3 Results

3.1 Guaranteed infidelity

Suppose ρ is permutation invariant. Then the probability that Alice accepts ρ reads

$$p_k(\rho) = \sum_{i=0}^{k} \binom{N}{i} \text{tr}\big(\big[\Omega^{\otimes(N-i)} \otimes \overline{\Omega}^{\otimes i} \otimes \mathbb{I}\big]\rho\big), \tag{3}$$

where $\overline{\Omega} := \mathbb{I} - \Omega$. Denote by σ_{N+1} the reduced state on the remaining system when at most k failures are observed. The fidelity between σ_{N+1} and the ideal state $|G\rangle$ reads $F_k(\rho) = f_k(\rho)/p_k(\rho)$ [assuming $p_k(\rho) > 0$], where

$$f_k(\rho) = \sum_{i=0}^{k} \binom{N}{i} \text{tr}\big(\big[\Omega^{\otimes(N-i)} \otimes \overline{\Omega}^{\otimes i} \otimes |G\rangle\langle G|\big]\rho\big). \tag{4}$$

The actual verification precision can be characterized by the following figure of merit with $0 < \delta \leq 1$,

$$\bar{\epsilon}_\lambda(k, N, \delta) := 1 - \min_\rho \{F_k(\rho) \,|\, p_k(\rho) \geq \delta\}, \tag{5}$$

where λ is determined by Eq. (1), and the minimization is taken over permutation-invariant states ρ on $\mathcal{H}^{\otimes(N+1)}$.

If Alice accepts the state prepared by Bob, then she can guarantee (with significance level δ) that the reduced state σ_{N+1} has infidelity at most $\bar{\epsilon}_\lambda(k,N,\delta)$ with $|G\rangle$. Consequently, the deviation of any measurement outcome probability from the ideal value is not larger than $\sqrt{\bar{\epsilon}_\lambda(k,N,\delta)}$. In the technical version we present the analytical formula and many useful properties of $\bar{\epsilon}_\lambda(k,N,\delta)$.

3.2 Verification with a fixed error rate

Now we set the number k to be proportional to the number of tests, that is, $k=\lfloor s\nu N\rfloor$, where $0\le s<1$ is the error rate. In this case, when Bob prepares i.i.d. states τ with infidelity $\epsilon_\tau<s$, the acceptance probability approaches one as N increases. In addition, we have the following theorems as proved in the technical version.

Theorem 1 *Suppose* $0<s,\lambda<1$, $0<\delta\le 1/4$. *Then*

$$s-\frac{1}{\nu N}<\bar{\epsilon}_\lambda(\lfloor\nu s N\rfloor,N,\delta)\le s+\frac{1}{\nu\lambda}\sqrt{\frac{s\ln\delta^{-1}}{N}}+\frac{\ln\delta^{-1}}{2\nu^2\lambda N}+\frac{2}{\lambda N}. \quad (6)$$

Theorem 1 implies that $\bar{\epsilon}_\lambda(\lfloor\nu s N\rfloor,N,\delta)$ converges to s when the number N of tests gets large. To achieve a given ϵ and δ, which means $\bar{\epsilon}_\lambda(\lfloor\nu s N\rfloor,N,\delta)\le\epsilon$, it suffices to set $s<\epsilon$ and choose a sufficiently large N.

Theorem 2 *Suppose* $0<\delta\le 1/2$, $0\le s<\epsilon<1$, *and* $0<\lambda<1$. *Then we have* $\bar{\epsilon}_\lambda(\lfloor\nu s N\rfloor,N,\delta)\le\epsilon$ *as long as*

$$N\ge\frac{\epsilon}{[\lambda\nu(\epsilon-s)]^2}\left(\ln\delta^{-1}+4\lambda\nu^2\right). \quad (7)$$

Notably, if the ratio s/ϵ is a constant, then the sampling cost is only $O(\epsilon^{-1}\ln\delta^{-1})$, which is optimal with respect to all parameters ϵ, δ, and the qudit number n.

3.3 Sampling complexity of robust verification

Let ρ be the state on $\mathcal{H}^{\otimes(N+1)}$ prepared by Bob and σ_{N+1} be the reduced state after Alice performs suitable tests and accepts the state ρ. To verify the target state within infidelity ϵ, significance level δ, and robustness r (with $0\le r<1$) entails the following two conditions.

1. (Soundness) If the infidelity of σ_{N+1} with the $|G\rangle$ is larger than ϵ, then the acceptance probability $<\delta$.
2. (Robustness) If $\rho=\tau^{\otimes(N+1)}$ with $\tau\in\mathcal{D}(\mathcal{H})$ and $\epsilon_\tau\le r\epsilon$, then the acceptance probability $\ge 1-\delta$.

Let k be the number of allowed failures; then the conditions of soundness and robustness can be expressed as

$$\bar{\epsilon}_\lambda(k,N,\delta)\le\epsilon,\qquad B_{N,k}(\nu r\epsilon)\ge 1-\delta. \quad (8)$$

Denote by $N_{\min}(\epsilon,\delta,\lambda,r)$ the minimum positive integer N such that Eq. (8) holds for some $k\le N-1$. Then $N_{\min}(\epsilon,\delta,\lambda,r)$ is the minimum number of tests required for robust verification; it can be calculated numerically by using Algorithm 1 presented in the technical version.

Our following theorem provides an informative upper bound for $N_{\min}(\epsilon,\delta,\lambda,r)$ and clarifies the sampling complexity of robust verification.

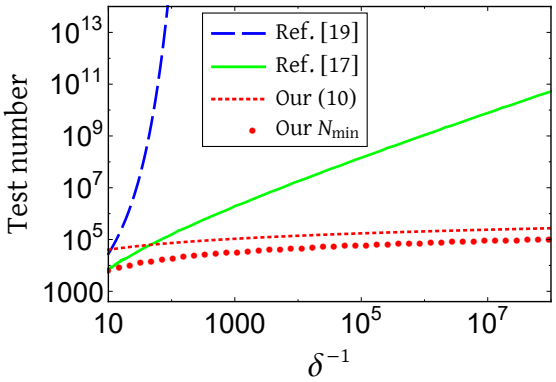


Figure 1: Number of tests required to verify a qudit graph state in the adversarial scenario within infidelity $\epsilon=0.01$, significance level δ, and robustness $r=1/2$. The red dots correspond to $N_{\min}(\epsilon,\delta,\lambda,r)$ with $\lambda=1/2$; the red dashed curve corresponds to the RHS of Eq. (10), which is an upper bound for $N_{\min}(\epsilon,\delta,\lambda,r)$. The blue curve corresponds to the protocol in [19]; and the green curve corresponds to the protocol in [17] with $\lambda=1/2$. The performances of the protocols in [21, 22] are not shown since the numbers of tests required are too large.

Theorem 3 *Suppose* $0<\lambda,\epsilon<1$, $0<\delta\le 1/2$, *and* $0\le r<1$. *Then the conditions in Eq.* (8) *hold as long as*

$$k=\left\lfloor\left(\frac{\lambda\sqrt{2\nu}+r}{\lambda\sqrt{2\nu}+1}\right)\nu\epsilon N\right\rfloor, \quad (9)$$

$$N\ge\left\lceil\left[\frac{\lambda\sqrt{2\nu}+1}{\lambda\nu(1-r)}\right]^2\frac{\ln\delta^{-1}+4\lambda\nu^2}{\epsilon}\right\rceil. \quad (10)$$

For given λ and r, the minimum number of tests is only $O(\epsilon^{-1}\ln\delta^{-1})$, which is independent of the qudit number n of $|G\rangle$ and achieves the optimal scaling behaviors with respect to the infidelity ϵ and significance level δ. If we choose $r=\lambda=1/2$ for example, then Theorem 3 implies that $N_{\min}(\epsilon,\delta,\lambda,r)\le\lceil 144\,\epsilon^{-1}(\ln\delta^{-1}+0.5)\rceil$, while numerical calculation shows $N_{\min}(\epsilon,\delta,\lambda,r)\le 67\,\epsilon^{-1}\ln\delta^{-1}$. Compared with previous protocols [17, 19, 27], our protocol improves the scaling behavior with respect to the significance level δ exponentially and even doubly exponentially, as illustrated in Fig. 1.

4 Discussion

We have proposed a highly robust and efficient protocol for verifying qudit (d is a prime) graph states in the adversarial scenario, which can be applied immediately to verifying blind MBQC. In addition to graph states, our protocol can also be used to verify many other important pure quantum states in the adversarial scenario (see the technical version for details), where the state preparation is controlled by a potentially malicious adversary, who can produce an arbitrary correlated or entangled state ρ on the whole system $\mathcal{H}^{\otimes(N+1)}$. Therefore, our verification protocol is of interest not only to blind MBQC, but also to many other tasks in quantum information processing that entail high-security.

References

[1] P. W. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th annual symposium on foundations of computer science* (IEEE, 1994) pp. 124-134.

[2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, U.K., 2000).

[3] J. Preskill, Quantum Computing in the NISQ Era and Beyond. *Quantum* **2**, 79 (2018).

[4] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking. *Nat. Rev. Phys.* **2**, 382-390 (2020).

[5] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller, Theoretical and Experimental Perspectives of Quantum Verification. *PRX Quantum* **2**, 010102 (2021).

[6] I. Šupić and J. Bowles, Self-testing of quantum systems: a review. *Quantum* **4**, 337 (2020).

[7] A. Gheorghiu, T. Kapourniotis, and E. Kashefi, Verification of quantum computation: An overview of existing approaches. *Theory of Computing Systems* **63**, 05 (2019).

[8] M. Kliesch and I. Roth, Theory of quantum system certification. *PRX Quantum* **2**, 010201 (2021).

[9] X.-D. Yu, J. Shang, and O. Gühne, Statistical methods for quantum state verification and fidelity estimation. *Adv. Quantum Technol.* **5**, 2100126 (2022).

[10] F. Arute *et al.*, Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505 (2019).

[11] H.-S. Zhong *et al.*, Quantum computational advantage using photons. *Science* **370**, 1460 (2020).

[12] R. Raussendorf and H. J. Briegel, A One-Way Quantum Computer. *Phys. Rev. Lett.* **86**, 5188 (2001).

[13] R. Raussendorf, D. E. Browne, and H. J. Briegel, Measurement-based quantum computation on cluster states. *Phys. Rev. A* **68**, 022312 (2003).

[14] D. L. Zhou, B. Zeng, Z. Xu, and C. P. Sun, Quantum computation based on d-level cluster state. *Phys. Rev. A* **68**, 062303 (2003).

[15] D. Gross, J. Eisert, N. Schuch, and D. Perez-Garcia Measurement-based quantum computation beyond the one-way model. *Phys. Rev. A* **76**, 052315 (2006).

[16] H. J. Briegel, W. Dür, R. Raussendorf, and M. Van den Nest, Measurement-based quantum computation. *Nature Physics* **5**, 19 (2009).

[17] H. Zhu and M. Hayashi, General framework for verifying pure quantum states in the adversarial scenario. *Phys. Rev. A* **100**, 062335 (2019).

[18] H. Zhu and M. Hayashi, Efficient verification of hypergraph states. *Phys. Rev. Appl.* **12**, 054047 (2019).

[19] M. Hayashi and T. Morimae, Verifiable Measurement-Only Blind Quantum Computing with Stabilizer Testing. *Phys. Rev. Lett.* **115**, 220502 (2015).

[20] T. Morimae, Y. Takeuchi, and M. Hayashi, Verification of hypergraph states. *Phys. Rev. A* **96**, 062321 (2017).

[21] Y. Takeuchi and T. Morimae, Verification of Many-Qubit States. *Phys. Rev. X* **8**, 021060 (2018).

[22] Y. Takeuchi, A. Mantri, T. Morimae, A. Mizutani, and J. F. Fitzsimons, Resource-efficient verification of quantum computing using Serfling's bound. *npj Quantum Inf.* **5**, 27 (2019).

[23] Q. Xu, X. Tan, R. Huang, and M. Li, Verification of blind quantum computation with entanglement witnesses. *Phys. Rev. A* **104**, 042412 (2021).

[24] T. Morimae and K. Fujii, Blind quantum computation protocol in which Alice only makes measurements. *Phys. Rev. A* **87**, 050301(R) (2013).

[25] Y. Takeuchi, T. Morimae, and M. Hayashi, Quantum computational universality of hypergraph states with Pauli-X and Z basis measurements. *Sci. Rep.* **9**, 13585 (2019).

[26] K. Fujii and M. Hayashi, Verifiable fault tolerance in measurement-based quantum computation. *Phys. Rev. A* **96**, 030301(R) (2017).

[27] H. Zhu and M. Hayashi, Efficient Verification of Pure Quantum States in the Adversarial Scenario. *Phys. Rev. Lett.* **123**, 260504 (2019).

[28] S. Pallister, N. Linden, and A. Montanaro, Optimal Verification of Entangled States with Local Measurements. *Phys. Rev. Lett.* **120**, 170502 (2018).

Reliability criteria for quantum data propagation on noisy quantum processors

Gaurav Saxena[1] * Ahmed Shalabi[1] † Thi Ha Kyaw[1] ‡

[1] *LG Electronics Toronto AI Lab, Toronto, Ontario M5V 1M3, Canada*

Abstract. The quantum imaginary time evolution algorithm is efficient in finding the ground state of a quantum Hamiltonian. This algorithm involves solving a system of linear equations in a classical computer. The solution is then used to propagate a quantum wavefunction across several timesteps. However, owing to the noisy nature of current quantum processors, we prove that such a quantum algorithm or the family of quantum algorithms that require classical computation of inverting a sparse matrix with high condition numbers will require very high fidelity single- and two-qubit gates. Failure to meet such criteria will result in erroneous quantum data propagation even for a relatively small quantum circuit ansatz.

Keywords: Quantum channels, quantum imaginary time evolution, noisy quantum hardware, probabilistic unitary errors

1 Introduction

Quantum computers promise quantum advantage in many applications such as factoring [1], linear systems solver [2], quantum simulations [3], and quantum chemistry/material discovery [4]. The aforementioned killer quantum applications sometimes require millions of physical qubits and gates, which current quantum hardware does not possess and hence hinders the direct experience of proposed practical quantum speedup [5]. In the current noisy intermediate-scale quantum (NISQ) era [6], many interesting practical applications are evaluated through hybrid quantum-classical quantum algorithms otherwise known as NISQ algorithms [7, 8, 9]. These algorithms are designed for short-depth quantum circuits with limited qubits and gates. One such important class of algorithms are variational quantum simulators (VQS). VQS simulate quantum systems by mapping the Hamiltonian of the system in question to a qubit/qudit Hamiltonian and simulating time evolution by trotterizing the time evolution operator into a sequence of one and two qubit gates [10].

In this work, we focus on the quantum imaginary time evolution algorithm that can efficiently find the ground state of a given Hamiltonian [11, 12, 13]. Imaginary time evolution or a Wick rotation [14] involves mapping to the time coordinate $\tau = it$. The technique establishes correspondences between Euclidean and Minkowskian geometry and between quantum mechanics and statistical mechanics. Imaginary time evolution has been a staple across many disciplines of physics like condensed matter, statistical mechanics, quantum field theory and cosmology, to name a few. While simulating imaginary time evolution can be done classically, the resources required scale exponentially with the size of the Hilbert space in question. On the other hand, quantum computing is a natural candidate for simulating quantum mechanics. However, since quantum computing involves performing unitary operations via quantum gates, quantum processors can not implement non-unitary time evolution as required for the imaginary time evolution.

To address these limitations, the quantum imaginary time evolution was proposed using a hybrid quantum-classical variational approach [11, 12, 13]. Imaginary time evolution of the Wick-rotated Schrödinger equation is implemented by encoding an initial quantum state as a parameterized quantum state $|\varphi(\tau)\rangle \approx |\phi(\theta(\tau))\rangle$ where θ is a real valued parameter vector $\theta(\tau) = (\theta_1(\tau), \theta_2(\tau), \ldots, \theta_N(\tau))$. To simulate non-unitary dynamics, an equivalent formulation of imaginary time evolution derived from a generalized McLachlan's variational principle is used to evolve the parameterized variational circuit [15]. The final step is to solve a linear equation of the form

$$M\dot{\theta} = Y$$

which involves inverting the matrix M to solve for the imaginary time evolution of the circuit. The matrix M and the vector Y depend on the ansatz circuit (used to evolve the parameters) and the given Hamiltonian [15].

The central question that we raise in our work is how well the quantum imaginary time evolution algorithm performs in the absence of any error correction or mitigation technique. In particular, what is the impact of errors incurred during each iteration on the classical matrix inversion process to solve for the vector of paramaterzing angles of the circuit, $\dot{\theta}$.

To answer the above question, we have allowed for probabilistic errors after the application of every gate. That is, we have considered the errors of the following form:

$$\mathcal{E}(\rho) := p\mathcal{N}(\rho) + (1-p)\rho\,. \tag{1}$$

By considering these types of errors, we have worked on the evolution of general mixed states under QITE. We further demand that the channel $\mathcal{N} \in \mathrm{CPTP}(A_0 \to A_1)$ is such that $\|\mathcal{N}(X)\|_2 \leq \|X\|_2 \ \forall X \in \mathfrak{L}(A_0)$ where $\|\cdot\|_2$ denotes the Frobenius norm [1], $\mathfrak{L}(A_0)$ denotes the

*gaurav.saxena@lge.com
†ahmed.shalabi@lge.com
‡thiha.kyaw@lge.com

[1] We denote $\|\cdot\|_2$ to indicate the Frobenius norm defined as $\|A\|_2 = \sqrt{\mathrm{Tr}(A^*A)}$, whereas $\|\cdot\|$ is used to indicate the matrix 2-norm defined as $\|A\| = \sqrt{\lambda_{\max} A^*A}$.

set of all linear operators in the Hilbert space A_0, and $\text{CPTP}(A_0 \to A_1)$[2] denotes the set of completely positive and trace preserving maps that map operators in system A_0 to operators in system A_1. By imposing this extra condition, we are avoiding replacement channels. Notice that the set of all channels $\mathcal{N}$ that obey $\|\mathcal{N}(X)\|_2 \leq \|X\|_2 \ \forall X \in \mathfrak{L}(A_0)$ contain the set of all probabilistic unitary channels, i.e., channels of the form

$$\mathcal{N}(\rho) = \sum_i p_i \mathcal{U}_i(\rho) = \sum_i p_i U_i \rho U_i^\dagger \tag{2}$$

where $p_i \geq 0$ and $\sum_i p_i = 1$. By allowing for such errors, we are accounting for all the probabilistic unitary errors that occur in NISQ devices. These probabilistic unitary errors also contain the set of random Pauli errors, the special cases of which are the (partially and completely) dephasing and depolarizing channels, which are considered very important in quantum computation and communication.

Considering the noise (of the form in Eq. (1)) after every gate in an arbitrary circuit with N parameters (this implies that the depth of the circuit is less than or equal to N), we derive analytical expressions for M and Y for the ideal and erroneous cases. Using the analytical expressions obtained, we derive tight upper bounds (given in Section 2) on the absolute error of the rate of change of $\dot{\theta}$. We show that the bound depends on the probability of error, the number of parameters, and the condition number of the matrix M. As noted previously, this bound is valid for all probabilistic unitary errors which include the important case of random Pauli errors, the partially dephasing errors and the partially depolarizing errors. Moreover, we were able to derive a tighter upper bound on the error incurred during each iteration in the presence of partially depolarizing noise.

2 Main results

For a given circuit that is used to evolve a parameterized quantum state with N parameters, we can find the matrix M and the vector Y (using that same circuit) that are used to update the vector of parameters $\dot{\theta}$ using the equation $M\dot{\theta} = Y$ [15]. To understand how the errors affect the evolution of the parameters, let $\mathfrak{D}(A_0)$ denote the set of all density matrices in A_0 and let $|\tilde{+}\rangle\langle\tilde{+}|$ denote the $N \times N$ matrix with all entries to be one. Then, by considering some finite probability of error occuring after every gate, we get the following results.

Theorem 1 *For probabilistic errors $\mathcal{E} \in \text{CPTP}(\text{A}_0 \to \text{A}_1)$ of the form:*

$$\mathcal{E}(\rho) = p\mathcal{N}(\rho) + (1-p)\rho \tag{3}$$

where $\rho \in \mathfrak{D}(A_0)$, $0 \leq p \leq 1$, and the channel $\mathcal{N}$ is such that $\|\mathcal{N}(X)\|_2 \leq \|X\|_2 \ \forall X \in \mathfrak{L}(A_0)$, the upper bound in the error $\epsilon := \|\Delta\dot{\theta}\|$ in computing $\dot{\theta}$ from $M\dot{\theta} = Y$ using a noisy quantum circuit is given by

$$\epsilon \leq \frac{\|\dot{\theta}\|\text{cond}(M)}{1 - \text{cond}(M)\left((1-(1-p)^{2N})\frac{\|\frac{1}{2}|\tilde{+}\rangle\langle\tilde{+}| - M\|}{\|M\|}\right)} \left((1-(1-p)^{2N})\frac{\|\frac{1}{2}|\tilde{+}\rangle\langle\tilde{+}| - M\|}{\|M\|} + (1-(1-p)^{N})\frac{\|\frac{1}{\sqrt{2}}|\tilde{+}\rangle - Y\|}{\|Y\|}\right) \tag{4}$$

if

$$\text{cond}(M) \leq \frac{\|M\|}{(1-(1-p)^{2N})\|\frac{1}{2}|\tilde{+}\rangle\langle\tilde{+}| - M\|} \tag{5}$$

where cond(M) represents the condition number of M and $|\tilde{+}\rangle\langle\tilde{+}|$ is a matrix whose entries are all 1.

Remark 1 *Note that the set of channels $\{\mathcal{N} : \|\mathcal{N}(X)\|_2 \leq \|X\|_2 \ \forall X \in \mathfrak{L}(A_0)\}$ contains the set of probabilistic unitary channels which contains the set of random Pauli errors and which in turn contains partially dephasing and partially depolarizing errors. One example of the set of channels that don't obey $\|\mathcal{N}(X)\|_2 \leq \|X\|_2$ are replacement channels.*

For partially depolarizing channels, we were able to find tighter bounds as given in the theorem below.

[2]We denote a system and its corresponding Hilbert space using an uppercase letter with a numerical subscript, like A_0 or A_1.

Theorem 2 *For partially depolarizing channels, i.e., channels of the form*

$$\mathcal{E}(\rho) = p\frac{I}{|A_0|} + (1-p)\rho \tag{6}$$

where $|A_0|$ denotes the dimension of the Hilbert space A_0, the upper bound in the error $\epsilon := \|\Delta\dot{\theta}\|$ is given by

$$\epsilon \leq \frac{\|\dot{\theta}\|\text{cond}(M)}{1 - \text{cond}(M)\left(1-(1-p)^{2N}\right)}\left(2 - (1-p)^{2N} - (1-p)^{N}\right) \tag{7}$$

if

$$\text{cond}(M) \leq \frac{1}{1-(1-p)^{2N}}\,. \tag{8}$$

2.1 Numerical analysis

We now provide some numerical analysis for the depolarizing error to understand the above bounds more clearly. We first plot in Fig. 1, the condition of Eq. (8)

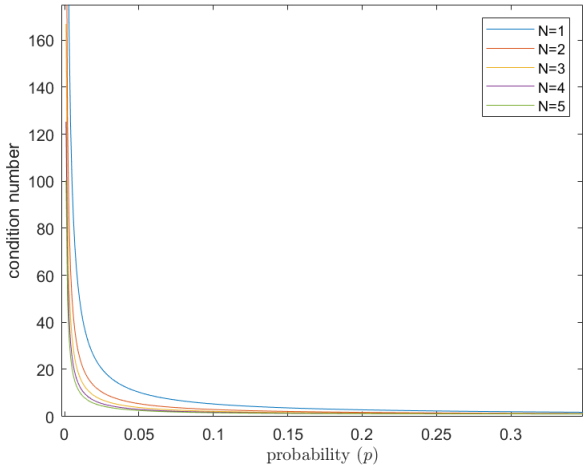


Figure 1: Maximum condition number allowed for a given probability of error (for different number of parameters)

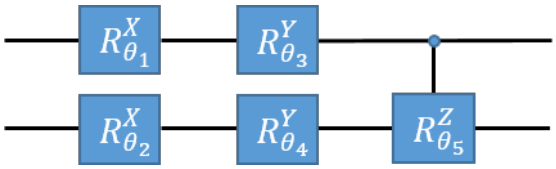


Figure 2: An example ansatz with 5 parameters

that bounds the condition number for which the upper bounds on error can be found given the probability of error. The maximum allowed condition number falls very rapidly as the probability is increased and approaches one as the probability tends to one. This implies that calculating the error bounds on $\|\Delta\dot{\theta}\|$ for circuits whose M have high condition number require very high fidelity gates.

Now, to investigate the error bound of Eq. (7), let us consider a very simple ansatz as given in Fig. 2. We randomly generate the angles (in radians): $\dot{\theta}_1 = 1.5249$, $\dot{\theta}_2 = 2.5142$, $\dot{\theta}_3 = 0.4457$, $\dot{\theta}_4 = 1.3250$, $\dot{\theta}_5 = 2.8769$. The vector $\dot{\theta} = (\dot{\theta}_1, \dot{\theta}_2, \dot{\theta}_3, \dot{\theta}_4, \dot{\theta}_5)^T$. Using these values, we found, $\|\dot{\theta}\| = 4.3447$ and the matrix M to be

$$M = \begin{pmatrix} 0.5000 & 0 & 0 & 0 & -0.0533 \\ 0 & 0.5000 & 0 & 0 & -0.1918 \\ 0 & 0 & 0.4958 & 0 & -0.0039 \\ 0 & 0 & 0 & 0.0483 & 0.1058 \\ -0.0533 & -0.1918 & -0.0039 & 0.1058 & 0.3816 \end{pmatrix}$$

Using the above M, we calculated $cond(M) = 66.9551$. Note that condition number of M is dependant on $\dot{\theta}$ and changes for different $\dot{\theta}$.

Now, to understand how the error would scale with probability and the number of parameters, we plot in Fig. 3, the error with respect to probability by keeping the condition number and the norm of $\dot{\theta}$ same. Note that the maximum probabilities for which we can bound the errors are different for different number of parameters. Also, notice how the error increases with the probability.

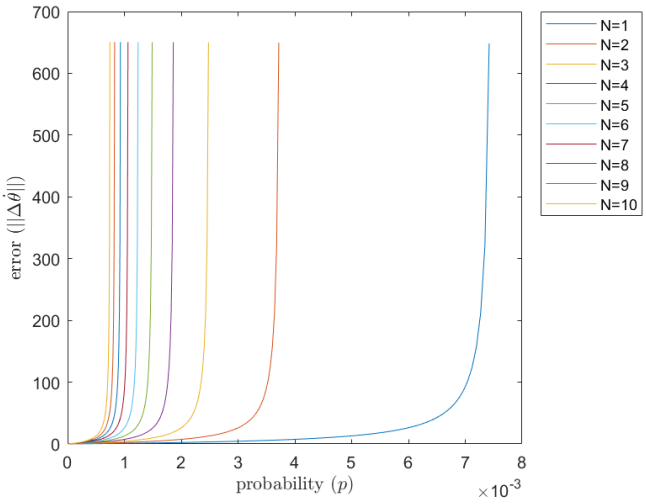


Figure 3: Plot of error vs probability under partially depolarizing errors for different number of parameters N and keeping $cond(M)$ and $\|\dot{\theta}\|$ fixed.

3 Conclusions

In this work we derived an upper bound on the error incurred during an iteration of quantum imaginary time evolution in the presence of probabilistic unitary errors. We assume that the error can occur after every gate in the ansatz circuit (assumed arbitrary) with a probability p. Furthermore, we derived a tighter upper bound for the errors incurred specifically due to partially depolarizing channels. Our results primarily show that to implement the variational imaginary time algorithm, we would require gates with very high fidelity due to the sensitivity of matrix inversion to perturbations. We also show that scalability is a problem with this algorithm because of how the error scales with the number of parameters in the circuit. Moreover, when implementing variational algorithms in general, one of the difficulties is choosing the appropriate ansatz circuits that can span the Hilbert space of the system in question while still maintaining a shallow circuit depth. While several metrics have been proposed to formalize choosing ansatz circuits [16], our work provides another important criterion for choosing effective ansatz circuits (since the matrix M is constructed from the structure of the ansatz circuit). Our work challenges the mainstream notion of hybrid quantum-classical quantum algorithms being able to perform certain computations under short-depth circuits and noisy environments while we show such algorithms in fact require very good quality quantum circuits to get reliable computational outcomes.

References

[1] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. pages 124–134, November 1994.

[2] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum Algorithm for Linear Systems of

Equations. *Phys. Rev. Lett.*, 103(15):150502, October 2009.

[3] I. M. Georgescu, S. Ashhab, and Franco Nori. Quantum simulation. *Rev. Mod. Phys.*, 86(1):153–185, March 2014.

[4] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, et al. Quantum Chemistry in the Age of Quantum Computing. *Chem. Rev.*, 119(19):10856–10915, October 2019.

[5] Thomas Hner Torsten Hoefler. Disentangling Hype from Practicality: On Realistically Achieving Quantum Advantage. May 2023. [Online; accessed 18. May 2023].

[6] John Preskill. Quantum Computing in the NISQ era and beyond. *arXiv*, January 2018.

[7] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, 94(1):015004, February 2022.

[8] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth, et al. The Variational Quantum Eigensolver: A review of methods and best practices. *Phys. Rep.*, 986:1–128, November 2022.

[9] M. Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nat. Rev. Phys.*, 3(9):625–644, September 2021.

[10] Roeland Wiersema, Cunlu Zhou, Yvette de Sereville, Juan Felipe Carrasquilla, Yong Baek Kim, and Henry Yuen. Exploring Entanglement and Optimization within the Hamiltonian Variational Ansatz. *PRX Quantum*, 1(2):020319, December 2020.

[11] Sam McArdle, Tyson Jones, Suguru Endo, Ying Li, Simon C. Benjamin, and Xiao Yuan. Variational ansatz-based quantum simulation of imaginary time evolution. *npj Quantum Inf.*, 5(75):1–6, September 2019.

[12] Mario Motta, Chong Sun, Adrian T. K. Tan, Matthew J. O'Rourke, Erika Ye, Austin J. Minnich, Fernando G. S. L. Brandão, and Garnet Kin-Lic Chan. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.*, 16(2):205–210, February 2020.

[13] Hirofumi Nishi, Taichi Kosugi, and Yu-ichiro Matsushita. Implementation of quantum imaginary-time evolution method on NISQ devices by introducing nonlocal approximation. *npj Quantum Inf.*, 7(85):1–7, June 2021.

[14] G. C. Wick. Properties of bethe-salpeter wave functions. *Phys. Rev.*, 96:1124–1134, Nov 1954.

[15] Xiao Yuan, Suguru Endo, Qi Zhao, Ying Li, and Simon C. Benjamin. Theory of variational quantum simulation. *Quantum*, 3:191, October 2019.

[16] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms. *Adv. Quantum Technol.*, 2(12):1900070, December 2019.

Robust one-sided self-testing of two-qubit states via quantum steering

Yukun Wang,[1,2] Xinjian Liu,[1] Shaoxuan Wang,[1] Haoying Zhang,[1] and Yunguang Han[3,*]
[1]*Beijing Key Laboratory of Petroleum Data Mining, China University of Petroleum, Beijing 102249, China*
[2]*State Key Laboratory of Cryptology, P.O. Box 5159, Beijing 100878, China*
[3]*College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China*

(Received 26 May 2022; revised 26 September 2022; accepted 27 September 2022; published 18 October 2022)

Entangled two-qubit states are the core building blocks for constructing quantum communication networks. Their accurate verification is crucial to the functioning of the networks, especially for untrusted networks. In this work we study the self-testing of two-qubit entangled states via steering inequalities, with robustness analysis against noise. More precisely, steering inequalities are constructed from the tilted Clauser-Horne-Shimony-Holt inequality and its general form, to verify the general two-qubit entangled states. The study provides a good robustness bound, using both local extraction map and numerical semidefinite-programming methods. In particular, optimal local extraction maps are constructed in the analytical method, which yields the theoretical optimal robustness bound. To further improve the robustness of one-sided self-testing, we propose a family of three measurement settings steering inequalities. The result shows that three-setting steering inequality demonstrates an advantage over two-setting steering inequality on robust self-testing with noise. Moreover, to construct a practical verification protocol, we clarify the sample efficiency of our protocols in the one-sided device-independent scenario.

DOI: 10.1103/PhysRevA.106.042424

I. INTRODUCTION

Quantum entangled states are the key resource of quantum information technologies, such as quantum networks [1], cryptography [2], computation [3], and metrology [4]. As we advance towards the second quantum revolution [5], the characterization and certification of quantum devices become extremely important topics in practical applications of quantum technologies [6,7].

To ensure the proper functioning of a quantum network, it is essential to certify the entangled state deployed in the network accurately and efficiently. Besides the traditional quantum state tomography method, various methods have been proposed to improve the efficiency and apply to different scenarios such as direct fidelity estimation [8], compressed sensing tomography [9], and shadow tomography [10]. In the past few years, quantum state verification (QSV) has attracted much attention by achieving remarkably low sample efficiency [11,12]. One drawback of the quantum state verification method is that it requires perfect characterization of the measurements performed by the quantum devices, and thus it is device dependent and not applicable to the untrusted quantum network. Self-testing [13,14] is a prominent candidate of quantum state certification in the device-independent (DI) scenario, in which all quantum devices are treated as black boxes. Taking advantage of Bell nonlocality [15], many important results on self-testing have been achieved such as self-testing various quantum entangled states [16–18], self-testing entangled quantum measurement [19,20], and parallel self-testing [21,22]. Self-testing has wide applications in device-independent quantum information tasks such as device-independent quantum random number generation [23,24] and quantum key distribution [25,26].

Lying between standard QSV and self-testing, there is a semi-device-independent (SDI) scenario [27] in which some parties are honest while some others may be dishonest. The certification in this scenario can be called SDI self-testing or SDI state verification. This scenario has wide applications in quantum information processing such as one-sided device-independent (1SDI) quantum key distribution [28], quantum random number generation [29], verifiable quantum computation [30], and anonymous communication [31–33]. Meanwhile, the certification in the SDI scenario is closely related to the foundational studies on quantum steering in the untrusted quantum networks [34–37]. However, not much is known about the quantum certification in the SDI scenario despite its significance. In [30,38] the authors studied the one-sided self-testing of a maximally entangled two-qubit state based on two-setting quantum steering inequality. In [39] the authors proposed various verification protocols for a Bell state based on multiple settings. For nonmaximally entangled two-qubit states, the authors in [40] realized the one-sided certification by combining fine-grained inequality [41] and analogous Clauser-Horne-Shimony-Holt (CHSH) inequalities [42], which is more complicated than traditional self-testing. In [27] the authors proposed a tilted steering inequality analogous to the tilted CHSH inequality [43] for one-sided self-testing of two-qubit states. Then they generalized the one-sided certification to general pure bipartite states by adopting the subspace method in the DI scenario [18]. In Ref. [44] a class of steering inequalities concentrating on the

*hanyunguang@nuaa.edu.cn

nonmaximally entangled bipartite-qudit state was constructed, where they achieve the bipartite-qudit state self-testing by performing only two measurements, while in Ref. [45] steering inequalities with $d+1$ measurement settings were used for self-testing the same states. However, the robustness analysis there follows the norm inequality method in [16,38] (if it is not missed); thus the result is quite weak. For the multipartite case, the studies of SDI certification are mainly focused on Greenberger-Horne-Zeilinger (GHZ) states as the generalization of the Bell state [39,46,47].

In this paper we focus on the robust one-sided self-testing of two-qubit entangled states. We construct two types of two-setting steering inequalities for general two-qubit entangled states based on tilted CHSH inequality and its general form. For the first type, an analytical and optimal robustness bound is obtained using the local extraction channel method introduced in [48]. For the second type, we get a nearly linear robustness bound using a numerical method based on the SWAP trick [17] and semidefinite programming (SDP). To put our work in perspective, we compare the robustness result in the 1SDI scenario with both DI and device-dependent scenarios. Our result can be applied to the certification of high-dimensional quantum devices as building blocks.

Furthermore, we construct three measurement settings steering inequalities for general two-qubit states, which is beyond the conventional one-sided self-testing based on two settings. In [39] the authors studied the optimal verification of the Bell state and GHZ states in the 1SDI scenario using multiple measurement settings. However, their study is limited to the maximally entangled state in the bipartite case. Based on the three-setting steering inequalities, it is shown that the robustness bound can be further improved. This opens the question of how much the resistance to noise can be improved using multiple measurement settings. Finally, to construct a practical verification protocol, we clarify the sample efficiency for our protocols in the 1SDI scenario. It is shown that approximately optimal sample efficiency can be obtained based on the steering inequalities we construct.

II. PRELIMINARIES

A. Steering scenario and steering inequalities

Let us start by recalling the steering theory. Two distant parties, Alice and Bob, are considered, and between them are many copies of the state $\rho_{AB} \in H_A \bigotimes H_B$. Bob performs two measurements, labeled y, on his particle and obtains the binary outcome b. Meanwhile, Alice receives the corresponding unnormalized conditional states $\rho_{b|y}$ and performs measurements randomly, labeled x, and obtains the binary outcome a. If Alice cannot explain the assemblage of received states by assuming preexisting states at her location and some preshared random numbers with Bob, she has to believe that Bob has steerability of her particle from a distance. To determine whether Bob has steerability of her, Alice asks Bob to run the experiment many times with her. Finally, they obtain the measurement statistics. If the statistics admit the description

$$p(a,b|x,y;\rho_{AB}) = \sum_{\lambda} p(\lambda)p(a|x,\rho_{\lambda})p(b|y,\lambda), \tag{1}$$

then Alice knows Bob does not have steerability of her. This nonsteerable correlation model is the so-called local hidden variable (LHV)–local hidden state (LHS) model [42]. The LHV-LHS decomposition is based on the idea that Bob's outcomes are determined by a local hidden random λ and Alice's outcomes are determined by local measurements on quantum state ρ_{λ}.

The combination of the statistics will give a steering inequality, where the LHV-LHS model can be used to establish local bounds for the steering inequality; violation of such inequalities implies steering. In Ref. [36] the authors introduced a family of steering inequalities for the Bell state

$$S_n \equiv \frac{1}{n}\sum_{k=1}^{n}\left\langle \hat{\sigma}_k^A B_k \right\rangle \leqslant C_n, \tag{2}$$

where C_n is the LHS bound

$$C_n = \max_{\{A_k\}}\left\{\lambda_{\max}\left(\frac{1}{n}\sum_{k=1}^{n}\hat{\sigma}_k^A B_k\right)\right\}, \tag{3}$$

with $\lambda_{\max}(\hat{O})$ denoting the largest eigenvalue of $\hat{O}$.

An approach to constructing this family of steering inequalities is transforming from Bell inequalities. Bell states are shown to maximally violate the analogous CHSH inequality [30,38,42]. For partially entangled two-qubit states, the authors in Ref. [27] constructed tilted steering inequalities from tilted CHSH inequalities [43]. In this paper we study the more general tilted steering inequalities constructed from tilted CHSH inequalities and study the robustness of one-sided self-testing based on analogous steering inequalities. Furthermore, we consider the construction of three-measurement-setting steering inequalities for general two-qubit states.

B. SDI certification and local extraction channel

In this paper we focus on a one-sided self-testing two-qubit entangled state based on the steering inequalities. To this end, we first review the concept of self-testing.

Self-testing was originally known as a DI state verification, where some observed statistics $p(a,b|x,y)$ from quantum devices can determine uniquely the underlying quantum state and the measurements, up to a local isometry. As an example, the maximal violation of CHSH inequality uniquely identifies the maximally entangled two-qubit state [14,16]. Usually, self- testing relies on the observed extremal correlations. If the quantum systems that achieve the extremal correlations are unique up to local isometries, we say the extremal correlations $p(a,b|x,y)$ self-test the target system $\{|\bar{\psi}\rangle, \bar{M}_{a|x}, \bar{N}_{b|y}\}$. Defining the local isometry as $\Phi = \Phi_{AA'} \otimes \Phi_{BB'}$, self-testing can be formally defined as

$$\begin{aligned}\Phi|\psi\rangle_{AB}|00\rangle_{A'B'} &= |\text{junk}\rangle_{AB}|\bar{\psi}\rangle_{A'B'},\\ \Phi M_{a|x}\otimes N_{b|y}|\psi\rangle_{AB}|00\rangle_{A'B'} &= |\text{junk}\rangle_{AB}\bar{M}_{a|x}\otimes\bar{N}_{b|y}|\bar{\psi}\rangle_{A'B'}.\end{aligned} \tag{4}$$

For the 1SDI scenario, only the existence of an isometry Φ_B on Bob's side is required,

$$\begin{aligned}\Phi|\psi\rangle_{AB}|0\rangle_{B'} &= |\text{junk}\rangle_{B}\otimes|\bar{\psi}\rangle_{AB'},\\ \Phi M_{b|y}|\psi\rangle_{AB}|0\rangle_{B'} &= |\text{junk}\rangle_{B}\otimes\bar{M}_{b|y}|\bar{\psi}\rangle_{AB'},\end{aligned} \tag{5}$$

where $M_{b|y}$ acts on $\mathcal{H}_B$ and $\bar{M}_{b|y}$ acts on $\mathcal{H}_{B'}$.

In addition to the above ideal definition of self-testing, it is essential to study the robustness of self-testing in the imperfect case when the obtained data deviate from the ideal value. There are two frameworks in the robustness analysis of self-testing. The first approach is based on the SWAP method by introducing an ancilla system. The desired state can be swapped out of the real quantum system and then the distance from the target state can be calculated. One way to calculate this closeness is based on the analytic method involving mathematical inequality techniques first proposed in [16]. The second one is the numerical method based on semidefinite programming combining the hierarchy strategy, which is proposed by Navascués, Pironio and Acín in Ref. [49] and called Navascués-Pironio-Acín (NPA) method. Usually, the numerical method gives much higher robustness.

The second approach is based on operator inequalities first introduced in Ref. [48], which is now widely used in the robustness analysis of self-testing. For a self-testing Bell state using CHSH inequality and a self-testing GHZ state using Mermin inequality, the operator inequalities give a nearly optimal bound. Robustness analysis of self-testing with operator inequalities can recur for local extraction map, which hinges on the idea that local measurements can be used to virtually construct a local extraction channel to extract the desired state from the real quantum system. The local extractability of the target ψ_{AB} from ρ_{AB} is quantified

$$\Xi(\rho_{AB} \to \psi_{AB}) := \max_{\Lambda_A, \Lambda_B} F((\Lambda_A \otimes \Lambda_B)(\rho_{AB}), \psi_{AB}), \quad (6)$$

where the maximum is taken over all possible local channels constructed with local measurements. For the 1SDI scenario, Alice's side is trusted, and thus the extraction channel on Alice's side is $\Lambda_A = I_A$. The lower bound of the fidelity between ρ and the target state under the observed steering inequality can be defined as one-sided extractability

$$F(\rho_{AB}, \psi_{AB}) := \inf_{\rho_{AB}: S(\rho) \geqslant S_{\text{obs}}} \max_{\Lambda_B} F(\Lambda_B(\rho_{AB}), \psi_{AB}), \quad (7)$$

where $S(\cdot)$ is the steering expression and S_{obs} is observed violation. To derive a linear bound of the fidelity about the observed steering inequality violation, real parameters s and τ must be fixed such that $F \geqslant sS_{\text{obs}} + \tau$. This is equivalent to finding Λ_B (constructed by Bob's local measurement operators M_y^b) to make

$$K \geqslant sS + \tau\mathbb{I}, \quad (8)$$

where $K := (I_A \otimes \Lambda_B^+)(\psi_{AB})$ and Λ^+ refers to the dual channel of quantum channel Λ. By taking the trace with the input state ρ_{AB} on both sides of Eq. (8), one can get $F \geqslant sS_{\text{obs}} + \tau$, in view of $\langle \Lambda_B^+(\psi_{AB}), \rho_{AB} \rangle = \langle \psi_{AB}, \Lambda_B(\rho_{AB}) \rangle$.

In the 1SDI scenario, Bob's side is untrusted; thus Eq. (8) is required to hold for Alice in two dimensions and Bob in arbitrary dimension. Since the measurements we consider in this paper is dichotomic, consideration in qubit space will be sufficient on Bob's side.

III. ONE-SIDED SELF-TESTING BASED ON TWO-SETTING STEERING INEQUALITIES

In the device-independent scenario, a general pure entangled two-qubit state

$$|\Phi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle \quad (9)$$

has been proven to be self-tested [50,51] by the maximal violation of tilted CHSH inequalities [43], which can be parametrized as

$$\hat{I}_\alpha = \alpha A_0 + A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1 \leqslant \alpha + 2, \quad (10)$$

where $\sin 2\theta = \sqrt{\frac{4-\alpha^2}{4+\alpha^2}}$. The maximum quantum value is $\sqrt{8+2\alpha^2}$. The quantum measurements used to achieve the maximal quantum violation are $\{\sigma_z; \sigma_x\}$ for Alice and $\{\cos\mu\sigma_z + \sin\mu\sigma_x; \cos\mu\sigma_z - \sin\mu\sigma_x\}$ for Bob, where $\tan\mu = \sin 2\theta$ and $\sigma_{x,z}$ are Pauli X and Z measurements.

Having $\alpha = 0$ corresponds to CHSH inequality and the state can be self-tested as a Bell state. The self-testing criterion based on this tilted CHSH inequality is robust against noise. The best robustness bound to date can be found in [48,51], in which the authors introduced the local extraction channel method. However, as claimed in [48], the theoretical optimal upper bound is not achievable. Theoretically, the optimal bound is tied to the maximum classical violation which starts to achieve nontrivial fidelity. The nontrivial fidelity that demonstrates entanglement for the target state is $F > \cos^2\theta$. Kaniewski guessed that it might be related to the fact that the quantum value of the CHSH inequality does not reach its algebraic limit of 4. Here in the 1SDI scenario, we will show that the theoretical optimal bound can be achieved.

To achieve the 1SDI self-testing criterion, we construct two types of two-setting steering inequalities, which are based on above tilted CHSH inequality by taking the measurements on Alice's side as trusted.

A. One-sided self-testing based on standard tilted CHSH steering inequality

Taking the measurements on Alice's side as trusted, the standard tilted CHSH inequality in Eq. (10) can be transformed to the analog of the tilted CHSH steering inequality

$$\begin{aligned}\hat{S}_\alpha &= \alpha A_0 + A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1 \\ &= \alpha Z + Z(B_0 + B_1) + X(B_0 - B_1) \\ &\leqslant \alpha + 2, \end{aligned} \quad (11)$$

which maintains the maximum quantum violation $S_\alpha^Q = \sqrt{8+2\alpha^2}$ as in the DI scenario. We prove that partially entangled two-qubit states can be self-tested using this analogous tilted CHSH steering inequality in a 1SDI manner. The proof is similar to DI self-testing using a tilted CHSH inequality, except that we can trust Alice's measurements now. The trustworthiness of Alice's side can simplify the proof as an advantage. Another advantage is that the theoretical optimal robustness bound can be obtained in the 1SDI scenario with this steering inequality. By contrast, the optimal bound cannot be achieved in DI self-testing with a tilted CHSH inequality. In the following we will show both the analytical proof and the robustness analysis.

1. Self-testing based on analogous tilted CHSH steering inequality

We provide the simple proof here. Though Alice's side is trustworthy, by definition only the existence of isometry on Bob's side will be sufficient to determine uniquely the state and the measurements. However, for simplicity,

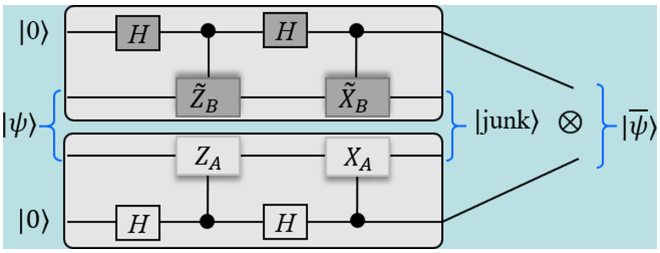


FIG. 1. The SWAP isometry applied on Alice and Bob's side, where the operators Z_A and X_A are exactly the Pauli Z and X operators.

we also introduce one isometry on Alice's side, which has been widely used in the DI scenario, shown in Fig. 1. As shown in below, with the sum of squares decomposition of a positive-semidefinite matrix [52], it is easy to find the algebraic relations that are necessarily satisfied by the target quantum state and measurements to complete the proof.

After the isometry, the systems will be

$$\begin{aligned}\Phi(|\psi\rangle) = \tfrac{1}{4}[&(I+Z_A)(I+\tilde{Z}_B)|\psi\rangle|00\rangle\\ &+ X_A(I+Z_A)(I-\tilde{Z}_B)|\psi\rangle|01\rangle\\ &+ \tilde{X}_B(I-Z_A)(I+\tilde{Z}_B)|\psi\rangle|10\rangle\\ &+ X_A\tilde{X}_B(I-Z_A)(I-\tilde{Z}_B)|\psi\rangle|11\rangle].\end{aligned} \tag{12}$$

To derive an underlying state $|\psi\rangle$ that is equivalent to the target one, the algebraic relations between the operators acting on the state should be given. We notice that the analogous tilted CHSH steering inequality $\hat{\mathcal{S}}_\alpha$ has the maximum quantum value $\mathcal{S}_\alpha^Q$. This implies that the operator $\hat{\mathcal{S}}_\alpha := \mathcal{S}_\alpha^Q\mathbb{I} - \hat{\mathcal{S}}_\alpha$ should be positive semidefinite (PSD) for all possible quantum states and measurement operators on Bob's side. This can be proven by providing a set of operators $\{P_i\}$ which are polynomial functions of A_x (Z_A and X_A) and B_y such that $\hat{\mathcal{S}}_\alpha = \sum_i P_i^\dagger P_i$ holds for any set of measurement operators satisfying the algebraic properties $A_x^2 = \mathbb{I}$ and $B_y^2 = \mathbb{I}$. The decomposition form of $\hat{\mathcal{S}}_\alpha = \sum_i P_i^\dagger P_i$ is called a sum of squares (SOS). By a SOS decomposition one can provide a direct certificate that the upper quantum bound of $\hat{\mathcal{S}}_\alpha$ is $\mathcal{S}_\alpha^Q$ from its PSD, as well as some relations between the projectors on the states, which will be used to give a self-testing statement. This method was first introduced in [50] for the family of CHSH-like Bell inequalities. Given SOS decompositions, if one observes the maximal quantum violation of the steering inequality (CHSH-like one) under state $|\psi\rangle$, then each squared term in SOS decompositions acting on $|\psi\rangle$ should be zero, i.e., $P_i|\psi\rangle = 0$. Then useful relations for the measurements operators acting on underlying state can be obtained from these zero terms.

Similar to the CHSH inequality scenario, two types of SOS decompositions for the analogous tilted CHSH operator in Eq. (11) can be given. The first one is

$$\hat{\mathcal{S}}_\alpha = \frac{1}{2\mathcal{S}_\alpha^Q}\{\hat{\mathcal{S}}_\alpha^2 + (\alpha X_A - S_0)^2\} \tag{13}$$

and the second one is

$$\begin{aligned}\hat{\mathcal{S}}_\alpha = \frac{1}{2\mathcal{S}_\alpha^Q}\Bigg\{&\left(2Z_A - \mathcal{S}_\alpha^Q\frac{B_0+B_1}{2} + \frac{\alpha}{2}S_1\right)^2\\ &+ \left(2X_A - \mathcal{S}_\alpha^Q\frac{B_0-B_1}{2} + \frac{\alpha}{2}S_2\right)^2\Bigg\},\end{aligned} \tag{14}$$

where

$$\begin{aligned}S_0 &= Z_A(B_0 - B_1) + X_A(B_0 + B_1),\\ S_1 &= Z_A(B_0 + B_1) - X_A(B_0 - B_1),\\ S_2 &= Z_A(B_0 - B_1) - X_A(B_0 + B_1).\end{aligned} \tag{15}$$

Based on the maximal violation of the analogous tilted CHSH inequality, the existence of the SOS decomposition for $\hat{\mathcal{S}}_\alpha$ implies that

$$Z_A|\psi\rangle - \tilde{Z}_B|\psi\rangle = 0, \tag{16}$$

$$\sin(\theta)X_A(I+\tilde{Z}_B)|\psi\rangle - \cos(\theta)\tilde{X}_B(I-Z_A)|\psi\rangle = 0, \tag{17}$$

where $\tilde{Z}_B := \frac{B_0+B_1}{2\cos\mu}$ and $\tilde{X}_B := \frac{B_0-B_1}{2\sin\mu}$. Then with the algebraic relations (16) and (17) and the fact that $Z_AX_A = -X_AZ_A$, Eq. (12) can be rewritten as

$$\Phi(|\psi\rangle) = |\text{junk}\rangle[\cos\theta|00\rangle + \sin\theta|11\rangle],$$

where $|\text{junk}\rangle = \frac{1}{2\cos\theta}(I+Z_A)|\psi\rangle$. This means the underlying state is unique to the target one up to local isometries, and thus completes the self-testing statement.

2. *Self-testing robustness*

Here we mainly focus on the self-testing of quantum states. For the self-testing of quantum measurements, the analysis can be related to quantum states according to Ref. [17]. The procedure is similar, starting with $\Phi M_B(|\psi\rangle)$ instead of $\Phi(|\psi\rangle)$. In this case, the figure of merit should quantify how $M_B|\psi\rangle$ is close to the ideal measurements acting on the target state.

As introduced in Sec. II B, to obtain a better self-testing robustness bound for the state, we should find the smallest value of s while keeping $K - s\hat{S} - \tau\mathbb{I}$ PSD. To this end, we first give the spectral decomposition of $\hat{\mathcal{S}}_\alpha$. Without loss of generality, we write Bob's measurements as

$$B_r = \cos\mu\sigma_z + (-1)^r\sin\mu\sigma_x, \tag{18}$$

with $r \in \{0, 1\}$ and $\mu \in [0, \pi/2]$. Then the spectral decomposition of $\hat{\mathcal{S}}_\alpha$ is

$$\hat{\mathcal{S}}_\alpha = \sum\lambda_i|\psi_i\rangle\langle\psi_i|, \quad i = 1, 2, 3, 4, \tag{19}$$

with $\lambda_1^2 + \lambda_2^2 = 8 + 2\alpha^2$, $\lambda_3 = -\lambda_2$, and $\lambda_4 = -\lambda_1$.

According to different value ranges of μ, the following two cases are discussed.

Case 1: $\cos 2\mu \geqslant \frac{\alpha^2}{4}$ *or equivalently* $\mu \in [0, \arcsin\sqrt{\frac{4-\alpha^2}{8}}]$. The eigenvalues of $\hat{\mathcal{S}}_\alpha$ have the form

$$\lambda_{1/2} = \pm\sqrt{\alpha^2 + 4\sin^2\mu} + 2\cos\mu.$$

The eigenvectors and the constraints for γ and μ are

$$\begin{aligned}|\psi_1\rangle &= \cos\gamma|00\rangle + \sin\gamma|11\rangle,\\ |\psi_2\rangle &= \sin\gamma|00\rangle - \cos\gamma|11\rangle,\end{aligned}$$

$$|\psi_3\rangle = \cos\gamma|01\rangle + \sin\gamma|10\rangle,$$
$$|\psi_4\rangle = -\sin\gamma|01\rangle + \cos\gamma|10\rangle,$$
$$\lambda_1\cos^2\gamma + \lambda_2\sin^2\gamma = \alpha + 2\cos\mu,$$
$$\lambda_2\cos^2\gamma + \lambda_1\sin^2\gamma = -\alpha + 2\cos\mu,$$
$$(\lambda_1 - \lambda_2)\cos\gamma\sin\gamma = 2\sin\mu,$$

with $\sin 2\gamma = \frac{2\sin\mu}{\sqrt{\alpha^2+4\sin^2\mu}}$.

To obtain the optimal robustness bound, we consider the following local extraction channel on Bob's side: With the probability of q_1, he performs the identity operation I on his qubit; with the probability of q_2, he performs σ_z on his qubit. By this local extraction channel, the ideal state is transformed into $K = q_1|\psi\rangle\langle\psi| + q_2\sigma_z|\psi\rangle\langle\psi|\sigma_z$. Denote $K - s\hat{I}_\alpha - \tau\mathbb{I}$ by G. The PSD condition of G requires that all the eigenvalues of it are non-negative, which gives

$$\frac{2\sin\mu s - C}{2\cos\theta\sin\theta} + \frac{1}{2} \leqslant q_1 \leqslant \frac{2\sin\mu s + C}{2\cos\theta\sin\theta} + \frac{1}{2}, \tag{20}$$

where

$$C = \sqrt{\cos^2\theta + [\beta_Q - (\alpha + 2\cos\mu)]s - 1} \times \sqrt{\sin^2\theta + [\beta_Q - (-\alpha + 2\cos\mu)]s - 1},$$

with $\beta_Q = \sqrt{8+2\alpha^2}$.

We can choose q_1 in the suitable range to saturate its upper bound, which makes G PSD. Meanwhile, we obtain the smallest value of s as

$$s = \frac{1-\cos^2\theta}{\beta_Q - (2+\alpha)} \tag{21}$$

and the corresponding value of τ is

$$\tau = 1 - \sqrt{8+2\alpha^2}s, \tag{22}$$

which is exactly equal to the theoretical optimal value. Thus we obtain the optimal robustness bound in the 1SDI scenario using the given extraction channel. Therefore, it gives the optimal robustness bound of self-testing based on the analogous tilted CHSH steering inequality

$$F = (\beta - \sqrt{8+2\alpha^2})s + 1 = (\beta - \sqrt{8+2\alpha^2})\frac{1 - \frac{\sqrt{2}\alpha}{\sqrt{4-\alpha^2}}}{2\sqrt{8+2\alpha^2} - (4+2\alpha)} + 1 \tag{23}$$

for observed violation β.

Case 2: $0 \leqslant \cos 2\mu \leqslant \frac{\alpha^2}{4}$ *or equivalently* $\mu \in (\arcsin\sqrt{\frac{4-\alpha^2}{8}}, \frac{\pi}{4}]$. The following is the local extraction channel in this case. Bob performs identity operation I with probability q_1 and performs σ_z with probability q_2. Then the ideal state is transformed into $K = q_1|\psi\rangle\langle\psi| + q_2\sigma_x|\psi\rangle\langle\psi|\sigma_x$. The PSD condition of $G := K - s\hat{I}_\alpha - \tau\mathbb{I} \geqslant 0$ gives

$$q_1 = \max\left\{0, \frac{4\sin^2\mu s^2 + (C_1 s + \tau)(C_2 s - \tau)}{(\beta_Q + 2\sin 2\theta\sin\mu + \cos^2\theta C_2 - \sin^2\theta C_1)s - 1}\right\},$$

where $\beta_Q = \sqrt{8+2\alpha^2}$. It also gives $s = \frac{1-\cos^2\theta}{\beta_Q - (2+\alpha)}$ and $\tau = 1 - \sqrt{8+2\alpha^2}s$, which turn out to obtain the same robustness bound as in case 1. (See Appendix A for details.)

In conclusion, the theoretical linear optimal robustness bound can be obtained for self-testing of two-qubit entangled states using the analogous tilted CHSH steering inequality. Different from self-testing in the DI scenario, theoretical optimal robustness bound can be obtained using the local extraction channel method. The reason might be that the extraction channel is needed only on one side in the steering scenario without coordination.

Comparison with the DI and device-dependent scenarios. To put our work into perspective, we compare the certification in the 1SDI scenario with both DI and device-dependent (DD) scenario.

In the DD scenario, the measurements on both sides are trusted and equal to the ideal measurements. In this case, we have

$$\hat{I}_\alpha = \alpha A_0 + A_0B_0 + A_0B_1 + A_1B_0 - A_1B_1 \tag{24}$$
$$= \alpha Z + 2\cos\mu ZZ + 2\sin\mu XX, \tag{25}$$

where $\sin 2\theta = \sqrt{\frac{4-\alpha^2}{4+\alpha^2}}$ and $\tan\mu = \sin 2\theta$. It could be shown that

$$|\Psi\rangle\langle\Psi| \geqslant \frac{\hat{I}_\alpha}{\sqrt{8+2\alpha^2}}. \tag{26}$$

Thus, in the trusted measurement scenario, we have the lower bound of the fidelity

$$F_{\mathrm{DD}} \geqslant \frac{\beta}{\sqrt{8+2\alpha^2}}. \tag{27}$$

In the DI scenario, the authors in [51] conjectured the lower bound of fidelity

$$F_{\mathrm{DI}} \geqslant s_\alpha\beta + \mu_\alpha, \tag{28}$$

with

$$s_\alpha = \frac{(\sqrt{8+2\alpha^2} + 2 + \alpha)(3\sqrt{8+2\alpha^2} - \sqrt{4-\alpha^2} - \alpha\sqrt{2})}{4(2-\alpha)^2\sqrt{8+2\alpha^2}}, \tag{29}$$

$$\mu_\alpha = 1 - s_\alpha\sqrt{8+2\alpha^2}. \tag{30}$$

Their comparison with the SDI scenario is given in Fig. 2. In the case of $\alpha = 0$, it corresponds to the CHSH inequality and the target state is a singlet. The two other cases correspond to the tilted CHSH inequality and partially entangled two-qubit states. Obviously, it has $F_{\mathrm{DD}} > F_{\mathrm{1SDI}} > F_{\mathrm{DI}}$ for all three cases. For $\alpha = 0$, the nontrivial fidelity bound of the singlet state is 0.5. The results show that the nontrivial fidelity bound can be obtained in the DI scenario when the quantum value is larger than 2.105, while for the 1SDI and DD scenarios the bounds are 2 and $\sqrt{2}$, respectively. For $\alpha = 0.5$, the nontrivial fidelity bound of the target state is 0.672. The results show that the nontrivial fidelity bound can be obtained in the DI scenario when the quantum value is larger than 2.655, while for the 1SDI and DD scenarios the bounds are 2.5 and 1.958, respectively. For $\alpha = 1$, the nontrivial fidelity bound of the

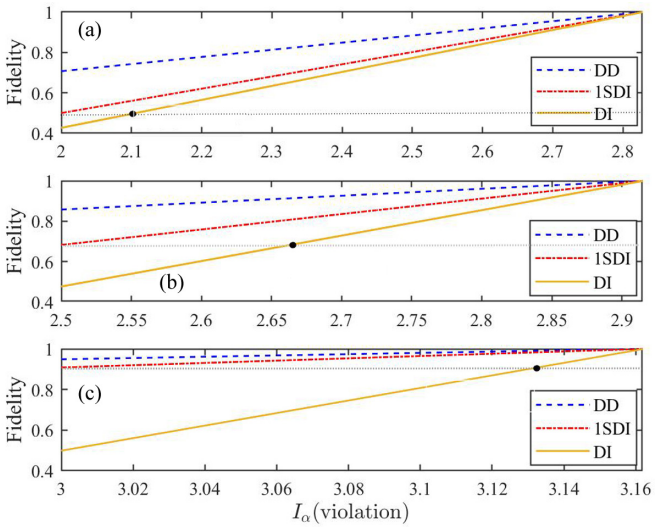


FIG. 2. Comparison of robustness bound between the DI (yellow solid line), 1SDI (red dash-dotted line), and DD (blue dotted line) scenarios for (a) $\alpha = 0$ and $\beta = 2.015$, (b) $\alpha = 0.5$ and $\beta = 2.655$, and (c) $\alpha = 0.1$ and $\beta = 3.103$.

target state is 0.816. The results show that the nontrivial fidelity bound can be obtained in the DI scenario when the quantum value is larger than 3.103, while for the 1SDI and DD scenarios the bounds are 3 and 2.581, respectively. It is shown that with the increase of α, especially for $\alpha = 1$, the 1SDI self-testing bound is much better than in the DI scenario and closer to the DD scenario. Thus our method achieves significant improvement in the 1SDI certification of less entangled two-qubit states, which is comparable to the device-dependent scenario.

B. One-sided self-testing based on general tilted CHSH inequality

In this section we construct two-setting steering inequalities from the general tilted CHSH inequality [43]

$$\hat{\mathcal{S}}_{\alpha,\beta} = \alpha A_0 + \beta A_0 B_0 + \beta A_0 B_1 + A_1 B_0 - A_1 B_1. \quad (31)$$

The maximal classical and quantum bounds are $\alpha + 2(1+\beta)$ and $\sqrt{(4+\alpha^2)(1+\beta^2)}$, respectively. The quantum bound can be achieved by pure two-qubit states (9) and corresponding measurements settings $\{\sigma_z; \sigma_x\}$ for Alice and $\{\cos\mu\sigma_z + \sin\mu\sigma_x; \cos\mu\sigma_z - \sin\mu\sigma_x\}$ for Bob, with $\sin 2\theta = \sqrt{\frac{4-\alpha^2\beta^2}{4+\alpha^2}}$ and $\tan\mu = \frac{\sin 2\theta}{\beta}$.

Taking the measurements on Alice's side as trusted, this Bell inequality can be transformed into

$$\hat{S}_{\alpha,\beta} = \alpha Z + \beta Z(B_0 + B_1) + X(B_0 - B_1), \quad (32)$$

which is a steering inequality. However, we can also introduce two other measurements to represent $B_0 + B_1$ and $B_0 - B_1$, thus rewriting the steering inequality as

$$S^{(1)}_{\alpha,\beta} = \alpha\langle Z\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle \leqslant \sqrt{1+(\alpha+\beta)^2}, \quad (33)$$

with $\beta > 0$. The maximal quantum violation is $\beta + \sqrt{1+\alpha^2} := S_Q$.

This form of steering inequality allows us to compare the construction with the one proposed in Ref. [27], which changes the marginal term to Bob's side,

$$S^{(2)}_{\alpha,\beta} = \alpha\langle B_0\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle \leqslant \alpha + \sqrt{1+\beta^2}, \quad (34)$$

with $\beta^2 = \alpha^2 + 1$, and keeps the quantum bound as in Eq. (33). It should be remarked that the constraints of β and α given in [27] can be relaxed to $\beta^2 \geqslant \alpha^2 + 1$, which we prove in Appendix D with SOS decomposition related to the steering operators.

Both of these steering inequalities of $S^{(1)}_{\alpha,\beta}$ and $S^{(2)}_{\alpha,\beta}$ can be used to self-test a pure partially entangled state with $\sin(2\theta) = \frac{1}{\sqrt{1+\alpha^2}}$. The only difference between our construction and the one in [27] is the exchanging roles of Alice and Bob. The advantage of our construction will be shown later. Before that, we should prove that the maximum violation of both $S^{(1)}_{\alpha,\beta}$ and $S^{(2)}_{\alpha,\beta}$ can be used to self-test a pure partially entangled state, though the proof for self-testing based on $S^2_{\alpha,\beta}$ was already given in (33). However, a different proof is provided here which is based on the SOS decomposition related to the steering inequality and the isometry given in Fig. 1. The benefit of this proof is that the constraints of $\beta^2 = \alpha^2 + 1$ can be relaxed (details are in Appendix D).

In the following we study the robustness of the self-testing based on these two steering inequalities. In Ref. [27] the robustness of one-sided self-testing was studied only for maximally entangled states based on operator inequalities. For the case $\alpha = 0$, when the violation of the steering inequality is $S = 2 - \epsilon$, the actual state is $24\sqrt{\epsilon} + \epsilon$, close to the target state (see also Ref. [13]). More precisely, the relation between the fidelity and the steering inequality value is

$$F \geqslant 1 - 24\sqrt{2-S} - (2-S), \quad (35)$$

which is quite loose. A nontrivial fidelity bound $f > \frac{1}{2}$ can only be obtained when the violation is larger than 1.999 57, which makes the robustness analysis in the one-sided self-testing impractical. Here we have improved this bound to be

$$F \geqslant \frac{S-2}{4-2\sqrt{2}} + 1, \quad (36)$$

which is the theoretical optimal linear bound. The local extraction channel to achieve this bound is constructed in Appendix B, and this channel coincides with the extraction channel in the DI scenario introduced in Ref. [48]. However, the reason this channel is used was not explained in Ref. [48]. Here we point out that the channel is the optimal local channel that the local party can take.

For the other case of α, we give the robustness analysis of one-sided self-testing based on the numerical method. The details are given in Appendix C. The method works for general, pure two-qubit states and the results show that the robustness bound is nearly linear.

The comparison of the robustness bound of self-testing based of Eqs. (33) and (34) is given in Fig. 3, where we take $\alpha = 1$ and $\beta = \sqrt{2}$ as an example. As shown, the one with trusted partial information gives a better robustness bound. The reason is that construction of the steering inequality (33) shows a smaller LHS bound compared with the inequality

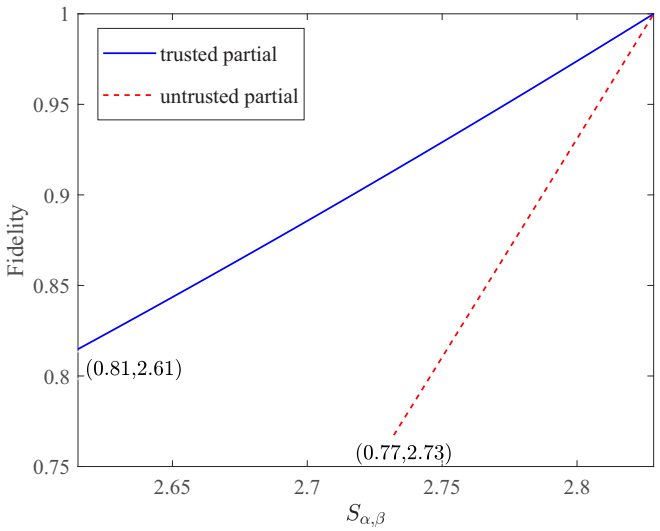


FIG. 3. Comparison of the robustness bound of self-testing based on the two-setting steering inequality $S_{\alpha,\beta}$ of Eqs. (33) and (34), where $\alpha = 1$ and $\beta = \sqrt{2}$.

(34); however, it keeps the quantum maximum bound. Thus the inequality (34) demonstrates an advantage for self-testing; it is more robust than using an untrusted party's partial measurement expectation. Actually, in addition to the advantage in self-testing, the steering inequality constructed with trusted partial expectation can also have fewer constraints on variants α and β, and thus could provide more reasonable steering inequalities (see Appendix D for details).

IV. ONE-SIDED SELF-TESTING BASED ON THREE-SETTING STEERING INEQUALITIES

So far the steering inequalities we have considered are all of two measurement settings. In this section we introduce more measurements settings in constructing steering inequalities. Later we show that adding more measurement settings can help increase the robustness in one-sided self-testing. We construct a family of three-setting steering inequalities

$$I_{\alpha,\beta} \equiv \alpha\langle Z\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle + \langle YB_2\rangle \leqslant \sqrt{2+(\alpha+\beta)^2}, \tag{37}$$

where $\beta \geqslant 0$. These inequalities can be viewed as a generalization of analogous tilted CHSH steering inequalities (34). A third measurement involving the Pauli Y measurement is added. Similar to the discussion of the two-setting scenario, the partial expectation in the construction can also be untrusted party Bob's measurement B_0. Thus $I_{\alpha,\beta} \equiv \alpha\langle B_0\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle + \langle YB_2\rangle$ is constructed. These two slightly different inequalities have a different LHS bound while keeping the same quantum bound (for a detailed discussion and their proof for self-testing a two-qubit partially entangled state see Appendix D).

Here we just consider the first case in the main text for simplicity and give its self-testing robustness bound. The LHS bound is the maximum violation that we can have, assuming Bob has a preexisting state known to Alice, rather than half of an entangled state shared with Alice. Bob's system may be derived from a classical system; thus we can denote his corresponding declared result by a random variable $B_k \in \{-1, 1\}$ for $k = 0, 1$. As shown in [36], it is easy to see that

$$I_{\mathrm{LHS}} = \max_{B_k} \lambda_{\max}(I_{\alpha,\beta}), \tag{38}$$

where $\lambda_{\max}(\hat{O})$ denotes the largest eigenvalue of $\hat{O}$. Then the LHS bound of Eq. (37) is shown to be $\sqrt{2+(\alpha+\beta)^2}$.

The maximum quantum bound is $\beta + \sqrt{4+\alpha^2} := S_Q$. This can be verified by the fact that $S_Q\mathbb{I} - I_{\alpha,\beta}$ is PSD. More precisely,

$$\begin{aligned} S_Q\mathbb{I} - \hat{I}_{\alpha,\beta} = {} & \frac{\beta}{2}(\mathbb{I} - ZB_0)^2 \\ & + \frac{\sqrt{\alpha^2+4}}{4}\left(\mathbb{I} - \frac{\alpha}{\sqrt{4+\alpha^2}}Z - \frac{2}{\sqrt{4+\alpha^2}}XB_1\right)^2 \\ & + \frac{\sqrt{\alpha^2+4}}{4}\left(\mathbb{I} - \frac{\alpha}{\sqrt{4+\alpha^2}}Z - \frac{2}{\sqrt{4+\alpha^2}}YB_2\right)^2. \end{aligned} \tag{39}$$

The quantum systems used to achieve the maximal quantum violation are $B_0 = Z$, $B_1 = X$, $B_2 = -Y$, and $|\Phi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$, with $\sin 2\theta = \frac{2}{\sqrt{4+\alpha^2}}$, which in turn can be self-tested when the maximum violation is reached (see Appendix D).

Here, for simplicity, we just consider the case of $\alpha = 0$ and $\beta = 1$. Assuming Bob's measurements are untrusted, without loss of generality, they can be written as $B_{0,1} = \cos\mu\sigma_z \pm \sin\mu\sigma_x$ and $B_2 = \cos\mu_1\cos\mu_2\sigma_z + \cos\mu_1\sin\mu_2\sigma_x + \sin\mu_1\sigma_y$. Due to the asymmetry of $I_{\alpha,\beta}$ introduced by the form of B_2, the spectral decomposition of it is not easy, which leads to the difficulty in constructing a local extraction channel making G PSD. We divide G into two parts. If each part is PSD, then the whole matrix G is PSD,

$$\begin{aligned} G &:= K - [s(ZB_0 + XB_1 + YB_2) + \tau\mathbb{I}] \\ &= K_1 - s(ZB_0 + XB_1) - \tau_1\mathbb{I} + K_2 - sYB_2 - \tau_2\mathbb{I}, \end{aligned} \tag{40}$$

where $K_1 + K_2 = K$ defines the two parts.

We consider the local extraction channel which ensures the parts of $G_1 := K_1 - s(ZB_0 + XB_1) - \tau_1\mathbb{I}$ and $G_2 := K_2 - sYB_2 - \tau_2\mathbb{I}$ PSD simultaneously (see Appendix F for details of the channel construction). The following robustness bound of self-testing in the three-setting steering scenario is obtained:

$$F \geqslant sS_{\mathrm{obs}} + \tau \geqslant \frac{3}{12-4\sqrt{2}}(S_{\mathrm{obs}} - 3) + 1. \tag{41}$$

It should be noted that here we did not get the expected robustness bound of $F \geqslant \frac{(S_{\mathrm{obs}}-3)}{2(3-\sqrt{3})} + 1$. This may be because the local extraction channel strategy we consider here is not optimal. It may be possible to find a better extraction strategy to obtain that bound. However, though the bound we give is optimal, it is still better than two-setting analogous CHSH steering scenarios.

For a straightforward comparison between different inequalities, we transform the steering inequalities into the games characterized by the guessing probability which belongs to the same interval $[\frac{1}{2}, 1]$. In the case of $\alpha = 0$, we have $P = \sum_{i=0,1} p(a = b|A_iB_i) = \frac{1}{2} + \frac{S}{2S_Q}$, which is the successful probability of the nonlocal game guessing the other

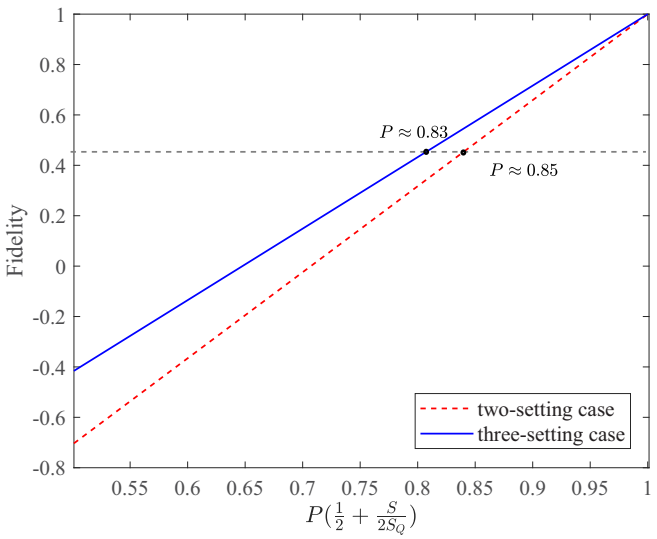


FIG. 4. Comparison of robustness bounds for one-sided self-testing of a singlet based on three-setting and two-setting steering inequalities.

party's outcomes. For the other case, we can also find a nonlocal game, namely, the guessing score is related to the inequalities (33) and (37), respectively. (See Appendix E for details.) We define the guessing probability as the probability for untrusted parties to successfully guess the trusted parties' outcomes, which is also important for the sample efficiency analysis in the next section. Based on the guessing probability, we can compare the robustness bound for one-sided self-testing of a singlet based on three-setting and two-setting steering inequalities. The result is shown in Fig. 4, where the three-setting steering inequality we constructed gives a better robustness bound. It is worth studying whether steering inequalities with more measurement settings can be constructed and further improve the robustness of one-sided self-testing.

V. SAMPLE EFFICIENCY

To construct a practical quantum verification protocol, it is crucial to study the sample efficiency [11,12,39,53]. Sample efficiency is used to study the performance of the self-testing criteria in the finite copy regime in a way that some of the state copies are measured to warrant the rest of the states being close to the target state.

Consider a quantum device producing the states $\rho_1, \rho_2, \ldots, \rho_N$ in N runs. Our task is to verify whether these states are sufficiently close to the target state $|\Phi\rangle \in \mathcal{H}$ on average. Here the one-sided extractability is a natural choice for quantifying the closeness in the one-sided self-testing scenario.

For the extraction channel method, we obtain a linear relation between the extractability and the observed value of the steering inequalities

$$F \geqslant sS_{\text{obs}} + \tau. \tag{42}$$

Since $\tau = 1 - sS_Q$, we have

$$s(S_Q - S_{\text{obs}}) \geqslant 1 - F. \tag{43}$$

The first step in constructing the verification protocol is to view the steering inequalities as testing games. (Details of the transformation of steering inequalities to testing games are shown in Appendix E.) Based on this, results of unmeasured copies can be guaranteed based on the measured copies. Define p as the guessing probability of the game for a single state. For the steering inequalities in Eqs. (33) and (37), when $\alpha = 0$, which corresponds to the singlet state, the testing game is straightforward based on the outcomes of the same Pauli measurements. When $\alpha > 0$, which corresponds to the nonmaximally entangled state, virtual testing games are constructed from the steering inequalities in Appendix E. For these testing games, we have

$$p = \frac{1}{4}\sum_{i=0,1} p(a = b|A_iB_i) = \frac{1}{2} + \frac{S}{2S_Q}. \tag{44}$$

This relation between the guessing probability and the violation of steering inequalities is essential for the study of sample efficiency. For the analogous CHSH steering inequality in Eq. (11), we have $p = \frac{1}{4}\sum_{a\otimes b=ij} p(a,b|A_iB_i) = \frac{1}{2} + \frac{S}{4}$. This probability corresponds to the successful probability to win the game of $a \otimes b = ij$ for Alice and Bob. For steering inequalities in Eq. (11) for $\alpha \neq 0$ and Eq. (34), we have not found corresponding testing games. One may resort to other theories to study its performance in the finite regime, such as [54].

Defining $\epsilon = 1 - F$ as the infidelity and combining Eqs. (43) and (44), we have

$$p \leqslant 1 - \frac{\epsilon}{2sS_Q}. \tag{45}$$

Defining $c = \frac{1}{2sS_Q}$, in general we have

$$p \leqslant 1 - c\epsilon. \tag{46}$$

Now for these inequalities which correspond to a testing game, we are ready to estimate the number of copies sufficient to exceed a certain bound on the average one-sided extractability. Suppose the states in the test are independently distributed. The goal is to guarantee that the average one-sided extractability of the states $\rho_1, \rho_2, \ldots, \rho_N$ is larger than $1 - \epsilon$ with significance level δ (confidence level $1 - \delta$). According to Ref. [53], the scaling of sample efficiency depends on whether the quantum bound and algebraic bound coincide for the games between participants. When the quantum bound and algebraic bound coincide, the number of copies satisfies

$$N \geqslant \frac{\ln \delta^{-1}}{\ln(1 - c\epsilon)^{-1}} \approx \frac{\ln \delta^{-1}}{c\epsilon}. \tag{47}$$

For all the steering inequalities we have considered in this paper, the two-setting inequality (33) and the three-setting inequality (37) satisfy this condition. In that case, the maximal guessing probability 1 can be obtained in the testing games according to the strategy given in Appendix E. Thus we obtain the approximately optimal sample efficiency for one-sided self-testing of general two-qubit states in both the two-setting and three-setting cases, which is comparable to the number needed in quantum state verification.

For the analogous CHSH steering inequality in Eq. (11), the quantum bound and algebraic bound are different. The

number of copies needed satisfies

$$N = O\left(\frac{\ln \delta^{-1}}{c^2\epsilon^2}\right), \tag{48}$$

according to Ref. [53].

In this section we studied the sample efficiency for one-sided self-testing of two-qubit entangled states. Based on the steering inequalities we constructed, approximately optimal sample efficiency can be obtained in the SDI scenario, which is comparable to the device-dependent scenario. For the general DI scenario, the scaling of testing number is usually in quadratic form. Thus our strategies demonstrate a significant advantage over DI self-testing in sample efficiency.

VI. CONCLUSION

In this paper we studied the one-sided self-testing of general, pure two-qubit states in the untrusted quantum network in which one party is not honest. The self-testing strategies are based on the violation of quantum steering inequalities. To achieve this goal, we first studied two setting scenarios, where the steering inequalities can be constructed from standard tilted CHSH inequalities and its general form. Based on these steering inequalities, we studied the robustness of one-sided self-testing using both the local extraction map method and the numerical semidefinite-programming method. In particular, the local extraction map method has been shown to provide the analytical and theoretical optimal linear bound. Our result also demonstrates an explicit approach to construct the local extraction channel. The comparison with the device-independent scenario and the device-dependent scenario shows clearly that the robustness of SDI certification lies in the middle. The numerical method involving SDP and the SWAP trick gives a nearly linear robustness bound for general, pure two-qubit states. To construct a practical certification protocol, we also clarified the sample efficiency of our 1SDI self-testing protocols. The results show that approximately optimal sample efficiency can be obtained based on the steering inequalities we constructed.

Furthermore, we constructed three-measurement-setting steering inequalities for general two-qubit states, for a partially entangled state. It was shown that the robustness bound can be further improved by introducing the third measurement setting. It is worth studying whether steering inequalities with more measurement settings can be constructed and further improve the robustness of one-sided self-testing. This question is also of interest in foundational studies on quantum steering. The improvement of the robustness bound in our work can be applied to the certification of high-dimensional quantum devices as building blocks. In the future, our results may be generalized to generic bipartite pure states, multipartite GHZ states, and other quantum states.

ACKNOWLEDGMENTS

This research was supported by National Nature Science Foundation of China (Grants No. 62101600, No. 62201252, and No. 61901218), Science Foundation of China University of Petroleum, Beijing (Grant No. 2462021YJRC008), State Key Laboratory of Cryptology (Grant No. MMKFKT202109), and Natural Science Foundation of Jiangsu Province, China (Grant No. BK20190407).

APPENDIX A: LOCAL EXTRACTION CHANNEL METHOD FOR SELF-TESTING BASED ON AN ANALOGOUS TILTED CHSH INEQUALITY

This Appendix provides the robust bound of the self-testing based on an analogous tilted CHSH inequality in case 2, i.e., $0 \leqslant \cos 2\mu \leqslant \frac{\alpha^2}{4}$ or equivalently $\mu \in (\arcsin\sqrt{\frac{4-\alpha^2}{8}}, \frac{\pi}{4}]$. In this case, the eigenvalues of the decomposition of $\hat{S}_\alpha = \sum \lambda_i |\psi_i\rangle\langle\psi_i|$ are $\lambda_{1,2} = \sqrt{\alpha^2 + 4\sin^2\mu} \pm 2\cos\mu$. The constraints between γ and μ are

$$\begin{aligned}
\lambda_1 \cos^2\gamma - \lambda_2 \sin^2\gamma &= \alpha + 2\cos\mu,\\
\lambda_2 \cos^2\gamma - \lambda_1 \sin^2\gamma &= \alpha - 2\cos\mu,\\
(\lambda_1 + \lambda_2)\cos\gamma \sin\gamma &= 2\sin\mu.
\end{aligned}$$

Still $\sin 2\gamma = \frac{2\sin\mu}{\sqrt{\alpha^2+4\sin^2\mu}}$.

The following is the local extraction channel in this case. Bob takes rotation operation I with probability q_1 and takes σ_z with probability q_2. Then the ideal state is transformed into $K = q_1|\psi\rangle\langle\psi| + q_2\sigma_x|\psi\rangle\langle\psi|\sigma_x$. The PSD requirement of $G := K - s\hat{I}_\alpha - \tau\mathbb{I} \geqslant 0$ gives

$$\begin{pmatrix}
q_1\cos^2(\theta) - C_1 s - \tau & 0 & 0 & q_1\frac{\sin 2\theta}{2} - 2\sin\mu s\\
0 & q_2\cos^2\theta - C_2 s - \tau & q_2\frac{\sin 2\theta}{2} - 2\sin\mu s & 0\\
0 & q_2\frac{\sin 2\theta}{2} - 2\sin\mu s & q_2\sin^2\theta + C_1 s - \tau & 0\\
q_1\frac{\sin 2\theta}{2} - 2\sin\mu s & 0 & 0 & q_1\sin^2\theta + C_2 s - \tau
\end{pmatrix} \geqslant 0, \tag{A1}$$

where $C_1 = \alpha + 2\cos\mu$ and $C_2 = \alpha - 2\cos\mu$. The eigenvalues of G are

$$\lambda_{1,2} = \frac{G_{11} + G_{44} \pm \sqrt{(G_{11} - G_{44})^2 + 4G_{14}^2}}{2}, \tag{A2}$$

$$\lambda_{3,4} = \frac{G_{22} + G_{33} \pm \sqrt{(G_{22} - G_{33})^2 + 4G_{23}^2}}{2}, \tag{A3}$$

which should be positive to make G PSD,

$$q_1 \geqslant \frac{4\sin^2\mu s^2 + (C_1 s + \tau)(C_2 s - \tau)}{(\beta_Q + 2\sin 2\theta \sin\mu + \cos^2\theta C_2 - \sin^2\theta C_1)s - 1},$$

$$q_2 \geqslant \frac{4\sin^2\mu s^2 + (C_2 s + \tau)(C_1 s - \tau)}{(\beta_Q + 2\sin 2\theta \sin\mu + \cos^2\theta C_1 - \sin^2\theta C_2)s - 1},$$

where $\beta_Q = \sqrt{8+2\alpha^2}$.

We can also set $s = \frac{1-\cos^2\theta}{\beta_Q-(2+\alpha)}$ and $\tau = 1-\sqrt{8+2\alpha^2}s$, keeping q_1 in the above range. This gives the same bound as in case 1. To this end, we take q_1 to be the maximum between 0 and the values which saturate the above two inequalities about q_1.

APPENDIX B: LOCAL EXTRACTION CHANNEL METHOD FOR SELF-TESTING BASED ON REVERSE CHSH INEQUALITY

The analogous CHSH steering operator $\hat{S} = ZB_0 + XB_1$ has the spectral decomposition

$$\hat{S} = \sum \lambda_i |\psi_i\rangle\langle\psi_i|, \tag{B1}$$

with $\lambda_1^2 + \lambda_2^2 = 4$, $\lambda_3 = -\lambda_2$, and $\lambda_4 = -\lambda_1$. Precisely,

$$\lambda_1 = \sqrt{2}(\cos\mu + \sin\mu), \quad \lambda_2 = \sqrt{2}(\cos\mu - \sin\mu), \tag{B2}$$

where Bob's measurements are written as $B_r = \cos\mu\sigma_z + (-1)^r \sin\mu\sigma_x$, with $r = 0, 1$. In the case of $\mu \in (0, \pi/4]$ we have $\lambda_1, \lambda_2 \geqslant 0$ and

$$\begin{aligned} |\psi_1\rangle &= \frac{|00_B\rangle + |11_B\rangle}{\sqrt{2}}, \quad |\psi_2\rangle = \frac{|00'_B\rangle + |11'_B\rangle}{\sqrt{2}}, \\ |\psi_3\rangle &= \frac{|01'_B\rangle - |10'_B\rangle}{\sqrt{2}}, \quad |\psi_4\rangle = \frac{|01_B\rangle - |10_B\rangle}{\sqrt{2}}, \end{aligned} \tag{B3}$$

where

$$\begin{aligned} 0_B &= \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle, \quad 1_B = \sin\frac{\pi}{8}|0\rangle - \cos\frac{\pi}{8}|1\rangle, \\ 0'_B &= \cos\frac{\pi}{8}|0\rangle - \sin\frac{\pi}{8}|1\rangle, \quad 1'_B = \sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle. \end{aligned}$$

We consider the following local extraction channel. Bob takes the rotation operation $R_1 = I$ on his qubit with the probability of q_1 and takes $R_2 = \sigma_z$ on his qubit with the probability of q_2. The ideal state is transformed into the mixture of the Bell operator eigenvectors $|\psi\rangle := q_1|\psi_1\rangle\langle\psi_1| + q_2|\psi_2\rangle\langle\psi_2|$. In this case, $G := K - s\hat{S} - \tau\mathbb{I}$ is diagonal and the PSD requirement gives

$$\begin{aligned} q_1 - s\lambda_1 - \tau &\geqslant 0, \\ q_2 - s\lambda_2 - \tau &\geqslant 0, \\ \mathrm{Tr}(\rho) = p_1 + p_2 &= 1, \\ \mathrm{Tr}(\rho\hat{S}) = \lambda_1 p_1 + \lambda_2 p_2 &= S, \end{aligned}$$

where we set $\tau = 1 - 2s$. By simplifying, we have $s\lambda_1 - 2s + 1 \leqslant q_2 \leqslant -s\lambda_2 + 2s$, which gives us $s \geqslant \frac{1}{4-(\lambda_1+\lambda_2)} \geqslant \frac{1}{4-2\sqrt{2}}$. This gives the following robustness bound of self-testing via the steering inequality:

$$F \geqslant sS + \tau \geqslant \frac{S-2}{4-2\sqrt{2}} + 1. \tag{B4}$$

In addition, we get the constraints on the rotation probability

$$(1+\sqrt{2})(\cos\mu + \sin\mu + 1) \leqslant q_1 \leqslant 1. \tag{B5}$$

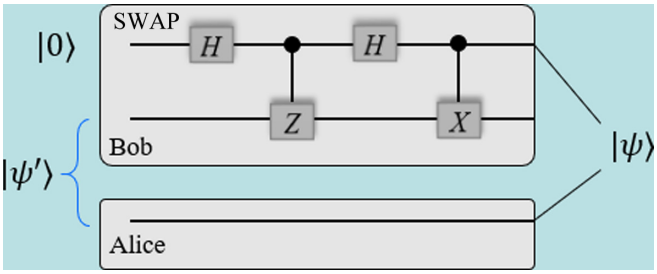


FIG. 5. One-sided SWAP isometry applied on Bob's side.

For the case of $\mu \in (\frac{\pi}{4}, \frac{\pi}{2})$, the local extraction channel is considered as follows. Bob takes the rotation $R_1 = I$ with the probability of q_1 and $R_2 = \sigma_x$ with the probability of q_2. This gives the same robustness bound.

Above we found that the optimal linear bound and nontrivial fidelity can be obtained as long as the steering inequality is violated. However, as shown in Ref. [48], that nontrivial fidelity bound could not be obtained for an inequality violation at 2, with this local extraction channel. The reason might be that to define the appropriate extraction channel, the two local sites need coordinating. In the DI scenario, both sides are not trusted. The decomposition of the Bell operator is related to both Alice's and Bob's local measurement directions.

Once Alice and Bob could inform each other what measurement directions they choose (do classical communication), it is possible for them to define the appropriate local rotation channel which could rotate the idea states to be the eigenvectors of the Bell operator with positive eigenvalues. It could make $G := K - s\hat{I} - \tau\mathbb{I}$ PSD. In this case, it is easy to find that s and t are the optimal ones. However, allowing communication is not usually device independent. Thus, in the DI scenario, when coordination is needed, the nontrivial fidelity could not be reached.

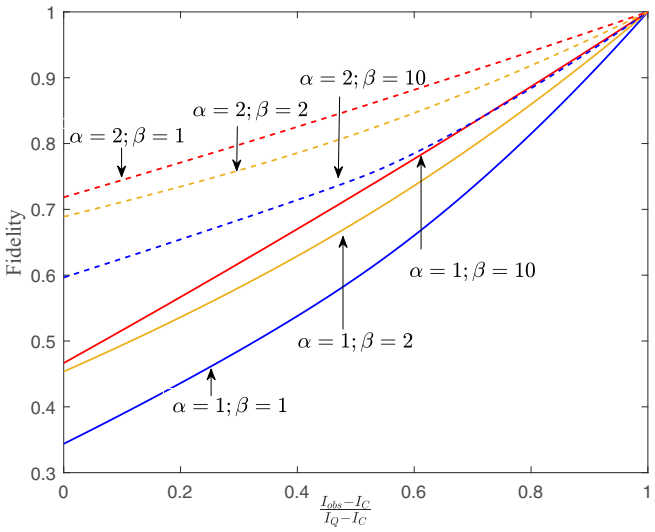


FIG. 6. Robustness bound of self-testing based on the three-setting steering inequality for six scenarios of (α, β), where $\alpha = 1, 2$ and $\beta = 1, 2, 10$.

APPENDIX C: NUMERICAL RESULTS UTILIZING THE SWAP ISOMETRY

In this Appendix we consider the numerical method based on SDP to show the robustness of the self-testing via steering inequality, which has been widely used in DI frameworks [17,55]. A detailed robustness analysis is given for three-setting steering inequalities. For two-setting scenarios, we only need to remove the third measurement in the code.

The target state is $|\psi\rangle = \cos\theta|00\rangle + \sin\theta|11\rangle$. Bob's measurements can be written as $B_0 = 2E_{0|0} - I$, $B_1 = 2E_{0|1} - I$, and $B_2 = 2E_{0|2} - I$, where $B_0^2 = B_1^2 = B_2^2$. After applying the isometry given in Fig. 5 to the physical state $|\psi'\rangle$, we obtain the state

$$|\psi'\rangle = E_{0|0}|\psi\rangle|0\rangle_{A'} + XE_{1|0}|\psi\rangle|1\rangle_{A'}. \tag{C1}$$

We trace the desired system out

$$\rho_{\text{SWAP}} = \text{tr}_A(|\psi'\rangle\langle\psi'|). \tag{C2}$$

Utilizing the SWAP isometry on Bob's side, the fidelity can be bounded as

$$\begin{aligned} f &= \langle\psi|\rho_{\text{SWAP}}|\psi\rangle \\ &= \cos^2\theta\langle 0|\text{tr}_A(E_{0|0}\rho_{AB})|0\rangle + \sin^2\theta\langle 1|\text{tr}_A(E_{1|0}\rho_{AB})|1\rangle + \frac{\sin 2\theta}{2}[\langle 0|\text{tr}_A(E_{1|0}XE_{0|0}\rho_{AB})|1\rangle + \langle 1|\text{tr}_A(E_{0|0}XE_{1|0}\rho_{AB})|0\rangle] \\ &= \cos^2\theta\langle 0|\text{tr}_A(E_{0|0}\rho_{AB})|0\rangle + \sin^2\theta\langle 1|(\rho_B - \text{tr}_A E_{0|0})|1\rangle + \sin 2\theta[\langle 0|\text{tr}_A(E_{0|1}E_{0|0} - E_{0|0}E_{0|1}E_{0|0})|1\rangle \\ &\quad + \langle 1|\text{tr}_A(E_{0|0}E_{0|1} - E_{0|0}E_{0|1}E_{0|0}\rho_{AB})|0\rangle] \\ &= \cos^2\theta\langle 0|\sigma_{0|0}|0\rangle + \sin^2\theta\langle 1|(\rho_B - \sigma_{0|0})|1\rangle + \sin 2\theta[\langle 0|(\sigma_{0|1,0|0} - \sigma_{0|0,0|1,0|0})|1\rangle + \langle 1|\sigma_{0|0,0|1} - \sigma_{0|0,0|1,0|0}|0\rangle]. \end{aligned} \tag{C3}$$

The goal is now to give a lower bound to f. The numerical method of minimizing the fidelity for a given steering inequality value is given by the SDP

$$\begin{aligned} &\text{minimize } f := \text{Tr}(M\Gamma) \\ &\text{subject to } \Gamma \geqslant 0, \quad I_{\alpha,\beta} = Q, \end{aligned} \tag{C4}$$

where M is a zero matrix (14,14), with $M_{2,2} = \sin^2\theta$, $M_{2,9} = M_{9,2} = \sin 2\theta$, $M_{3,3} = \cos^2\theta$, $M_{4,4} = -\sin^2\theta$, and $M_{9,10} = M_{10,9} = -\sin 2\theta$;

$$\Gamma = \begin{pmatrix} \rho_C & \sigma_{0|0} & \sigma_{0|1} & \sigma_{0|2} & \sigma_{0|1,0|0} & \sigma_{0|2,0|0} & \sigma_{0|2,0|1} \\ \sigma_{0|0} & \sigma_{0|0} & \sigma_{0|0,0|1} & \sigma_{0|0,0|2} & \sigma_{0|0,0|1,0|0} & \sigma_{0|0,0|2,0|0} & \sigma_{0|0,0|2,0|1} \\ \sigma_{0|1} & \sigma_{0|1,0|0} & \sigma_{0|1} & \sigma_{0|1,0|2} & \sigma_{0|1,0|0} & \sigma_{0|1,0|2,0|0} & \sigma_{0|1,0|2,0|1} \\ \sigma_{0|2} & \sigma_{0|2,0|0} & \sigma_{0|2,0|1} & \sigma_{0|2} & \sigma_{0|2,0|1,0|0} & \sigma_{0|2,0|0} & \sigma_{0|2,0|1} \\ \sigma_{0|0,0|1} & \sigma_{0|0,0|1,0|0} & \sigma_{0|0,0|1} & \sigma_{0|0,0|1,0|2} & \sigma_{0|0,0|1,0|0} & \sigma_{0|0,0|1,0|2,0|0} & \sigma_{0|0,0|1,0|2,0|1} \\ \sigma_{0|0,0|2} & \sigma_{0|0,0|2,0|0} & \sigma_{0|0,0|2,0|1} & \sigma_{0|0,0|2} & \sigma_{0|0,0|2,0|1,0|0} & \sigma_{0|0,0|2,0|0} & \sigma_{0|0,0|2,0|1} \\ \sigma_{0|1,0|2} & \sigma_{0|1,0|2,0|0} & \sigma_{0|1,0|2,0|1} & \sigma_{0|1,0|2} & \sigma_{0|1,0|2,0|1,0|0} & \sigma_{0|1,0|2,0|0} & \sigma_{0|1,0|2,0|1} \end{pmatrix}; \tag{C5}$$

and $I_{\alpha,\beta} = \alpha\langle Z\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle + \langle YB_2\rangle = \text{Tr}[(\alpha - \beta)Z\rho_C - (X+Y)\rho_C + 2\beta Z\sigma_{0|0} + 2X\sigma_{0|1} + 2Y\sigma_{0|2}]$ or $I_{\alpha,\beta} = \alpha\langle B_0\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle + \langle YB_2\rangle = \text{Tr}[-(\alpha + \beta Z + X + Y)\rho_C + (2\alpha I + 2\beta Z)\sigma_{0|0} + 2X\sigma_{0|1} + 2Y\sigma_{0|2}]$. We constrain Γ in the optimization to be positive semidefinite and note that each submatrix of Γ corresponding to something like an element of an assemblage is a valid quantum object. It actually turns out that all assemblages that satisfy no-signaling can be realized in quantum theory [56]. Discussion of this point is beyond the scope of this paper, as all we wish to do is give a lower bound on the value of G; therefore just imposing $\Gamma \geqslant 0$ gives such a bound. Based on the SDP of Eq. (C4), we show several robustness bounds of self-testing based on the three-setting steering inequality for six scenarios of (α, β), where $\alpha = 1, 2$ and $\beta = 1, 2, 10$ (see Fig. 6).

APPENDIX D: ANALYSIS OF DIFFERENT TYPES OF TWO-SETTING AND THREE-SETTING STEERING INEQUALITIES

Here we study the maximal quantum violation of the steering inequalities involved in the main text and show that the maximal violation of these inequalities can be used for self-testing. For the two-setting steering inequality

$$S^2_{\alpha,\beta} = \alpha\langle B_0\rangle + \beta\langle ZB_0\rangle + \langle XB_1\rangle \leqslant \alpha + \sqrt{1+\beta^2}. \tag{D1}$$

The maximum quantum bound is $\beta + \sqrt{1+\alpha^2} := S_Q$. This can be confirmed by showing $S_Q\mathbb{I} - \hat{S}^{(2)}_{\alpha,\beta} \geqslant 0$ to be true for all the possible underlying states and the measurements. To do so, we provide the following SOS decompositions of $S_Q\mathbb{I} - \hat{S}^{(2)}_{\alpha,\beta}$

to illustrate its PSD: The first SOS decomposition is

$$S_Q\mathbb{I}-\hat{S}^{(2)}_{\alpha,\beta}=\alpha_1^2(\mathbb{I}-cB_0-sX_AB_1)^2+\alpha_2^2(Z_A-B_0)^2+\alpha_3^2(-cB_1+sX_AB_0+Z_AB_1)^2+\alpha_4^2\big(S_Q\mathbb{I}-\hat{S}^2_{\alpha,\beta}\big)^2, \tag{D2}$$

where $c=\frac{\alpha}{\sqrt{1+\alpha^2}}$, $s=\frac{1}{\sqrt{1+\alpha^2}}$, $\alpha_4^2=\frac{1}{4\beta}$, $\alpha_3^2=\frac{\beta\sqrt{1+\alpha^2}}{1}\alpha_4^2=\frac{\sqrt{1+\alpha^2}}{4}$, $\alpha_1^2=(\frac{\beta\sqrt{1+\alpha^2}}{1}-\frac{1+\alpha^2}{1})\alpha_4^2$, and $\alpha_2^2=\frac{\beta-\sqrt{1+\alpha^2}}{4}$, and the second one is

$$\begin{aligned}S_Q\mathbb{I}-\hat{S}^{(2)}_{\alpha,\beta}&=\alpha_1^2(\mathbb{I}-cB_0-sX_AB_1)^2+\alpha_2^2(Z_A-B_0)^2\\&+\alpha_3^2[(\Delta+s^2)B_0-(\Delta+1)Z_A+cZ_AB_0-csX_AB_1]^2\\&+\alpha_4^2[-(\Delta+s^2)B_1+s(\Delta+1)X_A+\Delta cZ_AB_1-csX_AB_0)]^2,\end{aligned} \tag{D3}$$

where α_1 and α_2 are the same as the first SOS decomposition, $\alpha_3^2=\Delta\alpha_4^2$, $\alpha_4^2=\frac{S_Q}{4s\beta(\Delta^2+s^2)(\Delta^2+1)}$, and $\Delta=\frac{\beta}{\sqrt{1+\alpha^2}}$.

It is easy to verify that the left-hand sides of Eqs. (D2) and (D3) are equal to the SOS forms on the right. In addition, to make the SOS decompositions positive semidefinite, we should have $\alpha_i\geqslant 0$, and thus $\beta\geqslant\sqrt{1+\alpha^2}$. Apparently, S_Q is the upper bound of the steering inequality $S^2_{\alpha,\beta}$ under this constraint, although we do not know whether the quantum can reach the bound. Provided $B_0=Z$, $B_1=X$, and $|\Phi\rangle=\cos\theta|00\rangle+\sin\theta|11\rangle$ with $\sin 2\theta=\frac{1}{\sqrt{1+\alpha^2}}$ can make $S^2_{\alpha,\beta}$ achieve S_Q, we conclude that S_Q is the maximum quantum violation.

Next we show that the maximal violation of this steering inequality will self-test the partially entangled state. The local isometry used to determine the equivalence of the states is the same as in the main text, but with $\tilde{Z}_B=B_0$ and $\tilde{X}_B=B_1$. As shown in the main text, the relations required to show that this isometry works are

$$Z_A|\psi\rangle-B_0|\psi\rangle=0, \tag{D4}$$

$$\sin\theta X_A(I+B_0)|\psi\rangle-\cos\theta B_1(I-Z_A)|\psi\rangle=0. \tag{D5}$$

To obtain these relations, we let each side of Eqs. (D2) and (D3) take action on $|\psi\rangle$, a state that is supposed to reach the maximum violation of the steering inequality. Then seven terms of $P_i|\psi\rangle=0$ will be obtained; among them the second squared term in Eq. (D2) gives Eq. (D4), while the linear combination of the third squared term in Eq. (D2) and the fourth squared term in Eq. (D3) leads to Eq. (D4). Then, similar to the proof for the analog of tilted CHSH steering inequality given in the main text, by the isometry given in Fig. 1, we complete the self-testing statement via the two-setting steering inequality $S^2_{\alpha,\beta}$.

For the two-setting steering inequality

$$S^{(1)}_{\alpha,\beta}=\alpha\langle Z\rangle+\beta\langle ZB_0\rangle+\langle XB_1\rangle\leqslant\sqrt{1+(\alpha+\beta)^2}, \tag{D6}$$

which keeps the same maximal quantum violation as in Eq. (D1). For this steering inequality, three different types of SOS decompositions related to $S_Q\mathbb{I}-\hat{S}^{(1)}_{\alpha,\beta}$ can be given: The first one is

$$\frac{\beta}{2}(\mathbb{I}-Z_AB_0)^2+\frac{\sqrt{\alpha^2+1}}{2}(\mathbb{I}-cZ_A-sX_AB_1)^2, \tag{D7}$$

the second one is

$$\frac{1}{2S_Q}(-cX_A+sZ_AB_1+X_AB_0)^2+\frac{\beta\sqrt{\alpha^2+1}}{2S_Q}\big(S_Q\mathbb{I}-\hat{S}^{(1)}_{\alpha,\beta}\big)^2, \tag{D8}$$

and the third one is

$$\begin{aligned}&\alpha_1^2[(\Delta+s^2)Z_A-(\Delta+1)B_0+cZ_AB_0-csX_AB_1]^2\\&+\alpha_2^2[-(\Delta+s^2)X_A+s(\Delta+1)B_1+\Delta cX_AB_0-csZ_AB_1)]^2,\end{aligned} \tag{D9}$$

where $\alpha_1^2=\Delta\alpha_2^2$, $\alpha_2^2=\frac{(1+\alpha^2)^2}{2(\beta^2\sqrt{1+\alpha^2})+\beta(1+\alpha^2)+S_Q}$, and $\Delta=\frac{\beta}{\sqrt{1+\alpha^2}}$. The PSD requirements only require $\beta>0$. In addition, each squared term in Eqs. (D7)–(D9) acting on $|\psi\rangle$ being zero can lead to the relations our self-testing proofs heavily rely on, namely, Eqs. (D4) and (D5) [the first term in Eq. (D7) leads to Eq. (D4); the first term in Eq. (D7) and the second term in Eq. (D9) lead to Eq. (D5)]. Then we can complete the proof of self-testing based on $S^{(1)}_{\alpha,\beta}$.

For the three-setting scenario, the partial part expectation can be changed into the untrusted part's measurement. Thus there are two three-setting steering inequalities: the one in the main text,

$$I^{(1)}_{\alpha,\beta}\equiv\alpha\langle Z\rangle+\beta\langle ZB_0\rangle+\langle XB_1\rangle+\langle YB_2\rangle\leqslant\sqrt{2+(\alpha+\beta)^2}, \tag{D10}$$

and

$$I^{(2)}_{\alpha,\beta}\equiv\alpha\langle B_0\rangle+\beta\langle ZB_0\rangle+\langle XB_1\rangle+\langle YB_2\rangle\leqslant\alpha+\sqrt{2+\beta^2}. \tag{D11}$$

The advantage of this change is that its LHS bound is lower than using Alice's Z measurement in the three-setting inequality, while the quantum bound is maintained. It extends the gap between the LHS bound and steering bound, which is a benefit of the practical experiment. Denoting Bob's corresponding declared result by the random variable $B_k\in\{-1,1\}$ for $k=0,1$, it is easy to obtain the LHS bound $\alpha+\sqrt{2+\beta^2}$.

The quantum bounds of the both three-setting steering inequalities are the same, $\beta+\sqrt{4+\alpha^2}$. However, an extra condition should be satisfied for $I^{(2)}_{\alpha,\beta}$, that is, $\beta\geqslant\sqrt{4+\alpha^2}$. For $I^{(2)}_{\alpha,\beta}$ it only requires $\beta\geqslant 0$. This can be obtained from the following SOS: The first one is

$$\begin{aligned}(\beta+\sqrt{4+\alpha^2})\mathbb{I}-\hat{I}^{(2)}_{\alpha,\beta}&=\alpha_1^2(\mathbb{I}-cB_0-sX_AB_1)^2+\alpha_2^2(Z_A-B_0)^2\\&+\alpha_3^2(\mathbb{I}-cB_0-sY_AB_2)^2\\&+\alpha_4^2(-cB_1+sX_AB_0+Z_AB_1)^2\\&+\alpha_5^2(-cB_2+sY_AB_0+Z_AB_2)^2\\&+\alpha_6^2((\beta+\sqrt{4+\alpha^2})\mathbb{I}-I_{\alpha,\beta})^2\\&+\alpha_7^2(X_AB_1-Y_AB_2)^2,\end{aligned} \tag{D12}$$

where $c=\frac{\alpha}{\sqrt{4+\alpha^2}}$, $s=\frac{2}{\sqrt{4+\alpha^2}}$, $\alpha_6^2=\alpha_7^2=\frac{1}{4\beta}$, $\alpha_4^2=\alpha_5^2=\frac{\beta\sqrt{4+\alpha^2}}{2}\alpha_6^2=\frac{\sqrt{4+\alpha^2}}{8}$, $\alpha_1^2=\alpha_3^2=(\frac{\beta\sqrt{4+\alpha^2}}{2}-\frac{4+\alpha^2}{2})\alpha_6^2$, and

$\alpha_2^2 = \frac{\beta-\sqrt{4+\alpha^2}}{4}$, and the second one is

$$\begin{aligned}(\beta+\sqrt{4+\alpha^2})\mathbb{I} - \hat{I}^{(2)}_{\alpha,\beta} &= \alpha_1^2(\mathbb{I} - cB_0 - sX_AB_1)^2 + \alpha_2^2(Z_A - B_0)^2 \\ &+ \alpha_3^2(\mathbb{I} - cB_0 - sY_AB_2)^2 \\ &+ \alpha_4^2[(\Delta+s^2)B_0 - (\Delta+1)Z_A + cZ_AB_0 - csX_AB_1]^2 \\ &+ \alpha_5^2[(\Delta+s^2)B_0 - (\Delta+1)Z_A + cZ_AB_0 - csY_AB_2]^2 \\ &+ \alpha_6^2[-(\Delta+s^2)B_1 + s(\Delta+1)X_A \\ &+ \Delta cZ_AB_1 - csX_AB_0)]^2 \\ &+ \alpha_7^2[-(\Delta+s^2)B_2 + s(\Delta+1)Y_A \\ &+ \Delta cZ_AB_2 - csY_AB_0)]^2, \end{aligned} \tag{D13}$$

where $c = \frac{\alpha}{\sqrt{4+\alpha^2}}$, $s = \frac{2}{\sqrt{4+\alpha^2}}$, $\alpha_6^2 = \alpha_7^2 = \frac{1}{4s\Delta(\Delta^2+s)}$, $\alpha_4^2 = \alpha_5^2 = \Delta\alpha_6^2 = \frac{1}{4s(\Delta^2+s)}$, $\alpha_1^2 = \alpha_3^2 = \frac{1}{2S} - (\Delta+1)(\Delta+s^2)\alpha_6^2$, $\alpha_2^2 = \frac{\beta}{2} - \frac{\Delta^2+1}{s(\Delta+1)}$, and $\Delta = 1$.

Making the SOS decomposition positive semidefinite requires each $\alpha_i \geqslant 0$ and thus $\beta \geqslant \sqrt{4+\alpha^2}$. In addition, some squared terms in (D12) and (D13) acting on $|\psi\rangle$ being zero also can lead to the relations (D4) and (D5). Thus, with the isometry given in the main text, we can complete the proof of self-testing based on $S^{(2)}_{\alpha,\beta}$.

For the first three-setting steering inequality, three types of SOS decompositions can be given: The first one is

$$\begin{aligned}(\beta+\sqrt{4+\alpha^2})\mathbb{I} - \hat{I}^{(1)}_{\alpha,\beta} &= \frac{\beta}{2}(\mathbb{I} - Z_AB_0)^2 + \frac{\sqrt{\alpha^2+4}}{4}(\mathbb{I} - cZ_A - sX_AB_1)^2 \\ &+ \frac{\sqrt{\alpha^2-4}}{4}(\mathbb{I} - cZ_A - sY_AB_2)^2, \end{aligned} \tag{D14}$$

the second one is

$$\begin{aligned}(\beta+\sqrt{4+\alpha^2})\mathbb{I} - \hat{I}^{(1)}_{\alpha,\beta} &= \alpha_1^2(-cX_A + sZ_AB_1 + X_AB_0)^2 \\ &+ \alpha_2^2(-cY_A + sZ_AB_2 + Y_AB_0)^2 + \alpha_3^2\left(S_Q\mathbb{I} - \hat{I}^{(1)}_{\alpha,\beta}\right)^2, \end{aligned} \tag{D15}$$

where $\alpha_1^2 = \alpha_2^2 = \frac{\alpha^2+\beta^2+\beta\sqrt{4+\alpha^2}+3}{4S_Q}$ and $\alpha_3^2 = \frac{1}{2S_Q}$, and the third one is

$$\begin{aligned}(\beta+\sqrt{4+\alpha^2})\mathbb{I} - \hat{I}^{(1)}_{\alpha,\beta} &= \alpha_1^2[(\Delta+s^2)Z_A - (\Delta+1)B_0 + cZ_AB_0 - csX_AB_1]^2 \\ &+ \alpha_2^2[-(\Delta+s^2)X_A + s(\Delta+1)B_1 + \Delta cX_AB_0 - csZ_AB_1)]^2 \\ &+ \alpha_3^2[(\Delta+s^2)Z_A - (\Delta+1)B_0 + cY_AB_0 - csZ_AB_2]^2 \\ &+ \alpha_4^2[-(\Delta+s^2)Y_A + s(\Delta+1)B_2 + \Delta cY_AB_0 - csZ_AB_2)]^2, \end{aligned} \tag{D16}$$

where $\alpha_1^2 = \alpha_3^2 = \frac{\beta}{4(\Delta+s^2)(\Delta+1)}$, $\alpha_2^2 = \alpha_4^2 = \frac{(1}{2s(\Delta+s^2)(\Delta+1)}$, and $\Delta = \frac{\beta}{\sqrt{1+\alpha^2}}$.

The PSD condition requires $\beta \geqslant 0$. In addition, the first squared term in (D14) acting on $|\psi\rangle$ being zero ($|\psi\rangle$ is the state which maximally violates the steering inequality) gives the relations (D4), while the linear combination of the second squared term in (D15) and the first squared term in (D16) gives the relation (D5). Thus, with the isometry given in the main text, we can complete the proof of self-testing based on $S^{(1)}_{\alpha,\beta}$.

Self-testing for the measurements. Above we mainly focused on the states self-testing; the self-testing of the corresponding measurements (for analysis refer to [17]) will be similar. We start with $\Phi M_B(|\psi\rangle)$ instead of $\Phi(|\psi\rangle)$ and show it for one of the three measurements in three-setting steering inequality cases, for example. After the isometry, the systems will be

$$\begin{aligned}\Phi(\underline{\tilde{Z}_B}|\psi\rangle) = \tfrac{1}{4}[&(I+Z_A)(I+\tilde{Z}_B)\underline{\tilde{Z}_B}|\psi\rangle|00\rangle \\ &+ X_A(I+Z_A)(I-\tilde{Z}_B)\underline{\tilde{Z}_B}|\psi\rangle|01\rangle \\ &+ \tilde{X}_B(I-Z_A)(I+\tilde{Z}_B)\underline{\tilde{Z}_B}|\psi\rangle|10\rangle \\ &+ X_A\tilde{X}_B(I-Z_A)(I-\tilde{Z}_B)\underline{\tilde{Z}_B}|\psi\rangle|11\rangle]. \end{aligned} \tag{D17}$$

With the relations (D4) and (D5) and the fact that $Z_AX_A = -X_AZ_A$, we find $\tilde{Z}_B\tilde{X}_B|\psi\rangle = -\tilde{X}_B\tilde{Z}_B|\psi\rangle$. By using this anticommutation relation between Bob's two measurements, we move $\tilde{Z}_B$ to the left in the first, second, third, and fourth lines while changing the sign of the fourth line. The analysis is then the same as the state self-testing and the result is

$$\begin{aligned}\Phi(\underline{\tilde{Z}_B}|\psi\rangle) &= |\text{junk}\rangle[\cos\theta|00\rangle - \sin\theta|11\rangle] \\ &= |\text{junk}\rangle[\underline{(I\otimes\sigma_z)}\cos\theta|00\rangle + \sin\theta|11\rangle]. \end{aligned} \tag{D18}$$

In addition, from the SOS decomposition we can also find the relation $\sin\theta Y_A(I+B_0)|\psi\rangle - \cos\theta B_2(I-Z_A)|\psi\rangle = 0$. Thus we have $\tilde{Z}_B\tilde{Y}_B|\psi\rangle = -\tilde{Y}_B\tilde{Z}_B|\psi\rangle$. Following the above idea, we can finally conclude that the measurements on Bob's side are $B_0 = Z$, $B_1 = X$, and $B_2 = -Y$.

APPENDIX E: TRANSFORMATION OF A STEERING INEQUALITY INTO A GAME

In this Appendix we relate the constructed steering inequality to a game which two parties play to increase the score and build the relation between the quantum violation and success probability of the game defined. This is helpful for a direct comparison between different steering inequalities and it is necessary in the analysis of sample efficiency. For simplicity, here we only consider the three-setting steering inequality.

In principle, to obtain the maximum violation of the three-setting steering inequality (37), the state between Alice and Bob should be $\cos\theta|00\rangle + \sin\theta|11\rangle$, which can be further written as $\frac{1}{\sqrt{2}}(|\psi_0\rangle|+\rangle + |\psi_1\rangle|-\rangle)$, where we define $|\psi_0\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle$ and $|\psi_1\rangle = \cos(\theta)|0\rangle - \sin(\theta)|1\rangle$. We define two measurements on Alice's side $\{|\psi_0\rangle, |\psi_0^\perp\rangle; |\psi_1\rangle, |\psi_1^\perp\rangle\}$, which actually are measurements introduced to replace the measurements chosen in the main text in the real experiments. The measurements can also be written in the Pauli operator form $\{A_0 = \cos(2\theta)\sigma_z + \sin(2\theta)\sigma_x; A_1 = \cos(2\theta)\sigma_z - \sin(2\theta)\sigma_x\}$.

We notice that, if Bob gets $|+\rangle$, Alice takes A_0 and Bob can conclude that Alice's qubit must be projected into $|\psi_0\rangle$; if Bob gets $|-\rangle$, Alice takes A_1 and then Bob can conclude that Alice's qubit must be projected into $|\psi_1\rangle$. Since in the steering scenario Bob can send information to Alice, such measurements result. Thus, this allows us to define the success probability of Bob guessing Alice's measurement result as

$$P^x_{\text{virtual}} = p\big(A_0^0, B_1^0\big) + P\big(A_1^0, B_1^1\big), \tag{E1}$$

which actually is related to the operators in the three-setting steering inequality (34). More precisely, $\frac{\alpha}{2}Z + XB_1 = (\frac{\alpha}{2}Z + X)B_1^0 + (\frac{\alpha}{2}Z - X)B_1^1 = \frac{\sqrt{4+\alpha^2}}{2}(A_0B_1^0 + A_1B_1^1) = \frac{\sqrt{4+\alpha^2}}{2}(2A_0^0B_1^0 + 2A_1^0B_1^1 - I_B)$ for $\sin(2\theta) = \frac{2}{\sqrt{4+\alpha^2}}$. Thus P^x_{virtual} is related to $\alpha\langle Z\rangle + \langle XB_1\rangle$. Similarly, we can define P^y_{virtual} for the σ_y measurement scenario, which is related to $\frac{\alpha}{2}\langle Z\rangle + \langle YB_2\rangle$. Together with the guessing probability for (Z_A, B_0), we define the total average passing probability as

$$P_{\text{virtual}} = \frac{\sqrt{4+\alpha^2}\big(\frac{P^x_{\text{virtual}}+P^y_{\text{virtual}}}{2}\big) + \beta p(a=b|Z_A, B_0)}{\sqrt{4+\alpha^2}+\beta}. \tag{E2}$$

Thus we have

$$P_{\text{virtual}} = \frac{\sqrt{4+\alpha^2}+\beta+S}{2(\sqrt{4+\alpha^2}+\beta)} = \frac{1}{2} + \frac{S}{2S_Q}. \tag{E3}$$

This relation between the guessing probability and the violation holds for steering inequalities (30) and (34). Thus the steering inequalities are transformed to testing games.

APPENDIX F: ROBUST SELF-TESTING OF THREE-SETTING INEQUALITY

In this Appendix we provide an analytical robustness bound for self-testing via the three-setting steering inequality. We first consider the part of $G_1 := K_1 - s(ZB_0 + XB_1) - \tau_1\mathbb{I}$ for $\mu \in (0, \pi/4]$; the spectral decomposition is already given in Eq. (B3). To make $G_1 \geqslant 0$, we consider the following local extraction channel. Bob takes $R_1 = I$ with probability q_1 and $R_2 = \sigma_z$ with probability q_2; meanwhile, with the rest of the probability $1 - q_1 - q_2 := q_3$ Bob takes some other local extraction channel subject to the choice of B_2. Then we have

$$\begin{aligned} q_1 - s\lambda_1 - \tau_1 &\geqslant 0,\\ q_2 - s\lambda_2 - \tau_1 &\geqslant 0,\\ s\lambda_{(1/2)} - \tau_1 &\geqslant 0,\\ \text{Tr}(\rho) = q_1 + q_2 + q_3 &= 1,\\ \text{Tr}(\rho\hat{B}) = \lambda_1 q_1 + \lambda_2 q_2 + q_3\text{Tr}(\rho YB_2) &= S, \end{aligned}$$

where $\tau_1 = 1 - \gamma s$, with $\gamma \in [2, 3]$. In addition, τ_1 should be less than zero. We obtain $s \geqslant \frac{1+q_3}{2\gamma-(\lambda_1+\lambda_2)} \geqslant \frac{1+q_3}{2\gamma-2\sqrt{2}}$.

Next we determine the value of q_3 to make K_2 PSD. We notice $s\lambda_1 - \tau_1$ and $s\lambda_2 - \tau_1$, which, according to the coefficients of $|\psi_3\rangle$ and $|\psi_4\rangle$, are greater than zero. That is, if only the coefficients of $|\psi_1\rangle$ and $|\psi_2\rangle$ are greater than zero, the K_1 part will be PSD. Therefore, we put $|\psi_3\rangle$ and $|\psi_4\rangle$ into the K_2 part to make it PSD. Now the K_2 part becomes

$$\begin{aligned} G_2 :=\ & q_3\Lambda_B^+(\psi_1) + (s\lambda_1 - \tau_1)|\psi_3\rangle\langle\psi_3|\\ &+ (s\lambda_2 - \tau_1)|\psi_4\rangle\langle\psi_4| - sYB_2 - (\gamma-3)s\mathbb{I}, \end{aligned} \tag{F1}$$

which is equivalent to

$$\begin{aligned} G_2 :=\ & q_3\Lambda_B^+(\psi_1) + (s\lambda_1 - \tau_1)|\psi_3\rangle\langle\psi_3|\\ &+ (s\lambda_2 - \tau_1)|\psi_4\rangle\langle\psi_4|\\ &- s(\gamma-2)(U|\phi_1\rangle\langle\phi_1|U^T + U|\phi_2\rangle\langle\phi_2|U^T)\\ &+ s(4-\gamma)(U|\phi_3\rangle\langle\phi_3|U^T + U|\phi_4\rangle\langle\phi_4|U^T), \end{aligned} \tag{F2}$$

where $U = \left[\begin{smallmatrix} V & 0 \\ 0 & V \end{smallmatrix}\right]$ and $U^T = \left[\begin{smallmatrix} V^* & 0 \\ 0 & V^* \end{smallmatrix}\right]$, where

$$V = \begin{bmatrix} \frac{-\sin\mu_1 i - \cos\mu_1\sin\mu_2)}{\sqrt{2-2\cos\mu_1\cos\mu_2}} & \frac{-\sin\mu_1 i - \cos\mu_1\sin\mu_2}{\sqrt{2+2\cos\mu_1\cos\mu_2}} \\ \frac{\cos\mu_1\cos\mu_2 - 1}{\sqrt{2-2\cos\mu_1\cos\mu_2}} & \frac{\cos\mu_1\cos\mu_2+1}{\sqrt{2+2\cos\mu_1\cos\mu_2}} \end{bmatrix}, \tag{F3}$$

with $\phi_1 = [\frac{-1i}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0]$, $\phi_2 = [0, \frac{1i}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]$, $\phi_3 = [\frac{1i}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}, 0]$, and $\phi_4 = [0, \frac{-1i}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}]$. The requirement of $G_2 \geqslant 0$ gives ($\mathcal{O}$ denotes overlap)

$$\begin{aligned} &\frac{q_3(1+c)}{2} + (s\lambda_2 - \tau_1)\mathcal{O}^2(\psi_3, U^{-1}\phi_1)\\ &\quad + (s\lambda_1 - \tau_1)\mathcal{O}^2(\psi_4, U^{-1}\phi_1) - s(\gamma-2) \geqslant 0, \end{aligned} \tag{F4}$$

$$\begin{aligned} &\frac{q_3(1-c)}{2} + (s\lambda_2 - \tau_1)\mathcal{O}^2(\psi_3, U^{-1}\phi_2)\\ &\quad + (s\lambda_1 - \tau_1)\mathcal{O}^2(\psi_4, U^{-1}\phi_2) - s(\gamma-2) \geqslant 0, \end{aligned} \tag{F5}$$

that is,

$$\begin{aligned} &\frac{q_3(1-c)}{2}\\ &\quad + C_2\frac{\cos^2(\frac{\pi}{8})(\sin\mu_1 - 1)^2 + \cos^2\mu_1\sin^2(\frac{\pi}{8}+\mu_2)}{4}\\ &\quad + C_1\frac{\cos^2(\frac{\pi}{8})(\sin\mu_1 + 1)^2 + \cos^2\mu_1\sin^2(\frac{\pi}{8}-\mu_2)}{4}\\ &\quad - s(\gamma-2) \geqslant 0,\\ &\frac{q_3(1+c)}{2}\\ &\quad + C_2\frac{\sin^2(\frac{\pi}{8})(\sin\mu_1 - 1)^2 + \cos^2\mu_1\cos^2(\frac{\pi}{8}+\mu_2)}{4}\\ &\quad + C_1\frac{\sin^2(\frac{\pi}{8})(\sin\mu_1 + 1)^2 + \cos^2\mu_1\cos^2(\frac{\pi}{8}-\mu_2)}{4}\\ &\quad - s(\gamma-2) \geqslant 0, \end{aligned}$$

where $C_1 = s\lambda_1 - \tau_1$ and $C_2 = s\lambda_2 - \tau_1$. With this channel, we have

$$\frac{q_3(1+c)}{2} + \frac{2-\sqrt{2}}{8}(s + \gamma s - 1) - s(\gamma-2) \geqslant 0$$

and

$$\frac{q_3(1-c)}{2} + \frac{2+\sqrt{2}}{8}(s + \gamma s - 1) - s(\gamma-2) \geqslant 0,$$

which gives us $q_3 c = \frac{\sqrt{2}}{4}(s + \gamma s - 1)$ for $\gamma > 2$ and $q_3 \geqslant \frac{-5\gamma+2\sqrt{2}+9}{-\gamma+4\sqrt{2}-9}$. We can choose $\gamma = 3$, which gives $q_3 = \frac{1}{2}$; in addition, $s = \frac{3}{12-4\sqrt{2}} = 0.4730$ and $\tau = 1 - 3s$. Thus we give the following robustness bound of one-sided self-testing based

on the three-setting steering inequality:

$$F \geqslant sS_{\text{obs}} + \tau \geqslant \frac{3}{12-4\sqrt{2}}(S_{\text{obs}} - 3) + 1. \qquad \text{(F6)}$$

Although this does not reach the theoretical bound $s = \frac{1}{2(3-\sqrt{3})}$, the result is better than that of the two-setting inequality. This shows that adding more measurement settings can help increase the robustness in one-sided self-testing.

[1] H. J. Kimble, The quantum internet, Nature (London) **453**, 1023 (2008).

[2] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, Secure quantum key distribution with realistic devices, Rev. Mod. Phys. **92**, 025002 (2020).

[3] E. T. Campbell, B. M. Terhal, and C. Vuillot, Roads towards fault-tolerant universal quantum computation, Nature (London) **549**, 172 (2017).

[4] V. Giovannetti, S. Lloyd, and L. Maccone, Advances in quantum metrology, Nat. Photon. **5**, 222 (2011).

[5] I. H. Deutsch, Harnessing the power of the second quantum revolution, PRX Quantum **1**, 020101 (2020).

[6] J. Eisert, D. Hangleiter, N. Walk, I. Roth, D. Markham, R. Parekh, U. Chabaud, and E. Kashefi, Quantum certification and benchmarking, Nat. Rev. Phys. **2**, 382 (2020).

[7] M. Kliesch and I. Roth, Theory of quantum system certification, PRX Quantum **2**, 010201 (2021).

[8] S. T. Flammia and Y.-K. Liu, Direct Fidelity Estimation from Few Pauli Measurements, Phys. Rev. Lett. **106**, 230501 (2011).

[9] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Quantum State Tomography via Compressed Sensing, Phys. Rev. Lett. **105**, 150401 (2010).

[10] H.-Y. Huang, R. Kueng, and J. Preskill, Predicting many properties of a quantum system from very few measurements, Nat. Phys. **16**, 1050 (2020).

[11] S. Pallister, N. Linden, and A. Montanaro, Optimal Verification of Entangled States with Local Measurements, Phys. Rev. Lett. **120**, 170502 (2018).

[12] H. Zhu and M. Hayashi, Efficient Verification of Pure Quantum States in the Adversarial Scenario, Phys. Rev. Lett. **123**, 260504 (2019).

[13] I. Šupić and J. Bowles, Self-testing of quantum systems: A review, Quantum **4**, 337 (2020).

[14] D. Mayers and A. Yao, Self testing quantum apparatus, Quantum Inf. Comput. **4**, 273 (2004).

[15] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[16] M. McKague, T. H. Yang, and V. Scarani, Robust self-testing of the singlet, J. Phys. A: Math. Theor. **45**, 455304 (2012).

[17] T. H. Yang, T. Vértesi, J.-D. Bancal, V. Scarani, and M. Navascués, Robust and Versatile Black-Box Certification of Quantum Devices, Phys. Rev. Lett. **113**, 040401 (2014).

[18] A. Coladangelo, K. T. Goh, and V. Scarani, All pure bipartite entangled states can be self-tested, Nat. Commun. **8**, 15485 (2017).

[19] M. O. Renou, J. Kaniewski, and N. Brunner, Self-Testing Entangled Measurements in Quantum Networks, Phys. Rev. Lett. **121**, 250507 (2018).

[20] J.-D. Bancal, N. Sangouard, and P. Sekatski, Noise-Resistant Device-Independent Certification of Bell State Measurements, Phys. Rev. Lett. **121**, 250506 (2018).

[21] B. W. Reichardt, F. Unger, and U. Vazirani, Classical command of quantum systems, Nature (London) **496**, 456 (2013).

[22] X. Wu, J.-D. Bancal, M. McKague, and V. Scarani, Device-independent parallel self-testing of two singlets, Phys. Rev. A **93**, 062121 (2016).

[23] S. Pironio, A. Acín, S. Massar, A. B. de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Random numbers certified by Bell's theorem, Nature (London) **464**, 1021 (2010).

[24] Y. Liu, Q. Zhao, M.-H. Li, J.-Y. Guan, Y. Zhang, B. Bai, W. Zhang, W.-Z. Liu, C. Wu, X. Yuan, H. Li, W. J. Munro, Z. Wang, L. You, J. Zhang, X. Ma, J. Fan, Q. Zhang, and J.-W. Pan, Device-independent quantum random-number generation, Nature (London) **562**, 548 (2018).

[25] A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Device-Independent Security of Quantum Cryptography against Collective Attacks, Phys. Rev. Lett. **98**, 230501 (2007).

[26] U. Vazirani and T. Vidick, Fully Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[27] H. Shrotriya, K. Bharti, and L.-C. Kwek, Robust semi-device-independent certification of all pure bipartite maximally entangled states via quantum steering, Phys. Rev. Res. **3**, 033093 (2021).

[28] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, One-sided device-independent quantum key distribution: Security, feasibility, and the connection with steering, Phys. Rev. A **85**, 010301(R) (2012).

[29] E. Passaro, D. Cavalcanti, P. Skrzypczyk, and A. Acín, Optimal randomness certification in the quantum steering and prepare-and-measure scenarios, New J. Phys. **17**, 113010 (2015).

[30] A. Gheorghiu, P. Wallden, and E. Kashefi, Rigidity of quantum steering and one-sided device-independent verifiable quantum computation, New J. Phys. **19**, 023043 (2017).

[31] A. Unnikrishnan, I. J. MacFarlane, R. Yi, E. Diamanti, D. Markham, and I. Kerenidis, Anonymity for Practical Quantum Networks, Phys. Rev. Lett. **122**, 240501 (2019).

[32] F. Hahn, J. de Jong, and A. Pappa, Anonymous quantum conference key agreement, PRX Quantum **1**, 020325 (2020).

[33] Y. Wang, X. Li, Y. Han, and K. Zhang, Practical anonymous entanglement with noisy measurement, Quantum Inf. Process. **21**, 49 (2022).

[34] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, Rev. Mod. Phys. **92**, 015001 (2020).

[35] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[36] D. J. Saunders, S. J. Jones, H. M. Wiseman, and G. J. Pryde, Experimental EPR-steering using bell-local states, Nat. Phys. **6**, 845 (2010).

[37] D. Cavalcanti, P. Skrzypczyk, G. H. Aguilar, R. V. Nery, P. H. S. Ribeiro, and S. P. Walborn, Detection of entanglement in asym-

metric quantum networks and multipartite quantum steering, Nat. Commun. **6**, 7941 (2015).
[38] I. Šupić and M. J. Hoban, Self-testing through EPR-steering, New J. Phys. **18**, 075006 (2016)
[39] Y.-G. Han, Z. Li, Y. Wang, and H. Zhu, Optimal verification of the Bell state and Greenberger-Horne-Zeilinger states in untrusted quantum networks, npj Quantum Inf. **7**, 164 (2021).
[40] S. Goswami, B. Bhattacharya, D. Das, S. Sasmal, C. Jebaratnam, and A. S. Majumdar, One-sided device-independent self-testing of any pure two-qubit entangled state, Phys. Rev. A **98**, 022311 (2018).
[41] T. Pramanik, M. Kaplan, and A. S. Majumdar, Fine-grained Einstein-Podolsky-Rosen–steering inequalities, Phys. Rev. A **90**, 050305(R) (2014).
[42] E. G. Cavalcanti, C. J. Foster, M. Fuwa, and H. M. Wiseman, Analog of the Clauser-Horne-Shimony-Holt inequality for steering, J. Opt. Soc. Am. B **32**, A74 (2015).
[43] A. Acín, S. Massar, and S. Pironio, Randomness versus Nonlocality and Entanglement, Phys. Rev. Lett. **108**, 100402 (2012).
[44] S. Sarkar, J. J. Borkała, C. Jebarathinam, O. Makuta, D. Saha, and R. Augusiak, Self-testing of any pure entangled state with minimal number of measurements and optimal randomness certification in one-sided device-independent scenario, arXiv:2110.15176.
[45] P. Skrzypczyk and D. Cavalcanti, Maximal Randomness Generation from Steering Inequality Violations Using Qudits, Phys. Rev. Lett. **120**, 260401 (2018).
[46] A. Pappa, A. Chailloux, S. Wehner, E. Diamanti, and I. Kerenidis, Multipartite Entanglement Verification Resistant against Dishonest Parties, Phys. Rev. Lett. **108**, 260502 (2012).
[47] W. McCutcheon, A. Pappa, B. A. Bell, A. McMillan, A. Chailloux, T. Lawson, M. Mafu, D. Markham, E. Diamanti, I. Kerenidis, J. G. Rarity, and M. S. Tame, Experimental verification of multipartite entanglement in quantum networks, Nat. Commun. **7**, 13251 (2016).
[48] J. Kaniewski, Analytic and Nearly Optimal Self-Testing Bounds for the Clauser-Horne-Shimony-Holt and Mermin Inequalities, Phys. Rev. Lett. **117**, 070402 (2016).
[49] M. Navascués, S. Pironio, and A. Acín, A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations, New J. Phys. **10**, 073013 (2008).
[50] C. Bamps and S. Pironio, Sum-of-squares decompositions for a family of Clauser-Horne-Shimony-Holt-like inequalities and their application to self-testing, Phys. Rev. A **91**, 052111 (2015).
[51] T. Coopmans, J. Kaniewski, and C. Schaffner, Robust self-testing of two-qubit states, Phys. Rev. A **99**, 052123 (2019).
[52] H. Peyrl and P. A. Parrilo, Computing sum of squares decompositions with rational coefficients, Theor. Comput. Sci. **409**, 269 (2008).
[53] A. Gočanin, C. Šupić, and B. Dakić, Sample-Efficient device-independent quantum state verification and certification, PRX Quantum **3**, 010317 (2022).
[54] J.-D. Bancal, K. Redeker, P. Sekatski, W. Rosenfeld, and N. Sangouard, Self-testing with finite statistics enabling the certification of a quantum network link, Quantum **5**, 401 (2021).
[55] Y. Wang, X. Wu, and V. Scarani, All the self-testings of the singlet for two binary measurements, New J. Phys. **18**, 025021 (2016).
[56] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Phys. Lett. A **183**, 14 (1993).

A complete and operational resource theory of measurement sharpness

Francesco Buscemi[1] * Kodai Kobayashi[1] † Shintaro Minagawa[1] ‡

[1] *Graduate School of Informatics, Nagoya University, Chikusa-Ku, Nagoya 464-8601, Japan*

Abstract. We provide a resource theory of sharpness for positive operator-valued measures (POVMs) where free operations are quantum pre-processing channels and convex mixtures with POVMs consisting of identity operators. We show that the greatest elements are POVMs where all POVM elements have at least one real unit eigenvalue. Here, the sharpness is measured by the degree of perfect correlations between a POVM and all reference POVMs. We say that our framework is *complete* in that all measures provide a necessary and sufficient condition for the existence of a sharpness non-increasing operation, and *operational* in that the measures are experimentally accessible.

Keywords: Quantum measurement theory, measurement sharpness, resource theory, perfect correlation, quantum pre-processing

1 Introduction

One can regard the measurement described by projection-valued measure (PVM) as a good measurement for considering observable and thus, quantifying the closeness of measurement described by positive operator-valued measure (POVM), called "sharpness" is an important task [1, 2, 3, 4, 5, 6].

Given the recent development of quantum resource theories (for a review, see, e.g., [7]), one may naturally regard sharpness as a resource like Ref. [6]. However, this attempt left a problem about what the "free" operations are in the sense that such operations do not increase sharpness, which is the crucial concept to construct a resource theory of sharpness.

In response to these issues, our work clarifies what a free operation, i.e., a fuzzifying operation is. First, we define "sharp POVM" as the greatest element of our ordering of sharpness where all POVM elements have at least one eigenvalue 1, and this is equivalent to the existence of repeatable instruments. Next, we define a fuzzifying operation as a quantum pre-processing and making a convex mixture with a trivial measurement where all POVM elements are proportional to identity. Then we say that one POVM is sharper than another POVM if and only if there is a fuzzifying operation that can convert one into another.

In terms of sharpness measures, we adopt Ozawa's degree of measurement correlation [8, 9]. Then we give a kind of Blackwell's theorem [10] concerning the convertibility between POVMs by using a fuzzifying operation: one POVM is sharper than another POVM if and only if (this means that our condition is *complete*) all measures for the former POVM are greater than or equal to the one of the latter POVM.

*buscemi@i.nagoya-u.ac.jp

†kobayashi.kodai.z8@s.mail.nagoya-u.ac.jp

‡minagawa.shintaro@nagoya-u.ac.jp

2 Basic notions

Let us consider a quantum system A associated with a finite d_A-dimensional Hilbert space $\mathscr{H}_A$. We denote all linear operators of $\mathscr{H}_A$ as $\mathcal{L}(\mathscr{H}_A)$. Quantum states of A are *density operators* ρ_A on $\mathscr{H}_A$, which are positive semidefinite matrices and have unit trace $\mathrm{Tr}\,\rho_A = 1$. The pure states of A correspond to a unit vector described as $|\psi\rangle$.

We denote a set that contains all states of the system A as $\mathcal{S}(A)$. A family of positive semidefinite operators $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ with $\sum_{x\in\mathcal{X}} P_x = \mathbb{1}_A$ corresponds to a measurement on A, called a *POVM measurement*.

Next, we consider an evolution of a quantum state on A to another quantum state, in general, on a different quantum system B. We can describe this evolution as a linear and completely positive and trace-preserving (CPTP) map $\mathcal{E} : A \to B$ and call it a *quantum channel*. The Hilbert-Schmidt dual of quantum channels are *adjoint maps*, which is defined as $\mathrm{Tr}[\mathcal{E}^\dagger(Y)X] = \mathrm{Tr}[Y\mathcal{E}(X)]$ for all $X \in \mathcal{L}(\mathscr{H}_A)$ and $Y \in \mathcal{L}(\mathscr{H}_B)$. The adjoint maps are linear, CP, and unital that is, $\mathcal{E}^\dagger(\mathbb{1}_B) = \mathbb{1}_A$.

3 Sharp POVMs and fuzzifying operations

First, we define *sharp POVMs* as follows:

Definition 1 (sharp POVMs [11]) *A POVM* $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ *is* sharp *when all its elements contain*

eigenvalues 1, *i.e., there exist normalized vectors* $|\psi^x\rangle_A$ *such that* $P_A^x |\psi^x\rangle_A = |\psi^x\rangle_A$ *for all* $x \in \mathcal{X}$.

An adjoint map of a CPTP linear map $\mathcal{E}$, $\mathcal{E}^\dagger$ can convert a POVM to another POVM, which is, in general, in a different system from the former one. Note that one can regard the adjoint map as a quantum pre-processing before the measurement. Then we introduce a pre-processing ordering between two POVMs $\mathbf{P}_A = \{P_A^x\}_{x\in\mathcal{X}}$ and $\mathbf{Q}_B = \{Q_B^x\}_{x\in\mathcal{X}}$. We say that $\mathbf{P}_A$ is *post-processing cleaner* than $\mathbf{Q}_B$ if there is a CPTP linear map $\mathcal{E} : B \to A$ such that for all $x \in \mathcal{X}$, $\mathcal{E}^\dagger(P_A^x) = Q_B^x$ holds following the Ref. [12].

As the following theorem insists, sharp POVMs are the maximal elements of this ordering (for the proof, see Ref. [11]):

Theorem 2 (Theorem 1 in [11]) *A POVM* $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ *is sharp if and only if* $\mathbf{P}$ *is post-processing cleaner than any other POVM with the same outcome set* $\mathcal{X}$.

Next, we consider *sharpness non-increasing operations* or *fuzzifying operations* that corresponds to *free operations* in our resource theory on sharpness, and based on this, we introduce *sharpness preorder*. We say that a POVM is *trivial* if it consists of only POVM elements that are proportional to an identity operator. Since adjoint maps are unital, one can not convert trivial POVMs to non-trivial POVMs. Therefore, we regard trivial POVMs as the smallest elements of the sharpness preorder.

However, since adjoint maps can not convert one trivial POVM to another trivial POVM, which implies that the all smallest elements are not equivalent in the sharpness order. To avoid this problem, we also admit a convex mixture of trivial POVM as a sharpness non-increasing operation so that all trivial measurements are equivalent in our sharpness order. Based on these observations, we introduce the sharpness non-increasing operations and sharpness preorder:

Definition 3 (Sharpness preorder [11]) *Given two POVMs* $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ *and* $\mathbf{Q} = \{Q_B^x\}_{x\in\mathcal{X}}$, *possibly defined on different Hilbert spaces* $\mathscr{H}_A$ *and* $\mathscr{H}_B$ *but with the same outcome set* $\mathcal{X}$, *we say that* $\mathbf{P}$ *is* sharper *than* $\mathbf{Q}$, *and write*

$$\mathbf{P} \succeq_{\mathcal{X}}^{\text{sharp}} \mathbf{Q}\,, \tag{1}$$

whenever there exists a quantum channel $\mathcal{E} : B \to A$, *a trivial POVM* $\{p(x)\mathbb{1}_B\}_{x\in\mathcal{X}}$ *on* B, *and a number* $\mu \in [0, 1]$, *such that*

$$Q_B^x = \mu\mathcal{E}^\dagger(P_A^x) + (1-\mu)p(x)\mathbb{1}_B \quad (\forall x \in \mathcal{X})\,.$$

Note that the post-processing of a POVM is not a fuzzifying operation. Indeed, a rank-one POVM like

$$\left\{\frac{1}{2}|\psi\rangle\langle\psi^1|_A, \frac{1}{2}|\psi\rangle\langle\psi^1|_A, \frac{1}{2}|\psi\rangle\langle\psi^2|_A, \frac{1}{2}|\psi\rangle\langle\psi^2|_A, \dots\right\} \tag{2}$$

is post-processing clean but not sharp because POVM elements do not have eigenvalue 1 and we can make sharp POVM

$$\{|\psi\rangle\langle\psi^1|_A, |\psi\rangle\langle\psi^2|_A, \dots,\}$$

from a POVM Eq. (2) by combining outcomes together. This example simply shows the reason why the previous research [6] failed to capture the resource theory of sharpness by regarding post-processing as free operations.

4 Statistical comparison of measurement sharpness

In this section, we will give a Blackwell–like theorem which is a necessary and sufficient condition for one measurement to be sharper than the other.

Firstly, we will introduce the measure of measurement sharpness. We admit the degree of quantum perfect correlations in the maximally mixed state. Ozawa introduced these correlations to define EPR argument [13], that "two observables have the same value", in a given state [8, 9] as follows:

Definition 4 *Given a state* ρ_A *on* $\mathscr{H}_A$ *and two POVMs on* A *with same outcome set* $\mathcal{X}$, $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ *and* $\mathbf{R} = \{Z_A^x\}_{x\in\mathcal{X}}$, *we say that* $\mathbf{P}$ *and* $\mathbf{R}$ *are perfectly correlated in state* ρ_A *if*

1. *they are jointly distributed in* ρ_A *that is,* $\mathrm{Tr}[P_A^x Z_A^{x'} \rho_A] \geqslant 0$ *for all* $x, x' \in \mathcal{X}$, *and*
2. $\sum_{x\in\mathcal{X}} \mathrm{Tr}[P_A^x Z_A^x \rho_A] = 1$.

More generally, if two POVMs $\mathbf{P}$ *and* $\mathbf{R}$ *are jointly distributed in state* ρ_A, *their degree of correlation is defined as*

$$\kappa_\rho(\mathbf{P} : \mathbf{R}) := \sum_{x\in\mathcal{X}} \mathrm{Tr}[P_A^x Z_A^x \rho_A].$$

In definition 4, we take the state maximally mixed one $u = \frac{1}{d_R} I$ and any reference POVM $\mathbf{R} = \{Z_R^x\}_{x\in\mathcal{X}}$ on some reference system R with Hilbert space $\mathscr{H}_R$ and optimize this quantity over all fuzzifying operations $\mathcal{L}$ as follows

$$\begin{aligned}\kappa_u^*(\mathbf{R}|\mathbf{P}) :&= \max_{\mathcal{L}} \kappa_u(\mathcal{L}(\mathbf{P}) : \mathbf{R}) \\ &= \max_{\mathcal{L}} \frac{1}{d_R} \sum_{x\in\mathcal{X}} \mathrm{Tr}[\mathcal{L}(P_A^x) Z_R^x]\,,\end{aligned} \tag{3}$$

We admit Eq. (3) as the sharpness measure and call it *tuning degree* of $\mathbf{P}$ with respect to $\mathbf{R}$. Then, we introduce the following preorder with respect to the tuning degree.

Definition 5 (tuning preorder [11]) *Given a reference POVM $\mathbf{R} = \{Z_R^x\}_{x\in\mathcal{X}}$ and two POVMs $\mathbf{P} = \{P_A^x\}_{x\in\mathcal{X}}$ and $\mathbf{Q} = \{Q_B^x\}_{x\in\mathcal{X}}$, possibly defined on different Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ but with the same outcome set as the reference $\mathbf{R}$, we say that $\mathbf{P}$ is more tunable that $\mathbf{Q}$ with respect to $\mathbf{R}$, and write*

$$\mathbf{P} \succeq_{\mathbf{R}}^{\mathrm{t}} \mathbf{Q}\,, \tag{4}$$

whenever $\kappa_u^*(\mathbf{R}|\mathbf{P}) \geqslant \kappa_u^*(\mathbf{R}|\mathbf{Q})$.

Further, given two POVMs $\mathbf{P}$ and $\mathbf{Q}$ with the same outcome set $\mathcal{X}$, we say that $\mathbf{P}$ is always more tunable than $\mathbf{Q}$, and write

$$\mathbf{P} \succeq_{\mathcal{X}}^{t} \mathbf{Q}\,, \tag{5}$$

whenever $\mathbf{P} \succeq_{\mathbf{R}}^{\mathrm{t}} \mathbf{Q}$ for all reference POVMs $\mathbf{R}$ with outcome set $\mathcal{X}$.

Next, we give a necessary and sufficient condition for one measurement to be sharper than the other as follows (for the proof, see Ref. [11]):

Theorem 6 (Theorem 3 in [11]) *Given two POVMs $\mathbf{P}$ and $\mathbf{Q}$, possibly defined on different Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ but with the same outcome set $\mathcal{X}$, $\mathbf{P}$ can be transformed into $\mathbf{Q}$ by means of a fuzzifying operation, that is,*

$$\mathbf{P} \succeq_{\mathcal{X}}^{\mathrm{sharp}} \mathbf{Q}$$

if and only if $\mathbf{P}$ is always more tunable than $\mathbf{Q}$, that is,

$$\mathbf{P} \succeq_{\mathcal{X}}^{t} \mathbf{Q} \tag{6}$$

Moreover, the comparison (6) *can be restricted without loss of generality to reference POVMs defined on the same Hilbert space as $\mathbf{Q}$, i.e., $\mathscr{H}_B$.*

Hence, the tuning degrees $\kappa_u^*(\mathbf{R}|\mathbf{P})$, for varying reference POVM $\mathbf{R}$, provide a complete set of monotones for the resource theory of sharpness.

5 Summary

In this section, we summarize the main points of the resource theory of measurement sharpness which we introduced in [11].

- The *objects* of the resource theory are POVMs, and our resource theory does not depend on the numerical values of measurement outcomes.
- The *free operations* are fuzzifying operations, which we define as a convex combination of quantum pre-processing and trivial POVM (see Definition 3).
- The *sharpest measurements* in the resource theory are sharp POVMs, and they can be transformed into any other sharp POVM with the same outcome set by a free operation.
- The *most unsharp measurements* are trivial POVMs, i.e., POVMs whose elements are all proportional to the identity operator.
- The *sharpness measures* are given by the tuning degrees $\kappa_u^*(\mathbf{R}|\mathbf{P})$ with respect to reference POVM $\mathbf{R}$, defined in Eq. (3).
- A *Blackwell–like theorem* for sharpness, that is, a necessary and sufficient condition for transformation from one POVM to another one by a free operation holds.

Acknowledgements

F.B. acknowledges support from MEXT Quantum Leap Flagship Program (MEXT QLEAP) Grant No. JPMXS0120319794, from MEXT-JSPS Grant-in-Aid for Transformative Research Areas (A) "Extreme Universe" No. 21H05183, and from JSPS KAKENHI, Grants No. 20K03746 and No. 23K03230. S.M. would like to take this opportunity to thank the "Nagoya University Interdisciplinary Frontier Fellowship" supported by Nagoya University and JST, the establishment of university fellowships towards the creation of science technology innovation, Grant Number JPMJFS2120.

References

[1] Claudio Carmeli, Teiko Heinonen, and Alessandro Toigo. Intrinsic unsharpness and approximate repeatability of quantum measurements. *Journal of Physics A: Mathematical and Theoretical*, 40(6):1303, jan 2007.

[2] Serge Massar. Uncertainty relations for positive-operator-valued measures. *Phys. Rev. A*, 76:042114, Oct 2007.

[3] Paul Busch. On the sharpness and bias of quantum effects. *Foundations of Physics*, 39(7):712–730, 2009.

[4] Kyunghyun Baek and Wonmin Son. Unsharpness of generalized measurement and its effects

in entropic uncertainty relations. *Scientific Reports*, 6(1):30228, 2016.

[5] Yizhou Liu and Shunlong Luo. Quantifying unsharpness of measurements via uncertainty. *Phys. Rev. A*, 104:052227, Nov 2021.

[6] Arindam Mitra. Quantifying unsharpness of observables in an outcome-independent way. *International Journal of Theoretical Physics*, 61(9):236, 2022.

[7] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, Apr 2019.

[8] Masanao Ozawa. Perfect correlations between noncommuting observables. *Physics Letters A*, 335(1):11–19, 2005.

[9] Masanao Ozawa. Quantum perfect correlations. *Annals of Physics*, 321(3):744–769, 2006.

[10] David Blackwell. Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 1953.

[11] Francesco Buscemi, Kodai Kobayashi, and Shintaro Minagawa. A complete and operational resource theory of measurement sharpness. *arXiv:2303.07737*, 2023.

[12] Francesco Buscemi, Michael Keyl, Giacomo Mauro D'Ariano, Paolo Perinotti, and Reinhard F. Werner. Clean positive operator valued measures. *Journal of Mathematical Physics*, 46(8):082109, 2005.

[13] A. Einstein, B. Podolsky, and N. Rosen. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10):777–780, May 1935.

On the role of SIC structures in the data-driven approach to quantum statistical inference

Michele Dall'Arno[1 2 *]

[1]*Department of Computer Science and Engineering, Toyohashi University of Technology, Japan*
[2]*Yukawa Institute for Theoretical Physics, Kyoto University, Japan*

Abstract. Quantum tomography is the protocol typically adopted for the reconstruction of a given quantum device. For instance, if the device corresponds to a quantum measurement, a tomographic probe (an informationally complete set of states) is fed as input to the measurement, and the output statistics is collected. Based on the knowledge of both the probe and the statistics, the measurement is then reconstructed. This approach is problematic in that, while the statistics is directly accessible, the knowledge of the probe requires a tomographic reconstruction to be applied to the probe itself, and so on, in a never-ending chain of tomographic reconstructions.

Here, we introduce a protocol that breaks such a chain by performing the statistical inference of a quantum measurement solely based on the output statistic, without requiring any knowledge of the input probe. Hence our protocol, that we refer to as data-driven inference, can bootstrap quantum tomography. Similarly to Jaynes' MAXENT principle, data-driven inference is based on a minimality criterion according to which, among all the reconstructions that can explain the observed statistics, the one which explains as little else as possible should be preferred. We solve such a minimality problem by showing that the minimally-committal reconstruction for the unknown measurement is the one that would have generated the statistics upon the input of a symmetric, informationally complete (SIC) probe.

This presentation is based on Refs. [6, 7].

Keywords: quantum data-driven inference, quantum measurement, quantum statistical inference

Quantum Bayesianism [1, 2, 3, 4] (or QBism for short) is an interpretation of quantum theory, introduced by Fuchs and Schack, in which a fundamental role is played by the observer's ability to reconstruct reality by a process known as tomographic reconstruction. A major role in QBism is played by symmetric, informationally complete structures (SICs), which can be used as tomographic probes for such a reconstruction process. The importance of SICs' role in QBism is typically justified by the fact that the symmetry of such structures induces an analogous elegant symmetry in the formula for the tomographic reconstruction.

However, SICs are by far not the only structures suitable as probes for the tomographic reconstruction process, and indeed such a task can be accomplished by any structure within the broader class of informationally complete (non necessarily symmetric) structures. In fact, the possibility to be used as a universal probes for tomographic reconstruction is what defines (and justifies the name of) the class of informationally complete structures. Therefore, elegance aside, what is the actual criterion that should single out SICs as the "standard bureau of measurements" in quantum theory?

In this contribution, we answer this question by presenting a new type of reconstruction [5, 6], alternative to the usual tomographic reconstruction, and by showing that SICs play a pivotal role in such a reconstruction. Most importantly, we show that, in contrast to the case of tomographic reconstruction, SICs are the *unique* structures that play such a role: for instance, no other informationally complete structure can replace SICs for the reconstruction process we consider.

The reconstruction process we consider is an

*michele.dallarno.mv@tut.jp

instance of statistical inference. In contrast to tomographic reconstruction, which requires the a-priori knowledge of the tomographic probe, thus leading to a never-ending chain of reconstructions, the inferential process we propose is data-driven, that is, it is solely based on observable data. In this sense, data-driven inference breaks the chain of reconstructions implied by quantum tomography, and can be regarded as a bootstrap for the latter. Analogously to Jaynes' maximum entropy principle, data-driven inference is based on a minimality criterion that, among all possible quantum descriptions of the observed data, singles out the minimally committal one, that is, the one that explains the observed data and as little else as possible. More formally, the committal degree of any given measurements is given by the volume of its probability range, that is, it is a measure of how many observable data the measurement is consistent with.

The (unique) role played by SICs in data driven inference is two-fold. From the theoretical point of view (that is, one in which only observable data are accessible), we prove that the reconstruction process based on data-driven inference is equivalent to treating the observed data *as if* they had been generated by a SIC structure. From the experimental point of view (that is, one in which control over the devices is assumed), we show that the data-driven inference process produces the correct output if and only if SIC probes are utilized. Hence, the data-driven inference makes it explicit the unique role played by SICs in quantum theory, in a way that shares similarities, at least at the formal level, with the approach recently adopted by Szymusiak and Slomczyński in the study of morphophoric structures [8, 9].

References

[1] C. A. Fuchs, *QBism, the Perimeter of Quantum Bayesianism*, arXiv:1003.5209.

[2] C. A. Fuchs and R. Schack, *Quantum-Bayesian coherence*, Reviews of Modern Physics **85**, 1693 (2013).

[3] C. A. Fuchs, N. D. Mermin, and R. Schack, *An introduction to QBism with an application to the locality of quantum mechanics*, American Journal of Physics **82**, 749 (2014).

[4] C. A. Fuchs and B. C. Stacey, textit QBism: Quantum Theory as a Hero's Handbook, arXiv:1612.07308.

[5] F. Buscemi and M. Dall'Arno, *Device-Independent Inference of Physical Devices: Theory and Implementation*, New J. Phys. **21**, 113029 (2019).

[6] M. Dall'Arno, F. Buscemi, A. Bisio, and A. Tosini, *Data-driven inference, reconstruction, and observational completeness of quantum devices*, Phys. Rev. A **102**, 062407 (2020).

[7] M. Dall'Arno, *On the role of designs in the data-driven approach to quantum statistical inference*, arXiv:2304.13258.

[8] W. Slomczyński and A. Szymusiak, *Morphophoric POVMs, generalised qplexes, and 2-designs*, Quantum **4**, 338 (2020).

[9] A. Szymusiak and W. Slomczyński, *Can QBism exist without Q? Morphophoric measurements in generalised probabilistic theories*, arXiv:2302.04957.

Two-qutrit entanglement: a 56-years old algorithm challenges machine learning

Marcin Wieśniak[1,2,*]

[1]Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics, and Informatics, University of Gdańsk, 80-308 Gdańsk, Poland
[2]International Centre for Theory of Quantum Technologies, University of Gdańsk, ul. Bażyńskiego 1A, 80-309 Gdańsk, Poland
[*]marcin.wiesniak@ug.edu.pl

ABSTRACT

1 Introduction

Modern computing techniques could be employed to the problem of separability in specific cases. For example, there was a number of attempts to train neural networks to distinguish entangled states from separable ones[1–13]. This approach poses its own set of challenges. First, as every use of machine learning (ML) techniques is application-specific, relatively little can be said about the optimality of the topology of the neural network, the form of the activation function, etc.. The other issue is that the result of ML is dependent on the training data, which can be very limited for the entanglement-vs-separability problem. A network can become overtrained with a fixation on the presented pattern. Finally, neural networks are not designed to estimate the missing continuous value of function, which makes them less capable of providing the interpolation of the amount of entanglement,Algorithm although there were some attempts of such tasks.

Recently, other strategies for classifying states as entangled or separable, based on variational algorithms[14] and semi-definite programming[15], were proposed.

In this articles we compare the algorithmic approach described in details below with ML results obtained by Hiesmayr[5], who focused on "the magic simplex", or Bell-diagonal states of three qubits. The task is to classify states as free entangled, bound entangled, or separable. Our methodology is described below, followed by a presentation of results and conclusions.

2 Hilbert-Schmidt distance and Gilbert's algorithm

The Hilbert-Schmidt distance is defined simply as

$$\begin{aligned} D_{HS}(A,B) = & \sqrt{\mathrm{Tr}((A-B)(A-B)^{\dagger})}, \\ = & \sqrt{\sum_{i,j}|(A-B)_{i,j}|^2}. \end{aligned} \tag{1}$$

Unsurprizingly, it was suggested as a core of entanglement measure by Witte and Trucks[16]. This entanglement quantifier shall be defined as

$$D_{HS}(\rho) = \min_{\sigma \in SEP} D_{HS}(\rho,\sigma), \tag{2}$$

where the maximum is taken over all separable states.

In 1966 Gilbert[17] has introduced an algorithm to estimate the distance between a given point and a convex set. The algorithm can be outlined as follows.

Algorithm 1 *(bipartite case):*
Input: *Test state ρ_0, initial separable state ρ_1.*
Output: *Approximation of CSS ρ_1, list of squared distances to subsequent CSS approximations l.*

1. *Take a random pure $\rho_2 = |\varphi_A\rangle|\varphi_B\rangle\langle\varphi_A|\langle\varphi_B|$, that will be referred to as a trial state.*

2. *If the the preselection criterion, $Tr(\rho_0-\rho_1)(\rho_2-\rho_1)>0$ is not met, go to step 1 or abort if the HALT condition is met.*

3. *Maximize $Tr(\rho_0-\rho_1)(\rho_2-\rho_1)>0$ with local unitary transformation (run Algorithm 3)*

4. *Update $\rho_1 \leftarrow p\rho_1+(1-p)\rho_2$ for p minimizing $D(\rho_0, p\rho_1+(1-p)\rho_2)$.*

5. *Every 50 corrections append $D(\rho_0,\rho_1)$ to l.*

6. *If the HALT condition is not met, go to step 1, otherwise quit.*

The algorithm provides three pieces of information that can be used here to classify states as entangled or separable. The first is, rather trivially, the last squared distance D^2_{Last} found within a fixed number of corrections. The second indicator is the distance decay estimate, D^2_{Est}.

The third and final figure of merit considered here is the witness distance estimate D_{Wit},

$$W=\rho_0-\rho_1-\mathbb{1}\max_{|\phi_1\rangle,|\phi_2\rangle}\langle\phi_1|\langle\phi_2|\rho_0-\rho_1|\phi_1\rangle|\phi_2\rangle, \tag{3}$$

with the maximum taken over all product states. Then

$$D_{\text{Wit}}=\max\left(0,\frac{\text{Tr}W\rho_0}{\sqrt{\text{Tr}(\rho_0-\rho_1)^2}}\right). \tag{4}$$

3 "Magic Simplex" states

Let us consider a maximally entangled state of two qutrits,

$$|\psi_{00}\rangle=\frac{1}{\sqrt{3}}\sum_{i=0}^{2}|ii\rangle, \tag{5}$$

and two Weyl operators, $X=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$ and $Z=\begin{pmatrix}1&0&0\\0&\alpha&0\\0&0&\alpha^2\end{pmatrix}$ with $\alpha=e^{2\pi\imath/3}$. The Bell basis is given by

$$\{|\psi_{ij}\rangle\}_{i,j=0}^2=\{\mathbb{1}\otimes X^iZ^j|\psi_{00}\rangle\}_{i,j=0}^2. \tag{6}$$

Then the Bell-diagonal states are given by

$$\begin{aligned}&\rho=\sum_{i,j=0}^{2}p_{ij}|\psi_{ij}\rangle\langle\psi_{ij}|,\\&\sum_{i,j=0}^{2}p_{ij}=1, p_{ij}\leq 0.\end{aligned} \tag{7}$$

In particular, Hiesmayr focused on four families of states. Family A was given by

$$\begin{aligned}\rho(\alpha,\beta,\gamma)=\quad&(1-\alpha-\beta-\gamma)\frac{\mathbb{1}}{9}\\+\quad&\alpha|\psi_{00}\rangle\langle\psi_{00}|\\+\quad&\beta|\psi_{01}\rangle\langle\psi_{01}|\\+\quad&\gamma|\psi_{02}\rangle\langle\psi_{02}|,\end{aligned} \tag{8}$$

and it was considered for $\gamma=0$. The other three families are a part of Family B, which includes all the states in form

$$\begin{aligned}\rho(\alpha,\beta,\gamma,\delta)\quad&=(1-\alpha-\beta-\gamma-\delta)\frac{\mathbb{1}}{9}\\+\quad&\alpha|\psi_{00}\rangle\langle\psi_{00}|+\frac{\beta}{2}(|\psi_{01}\rangle\langle\psi_{01}|+|\psi_{02}\rangle\langle\psi_{02}|)+\frac{\gamma}{3}\sum_{j=0}^{2}|\psi_{1j}\rangle\langle\psi_{1j}|+\frac{\delta}{3}\sum_{j=0}^{2}|\psi_{2j}\rangle\langle\psi_{2j}|.\end{aligned} \tag{9}$$

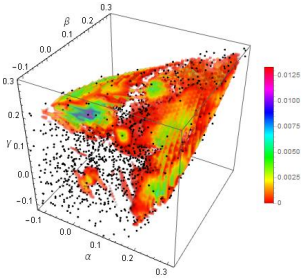


Figure 1. 3D plot of D_{Last} for PPT states from family A. Black dots represent studied states.

4 Results

4.1 Family A

We first focus on Family A, in which we studied 1485 PPT states, each with up to 4000 corrections (the algorithm HALTS at $D^2(\rho_0,\rho_1) < 10^{-7}$). Figure 1 presents a 3D plot of D_{EST}.

4.2 Family B

Subsequently, we have conducted the analysis for the remaining three families. Within each family we generated 300 states. For Family B_1, B_2, and B_3 we conducted up to 10000, 8000 and 10000 corrections respectively. The results are presented in Figure 3 and they seem to adequately reproduce the plots in Ref.[5].

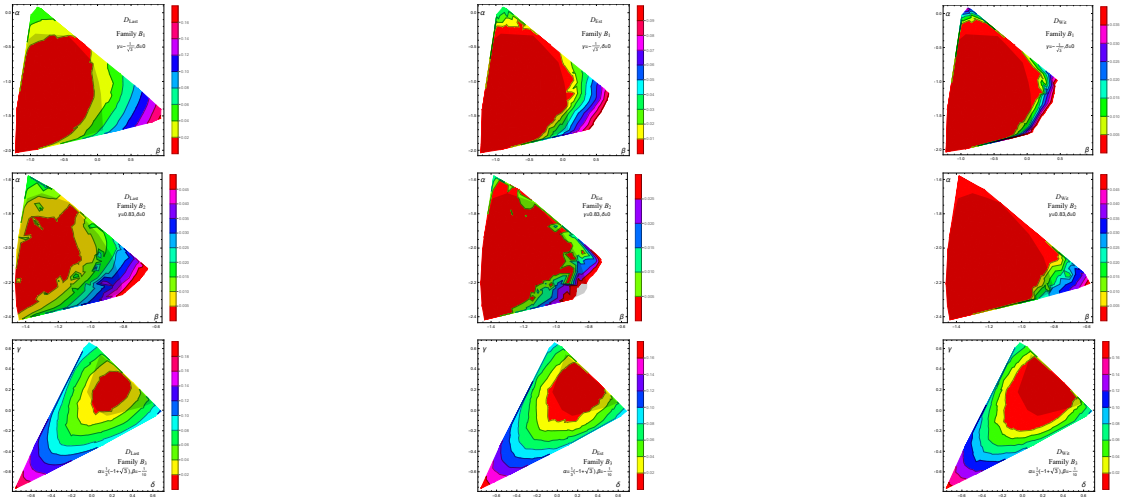


Figure 2. The interpolated plots of D_{Last} (left), D_{Est} (center), and D_{Wit} (right) for Families B_1 (top), B_2 (middle) and B_3 (bottom) Shaded areas represent PPT states.

5 Conclusions

The presented technique of "cartography of entanglement" can be used universally applied to identify or estimate the boundary between separable and entangled states, regardless of the dimension, the number of subsystems, or a type of quantum correlations in question. For example, it can find a variety of applications in solid state models. Importantly, the algorithm has not been fed with any information other than the input state. It is irrelevant if the state has free or bound entanglement. The technique could be combined with machine learning and interpolation techniques, but it can also generate useful results on its own. In contrast to FREE/SEP/BOUND categorization, it provides a qualitative information about entanglement. While new bound entangled states can be easily detected, it is the question if their nonclassicality can be meaningful in an experimental realization.

References

1. Wang, B. Learning to detect entanglement. *arXiv preprint arXiv:1709.03617* (2017).
2. Gray, J., Banchi, L., Bayat, A. & Bose, S. Machine-learning-assisted many-body entanglement measurement. *Phys. review letters* **121**, 150503 (2018).
3. Chen, Y., Pan, Y., Zhang, G. & Cheng, S. Detecting quantum entanglement with unsupervised learning. *Quantum Sci. Technol.* **7**, 015005 (2021).
4. Harney, C., Paternostro, M. & Pirandola, S. Mixed state entanglement classification using artificial neural networks. *New J. Phys.* **23**, 063033 (2021).
5. Hiesmayr, B. C. Free versus bound entanglement, a np-hard problem tackled by machine learning. *Sci. Reports* **11**, 19739 (2021).
6. Girardin, A., Brunner, N. & Kriváchy, T. Building separable approximations for quantum states via neural networks. *Phys. Rev. Res.* **4**, 023238 (2022).
7. Roik, J., Bartkiewicz, K., Černoch, A. & Lemr, K. Entanglement quantification from collective measurements processed by machine learning. *Phys. Lett. A* **446**, 128270 (2022).
8. Lin, X., Chen, Z. & Wei, Z. Quantifying unknown quantum entanglement via a hybrid quantum-classical machine learning framework. *arXiv preprint arXiv:2204.11500* (2022).
9. Scala, F., Mangini, S., Macchiavello, C., Bajoni, D. & Gerace, D. Quantum variational learning for entanglement witnessing. In *2022 International Joint Conference on Neural Networks (IJCNN)*, 1–8 (IEEE, 2022).
10. Vintskevich, S., Bao, N., Nomerotski, A., Stankus, P. & Grigoriev, D. Classification of four-qubit entangled states via machine learning. *arXiv preprint arXiv:2205.11512* (2022).
11. Pawłowski, J. & Krawczyk, M. Quantification of entanglement with siamese convolutional neural networks. *arXiv preprint arXiv:2210.07410* (2022).
12. Ayachi, F. E. & Baz, M. E. General classification of entanglement using machine learning. *arXiv preprint arXiv:2210.07711* (2022).
13. Chen, Z., Lin, X. & Wei, Z. Certifying unknown genuine multipartite entanglement by neural networks. *arXiv preprint arXiv:2210.13837* (2022).
14. Consiglio, M., Apollaro, T. J. & Wieśniak, M. Variational approach to the quantum separability problem. *Phys. Rev. A* **106**, 062413 (2022).
15. Ohst, T.-A., Yu, X.-D., Gühne, O. & Nguyen, H. C. Certifying quantum separability with adaptive polytopes. *arXiv preprint arXiv:2210.10054* (2022).
16. Witte, C. & Trucks, M. A new entanglement measure induced by the hilbert–schmidt norm. *Phys. Lett. A* **257**, 14–20 (1999).
17. Gilbert, E. G. An iterative procedure for computing the minimum of a quadratic form on a convex set. *SIAM J. on Control.* **4**, 61–80 (1966).

6 Acknowledgements

This work is a part of NCN Grant No. 2017/26/E/ST2/01008. MW acknowledges partial support by the Foundation for Polish Science (IRAP project, ICTQT, Contract No. 2018/MAB/5, co-financed by EU within Smart Growth Operational Programme).

7 Author contributions statement

MW fully conducted the research in all aspects.

8 Data availability

The source code is available at https://www.github.com/wiesnim9/CSSFinder. The data are available from the Author upon a reasonable request.

9 Conflict of interests

The Author declares no conflicting interests.

Perturbative Tools for Analyzing Quantum Error Amplification Circuits

Takanori Sugiyama[1 2 *]

[1] *Quantum Laboratory, Fujitsu Limited. Nakahara-ku, Kawasaki, Kanagawa 211-8588, Japan.*
[2] *RIKEN RQC-FUJITSU Collaboration Center, RIKEN. Wako, Saitama 351-0198, Japan.*

Abstract. Error amplification circuit (EAC) is a repetition of a sequence of quantum gates, and it is recently used in advanced tomographic characterization methods. Purpose of the use is to suppress bias originated from a mismatch of our model on state preparation and measurement (SPAM) errors. In general, effects of EAC, e.g., how and which part of gate error is amplified (or not amplified) by the given EAC, was not clear except for specific Hamiltonian dynamics. Difficulty for analyzing more general settings is mainly originated from (i) co-existence of Hamiltonian and decoherence dynamics, (ii) non-commutativity between different gates or between ideal and error parts of a gate, and (iii) periodic (non-linear) properties of ideal gates. Here we develop theoretical tools for analyzing effects of EAC for arbitrary finite-dimensional quantum system, which is based on the first order perturbation with respect to generator (Hamiltonian or Lindbladian) errors of gates, taking into account all of (i), (ii), and (iii). We numerically show that, for typical quantum gates, some of our tools give more accurate approximation than the Baker–Campbell–Hausdorff (BCH) formula. By combining the tools, we can calculate a set of matrices that quantify the effects of EAC. These results indicate that the tools and method proposed contribute to deeper and more accurate theoretical understanding of EAC.

Keywords: Error Amplification, Matrix Perturbation Theory, Quantum Tomography

1 Introduction

Development toward realizing practical quantum computer is accelerated in the past decade since accuracies of one- and two- qubit gates arrived around surface code' threshold at a superconducting quantum circuit [1] in 2014, and recently, achievement of break-even point for surface code with distance-3 and 5 is reported in the system [2]. As shown in these achievements, accuracies of elementary quantum operations is getting improved, but more than 1-digit improvements is necessary for execution of a surface code with practical size [3, 4, 5].

Characterization methods for elementary quantum operations such as quantum tomography and randomized benchmarking are used for improving accuracies, and take a role to obtain information of errors of the operations. Tomographic methods are suitable for obtaining detailed information of the errors, but its standard protocol suffers from not-negligible systematic errors originated from mismatch of our model values on states and measurements, which is called state-preparation-and-measurement (SPAM) error. Error amplification circuit (EAC) consists of a repetition of a sequence of quantum gates (Figure 1). It is used to suppress such SPAM error in advanced tomographic methods such as gate-set tomography (GST) [6], idle tomography (IT) [7], and Hamiltonian-Error Amplifying tomography (HEAT) [8].

Generally speaking, effects of EAC was not clear except for specific Hamiltonian dynamics [7, 8]. In this talk, we intoduce theoretical tools for analyzing effects of EAC on generator (Hamiltonian or Lindladian) error of gate, which is based on the first order perturbation theory. We derive new formulae for composition and decomposition of matrix exponential, and numerically show that these are more accurate than the Baker–Campbell–Hausdorff (BCH) formula for typical gates. We propose a systematic method for deriving matrices that describe the effects of EAC on generator error quantitatively.

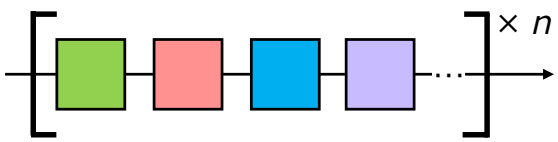


Figure 1: Quantum circuit diagram of an error amplification circuit. The superscript, "$\times n$", means n times repetition of the braketed gate sequence.

*sugiyama-taka@fujitsu.com

2 Results

We introduce elemental tools for analyzing EAC. We consider a square complex matrix $A, B \in \mathbb{C}^{m\times m}$. A corresponds to an ideal matrix representation of generator of a gate, e.g., ideal Hamiltonian with imaginary factor $-iH^{\rm ideal}$ or ideal Lindbladian $A = L^{\rm ideal}$. B does to a small perturbation added to A representing an error on the generator, e.g., $B = -i\delta H$ or $B = \delta L$. $A + B$ does to an implemented generator. Note that here terms "Hamiltonian" and "Lindbladian" are not used for infinitesimal generator of the dynamics, but are used for accumulated generator during gate operation. Details of matrix representation of accumulated generator is explained in [9].

Suppose that A is diagonalisable (not necessarilly unitarilly diagonalisable), and $A = \sum_i a_i P_i$ is the spectral decomposition, where a_i are eigenvalues ($a_i \neq a_j$ if $i \neq j$) and P_i are projections satisfying $P_i P_j = \delta_{ij} P_i$.

2.1 Decomposition and Composition

We define complex values ℓ_{jk} for A as

$$\ell_{jk} := \begin{cases} 1 & (j = k) \\ (e^{a_j - a_k} - 1)/(a_j - a_k) & (j \neq k) \end{cases}. \quad (1)$$

Let us introduce the following four linear maps.

$$dcl_A(B) := \sum_{j,k} \ell_{jk} P_j B P_k, \quad (2)$$

$$dcr_A(B) := \sum_{j,k} \ell_{kj} P_j B P_k, \quad (3)$$

$$cml_A(B) := \sum_{j,k} \frac{1}{\ell_{jk}} P_j B P_k, \quad (4)$$

$$cmr_A(B) := \sum_{j,k} \frac{1}{\ell_{kj}} P_j B P_k. \quad (5)$$

We derived the first order perturbation formulae with these maps for decomposition and composition of matrix exponentials.

Theorem 1 *For A and B mentioned above, the following decomposition formulae hold.*

$$e^{A+B} = e^{{\rm dcl}_A(B)} e^A + O(\|B\|^2), \quad (6)$$

$$e^{A+B} = e^A e^{{\rm dcr}_A(B)} + O(\|B\|^2), \quad (7)$$

If A satisfies $e^{a_j - a_k} \neq 1$ for any j, k ($j \neq k$), the following composition formulae hold.

$$e^B e^A = e^{A + cml_A(B)} + O(\|B\|^2), \quad (8)$$

$$e^A e^B = e^{A + cmr_A(B)} + O(\|B\|^2), \quad (9)$$

where $\|\cdot\|$ denote the Frobenius norm.

For proving Theorem 1, we used the integral formula of matrix exponential derivative [10, 11].

A major difference between our result (Theorem 1) and the BCH formula is a treatment of series expansion orders of A and B. In the BCH formula, its series expansion order is for both of A and B, but int our result the order ris only for B (up to the 1st order). Our result can be interpreted as the BCH formula with order of A up to the infinity and that of B up to one. Additionally, in the BCH formula, both of A and B are assumed to be sufficiently small. There are several sufficient condition for the convergence of BCH series expantion, and an explicit inequality [12] is

$$\|A\| + \|B\| \leq \ln 2 \approx 0.693147\ldots. \quad (10)$$

Unfortunately, typical quantum gates like $\pi/4$-, $\pi/2$-, and π-pulse gates do not satisfy Eq. (10). For example, in the cases of ideal Lindbladians of 1-qubit X90 gate ($A = L^{\rm ideal}_{X90}$) and 2-qubit ZX90 gate ($A = L^{\rm ideal}_{ZX90}$), which is a popular 2-qubit gate used at fixed-frequency superconducting qubit [13]) are

$$\|A\| = \begin{cases} \pi/\sqrt{2} \approx 2.22144\ldots & ({\rm X90}) \\ \sqrt{2}\pi \approx 4.44288\ldots & ({\rm ZX90}) \end{cases}, \quad (11)$$

and both exceed the R.H.S. of Eq. (10). On the other hand, in our result the size of A is arbitrary. We numerically evaluated the approximation errors of Eqs. (6) to (9) for 1-qubit, 2-qubit, and 1-qutrit gates, and compared Eqs. (8) and (9) with the BCH formula. Fig. 2 shows the results for Eq. (9) with Lindladians of (a) 1-qubit X90 gate and (b) 2-qubit ZX90 gate. As expected from the discussion above, our result has much smaller approximation error than the BCH formula. These examples indicate that our result is more appropreate for treating typical quantum gates than the BCH formula.

By combining Eqs. (6) to (9), which are for ideal and error parts of a gate, we obtain a composition formula for two gates.

Theorem 2 *Suppose that A and $A' \in \mathbb{C}^{m\times m}$ are diagonalizable. Let B and B' are small perturbations added to A and A', respectively. Let C denote the ideal generator of the composed gate, i.e.,*

$$C := \ln(e^A e^{A'}) \quad (12)$$

Then, the following composition formula holds.

$$e^{A+B} e^{A'+B'} = e^{C + cml_C \circ dcl_A(B) + cmr_C \circ dcr_{A'}(B')} + O(\|B\|^2, \|B\|\|B'\|, \|B'\|^2), \quad (13)$$

where $\circ$ denote the composition of maps.

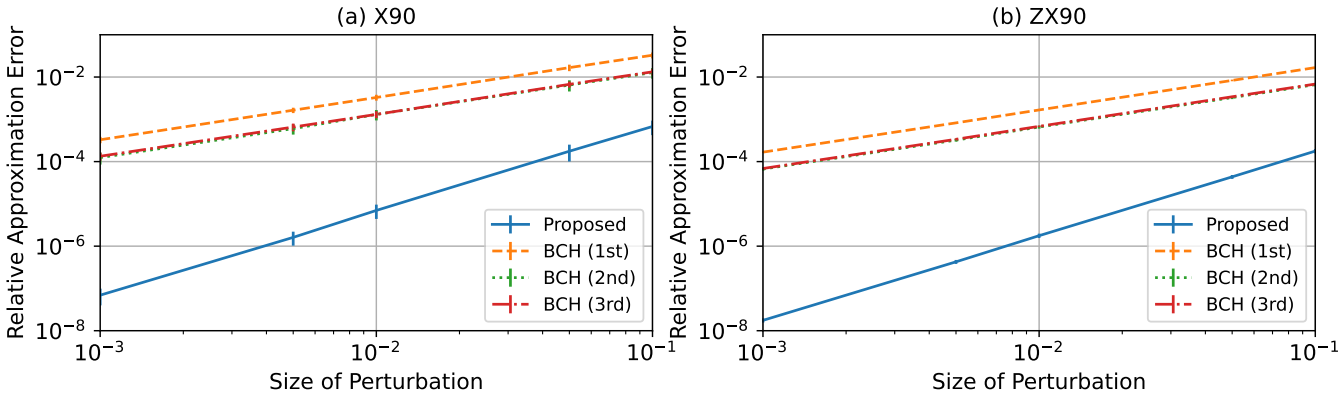


Figure 2: Numerical comparison between our tool and the BCH formula for Lindbladians of (a) X90 and (b) ZX90 gates. For both panels, the vertical axis is the relative approximation error, and the horizontal axis is the size of perturbation, $\|B\| = \|\delta L\|$. Blue solid lines are for our tool (Eq. (9)). Yellow dashed, green dotted, and red dashed-and-dotted lines are for the 1st, 2nd, 3rd order BCH formulae, respectively. Error bars are the standard deviations calculated with 100 randomly generated Lindbladian errors δL for each fixed size, 0.001, 0.005, 0.01, 0.05, and 0.1.

Theorem 2 clarifies how generator errors B and B' are transformed by the composition of two gates up to the first order of them, i.e.,

$$B \rightarrow cml_C \circ dcl_A(B), \tag{14}$$
$$B' \rightarrow cmr_C \circ dcr_{A'}(B'). \tag{15}$$

We can treat composition of more than two gates by applying Eq. (13) to a gate sequence recursively.

2.2 Repetition

Next, we introduce tools for analyzing effect of gate repetition on generator error. A major difficulty of the analysis is the periodic properties of ideal generator of typical quantum gates. For example, in the case of $\pi/2$-pulse gates, whose action is a 90-degree rotation along with an axis, there is a periodicity with period four, e.g.,

$$\left[e^A\right]^5 = e^{5A} = e^A \tag{16}$$

holds, but this periodicity does not hold directly when an error exists, e.g.,

$$\left[e^{A+B}\right]^5 = e^{5(A+B)} \neq e^{A+5B}, \tag{17}$$

because of the possible non-commutativity between A and B.

In order to clarify the effect of gate repetition taking into account the non-commutativity above, we introduce two linear maps.

$$ssp_A(B) := \sum_j P_j B P_j, \tag{18}$$
$$sspc_A(B) := \sum_{j,k(j\neq k)} P_j B P_k \tag{19}$$

Then we have proved the following theorem.

Theorem 3 *Let n denote a repetition number of a gate, which is a positive integer. Suppose that $A + B$ is diagonalizable, and $\|B\|$ is sufficiently small. We assume that there exists an integer c satisfying* $\exp(nA) = \exp(cA)$. *Then the following equality holds,*

$$\left[e^{A+B}\right]^n = e^{cA+c\cdot sspc_A(B)+n\cdot ssp_A(B)} + O(\|B\|^2). \tag{20}$$

In the case of the $\pi/2$-pulse gate example above, $c = 1$ for $n = 5$, but Theorem 2 is applicable for more general cases. Theorem 2 clarifies that the amplified part, which is proportional to n, by the repetition is $ssp_A(B)$, and non-amplified part is $sspc_A(B)$. The former corresponds to change of eigenvalues by addition of error B, and the later does to that of eigenvectors.

2.3 Error Amplification Circuit

An EAC consists of a sequence of gates and its repetition. Therefore, by combining Theorems 2 and 3, we can clarify how and which part of a generator error is amplified or not-amplified. Such effect is characterized by six linear maps defined by Eqs. (6), (7), (8), (9), (18), and (19). The action of a linear map can be described by its matrix representation, and the matrix representation is useful for the analysis of EAC because composition of maps becomes matrix multiplication in the representation.

References

[1] R. Barends et al., Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature* 508, 500 (2014).

[2] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit. *Nature* 614, 676 (2023).

[3] A. G. Fowler et al., Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A* 86, 032324 (2012).

[4] N. Yoshioka et al., Hunting for quantum-classical crossover in condensed matter problems. arXiv:2210.14109 [quant-ph]

[5] M. E. Beverland et al., Assessing requirements to scale to practical quantum advantage. arXiv:2211.07629 [quant-ph]

[6] R. Blume-Kohout et al., Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography. *Nature Commun.* 8, 14485 (2017).

[7] R. Blume-Kohout, Idle Tomography. United States (2019) https://www.osti.gov/servlets/purl/1581878

[8] N. Sundaresan et al., Reducing Unitary and Spectator Errors in Cross Resonance with Optimized Rotary Echoes. *PRX Quantum* 1, 020318 (2020).

[9] T. Sugiyama, S. Imori, and F. Tanaka, Self-consistent quantum tomography with regularization. *Phys. Rev. A* 103, 062615 (2021).

[10] R. M. Wilcox, Exponential Operators and Parameter Differentiation in Quantum Physics. *J. Math. Phys.* 8, 962 (1967).

[11] M. Hayashi, Quantum Information Theory. (2nd eds.), Springer (2016).

[12] W. Rossmann, Lie Groups: An Introduction through Linear Groups. Oxford University Press (2006).

[13] A. Blais et al., Circuit quantum electrodynamics. *Rev. Mod. Phys.* 93, 025005 (2021).

A duplication-free quantum neural network for universal approximation

Xiaokai Hou[1] Guanyu Zhou[1] * Qingyu Li[1] Shan Jin[1] Xiaoting Wang[1] †

[1] *Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu, 610051, China*

Abstract. The universality of a quantum neural network (QNN) refers to the ability to approximate arbitrary functions. However, conventional universal QNNs may result in a huge quantum register that is challenging to implement due to noise on a near-term device. To address this, we propose a duplication-free quantum neural network whose universality can be rigorously proved. Our method relies on a single quantum register combined with multiple activation functions to achieve universality. Accordingly, our proposal requires significantly fewer qubits with shallower circuits, and hence substantially reduces the resource overhead and the noise effect, illustrating a great potential in solving larger-scale learning problems on near-term devices.

Keywords: quantum computing, quantum machine learning, quantum neural network, universality

1 Introduction

Machine learning (ML) is a powerful data-analyzing tool that has generated a series of impactful results. In the meanwhile, quantum machine learning has become an emerging interdisciplinary subject that combines machine learning with quantum computing. It studies two fundamental questions [2], one on applications of classical ML to quantum problems [4, 8, 10, 29, 9, 21], and the other on implementations of ML algorithms on a quantum processor. For the latter, a crucial question lies in how to implement neural networks on quantum devices while ensuring their performance is equivalent to that of classical neural networks. A neural network (NN) is a parameterized composite mapping comprised of activation functions and excels a great power in data fitting. The quantum neural network (QNN), including the variational QNN [19, 26, 6, 14, 17], is an NN implemented on a quantum device [18, 24, 11, 34, 31]. Its success in accurately solving the learning problems [35, 7] relies on several important intrinsic properties of the QNN. One property is the *trainability*, which focuses on how to avoid the occurrence of barren plateaus that cause the optimization to fail [18]. Another property is the *universality*, describing its ability to approximate arbitrary nonlinear functions [12, 16]. For classical NNs, universality is easy to achieve; for QNNs, a smart design is required to achieve nonlinearity and universality. Proposals for universality include the data re-uploading approach [22], the Fourier series method [28] and the construction of quantum neurons [3, 36, 30, 15, 32, 1]. Another well-known method for universality is to duplicate the quantum data into a tensor product of multiple copies [27, 33, 19, 26]. For the duplication-based method, approximating a highly-nonlinear function would require a tensor product of many data-encoding subsystems, resulting in a large overall system size with a considerable circuit complexity, conflicting with the principle of NISQ computing where a relatively small quantum system with a shallow circuit is preferred.

To address this problem, we propose a duplication-free quantum neural network (DQNN) based on a variational quantum circuit and rigorously prove its universality. Our model utilizes the classical sigmoid function to generate nonlinearity without duplicating the quantum data into a tensor product of multiple copies. Compared with the CCQ or QCL algorithms, our DQNN significantly reduces the system size and gate complexity, and hence the overall noise effect. Therefore, it is more likely to be implemented on near-term devices. Numerical simulations show that our DQNN outperforms the other two variational QNN algorithms with better performance on typical regression and classification problems and is more robust against coherent and decoherence noise. In addition, through solving a broad range of classical and quantum learning problems, our model has well demonstrated its wide application potential.

2 DQNN Model

Our DQNN model consists of three parts, a quantum processor, a classical processor and a classical optimizer. First, each data point $\boldsymbol{x} = [x_1, x_2, \cdots, x_d]^T \in \mathbb{R}^d$ is encoded into a state of an n-qubit register ($n = \lceil \log d \rceil$):

$$|\bar{\boldsymbol{x}}\rangle = \frac{1}{\gamma}[x_1, x_2, \cdots, x_d, \tilde{x}, 0, \cdots 0]^T \in \mathbb{C}_2^{\otimes n}, \qquad (1)$$

using the amplitude encoding method [23, 20, 25], where $\tilde{x} \equiv \frac{\|\boldsymbol{x}\|}{1+\|\boldsymbol{x}\|}$ is a padding term chosen by the user and γ is a normalization factor. After initial state preparation, we apply to $|\bar{\boldsymbol{x}}\rangle$ a series of variational quantum circuits $\{U^{(j)}(\boldsymbol{\theta}^{(j)})\}_{j=1}^{n_{\rm cir}}$ according to $|\bar{\boldsymbol{x}}_f^{(j)}\rangle = U^{(j)}(\boldsymbol{\theta}^{(j)})|\bar{\boldsymbol{x}}\rangle$, followed by measurements of a set of observables $\{B_i\}_{i=1}^{n_{\rm obs}}$, to obtain the outcomes $\langle B_i\rangle_{\boldsymbol{\theta}^{(j)},\bar{\boldsymbol{x}}} = {\rm Tr}(|\bar{\boldsymbol{x}}_f^{(j)}\rangle\langle\bar{\boldsymbol{x}}_f^{(j)}|B_i)$. In particular, $\{B_i\}$ can be chosen from a generalized k-local Pauli basis $\{P_l\}$, where $P_l = A_1^{(l)} \otimes A_2^{(l)} \otimes \cdots \otimes A_n^{(l)}$, $A_i^{(l)} \in \{X, Y, Z, I\}$, with at most k sites with nontrivial $A_k^{(l)}$. Such choices of $\{B_i\}$ will help avoid the occurrence of the barren plateaus [5]. $U^{(j)}(\boldsymbol{\theta}^{(j)})$ comprises L-layers of quantum circuits. Each layer consists of n parameterized R-rotations (with one on each qubit), and n param-

*zhoug@uestc.edu.cn
†xiaoting@uestc.edu.cn

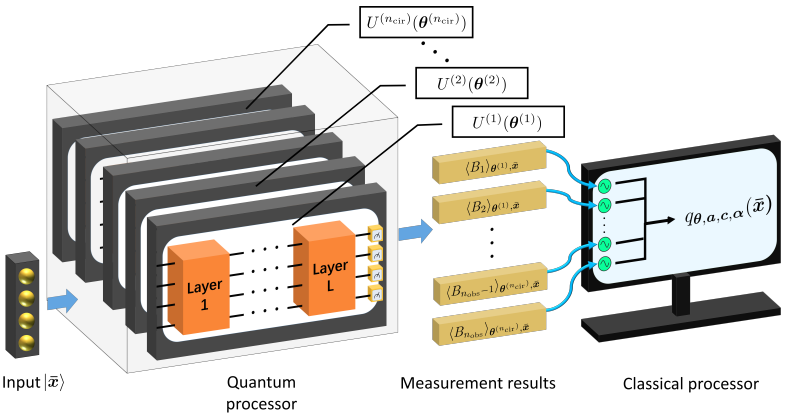


Figure 1: Framework of DQNN and the transformation from $|\bar{\boldsymbol{x}}\rangle$ to $q_{\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}})$. Yellow balls represent the quantum qubits; green circles represent the classical sigmoid function.

eterized controlled-R gates, as shown in Fig. 2, with

$$R = R(\theta_1, \theta_2, \theta_3) = \begin{pmatrix} e^{i\theta_2}\cos\theta_1 & e^{i\theta_3}\sin\theta_1 \\ -e^{-i\theta_3}\sin\theta_1 & e^{-i\theta_2}\cos\theta_1 \end{pmatrix}.$$

Next, based on $\langle B_i\rangle_{\boldsymbol{\theta}^{(j)},\bar{\boldsymbol{x}}}$ and the sigmoid function $\sigma(x) = 1/(1+e^{-x})$, the classical processor computes and obtains the output:

$$q_{\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}}) \equiv \sum_{j=1}^{n_{\text{cir}}}\sum_{i=1}^{n_{\text{obs}}} \alpha_i^{(j)}\sigma(a_i^{(j)}(\langle B_i\rangle_{\boldsymbol{\theta}^{(j)},\bar{\boldsymbol{x}}} - c_i^{(j)})) \quad (2)$$

with $a_i^{(j)} > 0$, $c_i^{(j)} \in [0,1]$ where $\boldsymbol{\theta}$ and $(\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha})$ are parameters to be trained. The entire process from $|\bar{\boldsymbol{x}}\rangle$ to $q_{\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}})$ is summarized in Fig. 1. Finally, one can find the optimal values of $(\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha})$ to solve the given learning problem through gradient-based optimization using algorithms.

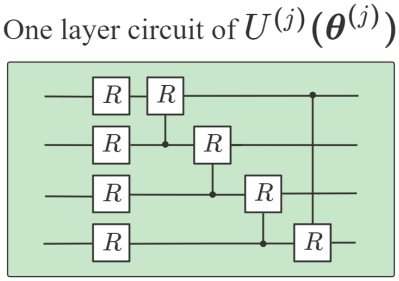


Figure 2: One layer variational quantum circuit of $U^{(j)}(\boldsymbol{\theta}^{(j)})$

3 Universality of DQNN

Universality of neural networks refers to the ability of to approximate any arbitrary function. For our DQNN, its universality can be proved to be universal based on L^2 approximation, which is summarized in the following theorem:

Theorem 1. *1 Let $\bar{G}$ be a subset of the complex sphere $\mathbb{S}$ in $\mathbb{C}_2^{\otimes n}$, and $f(\bar{\boldsymbol{x}}) : \bar{G} \to \mathbb{R}$ be an arbitrary square-integrable function on $\bar{G}$. We define the following parameterized functions:*

$$q_{\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{cir}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}}) \equiv \sum_{j=1}^{n_{cir}} \alpha_j\sigma(a_j(|\langle\bar{\boldsymbol{x}}|\boldsymbol{z}_j\rangle|^2 - c_j)) \quad (3)$$

and denote $Q(\bar{G}) \equiv \{q_{\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{cir}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}\}$ as the set of all such functions, where $n_{cir} \in \mathbb{N}$, $\{|\boldsymbol{z}_j\rangle\}_{j=1}^{n_{cir}} \subset \mathbb{S}$, $\boldsymbol{a} \in \mathbb{R}_+^{n_{cir}}$, $\boldsymbol{c} \in [0,1]^{\otimes n_{cir}}$, and $\boldsymbol{\alpha} \in \mathbb{R}^{n_{cir}}$. Then $Q(\bar{G})$ is dense in $L^2(\bar{G})$ in the following sense: for any $\epsilon > 0$,

$$\int_{\bar{G}} |q_{\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{cir}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}}) - f(\bar{\boldsymbol{x}})|^2 d\mu < \epsilon, \quad (4)$$

where μ is some measure defined on $\bar{G}$.

It turns out that any $q_{\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{\text{cir}}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}} \in Q(\bar{G})$ in Eq. (3) can be generated as the output of the DQNN in Eq. (2). Specifically, given the parameters $(\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{\text{cir}}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha})$, we can design a DQNN with a single observable $B \equiv |b\rangle\langle b|$, $n_{\text{obs}} = 1$, and a series of quantum circuits $\{U^{(j)}(\boldsymbol{\theta}^{(j)})\}_{j=1}^{n_{\text{cir}}}$ such that $U^{(j)}(\boldsymbol{\theta}^{(j)})|b\rangle = |z_j\rangle$. Then the output of the DQNN in Eq. (2) is reduced to:

$$\begin{aligned} q_{\boldsymbol{\theta},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}}) &= \sum_{j=1}^{n_{\text{cir}}} \alpha_j\sigma(a_j(\langle B\rangle_{\boldsymbol{\theta}^{(j)},\bar{\boldsymbol{x}}} - c_j)) \\ &= \sum_{j=1}^{n_{\text{cir}}} \alpha_j\sigma(a_j(|\langle\bar{\boldsymbol{x}}|\boldsymbol{z}_j\rangle|^2 - c_j)) \\ &= q_{\boldsymbol{z}_1,\cdots,\boldsymbol{z}_{n_{\text{cir}}},\boldsymbol{a},\boldsymbol{c},\boldsymbol{\alpha}}(\bar{\boldsymbol{x}}). \end{aligned} \quad (5)$$

Thus we have proved the following corollary:

Corollary 2. *The DQNN designed in Fig. 1 with the output in Eq. (2) is universal.*

4 Comparison among QNN models

To demonstrate that our DQNN has advantage in solving classical ML problems and in reducing circuit complexity, we apply DQNN and two well-known duplication based QNN models, circuit centric quantum

Table 1: A comparison of the performance among DQNN, QCL and CCQ in regression problem. n_{copy} indicates the number of duplications of the data qubits storing the classical data. n_{tot} indicates the number of qubits in each QNN model. C is the complexity of each QNN model.

Model	n_{copy}	n_{tol}	n_{obs}	C	Relative error
DQNN_1	1	2	4	48	6.79%
DQNN_4	1	2	1	48	5.46%
QCL	1	2	1	48	8.21%
CCQ	1	2	1	51	12.92%
QCL	2	4	1	120	7.35%
CCQ	2	4	1	99	4.74%

classifier (CCQ) [26] and the quantum circuit learning (QCL) [19], to two typical supervised learning problems, a regression problem and a classification problem, and compare their performance. To quantify the complexity of QNN, we define it as $C \equiv n_{\text{gate}} n_{\text{obs}}$ where n_{gate} is the number of quantum gates in the variational quantum circuit and n_{obs} is the number of observables.

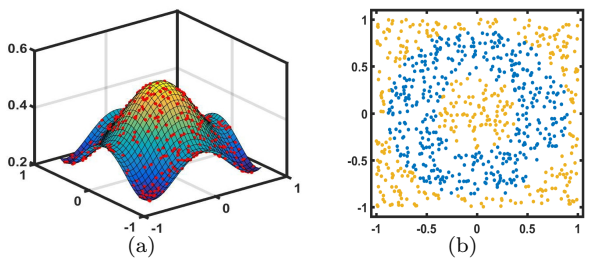


Figure 3: **(a)** The regression data set generated by a polynomial function. **(b)** The classification data set with the ring-shaped boundaries. The points in the yellow part are labeled 0; the others are labeled 1.

The first problem is learning from a data set $\{\vec{x}^{(i)}, y^{(i)}\}$ to approximate a highly nonlinear polynomial $f(\vec{x}) = (0.7156 - 1.0125x_1^2 + x_1^4)(0.7156 - 1.0125x_2^2 + x_2^4)$ with $x_1, x_2 \in [-0.8, 0.8]^2$ (Fig. 3(a)). We compare the optimal relative errors achieved for each model. We choose two types of DQNN in solving the regression problem. One has a single-layer variational quantum circuit and multiple observables. We name it as DQNN_1. The other has 4 single-layer variational quantum circuits and one observable. We name it as DQNN_4. We find that, with no duplicate, $n_{\text{copy}} = 1$, the relative error achieved by DQNN is substantially lower than those achieved by QCL and CCQ, as shown in Table 1. If we add one more duplicate to both the QCL and the CCQ models, their optimal relative error will decrease, but their complexity will increase and surpass that of the DQNN.

We apply the three models to a binary classification problem on a ring-shaped data set in the second task. The boundaries of the two sub-datasets are determined by six curves $x_1^2 + x_2^2 = 0.16$, $x_1^2 + x_2^2 = 0.81$, and $x_1 = x_2 = \pm 1$ (Fig. 3(b)). Analogous to the first task, we choose the DQNN with one variational quantum circuit and multiple observables (DQNN_1) and the DQNN with 4 variational quantum circuits and one observable (DQNN_4) for the classification task. After optimizing the three QNN models, we find that accuracy achieved by the DQNN is substantially higher than those achieved by QCL and CCQ. In addition, increasing the number of duplicates does help QCL and CCQ to improve their optimal accuracy, but even a 5-copy model (with 12 qubits) for QCL or CCQ cannot perform equally well as a DQNN with only two qubits (Table 2). Simulation results suggest that DQNN can prominently reduce the circuit complexity compared to QCL and CCQ, but maintain a good performance in solving classification problems.

Table 2: A comparison of the performance among DQNN, QCL and CCQ in a classification problem defined on a ring-shaped dataset.

Model	n_{copy}	n_{tol}	# n_{obs}	C	Accuracy
DQNN_1	1	2	20	240	91.40%
DQNN_4	1	2	1	48	97.63%
QCL	2	4	2	240	74.18%
CCQ	2	4	1	243	75.20%
QCL	6	12	2	1440	76.93%
CCQ	6	12	1	723	80.63%

Since the DQNN can reduce the circuit complexity, we expect that it can achieve a better performance in the presence of noise, and hence is easier to be implemented on NISQ devices. We further apply above QNN models to solve these two learning problems under noisy environment. We consider two types of noise in this work: one is the coherent noise caused by the imprecision of the classical control on the parameter values in the QNN circuits; the other is the decoherence generated by interactions between the quantum register and its environment. Numerical results demonstrate that the DQNN can decrease the influence of noise accumulation in the training process and is more likely to be implemented on near-term devices. More details can be found in Ref. [13].

5 Conclusion

In this work, we present a duplication-free quantum neural network model and provide it with the universal guarantee to approximate any arbitrary continuous function using several variational quantum circuits, multiple measurement observables and the classical parameterized sigmoid function. We can enhance its expressibility by increasing the number of quantum circuits and the number of observables without requiring auxiliary qubits. This property for the number of qubits makes DQNN more likely to be implemented on NISQ devices. Without duplicates, DQNN significantly reduces the number of required qubits and decreases circuit complexity compared with two well-known QNN models and hence weaken the influence of the circuit noise. These results indicate that the DQNN is an efficient QNN model which can find the patterns hidden in the classical and quantum data sets.

Acknowledgements

The authors gratefully acknowledge the grant from the National Natural Science Foundation of China (Grant No. 92265208) and the National Key R&D Program of China (Grant No. 2018YFA0306703). We also thank Chu Guo, Bujiao Wu, Yusen Wu, Shaojun Wu, Yuhan Huang, Donghong Han, Yingli Yang and Yi Tian for helpful discussions.

References

[1] K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, and R. Wolf. Training deep quantum neural networks. *Nat. Commun.*, 11(1):1–6, 2020.

[2] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[3] Y. D. Cao, G. G. Guerreschi, and A. Aspuru-Guzik. Quantum neuron: an elementary building block for machine learning on quantum computers, 2017. arXiv e-prints, ArXiv:1711.11240.

[4] G. Carleo and M. Troyer. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(6325):602–606, 2017.

[5] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nat. Commun.*, 12(1):1791, 2021.

[6] I. Cong, S. Choi, and M. D. Lukin. Quantum convolutional neural networks. *Nat. Phys.*, 15(12):1273–1278, 2019.

[7] D.-L. Deng. Quantum enhanced convolutional neural networks for nisq computers. *Sci. China-Phys. Mech. Astron.*, 64(10):100331, 2021.

[8] D. L. Deng, X. P. Li, and S. Das Sarma. Quantum entanglement in neural network states. *Phys. Rev. X*, 7:021021, May 2017.

[9] X. Gao and L. M. Duan. Efficient representation of quantum many-body states with deep neural networks. *Nat. Commun.*, 8(1):1–6, 2017.

[10] J. A. Garrido Torres, V. Gharakhanyan, N. Artrith, T. H. Eegholm, and A. Urban. Augmenting zero-kelvin quantum mechanics with machine learning for the prediction of chemical reactions at high temperatures. *Nat. Commun.*, 12(1):1–9, 2021.

[11] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, 3:010313, Jan 2022.

[12] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2:359–366, 1989.

[13] X. Hou, G. Zhou, Q. Li, S. Jin, and X. Wang. A duplication-free quantum neural network for universal approximation. *Sci. China Phys. Mech.*, 66(7):270362, 2023.

[14] H.-L. Huang, X.-Y. Xu, C. Guo, G. Tian, S.-J. Wei, X. Sun, W.-S. Bao, and G.-L. Long. Near-term quantum computing techniques: Variational quantum algorithms, error mitigation, circuit compilation, benchmarking and classical simulation. *Sci. China-Phys. Mech. Astron.*, 66:260301, 2023.

[15] L. B. Kristensen, M. Degroote, P. Wittek, A. Aspuru-Guzik, and N. T. Zinner. An artificial spiking quantum neuron. *npj Quantum Inf.*, 7(1):1–7, 2021.

[16] M. Leshno, V. Y. Lin, A. Pinkus, and S. Schocken. Multilayer feedforward networks with a nonpolynomial activation function can approximate any function. *Neural Networks*, 6(6):861–867, 1993.

[17] J. Liu, K. H. Lim, K. L. Wood, W. Huang, C. Guo, and H.-L. Huang. Hybrid quantum-classical convolutional neural networks. *Sci. China-Phys. Mech. Astron.*, 64(9):290311, 2021.

[18] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.*, 9(1):4812, 2018.

[19] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Phys. Lett. A*, 98(3):032309, 2018.

[20] K. Nakaji, S. Uno, Y. Suzuki, R. Raymond, T. Onodera, T. Tanaka, H. Tezuka, N. Mitsuda, and N. Yamamoto. Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicators. *Phys. Rev. Res*, 4(2):023136, 2022.

[21] F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti, and L. Emsley. Chemical shifts in molecular solids by machine learning. *Nat. Commun.*, 9(1):1–10, 2018.

[22] A. Pérez-Salinas, A. Cervera-Lierta, E. Gil-Fuster, and J. I. Latorre. Data re-uploading for a universal quantum classifier. *Quantum*, 4:226, Feb. 2020.

[23] M. Plesch and Č. Brukner. Quantum-state preparation with universal gate decompositions. *Phys. Rev. A*, 83:032302, Mar 2011.

[24] P. Rebentrost, T. R. Bromley, C. Weedbrook, and S. Lloyd. Quantum hopfield neural network. *Phys. Rev. A*, 98:042308, Oct 2018.

[25] M. Schuld. Supervised quantum machine learning models are kernel methods. *arXiv:2101.11020*, 2021.

[26] M. Schuld, A. Bocharov, K. M. Svore, and N. Wiebe. Circuit-centric quantum classifiers. *Phys. Lett. A*, 101(3):032308, 2020.

[27] M. Schuld, I. Sinayskiy, and F. Petruccione. Simulating a perceptron on a quantum computer. *Phys. Lett. A*, 379(7):660–663, 2015.

[28] M. Schuld, R. Sweke, and J. J. Meyer. Effect of data encoding on the expressive power of variational quantum-machine-learning models. *Phys. Rev. A*, 103:032430, Mar 2021.

[29] K. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer. Unifying machine learning and quantum chemistry with a deep neural network for molecular wavefunctions. *Nat. Commun.*, 10(1):1–10, 2019.

[30] F. Tacchino, C. Macchiavello, D. Gerace, and D. Bajoni. An artificial neuron implemented on an actual quantum processor. *npj Quantum Inf.*, 5(1):1–8, 2019.

[31] S. Thanasilp, S. Wang, N. A. Nghiem, P. J. Coles, and M. Cerezo. Subtleties in the trainability of quantum machine learning models. *arXiv:2110.14753*, 2021.

[32] E. Torrontegui and J. J. García-Ripoll. Unitary quantum perceptron as efficient universal approximator. *Europhys. Lett.*, 125(3):30004, mar 2019.

[33] K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. Kim. Quantum generalisation of feedforward neural networks. *npj Quantum Inf.*, 3(1):1–8, 2017.

[34] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles. Noise-induced barren plateaus in variational quantum algorithms. *Nat. Commun.*, 12(1):6961, 2021.

[35] S. Wei, Y. Chen, Z. Zhou, and G. Long. A quantum convolutional neural network on nisq devices. *AAPPS Bulletin*, 32:1–11, 2022.

[36] S. Yan, H. Qi, and W. Cui. Nonlinear quantum neuron: A fundamental building block for quantum neural networks. *Phys. Lett. A*, 102(5):052421, 2020.

Simple and high-precision Hamiltonian simulation by compensating Trotter error with linear combination of unitary operations

Pei Zeng Jinzhao Sun Liang Jiang Qi Zhao

Abstract. Trotter and linear-combination-of-unitary (LCU) are two popular Hamiltonian simulation methods. We propose Hamiltonian simulation algorithms using LCU to compensate Trotter error, which enjoy both of their advantages. By adding few gates after the Kth-order Trotter formula, we realize a better time scaling than 2Kth-order Trotter. Our first algorithm exponentially improves the accuracy scaling of the Kth-order Trotter formula. In the second algorithm, we consider the detailed structure of Hamiltonians and construct LCU for Trotter errors with commutator scaling. Consequently, for lattice Hamiltonians, the algorithm enjoys almost linear system-size dependence and quadratically improves the accuracy of the Kth-order Trotter.

Keywords: Quantum simulation, Trotter, linear-combination-of-unitary operations, random sampling

The complete version of this work can be found in arXiv:2212.04566.

1 Introduction

Hamiltonian simulation, i.e., to simulate the real-time evolution $U(t) = e^{-iHt}$ of a physical Hamiltonian $H = \sum_l H_l$, is considered to be a natural and powerful application of quantum computing [1]. To pursue real-world applications of Hamiltonian simulation with near-term quantum devices, we need to design feasible algorithms requiring small space complexity (i.e., qubit number) and time complexity (i.e., circuit depth and gate number). The most natural Hamiltonian simulation method is based on Trotter formulas [2, 3, 4, 5, 6, 7, 8, 9], which approximate the real-time evolution operator $U(t)$ by the product of the evolution of the summands $e^{-iH_l t}$. Besides its prominant advantage of simple realization without ancillas, Trotter methods are recently shown to enjoy commutator scaling [8, 9], i.e., the Trotter error is only related to the number of commutators related to the Hamiltonian summands $\{H_l\}$. This is very helpful for the Hamiltonians with strong locality constraints. Hereafter, we refer to the error analysis utilizing the detailed structure of Hamiltonians as commutator scaling analysis. For example, when we consider n-qubit lattice Hamiltonians, the gate cost of high-order Trotter methods is almost linear to the system size n, which is nearly optimal [8]. The major drawback of Trotter methods is its polynomial gate cost to the inversed accuracy $1/\varepsilon$, Poly$(1/\varepsilon)$. In many applications where the high-precision simulation is demanded to obtain practical advantages over the existing classical algorithms [10], the gate cost of Trotter methods is large.

In recent years, we have seen developments of "post-Trotter" algorithms with exponentially improved accuracy dependence [11, 12, 13, 14]. These advanced algorithms, however, requires the implementation of linear-combination-of-unitary (LCU) formulas [15] or block encoding of Hamiltonians [13] which often costs many ancillary qubits and multi-controlled Toffoli gates. This is unfavorable in a near-term or early fault-tolerant quantum computer [16]. On the other hand, unlike Trotter methods, these "post-Trotter" methods are not able to enjoy the specific structure of Hamiltonians without commutator scaling analysis. Consequently, for instance, for n-qubit lattice Hamiltonians, their gate complexities are $\mathcal{O}(n^2)$, which is worse than those in Trotter algorithms $\mathcal{O}(n^{1+o(1)})$.

In this work, we develop the theory of composite algorithms that combine the inherent advantages of Trotter and LCU methods — easy implementation, high precision, and commutator scaling — by performing the Trotter method and then compensating the Trotter error with the LCU methods. We propose two types of composite algorithms with unique advantages in various regions. Our first algorithm is generic, with the random-sampling implementation, we prove that with **only one ancillary qubit**, one can achieve a gate cost of $\mathcal{O}((\lambda t)^{1+o(1)}(L + \log(1/\varepsilon)))$ with $\lambda = \sum_l \|H_l\|$, which is almost linear to the time t and logarithmically dependent on the inversed accuracy $1/\varepsilon$. In our second algorithm, we consider the detailed structure of Hamiltonians and propose a modified composite algorithms with commutator scaling utilizing nested-commutator analysis in the Trotter algorithms. Consequently, for the lattice Hamiltonians, our algorithm enjoys almost linear time and system-size dependence, and achieves a higher accuracy than $2K$th-order Trotter only with Kth-order Trotter's gate complexity.

2 Trotter-LCU algorithms

The algorithm is based on a series connection of Trotter and LCU algorithms. We decompose the time evolution $U(t) = e^{-iHt}$ to ν segments, each with a small evolution time $x = t/\nu$. In each segment, we first perform the Kth-order Trotter formula $S_K(x)$ ($K = 0, 1, 2k, k \in \mathbb{N}_+$). For consistency, we denote the 0th-order Trotter formula as $S_0(x) = I$. Then, we compensate the multiplicative Trotter remainder $V_K(x) := U(x)S_K(x)^\dagger$ by implementing the LCU formula of $V_K(x)$. A (μ, ε)-LCU formula of an operator V is defined to be a set of probabilities and unitaries $\{p_i, V_i\}$ with $\tilde{V} = \mu \sum_{i=0}^{\Gamma-1} p_i V_i$, such that the

spectral norm distance $\|V-\tilde{V}\| \leq \varepsilon$. We call $\mu > 0$ the 1-norm of this LCU formula.

We consider two ways to implement the LCU formula. In the random-sampling implementation [17, 18], we sample the elementary unitaries $\{V_i\}$ based on the corresponding probability $\{p_i\}$ and embed the realization of the LCU formula into a Hadamard test [19]. In the coherent implementation [11, 12], we introduce $m = \lceil \log_2 \Gamma \rceil$ ancillary qubits to realize the superposition of different $\{V_i\}$. Our major focus is on the random-sampling implementation, where we can estimate the properties of the target state $U(t)|\psi_0\rangle$ with only one ancillary qubit. Given a (μ, ε)-LCU formula, we can estimate arbitrary observable with ε accuracy using $\mathcal{O}(\mu^4/\varepsilon^2)$ sampling resource which owns an extra μ^4 overhead compared to the normal Hamiltonian simulation algorithms.

In the naive LCU-construction based on Taylor-series expansion of $U(t)$ [12], μ grows exponentially. However, in our composite algorithms, we only randomly implement the remainder terms in Kth-order Trotter errors, which makes μ a small constant. For a small time slice x, we can suppress μ of each segment to the order of $1 + O((\lambda x)^{K+1})$. Then we further suppress μ by pairing the terms by constructing Pauli rotation unitaries [18, 20] due to the fact $I + yP = \sqrt{1+y^2} e^{i\theta P}$ with $\theta = \tan^{-1}(x)$, which could double the x-order, $\mu = 1 + O((\lambda x)^{2K+2})$. See Theorem 1 and Section VI in Technical manuscript for more details.

Theorem 1 (Informal) In a Kth-order Paired Taylor-series compensation (PTSc) algorithm ($K = 0$ or $2k$), the single-round gate complexity is $\mathcal{O}\left((\lambda t)^{1+\frac{1}{2K+1}}(\kappa_K L + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)})\right)$, where $\lambda = \sum_l \|H_l\|$, $\kappa_K = 0$ when $K = 0$, $\kappa_K = 2 \cdot 5^{K/2-1}$ otherwise.

Note that our PTSc algorithms are generic and applicable for any Hamiltonian. When we consider the detailed structure of Hamiltonians and Trotter errors, we could make the compensation algorithms more efficient. We propose a Kth-order Nested-commutator compensation (NCc) algorithm that gives the detailed forms of Trotter error terms from $K+1$ order to $2K$ order. A key difference between these two types of algorithms is that in PTSc algorithms, we compensate the terms in the remainder up to arbitrary order. While in NCc algorithms, we only compensate Trotter error terms from $(K+1)$th order to $(2K+1)$th order, which shrinks the error from $\mathcal{O}(x^{K+1})$ to $\mathcal{O}(x^{2K+2})$ in one slice with the sampling cost $\mu = 1 + O(\kappa_K \alpha_c x^{2K+2})$ where α_c is some summation of 1-norm of nested commutators. We do not choose to compensate the higher-order terms because of the super-exponential growth of the commutator α_c. The gate complexity estimation problem is converted to the calculation of α_c. For instance, n-qubit lattice Hamiltonians, $\alpha_c = \mathcal{O}(n)$ provides a simple and accurate implementation of NCc algorithms. See Section VII in Technical manuscript and Theorem 2 for general formulas and more details.

Theorem 2 (Informal) In a Kth-order nested-commutator compensation (NCc) algorithm ($K = 1$ or $2k$) with n-qubit lattice Hamiltonians, the single-round gate complexity is $\max\{\mathcal{O}(n^{1+\frac{1}{K}} t^{1+\frac{1}{K}}), \mathcal{O}(n^{1+\frac{2}{2K+1}} t^{1+\frac{1}{2K+1}} \varepsilon^{-\frac{1}{2K+1}})\}$.

In Table 1, we summarize the advantages and disadvantages of 0th-order PTSc, Kth-order PTSc, Kth-order NCc algorithms with random sampling implementation of LCU. The 0th-order PTSc algorithm is L-independent (the number of terms in Hamiltonians), which is especially useful for many quantum chemistry problem with many terms; the Kth-order PTSc algorithm owns the best time dependence of $\mathcal{O}(t^{1+\frac{1}{2K+1}})$ over all random-sampling based algorithms, which is suitable for a long-time high-precision simulation; the Kth-order NC algorithm owns almost linear n and t dependence for the lattice models.

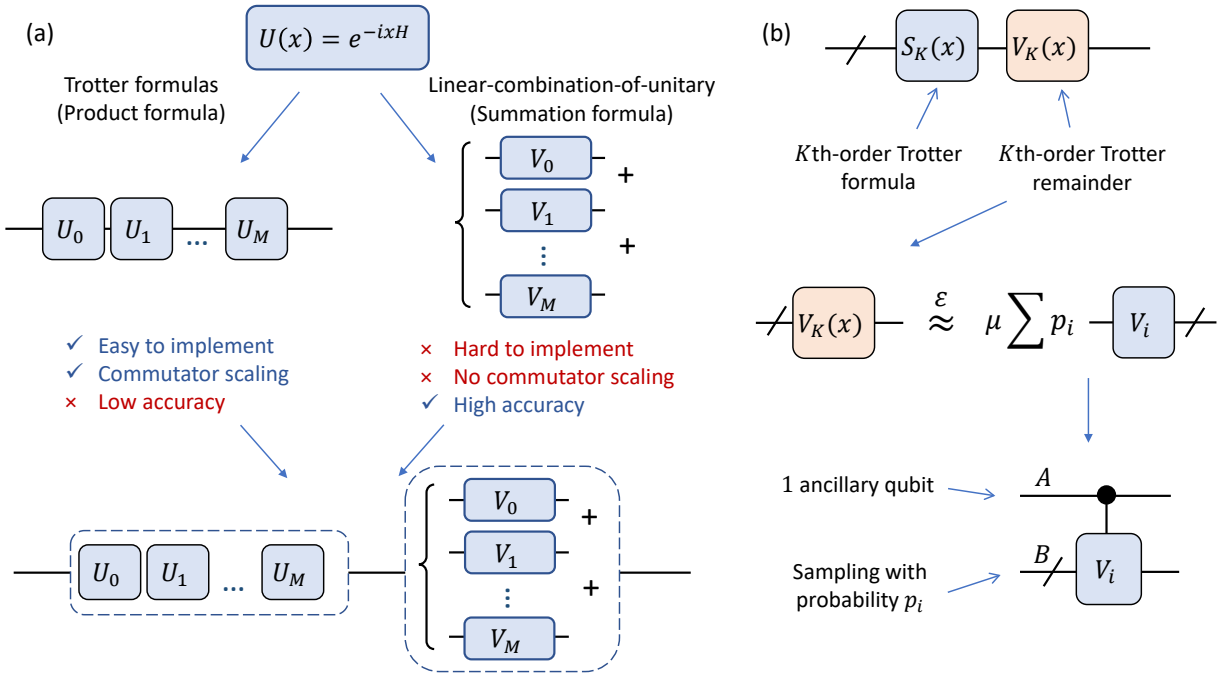
3 Numerical results

We estimate the gate cost and compare our results of PTSc algorithms with the advanced methods, including Trotter algorithms and quantum signal processing (QSP). We choose the 4th-order Trotter formula since it performs the best in all Trotter methods. We first consider the simulation of generic L-sparse Hamiltonian without the usage of commutator information, for which we choose the 2-local Hamiltonian with power-law decay interactions, $H = \sum_{i,j} J_{ij} X_i X_{i+1} + \sum_i Z_i$ with $J_{ij} = |i-j|^{-3}$, as an example. Figs. 2(a,b) show the CNOT and T gate counts for the 4th-order Trotter formula, QSP, and our results with different orders with an increasing system size. The CNOT gate counts of our 0th- and 2nd-order PTSc algorithms and the T gate counts of our 2nd and 4th PTSc algorithms are better than the previous best QSP algorithm, while they only require 1 ancillary qubit without complicated quantum circuits of block encoding.

Next, we compare the CNOT gate count for simulating the Heisenberg Hamiltonian $H = J\sum_i \vec{\sigma}_i \vec{\sigma}_{i+1} + h\sum_i Z_i$, where $\vec{\sigma}_i := (X_i, Y_i, Z_i)$ is the vector of Pauli operators on the ith qubit, using the nested commutator bounds. In Fig. 2(c), we choose $n = t = 12$ and 50 and show the gate number with respect to the accuracy requirement ε. While enjoying near-optimal system-size scaling similar to the 4th-order Trotter algorithm which is currently the best one for lattice Hamiltonians [9, 8], our 2nd-order NCc algorithm shows better accuracy dependence than 4th-order Trotter algorithm. Especially, using the same gate number as the 4th-order Trotter, we are able to achieve a 3 to 4 orders of magnitudes higher accuracy ε.

Why AQIS? We set up a unified framework for the Trotter-LCU algorithms with a series connection of Trotter and LCU algorithms. These composite algorithms are able to enjoy the advantages of Trotter and LCU algorithms simultaneously, providing simple and high-accuracy implementations for Hamiltonian simulation. Our developed framework represents a significant advance in designing quantum algorithms and can inspire other composite or compensation quantum algorithms by leveraging the advantages of the current advanced algorithms.

Figure 1: (a) By a series connection of Trotter and LCU formulas, we aims to enjoy the advantages of both algorithms. (b) Random-sampling implementation of Trotter-LCU algorithms. In each segment, we first perform Kth-order Trotter formula, and then randomly sample the elementary unitaries V_i based on the LCU formula.



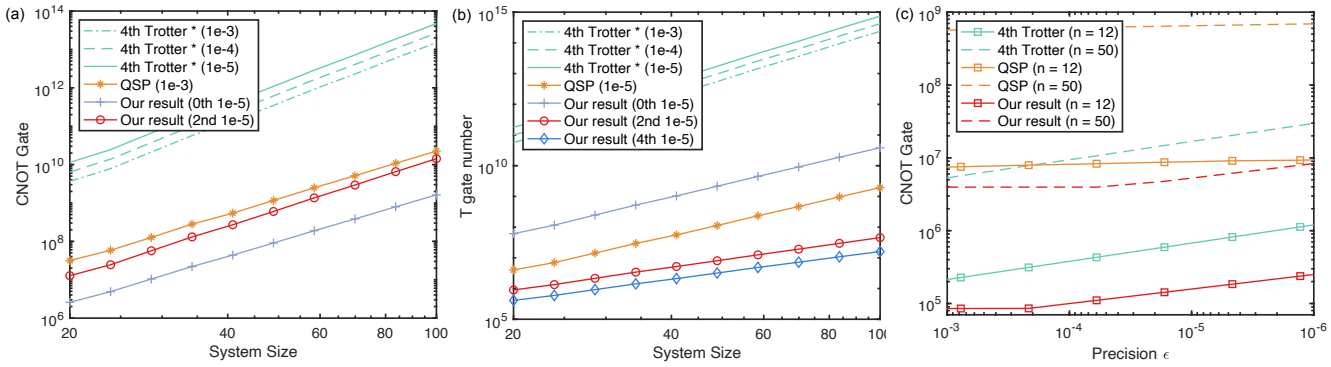


Figure 2: CNOT gate count for simulating real-time dynamics with an increasing system size n, and $t = n$. (a,b) The CNOT gate count (a) and T gate count (b) for 2-local Hamiltonian with power law decay interactions. (c) The results for Heisenberg model using the nested commutator bound with $n = 12$ and $n = 50$.

Algorithm	Implementation hardness	Simulation accuracy	Commutator scaling
Kth-order Trotter [3]	Easy	Low, $\mathcal{O}(\varepsilon^{-1/K})$	Yes, $\mathcal{O}(n)$
Post-Trotter [12, 13]	Hard	High, $\mathcal{O}(\widetilde{\log}(1/\varepsilon))$	No, $\mathcal{O}(n^2)$
0th-order PTSc	Easy	High, $\mathcal{O}(\widetilde{\log}(1/\varepsilon))$	No, $\mathcal{O}(n^2)$
Kth-order PTSc	Easy	High, $\mathcal{O}(\widetilde{\log}(1/\varepsilon))$	No, $\mathcal{O}(n^2)$
Kth-order NCc	Easy	Medium, $\mathcal{O}(\varepsilon^{-1/(2K+1)})$	Yes, $\mathcal{O}(n)$

Table 1: Compare Trotter-LCU methods with former algorithms. The implementation hardness refers to whether one needs to implement multi-controlled gates with plenty of ancillary qubits. The algorithms with commutator scaling can use advanced analysis to reduce the gate complexity of n-qubit lattice Hamiltonians from $\mathcal{O}(n^2)$ to $\mathcal{O}(n)$. More details of the comparisons and gate complexities are in Technical Manuscript.

References

[1] Richard P Feynman et al. Simulating physics with computers. *Int. j. Theor. phys*, 21(6/7), 1982.

[2] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[3] Masuo Suzuki. Fractal decomposition of exponential operators with applications to many-body theories and monte carlo simulations. *Physics Letters A*, 146(6):319–323, 1990.

[4] Masuo Suzuki. General theory of fractal path integrals with applications to many-body theories and statistical physics. *Journal of Mathematical Physics*, 32(2):400–407, 1991.

[5] Dominic W Berry, Graeme Ahokas, Richard Cleve, and Barry C Sanders. Efficient quantum algorithms for simulating sparse hamiltonians. *Communications in Mathematical Physics*, 270(2):359–371, 2007.

[6] Earl Campbell. Random compiler for fast hamiltonian simulation. *Phys. Rev. Lett.*, 123:070503, Aug 2019.

[7] Andrew M. Childs, Aaron Ostrander, and Yuan Su. Faster quantum simulation by randomization. *Quantum*, 3:182, sep 2019.

[8] Andrew M. Childs and Yuan Su. Nearly optimal lattice simulation by product formulas. *Phys. Rev. Lett.*, 123:050503, Aug 2019.

[9] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of trotter error with commutator scaling. *Phys. Rev. X*, 11:011020, Feb 2021.

[10] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the National Academy of Sciences*, 114(29):7555–7560, 2017.

[11] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Exponential improvement in precision for simulating sparse hamiltonians. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*. ACM, may 2014.

[12] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating hamiltonian dynamics with a truncated taylor series. *Phys. Rev. Lett.*, 114:090502, Mar 2015.

[13] Guang Hao Low and Isaac L. Chuang. Hamiltonian Simulation by Qubitization. *Quantum*, 3:163, July 2019.

[14] Guang Hao Low and Isaac L. Chuang. Optimal hamiltonian simulation by quantum signal processing. *Phys. Rev. Lett.*, 118:010501, Jan 2017.

[15] Andrew M Childs and Nathan Wiebe. Hamiltonian simulation using linear combinations of unitary operations. *Quantum Information and Computation*, 12(11-12), Nov 2012.

[16] Lin Lin and Yu Tong. Heisenberg-limited ground state energy estimation for early fault-tolerant quantum computers, 2021.

[17] Paul K. Faehrmann, Mark Steudtner, Richard Kueng, Maria Kieferova, and Jens Eisert. Randomizing multi-product formulas for improved hamiltonian simulation, 2021.

[18] Kianna Wan, Mario Berta, and Earl T. Campbell. Randomized quantum algorithm for statistical phase estimation. *Phys. Rev. Lett.*, 129:030503, Jul 2022.

[19] A. Yu. Kitaev. Quantum measurements and the abelian stabilizer problem, 1995.

[20] Yongdan Yang, Bing-Nan Lu, and Ying Li. Accelerated quantum monte carlo with mitigated error on noisy quantum computer. *PRX Quantum*, 2:040361, Dec 2021.

Experimental quantum state transfer of an arbitrary single-qubit state on a cycle with four vertices using a coined quantum random walk

Gayatri Singh[1] * Kavita Dorai[1] † Arvind[1 2] ‡

[1] *Department of Physical Sciences, Indian Institute of Science Education & Research Mohali, Sector 81 SAS Nagar, Manauli PO 140306 Punjab India*

[2] *Vice Chancellor, Punjabi University Patiala, 147002, Punjab, India*

Abstract. We experimentally demonstrate transfer [1] of an unknown single-qubit state from Alice to Bob using two-step discrete-time quantum random walk (QRW) on a cycle with four vertices on NMR QIP. The QRW carried out in two-qubit gaming arena involving Alice and Bob, each having their own coin qubits. In this scheme, required entangled state is naturally generated through conditional shift operators during the quantum-walk, eliminating need for prior preparation. By incorporating controlled-unitary operations based on measurements of Alice's coin qubit and Arena qubits, we reconstructed Alice's randomly generated state at Bob's end. Our results highlight the efficacy of QRW and high-fidelity state transfer.

Keywords: Quantum random walk (QRW), Quantum state transfer, Quantum teleportation, Quantum state and process tomography

1 Introduction

Quantum analogs of classical random walks are a versatile tool to perform quantum information processing tasks such as universal quantum computing, quantum search and quantum simulation etc. Standard quantum teleportation schemes [2] requires a prior entangled state shared between the two parties. However, conditional shift operator in quantum random walk can introduce entanglement between position space and coin space during steps of walk. Using this entanglement resource as quantum channel to perform teleportation, a new scheme was developed [3].

Since in the NMR scenario, the state is teleported to another location within the same molecule, we construed this scheme as not strictly being teleportation but as being akin to state transfer. We hence recast the entire theoretical scheme in terms of achieving quantum state transfer instead of quantum teleportation. In this work, we experimentally demonstrate the transfer of an unknown single-qubit quantum state between two parties (Alice and Bob), via a quantum random walk on a cycle with four vertices, on a four-qubit NMR quantum information processor.

2 Basic Theoretical Framework

Coined Quantum Random Walk on Closed Cycle: At each step of the quantum walk, the coin is flipped, and then move coherently depending upon the state of the coin. The unitary operator for one step of walk is defined as $U = S(I \otimes C)$ [4], where C is the coin flipping operator, I is the identity operator (do nothing) and S is the conditional shift operator, defined on cycle with n-vertices as:

$$\begin{aligned} S_{\pm} &= |i \pm 1 \bmod n\rangle\langle i| \\ \hat{S} &= \sum_{i=0}^{n-1} (S_+ \otimes |0\rangle_c\langle 0| + S_- \otimes |1\rangle_c\langle 1|) \end{aligned} \quad (1)$$

In this formulation, if the coin is in the state $|0\rangle$, the walker moves in the anticlockwise direction and if the coin is in the state $|1\rangle$, the walker moves in the clockwise direction.

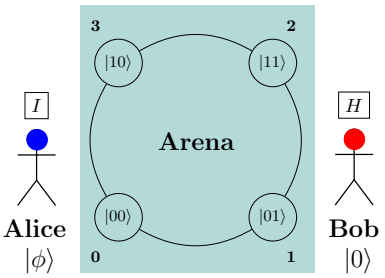


Figure 1: The two-player 'Game Arena' is depicted by a closed cycle with four vertices representing state $|00\rangle, |01\rangle, |10\rangle$ and $|11\rangle$. During the game, Alice employs the identity operator I as her coin operator in first step, while Bob utilizes the Hadamard gate H as his coin operator during second step. If the coin qubit is in $|0\rangle$ state, walker moves to the right and if it's in $|1\rangle$ state, walker moves to the left.

Transferring a single qubit state on a cycle with four vertices: As mentioned in the introduction, the protocol is reformulated to transfer the coin state from Alice to Bob instead of teleportation. The schematic diagram for QRW on a closed cycle with four vertices, involving two coins is depicted in Figure 1.

Initial state of Alice's coin-Arena-Bob's coin system to transfer state $|\phi\rangle$ from Alice to Bob is:

$$|\psi\rangle_{\text{in}} = |\phi\rangle \otimes |00\rangle \otimes |0\rangle \quad (2)$$

*ph20015@iisermohali.ac.in
†kavita@iisermohali.ac.in
‡arvind@iisermohali.ac.in

Where, $|\phi\rangle = a|0\rangle + b|1\rangle$ with $|a|^2 + |b|^2 = 1$. Following two steps of the walk, Alice's coin qubit is measured in $\{|+\rangle, |-\rangle\}$ basis, while 'Arena' qubits are measured in $\{|0\rangle, |1\rangle\}$ basis. We apply Hadamard gate H to Alice's coin qubit so that all the measurements can be performed in $\{|0\rangle, |1\rangle\}$ basis:

$$|\psi\rangle_f = \frac{1}{2}\{|000\rangle \otimes (a|1\rangle + b|0\rangle) + |100\rangle \otimes (a|1\rangle - b|0\rangle) + |011\rangle \otimes (a|0\rangle + b|1\rangle) + |111\rangle \otimes (a|0\rangle - b|1\rangle\} \quad (3)$$

Table 1: Measurement results on Alice's coin qubit,'Arena' qubits (A_c) and the corresponding controlled operations M on Bob's coin qubit (B_c); Z and X denote the Pauli matrices σ_z and σ_x respectively, while I is the Identity operator ('do nothing').

Measurement results on A_c	'Arena' qubits	Revised Operation M
0	11	I
1	11	Z
0	00	X
1	00	ZX

After two step of walk, Bob applies controlled unitary operations M (Table 1) based on measurement results, to recover the transferred state. Then the state becomes:

$$|\Psi\rangle = \{|000\rangle + |100\rangle + |011\rangle + |111\rangle\} \otimes |\phi\rangle \quad (4)$$

The quantum circuit to experimentally realize the transfer of an unknown single-qubit state using a two-step QRW on a cycle with four vertices is shown in Figure 2.

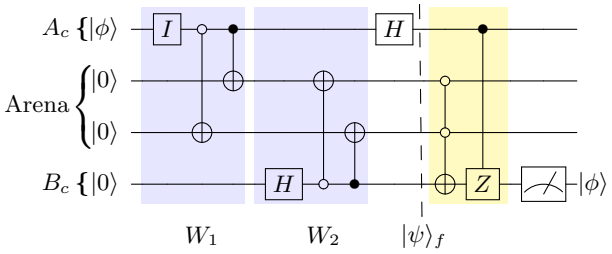


Figure 2: Quantum circuit to transfer an unknown single-qubit state $|\phi\rangle$ using a two-step quantum random walk on a cycle with four vertices. A_c and B_c denote the coin qubits of Alice and Bob, respectively, while the two middle qubits are labeled as the 'Arena' qubits. W_1 and W_2 denote the first and second steps of the quantum walk. The yellow shaded box represents the implementation of unitaries corresponding to the controlled operations.

3 Experimental Implementation and Experimental Results

We used four ^{13}C nuclei of Trans-crotonic acid dissolved in acetone-D6 as the four qubit system. All the experiments were performed at ambient temperature ($\approx$ 300K) on a Bruker DRX Avance III 600 MHz NMR spectrometer equipped with a standard 5 mm QXI probe. The methyl group and other proton spins were decoupled throughout the experiments. Pseudo-pure state [5] is prepared with total pulse sequence duration $\approx$ 41 ms with an experimental fidelity 0.9812 ± 0.0020.

Experimental quantum state transfer using revised measurements: We choose C_1 as Bob's coin, C_3 as Alice's coin and C_2,C_4 as 'Arena' qubits. After implementing the entire NMR pulse sequence, as illustrated in Figure 3, for state transfer, with a total time duration of 156 ms, the tomography process is solely performed on Bob's qubit to retrieve the transferred state. The experimentally reconstructed transferred state ρ_{C_1} for the input state $|\phi\rangle = |-\rangle = \frac{|0\rangle+\iota|1\rangle}{\sqrt{2}}$ is shown in Figure 4. The average experimental fidelities for the input states $|0\rangle, |1\rangle, |+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|-\rangle = \frac{|0\rangle+\iota|1\rangle}{\sqrt{2}}$ are $0.9918 \pm 0.0006, 0.9924 \pm 0.0012, 0.96 \pm 0.0027$ and 0.9896 ± 0.0007, respectively.

Efficacy and Characterization of Quantum State Transfer Scheme: We repeated the experimental protocol for a set of different input states $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$. The NMR pulse sequence after PPS preparation and including the box I (Figure 3) is applied with a total gate implementation time of $\approx$ 124 ms. By utilizing the normalized reduced density matrix $\sigma_{C_1}^{ijk}$ and employing the appropriate controlled operations M listed in Table 1, we reconstructed the transferred state at Bob's qubit, denoted as $\rho_{C_1}^{ijk}$ [6]:

$$\sigma_{C_1}^{ijk} = \frac{\mathrm{Tr}_{C_2C_3C_4}[P_{ijk}\rho_{\mathrm{expt}}P_{ijk}^\dagger]}{\mathrm{Tr}[P_{ijk}\rho_{\mathrm{expt}}]}$$
$$\rho_{C_1}^{ijk} = M.\sigma_{C_1}^{ijk}.M^\dagger \quad (5)$$

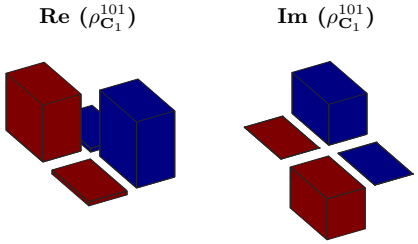


Figure 5: Real (left) and imaginary (right) parts of transferred state reconstructed at Bob's qubit $\rho_{C_1}^{101}$, with a fidelity of 0.9785 ± 0.002 for the input state $|\phi\rangle = |-\rangle$.

Where, $P_{ijk} = I \otimes |ijk\rangle\langle ijk|$ are projectors onto the basis states $|\ C_2C_3C_4\rangle = |ijk\rangle = \{|000\rangle, |010\rangle, |101\rangle, |111\rangle\}$. Figure 5 depicts the experimentally reconstructed transferred state $\rho_{C_1}^{101}$ at Bob's end for the input state $|\phi\rangle = |-\rangle$. We use a witness operator $W_{|\psi\rangle} = \frac{1}{2}I - |\psi\rangle\langle\psi|$, to certify the presence of genuine quadripartite entanglement in ρ_{expt}. For input state $|+\rangle$, $\mathrm{Tr}[\mathcal{W}_{|\psi\rangle}\rho_{\mathrm{expt}}] = -0.4358 \pm 0.0018 < 0$ and for input state $|-\rangle$, $\mathrm{Tr}[\mathcal{W}_{|\psi\rangle}\rho_{\mathrm{expt}}] = -0.4183 \pm 0.0028 < 0$, clearly

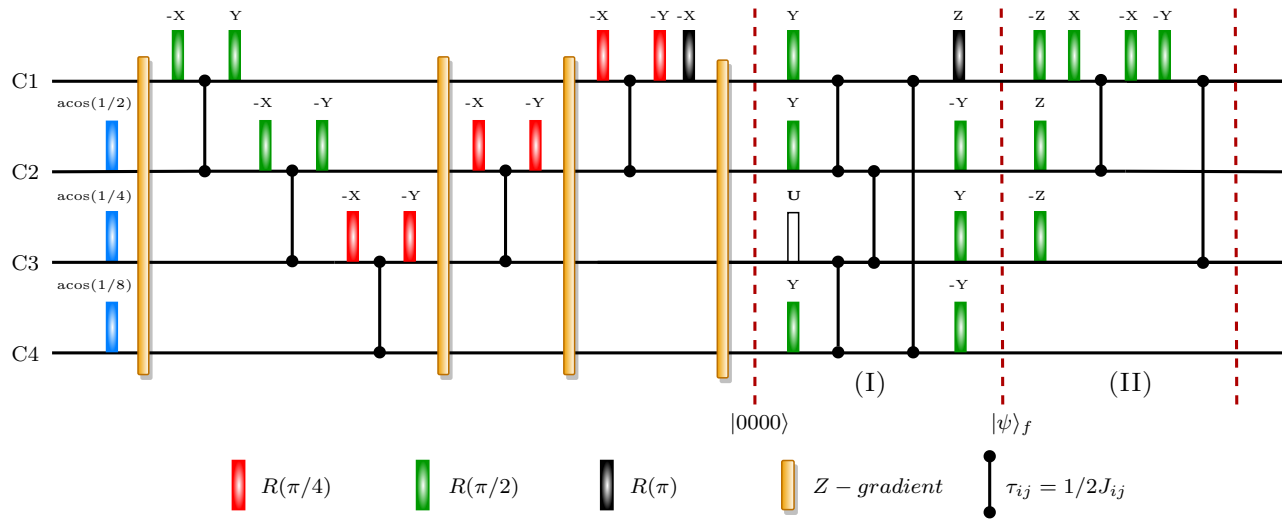


Figure 3: NMR pulse sequence to transfer of arbitrary state $|\phi\rangle$ using two-step QRW. The sequence of rf pulses, z-gradients and time evolution periods upto the first dashed line prepares the system in the $|0000\rangle$ PPS, starting from thermal equilibrium. The sequence given in the box labeled I implements the two-step quantum walk, while the box labeled II implements the unitaries corresponding to the controlled operations. $\tau_{ij} = 1/2J_{ij}$ corresponds to the free evolution period and the unitary U is used to prepare the state $|\phi\rangle$.

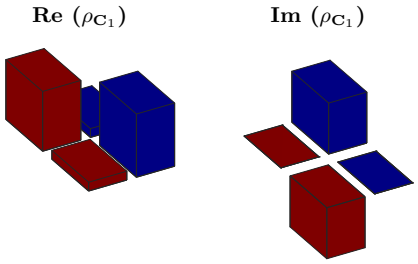


Figure 4: Real (left) and imaginary (right) parts of the experimentally tomographed transferred state at Bob's qubit ρ_{C_1} with a fidelity of 0.9896 ± 0.0007 for the input state $|\phi\rangle = |-\rangle$.

indicates that the state has quadripartite entanglement. We also performed constrained convex optimization [7] based quantum process tomography to fully characterize the state transfer protocol. The tomographed results for process matrix are depicted in Figure 6

The fidelities $\mathcal{F}^{ijk}$ of experimentally constructed process matrix χ^{ijk}_{expt} in Pauli operator basis $\{I, \sigma_x, \iota\sigma_y, \sigma_z\}$ with respect to the theoretical process matrix, for different projections are $\mathcal{F}^{101} = 0.9682 \pm 0.0021$, $\mathcal{F}^{111} = 0.9658 \pm 0.0015$, $\mathcal{F}^{000} = 0.9842 \pm 0.0023$, and $\mathcal{F}^{010} = 0.9450 \pm 0.0015$.

4 Conclusions

The scheme was able to achieve near perfect transfer of an arbitrary single-qubit state, with high state fidelity. Furthermore, the experimental circuit along with a few local unitaries can be used to generate four-qubit cluster states, which are of much interest in quantum information processing.

Our experimental schemes are general and can be easily extended to transferring states of a larger qubit register size via multi-coin, multi-walker setups. Our results demonstrate that coined quantum random walk schemes can be used to achieve robust transfer of an arbitrary quantum state and pave the way for wider applications in quantum communication and quantum information processing.

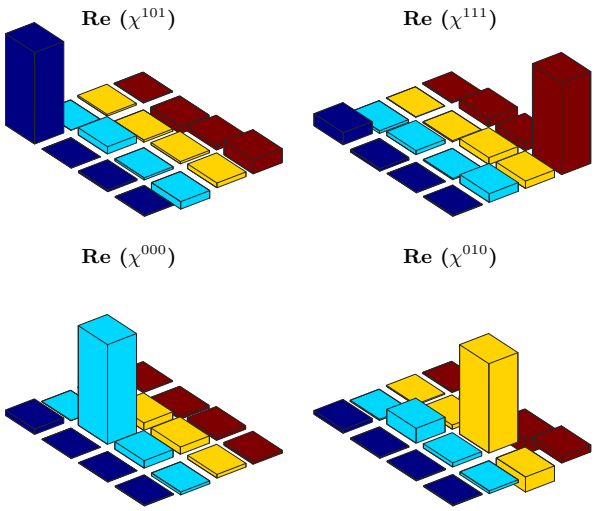


Figure 6: Real parts of experimentally tomographed process matrix χ for different projections of ρ_{expt} onto the basis state $|000\rangle, |010\rangle, |101\rangle$ and $|111\rangle$

References

[1] G. Singh, K. Dorai, Arvind. Experimental quantum state transfer of an arbitrary single-qubit state on a cycle with four vertices using a coined quantum random walk. *arXiv preprint*, arXiv:2305.02106, 2023.

[2] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Phys. Rev. Lett.*, 70:1895-1899, 1993.

[3] Y. Wang, Y. Shang, and P. Xue. Generalized teleportation by quantum walks. *Quant. Inf Proc.*, 16(9):221, 2017.

[4] N. Shenvi, J. Kempe, and K. B. Whaley. Quantum random-walk search algorithm. *Phys. Rev. A*, 67:052307, 2003.

[5] G. Bhole. *Coherent control for quantum information processing.* PhD thesis, Oxford University, 2020.

[6] M. Baur, A. Fedorov, L. Steffen, S. Filipp, M. P. da Silva, and A. Wallraff. Benchmarking a quantum teleportation protocol in superconducting circuits using tomography and an entanglement witness. *Phys. Rev. Lett.*, 108:040502, 2012.

[7] A. Gaikwad, Arvind, and K. Dorai. True experimental reconstruction of quantum states and processes via convex optimization *Quant. Inf Proc.* 20(1):19,2021

Quantum Enhanced Inference of Conditional Future Probabilities in Stochastic Processes

Jianjun Chen[1] * Chengran Yang[2] † Mile Gu[1 2] ‡

[1] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University*
[2] *Centre for Quantum Technologies, National University of Singapore*

Abstract. Due to the widespread prevalence of stochastic processes, accelerated probability estimation of events in stochastic processes can have far-reaching benefits. This speedup is particularly important for rare events, such as financial crashes or earthquakes, which have disastrous consequences. In this paper, we propose a new pipeline that builds on both the quantum models [1] and quantum amplitude estimation (QAE) protocols [2, 3] to improve the probability estimation of specific events in stochastic processes. This new pipeline is tested on diverse stochastic processes and consistently demonstrated up to a quadratic speedup in the error convergence compared to classical Monte Carlo methods.

Keywords: Quantum computing, stochastic processes, quantum amplitude estimation, quantum models

1 Introduction

Time series processes, prevalent across domains from finance to weather, contain patterns that, if understood and modelled accurately, can predict specific future events. The importance of this prediction increases with rare, impactful events such as financial crashes or earthquakes. The efficacy of any model hinges on memory cost and prediction speed. Lower memory usage is one area where quantum models have an advantage over classical models [1, 4, 5, 6]. Furthermore, using quantum models allows for leveraging quantum algorithms, such as the quantum amplitude estimation (QAE) protocol, which can accelerate predictions for these stochastic processes [2], especially for rare events where classical sampling methods prove inefficient. However, the original QAE protocol is unsuitable for near-term Noisy Intermediate Scale Quantum (NISQ) computers. Recent advancements have addressed this, making QAE implementation possible on NISQ-era quantum computers [3, 7, 8].

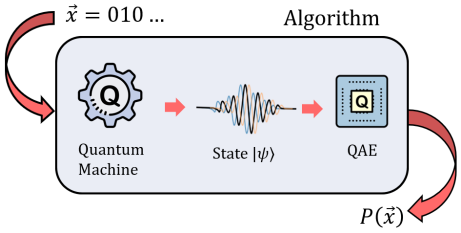


Figure 1: Illustration of the proposed algorithm. It takes in a future sequence $\vec{x}$ and outputs the estimated probability of this sequence at the end.

This paper introduces a new algorithm that merges quantum models with the QAE protocol to enhance probability estimation in stochastic processes, illustrated in Figure 1.

*chen1554@e.ntu.edu.sg
†yangchengran92@gmail.com
‡ceptryn@gmail.com

2 Framework

We focus on a random time series process, where at each time step, the process outputs an observable $x_t \in \mathcal{X}$, where $\mathcal{X}$ is the set of all observable. Observing this process over time will generate a sequence $\overleftrightarrow{X}$. This sequence can be divided into past and future components, and given the past, we want to build a model to predict the future. Classically, the model with the optimal memory efficiency is the ϵ-machine [9, 10]. However, its memory complexity is often higher than the theoretical lower bound [11]. It has been proven that given an ϵ-machine, we can build a quantum machine, also called q-machine, with a lower memory complexity [1, 4]. The q-machine comprises a memory and an output system, with transitions between quantum causal states $|S_i\rangle$ defined by a unitary operator U. Upon initialising the memory qubits into a given causal state, the unitary operator interacts with both the memory and an ancilla qubit at each time step, generating a sequence representing a potential future of the stochastic process. This process is illustrated in figure 2.

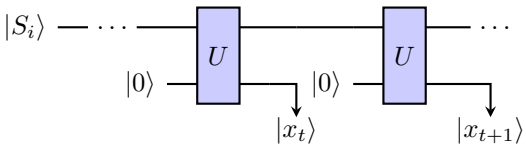


Figure 2: A q-machine comprises a quantum memory ($|S_i\rangle$) and ancillary system initialised in a computational basis. At each time step t, the unitary U operates on the memory register and an ancillary qubit. The resulting output ($|x_t\rangle$) can be immediately measured to create a sequence for the stochastic process defined by U or forwarded to a quantum circuit for additional processing.

This quantum machine can be combined with QAE, potentially providing up to a quadratic speedup for estimating any future sequence of interest compared to classical Monte Carlo sampling. The specific QAE algorithm used is the Maximum Likelihood Amplitude Es-

timation (MLAE) by Suzuki et al. [3]. This algorithm incorporates Grover iterators and classical processing, significantly reducing the use of controlled operations and thereby making it more relevant for NISQ computers.

In the MLAE algorithm, a *schedule* will define a set of experiments to conduct, with the kth experiment having m_k Grover iterators and N_k measurement shots, $k = 0, \ldots, M-1$, where M is the number of experiments. Suzuki et al. provided two schedules: Linear Increment Schedule (LIS = $\{m_k = k, N_k = constant\}$) and Exponential Increment Schedule (EIS= $\{m_k = 2^k, N_k = constant, k = 0, \ldots, M-1\}$). A maximum likelihood estimation can then be performed across the results obtained to get the probability of the future sequence of interest.

The standard error of the estimate, or ϵ, depends on the schedule used. The LIS and EIS schedule provide an error convergence of $O(N_q^{-0.75})$ and $O(N_q^{-1})$ respectively, as compared to the classical convergence of $O(N_q^{-0.5})$, where N_q is the number of oracle calls [3]. We also included a new r-polynomial increment schedule (PIS = $\{m_k = r^k, N_k = constant\}$, a subset of a schedule provided in [12]) which provide a different convergence speed of $O(N_q^{-(2r+1)/(2r+2)})$.

3 Algorithm

We developed a new algorithm to construct a specific Grover iterator based on a stochastic process defined by an arbitrary ϵ-machine. This algorithm converts classical machines into implementable quantum circuits for quantum-enhanced probability estimation. The Grover iterator is defined as:

$$\mathbf{Q} = -\mathcal{A}S_0\mathcal{A}^\dagger S_\chi \tag{1}$$

[2]. The algorithm will generate two components essential for the operator $\mathcal{A}$: the initialiser operator U_{ini} that rotates the memory register to the corresponding quantum causal state and the unitary operator based on the stochastic process. These two operators are assembled as in Figure 3 to create the operator $\mathcal{A}$.

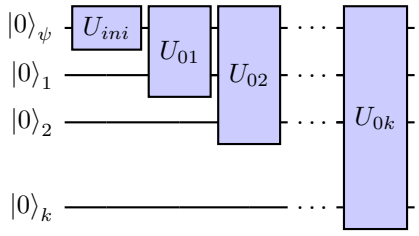


Figure 3: The quantum circuit that sets up the initial superposition. $|0\rangle_\psi$ is the quantum memory register and is normally initialised in the zero state at the start of the circuit. This whole set of gates also forms the algorithm operator $\mathcal{A}$

.

Since the oracle and phase shift operators are easily defined for a search scenario, this completes the components of the Grover iterator. This Grover iterator can then be used as a component in constructing the quantum circuit (Figure 4) for estimating the probability of any specific future sequence of interest. We also created a new schedule for this algorithm, named Inverse Decreasing Shots Schedule (IDSS = $\{m_k = k, N_k = N_{max}/(k+1)\}$), with an error convergence of $O(N_q^{-1})$, matching that of the EIS schedule.

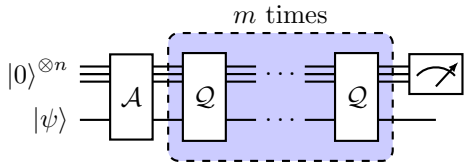


Figure 4: The quantum circuit of one instance of the proposed algorithm with m number of Grover iterators applied. C refers to the classical registers to store the results of the measurements.

4 Results

This new algorithm was applied to four different models of increasing complexity: the perturbed coin process, the dual Poisson process, the Nemo process and the autoregressive model. All processes considered have an output alphabet size $|\mathcal{X}| = 2$, i.e. $x_t \in \{0, 1\}$. We focused on sequences with a length of 4, i.e. $\overrightarrow{x} = x_0x_1x_2x_3$.

Since most of the processes show similar trends, we will elaborate only on the results for the autoregressive model. The autoregressive model is a simple model frequently used to represent time-series processes, including financial processes. The general autoregressive model of order p (AR(p)) is given as:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \epsilon_t \tag{2}$$

, where X_t is the current value of the series at time t, $\phi_i \in \mathbb{R}$ is the set of weights on the previous time steps, and ϵ_t is the white noise at time t [13]. This paper considered the case where the Markov order $p = 2$, giving the AR(2) model. The model is further discretised such that $X_t \in 0, 1$, resulting in the ϵ-machine shown in FIG. 5.

The algorithm is applied to estimate the future sequence of $\overrightarrow{x} = 1001$ from the initial causal state S_0. The parameters ϕ_1 and ϕ_2 are set to 0.5 and 0.4, respectively. With these initial conditions, the probability of the future sequence $P(\overrightarrow{x} = 1001|S_0) \approx 0.00842$. Referencing Figure 6a, it is evident that as the number of Grover iterators increases, the standard error of the probability estimate decreases. Moreover, the quantum estimation error declines at a faster rate compared to the classical estimation. Figure 6b clearly shows the power-law error trend of both the classical and quantum algorithms, with the quantum algorithm having a polynomial advantage compared to the classical case. A best-fit line of the

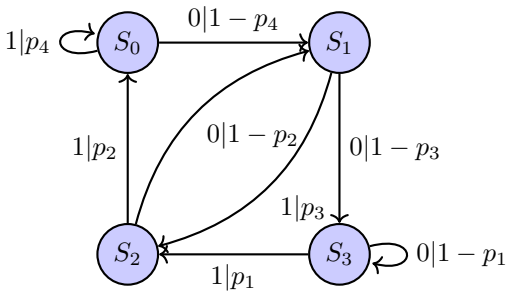


Figure 5: ϵ-machine of the discretised AR(2) process. The value of $p_i, i \in \{1,2,3,4\}$ are parameterised by ϕ_1 and ϕ_2.

following form is then considered:

$$\text{best-fit: } \epsilon = bN_q^a, \tag{3}$$
$$a, b \in \mathbb{R}, N_q = \text{ number of oracle calls}$$

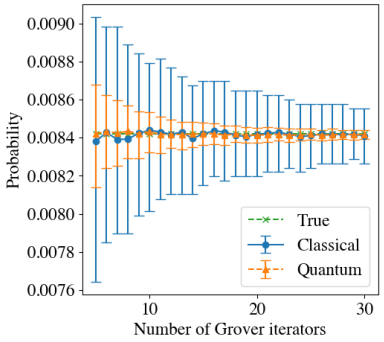
Table 1: Table of the exponents of the best-fit lines over different processes and schedules. $\overline{a}$ refers to the average exponent across all the processes. 2-PIS refers to the PIS with the polynomial degree set as 2.

Schedule	a_{min}	a_{max}	$\overline{a}$
Classical	-0.507	-0.491	-0.497
LIS	-0.759	-0.748	-0.752
EIS	-1.00	-0.875	-0.962
2-PIS	-0.887	-0.831	-0.846
IDSS	-0.931	-0.916	-0.923

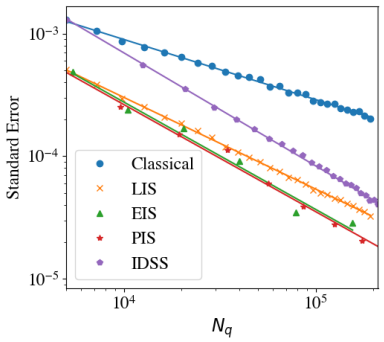
Referring to Table 1, the value of a is close to the lower bound set by the Fisher information, especially for the LIS schedule. The value of a also appears to be process-independent, varying primarily with the chosen schedule. The 2-PIS (PIS with polynomial degree two) schedule offers a speedup faster than the LIS schedule but slower than the EIS schedule. The new IDSS schedule exhibited an increased speedup compared to the PIS schedule, delivering a quantum speedup that is only slower than the exponential schedule (EIS).

5 Discussion

This paper presents a novel algorithm combining the quantum model creation [1] and probability estimation algorithm MLAE [3]. We demonstrated its effectiveness in estimating the probability of rare events in stochastic processes. Both Markovian and Non-Markovian stochastic models were tested, and the new algorithm consistently provided a speedup of up to a quadratic factor over classical methods, highlighting the potential benefits of quantum algorithms in probability estimation. Additionally, the algorithm's ability to convert ϵ-machines to q-machines makes it applicable to many stochastic processes. The quantum speedup in probability estimation has the potential to enhance accuracy, particularly for



(a) Convergence



(b) Standard error trend

Figure 6: (a) This plot shows the convergence of the mean probability estimate as the number of Grover iterators increases for the AR(2) process. The convergence of the classical algorithm is plotted for comparison. The error bar shows the 75th and 25th percentile of the estimates over 1000 samples. (b) This plot shows the error trend of the different schedules as the number of shots the algorithm uses for estimation increases.

rare events where classical Monte Carlo methods require large sample sizes. The algorithm offers advantages over importance sampling by not relying on detailed knowledge of the underlying distribution and only requiring knowledge of the classical transition matrix to exploit the quantum speedup.

The paper also introduces the new IDSS schedule, which reduces the number of shots for higher-depth circuits while maintaining close to a quadratic speedup over classical sampling. This schedule effectively reduces the computational cost, especially considering that higher-depth circuits are typically resource-intensive on quantum computers, moving us closer to achieving quantum advantage on real quantum hardware.

Future research directions include implementing the algorithm on actual quantum computers like the IBM Quantum systems to assess its performance under noisy conditions. As studies have shown that dimensionally-reduced quantum models outperform classical models [14], another exciting avenue is to explore the trade-off between model approximation and probability estimation accuracy.

References

[1] F. C. Binder, J. Thompson, and M. Gu, "Practical Unitary Simulator for Non-Markovian Complex Processes," *Physical Review Letters*, vol. 120, no. 24, 2018.

[2] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, "Quantum amplitude amplification and estimation," 2002.

[3] Y. Suzuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto, "Amplitude estimation without phase estimation," *Quantum Information Processing*, vol. 19, pp. 1–17, 1 2020.

[4] M. Gu, K. Wiesner, E. Rieper, and V. Vedral, "Quantum mechanics can reduce the complexity of classical models," *Nature Communications*, vol. 3, pp. 1–5, mar 2012.

[5] T. J. Elliott, C. Yang, F. C. Binder, A. J. Garner, J. Thompson, and M. Gu, "Extreme Dimensionality Reduction with Quantum Modeling," *Physical Review Letters*, vol. 125, no. 26, 2020.

[6] C. Aghamohammadi, S. P. Loomis, J. R. Mahoney, and J. P. Crutchfield, "Extreme Quantum Memory Advantage for Rare-Event Sampling," *Physical Review X*, vol. 8, no. 1, 2018.

[7] S. Aaronson and P. Rall, "Quantum Approximate Counting, Simplified," in *Symposium on Simplicity in Algorithms*, pp. 24–32, 2020.

[8] D. Grinko, J. Gacon, C. Zoufal, and S. Woerner, "Iterative quantum amplitude estimation," *npj Quantum Information*, vol. 7, pp. 1–6, mar 2021.

[9] C. R. Shalizi and J. P. Crutchfield, "Computational mechanics: Pattern and prediction, structure and simplicity," *Journal of Statistical Physics*, vol. 104, no. 3-4, pp. 817–879, 2001.

[10] J. P. Crutchfield and K. Young, "Inferring statistical complexity," *Physical Review Letters*, vol. 63, pp. 105–108, 7 1989.

[11] J. P. Crutchfield, C. J. Ellison, and J. R. Mahoney, "Time's barbed arrow: Irreversibility, crypticity, and stored information," *Phys. Rev. Lett.*, vol. 103, p. 094101, Aug 2009.

[12] T. Giurgica-Tiron, I. Kerenidis, F. Labib, A. Prakash, and W. Zeng, "Low depth algorithms for quantum amplitude estimation," *Quantum*, vol. 6, p. 745, June 2022.

[13] R. H. Shumway and D. S. Stoffer, *Time Series Analysis and Its Applications*. Springer Texts in Statistics, Springer New York, 2000.

[14] C. Yang, A. Garner, F. Liu, N. Tischler, J. Thompson, M.-H. Yung, M. Gu, and O. Dahlsten, "Provable superior accuracy in machine learned quantum models," 5 2021.

Quantum Chernoff Bound for Quantum Reading Using the Quasi-Bell State

Tiancheng Wang[1 2 *] Tsuyoshi Sasaki Usuda[2 †]

[1] *Faculty of Informatics, Kanagawa University*
3-27-1, Rokkakubashi, Kanagawa-ku, Yokohama-shi, Kanagawa, 221-8686, Japan.
[2] *School of Information Science and Technology, Aichi Prefecture University*
1522-3, Ibaragabasama, Nagakute-shi, Aichi, 480-1198, Japan.

Abstract. Quantum reading is a quantum sensing protocol for reading out binary information from an optical disk using an entangled state. The information is encoded by two different reflectivities, and the problem is how to discriminate between them accurately. Several previous studies have shown that the two-mode squeezed vacuum (TMSV) state offers an advantage over a classical light source under some constraints. In this paper, we evaluate and discuss the performance achieved using another type of entangled state called the quasi-Bell state. We show that this state can outperform the TMSV state.

Keywords: Quasi-Bell state, Quantum reading, Quantum Chernoff bound

1 Introduction

In recent years, the diffusion of big data, cloud computing, and other technologies has made cold storage devices, that is, optical disks such as DVDs, popular again. Optical disks store information in memory cells with different material reflectivities, and the information is read out by a laser beam emitted from an optical head. The physically recorded information can be stored for centuries, but how accurately the information is read out, that is, how accurately the reflectivities are discriminated, is important to its performance.

In 2011, S. Pirandola treated the task of reflectivity discrimination as a quantum channel discrimination problem and proposed a quantum reading model [1] using entangled states to read the binary information encoded by reflectivities $\{r_0, r_1\}$, which may be more accurate than the classical light source approach. The results showed that the two-mode squeezed vacuum (TMSV) state performs better than a classical light source under some constraints. In the same year, O. Hirota proposed a model using the quasi-Bell state [2], another type of entangled state, to read the binary information encoded by phases $\{0, \pi\}$ and showed that this model achieves error-free performance using little input energy [3]. These important quantum reading models have different applications, such as symmetric quantum communication systems [4–7], and therefore the performance of these models given various entangled states needs to be clarified.

In this study, we aim to clarify the performance of S. Pirandola's quantum reading model when the quasi-Bell state is used. As a first step, in this paper, we focus on an ideal memory in which r_1 is fixed to 1 and derive the analytical expression of the quantum Chernoff bound for quantum reading using the quasi-Bell state. We also evaluate and compare its performance with the TMSV state and a classical light source.

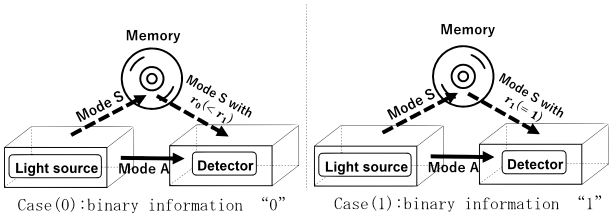


Figure 1: Model of quantum reading

2 Problem setup

2.1 Quantum reading

In this section, we explain the quantum reading protocol as proposed in [1]. The quantum reading model is illustrated in Fig. 1. The dashed and solid arrows represent mode S (signal mode) and mode A (ancilla mode) of the entangled state, respectively. The protocol is as follows:

1. The light beam of mode A of the entangled state is irradiated to the detector.
2. The light beam of mode S is irradiated toward the memory cell in which binary information 0 or 1 is recorded. If the recorded information is 0 (or 1), the light beam, subject to reflectivity r_0 (or r_1), is reflected.
3. Both the light beam of mode S reflected from the memory and that of mode A are input to the detector.
4. The detector decodes the binary information by

*wang@kanagawa-u.ac.jp
†usuda@ist.aichi-pu.ac.jp

performing an optimum quantum measurement on both light beams.

In this paper, we set $r_1 = 1$ and assume that the binary information has equal *a priori* probabilities.

2.2 Classical discrimination bound

When considering the input energy constraints, such as the average number of photons N_c, the lower bound on the error probability using any classical light source for the discrimination of r_0 and $r_1(= 1)$ is as follows [1]:

$$P_{LB}^{(Classical)} = \frac{1 - \sqrt{1 - \mathrm{e}^{-N_c(1-\sqrt{r_0})^2}}}{2}. \tag{1}$$

2.3 Quantum Chernoff bound

In quantum information science fields such as quantum sensing [1, 8, 9], the quantum Chernoff bound is a common criterion for evaluating system performance. The quantum Chernoff bound [10] is a mathematical method that obtains the upper bound on probability of error $P_e^{(M)}$ when discriminating M copies of a quantum state ρ_0 from that of a quantum state ρ_1. The quantum Chernoff bound P_{QCB} is expressed as

$$P_e^{(M)} \le \frac{1}{2} C^M =: P_{QCB}(M), \quad C = \inf_{s\in(0,1)} \mathrm{Tr}\left(\rho_0^s \rho_1^{1-s}\right), \tag{2}$$

where F is the fidelity and $C = F = \langle\psi|\rho_0|\psi\rangle$ can be obtained when $\rho_1 := |\psi\rangle\langle\psi|$ is a pure state [10, 11].

2.4 Quasi-Bell state

One class of entangled state constructed using a set of nonorthogonal quantum states such as coherent states is called the quasi-Bell state [2]. A quasi-Bell state is represented by

$$|\Psi\rangle_{\mathrm{SA}} = h\left(|\alpha\rangle_{\mathrm{S}}|\alpha\rangle_{\mathrm{A}} - |-\alpha\rangle_{\mathrm{S}}|-\alpha\rangle_{\mathrm{A}}\right), \tag{3}$$

where

$$h = \frac{1}{\sqrt{2(1-\kappa^2)}}, \quad \kappa = \mathrm{e}^{-2|\alpha|^2}, \quad \alpha \in \mathbb{R}_+, \tag{4}$$

and $|\alpha\rangle_{\mathrm{S}}$ and $|\alpha\rangle_{\mathrm{A}}$ are coherent states of mode S and mode A, respectively. The average number of photons of the signal light, i.e., the input energy, is denoted by $N_q = |\alpha|^2 \coth(2|\alpha|^2)$. The minimum N_q is 0.5.

In this paper, we derive the representation of the received quantum states $\rho_0^{(q)}$ and $\rho_1^{(q)}(= |\Psi\rangle_{\mathrm{SA}}\langle\Psi|)$ corresponding to r_0 and $r_1(= 1)$, respectively, in the 8-dimensional subspace of a Hilbert space using the method in [4, 5]. By substituting the representation into the fidelity ${}_{\mathrm{SA}}\langle\Psi|\rho_0^{(q)}|\Psi\rangle_{\mathrm{SA}}$, we obtain the quantum Chernoff bound, which can be expressed as

$$P_{QCB}^{(Quasi-Bell)}(M) = \frac{1}{2}\left(\frac{(1+L)(A_- - A_+\kappa)^2}{2(\kappa^2-1)^2}\right)^M, \tag{5}$$

where $A_\pm = \mathrm{e}^{-\frac{1}{2}(\sqrt{r_0}\pm 1)^2|\alpha|^2}$ and $L = \mathrm{e}^{-2(1-r_0)|\alpha|^2}$.

2.5 Two-mode squeezed vacuum state

A TMSV state, also known as an Einstein–Podolsky–Rosen state, is represented by

$$|\psi\rangle_{\mathrm{SA}} = \sum_{n=0}^{\infty} \sqrt{\frac{N_t^n}{(N_t+1)^{n+1}}}|n\rangle_{\mathrm{S}}|n\rangle_{\mathrm{A}}, \tag{6}$$

where N_t is the average number of photons of the signal light.

For $\rho_0^{(t)}$ and $\rho_1^{(t)} = |\psi\rangle_{\mathrm{SA}}\langle\psi|$ corresponding to r_0 and $r_1(= 1)$, respectively, the quantum Chernoff bounds are obtained using the fidelity with a covariance matrix representation as follows [11–13]:

$$P_{QCB}^{(TMSV)}(M) = \frac{1}{2}\left(\frac{1}{(1+N_t-N_t\sqrt{r_0})^2}\right)^M. \tag{7}$$

3 Performance comparison

In [1], information gain G, expressed as

$$G(Q,R) = 1 - H(Q) - [1 - H(R)], \tag{8}$$

was introduced as a performance evaluation index, where $H(p) = -p\log_2(p) - (1-p)\log(1-p)$. This index represents the gain in bits obtained by the light source corresponding to Q over that corresponding to R for each memory cell. In this paper, G is also employed to evaluate the gain achieved with each light source.

3.1 Information gain for $M = 1$

Figs. 2 and 3 respectively plot the information gain $G\left(P_e^{(Quasi-Bell)}, P_{LB}^{(Classical)}\right)$, and $G\Big(P_e^{(Quasi-Bell)}, P_{QCB}^{(TMSV)}(1)\Big)$ versus the average number of photons N and reflectivity r_0 when $M = 1$. Note that we evaluate the information gain of the quasi-Bell state using the minimum error probability $P_e^{(Quasi-Bell)}$, which is obtained from [4, 5]. Moreover, the gains achieved by the quasi-Bell state over any classical light source and the TMSV state are shown in Figs. 2 and 3, respectively.

The results in Fig. 2 show that the performance of the quasi-Bell state is always substantially superior to that of any classical light source, regardless of increases in N or changes in r_0. In contrast, as shown in Fig. 3, the performance of the quasi-Bell state is superior to that of the TMSV state when r_0 is low and vice versa.

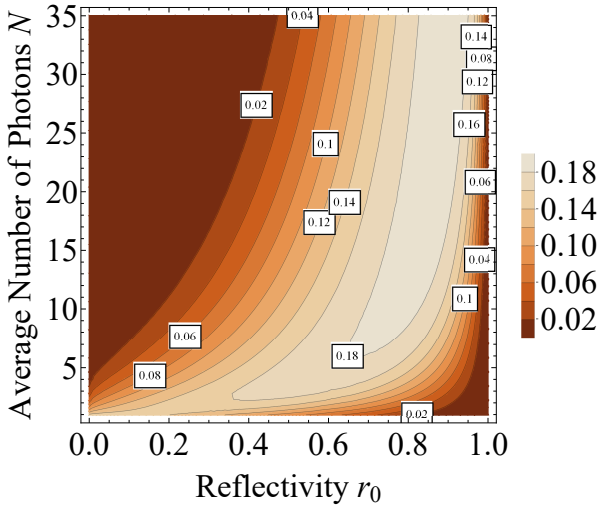


Figure 2: Gain by quasi-Bell state over classical LB when $M = 1$

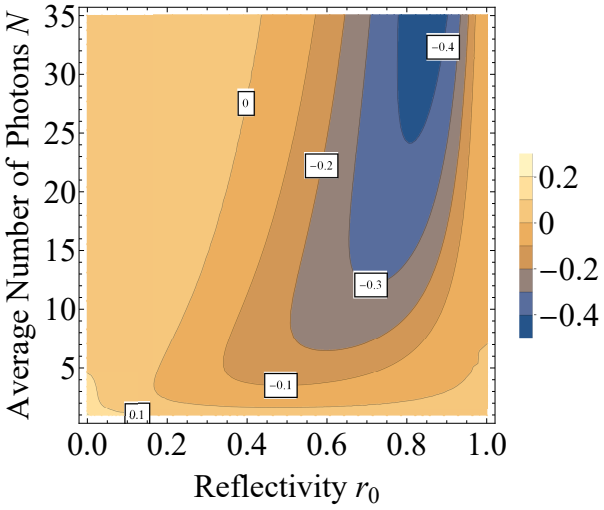


Figure 3: Gain by quasi-Bell state over TMSV state when $M = 1$

Ref. [1] proved that there is a minimum M such that $P_{QCB}^{(TMSV)}(M) < P_{LB}^{(Classical)}$. In this paper, we show that the minimum M is 1 when using the quasi-Bell state, even though there is a regime under which the TMSV state is superior.

3.2 Information gain for $M > 1$

Figs. 4 and 5 plot the information gain $G\left(P_{QCB}^{(Quasi-Bell)}(M), P_{LB}^{(Classical)}\right)$ and $G\left(P_{QCB}^{(Quasi-Bell)}(M), P_{QCB}^{(TMSV)}(M)\right)$, respectively, using the same setting as Figs. 2 and 3 with the exception that $M = 15$. Note that $N = N_c = N_t M = N_q M \geq 0.5M$.

The results in Fig. 4 show that the performance of the quasi-Bell state can be substantially improved by dividing N into M copies. In contrast, as shown in Fig. 5, the use of the quasi-Bell state leads to more gain than the TMSV state in the regime of low N and vice versa. A comparison of Figs. 3 and 5 shows that the region in which the quasi-Bell outperforms the TMSV state remains despite increases in M.

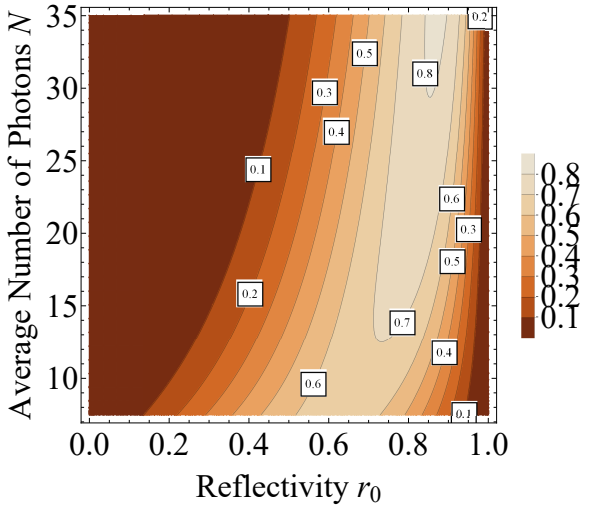


Figure 4: Gain by quasi-Bell state over classical LB when $M = 15$

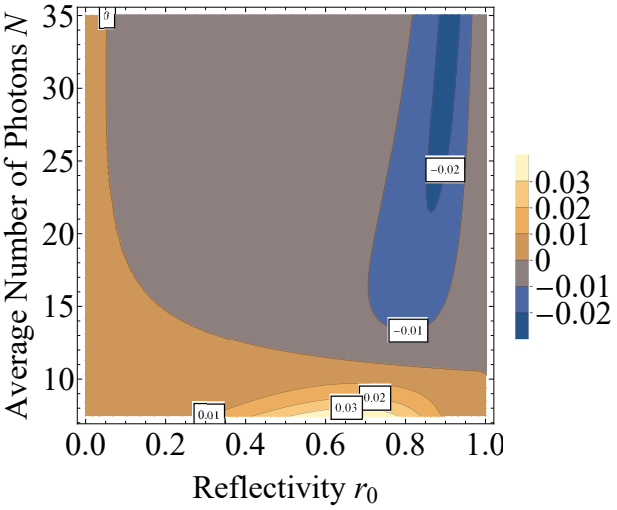


Figure 5: Gain by quasi-Bell state over TMSV state when $M = 15$

4 Conclusion

In this paper, we evaluated and compared the performance of quantum reading when using the TMSV state, quasi-Bell state, and classical light source with respect to the criterion of the quantum Chernoff bound. In particular, we derived an analytical expression for the quantum Chernoff bound when the quasi-Bell state is used. We also clarified that to exceed the performance of any classical light source, the minimum number of copies M is 1 when the quasi-Bell state is used. When $M > 1$, we showed that there is a region in which the quasi-Bell outperforms the TMSV state. This result suggests that the appropriate type of entangled state should be chosen for the required M and r_0.

Acknowledgments This work has been supported in part by JSPS KAKENHI Grant Number 20H00581, 20K20397, 21K04064, 22K20437, and The Nitto Foundation. We thank Kimberly Moravec, PhD, from Liwen Bianji (Edanz) for editing the English text of a draft of this manuscript.

References

[1] S. Pirandola. Quantum reading of a classical digital memory. *Phys. Rev. Lett.*, Vol.106, 090504, 2011.

[2] O. Hirota and M. Sasaki. Entangled state based on nonorthogonal state. In *Proc. of Quantum Communication, Computing, and Measurement 3*, pages 359-366, 2001.

[3] O. Hirota. Error free quantum reading by quasi Bell state of entangled coherent states. *Quantum Meas. Quantum Metrol.*, Vol.4, pages 70-73, 2017. arXiv:quant-ph/1108.4163v2 (2011)

[4] T. Wang, S. Takahira, and T. S. Usuda. Error performance of ASK-type asymmetric quantum communication. In *Proc. of AQIS2021 (poster sessions)*, pages 200-203, 2021.

[5] T. Wang and T. S. Usuda. Error performance of amplitude shift keying-type asymmetric quantum communication systems. *Entroy*, Vol.24, Issue 5, 708, 2022.

[6] S. Sameshima, T. Wang, S. Usami, and T. S. Usuda. PSK-type asymmetric quantum communication and its attenuation characteristics. In *Proc. of ISITA2022*, pages 241-245, 2022.

[7] S. Sameshima, T. Wang, S. Usami, and T. S. Usuda. A PSK-type asymmetric quantum communication and its error performance in attenuated environments. *IEICE Trans. Commun.*, Vol.J106-B, No.03, pages 112-125, 2023. (In Japanese)

[8] S. Lloyd. Enhanced sensitivity of photodetection via quantum illumination. *Science*, Vol.321, pages 1463-1465, 2008.

[9] S. H. Tan, B. I. Erkmen, V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, S. Pirandola, and J. H. Shapiro. Quantum illumination with Gaussian states. *Phys. Rev. Lett.*, Vol.101, 253601, 2008.

[10] K. M. R. Audenaert, J. Calsamiglia, L. Masanes, R. Muñoz-Tapia, A. Acin, E. Bagan, and F. Verstraete. Discriminating states: The quantum Chernoff bound. *Phys. Rev. Lett.*, Vol.98, 160501, 2007.

[11] G. Spedalieri, C. Weedbrook, and S. Pirandola. A limit formula for the quantum fidelity. *J. Phys. A Math. Theor.*, Vol.46, 025304, 2013.

[12] G. Spedalieri and S.L. Braunstein. Asymmetric quantum hypothesis testing with Gaussian states. *Phys. Rev. A*, Vol.90, 052307, 2014.

[13] G. Spedalieri. Cryptographic aspects of quantum reading. *Entroy*, Vol.17, Issue 4, papes 2218-2227, 2015.

Adiabatic quantum computing with parameterized quantum circuits

Ioannis Kolotouros[1] * Ioannis Petrongonas[2] † Miloš Prokop[1] ‡ Petros Wallden[1] §

[1] *University of Edinburgh, School of Informatics, EH8 9AB Edinburgh, United Kingdom*
[2] *Heriot-Watt University, EH14 4AS Edinburgh, United Kingdom*

Abstract. In this work, we start by analyzing how small perturbations of a Hamiltonian affect the parameters that minimize the energy within a family of parameterized quantum states. By measuring a series of observables in the unperturbed system, we derive a constrained linear system of equations whose solution provides the new minimum. Then, we propose a NISQ-friendly discrete version of adiabatic quantum computing, with a proven lower bound of discretization steps to guarantee success, that is insensitive to parameter initialization and requires no energy minimization. We show that our algorithm outperforms VQE on MaxCut, Number-Partitioning, and on Transverse-Field Ising Chain model.

Keywords: Adiabatic quantum computing, NISQ

We are currently in the Noisy Intermediate-Scale Quantum Computing era [1] where small quantum computers are dominated by noise, small coherence times and limited connectivity. Conventional techniques like *Adiabatic Quantum Computing* (AQC) [2] with proven theoretical guarantees require quantum circuits with depth that is unreachable for the existing hardware. Variational Quantum Algorithms (VQAs) [3, 4] offer a promising framework for practical quantum advantage in this era. In this approach, a quantum computer works in parallel with a classical computer in an iterative feedback loop with the goal of solving a problem-specific task. Specifically, in VQAs the problem at hand is mapped into an interacting qubit Hamiltonian H_C whose minimum eigenvalue corresponds to the (optimal) solution of the problem. The quantum computer then prepares and measures a parameterized quantum state (usually referred to as "ansatz") and the classical computer iteratively updates the parameters using a classical optimization algorithm until convergence, possibly to a local minimum, is achieved.

This approach, however, is hindered by a series of limitations that make these algorithms less likely to offer a practical advantage. First of all, the emerging cost landscapes are filled with a vast number of local minima [5] where local optimization algorithms can falsely converge to. As a result, a bad initialization will have an immediate effect on the quality of the solution. Finding a parameterized quantum circuit that contains the solution may sometimes be hard and on top of that, utilizing highly expressive ansatz families that span a large fraction of the total Hilbert space lead to phenomena like *barren plateaux* [6, 7] where exponentially vanishing gradients make the optimization intractable. It is thus crucial to further analyze the geometry of the underlying non-convex landscapes and subsequently understand the limitations of variational quantum algorithms. We present the main contributions of our work that lead to an alternative approach to NISQ algorithms than that offered by VQAs:

- We study how small perturbations of the Hamiltonian affect the optimization landscape and derive a theorem that enables us to calculate how much the global minimum is shifted under these perturbations.
- We utilize the aforementioned theorem and adiabatic quantum computing and derive an algorithm (AQC-PQC) that returns the best approximation of the ground state of a Hamiltonian within a family of parameterized quantum states that (i) can be applied in the NISQ setting, (ii) is not sensitive to the initialization points, (iii) requires fixed calls to the quantum computer with theoretical guarantees on the performance (iv) requires no energy minimization.
- We derive a lower bound on the number of discretization steps needed and compare our algorithm with the Variational Quantum Eigsolver (VQE) in two classical optimization problems, namely MaxCut and Number Partitioning and one quantum spin-interacting problem namely the Transverse-Field Ising Chain (TFIC) model.

Our main theorem (outlined in Theorem 1) aims to quantify how much must we shift the optimal angles that minimize a Hamiltonian H_1 if we perturb the Hamiltonian by a small amount λH_2 with $\lambda \ll 1$.

Theorem 1 *Consider a parameterized quantum circuit defined via the unitaries $U(\boldsymbol{\theta})$, and the corresponding states $|\psi(\boldsymbol{\theta})\rangle = U(\boldsymbol{\theta})|0\rangle$. We are given a Hamiltonian H_0 and the angles $\boldsymbol{\theta^*}$ that minimize its energy, i.e. $\boldsymbol{\theta^*} = \arg\min_{\boldsymbol{\theta}} \langle\psi(\boldsymbol{\theta})| H_0 |\psi(\boldsymbol{\theta})\rangle$. If we perturb the Hamiltonian H_0 by a small amount λH_2 with $\lambda \ll 1$, and $\|H_0\| \approx \|H_2\|$, then there exists a shift vector $\boldsymbol{\epsilon}$ such that, with high probability, the state $|\psi(\boldsymbol{\theta^*} + \boldsymbol{\epsilon})\rangle$ is the ground state of the perturbed Hamiltonian $H_\lambda = H_0 + \lambda H_2$ and the shift vector is the solution of the following mathemat-*

*i.kolotouros@sms.ed.ac.uk
†ip2004@hw.ac.uk
‡m.prokop@sms.ed.ac.uk
§petros.wallden@ed.ac.uk

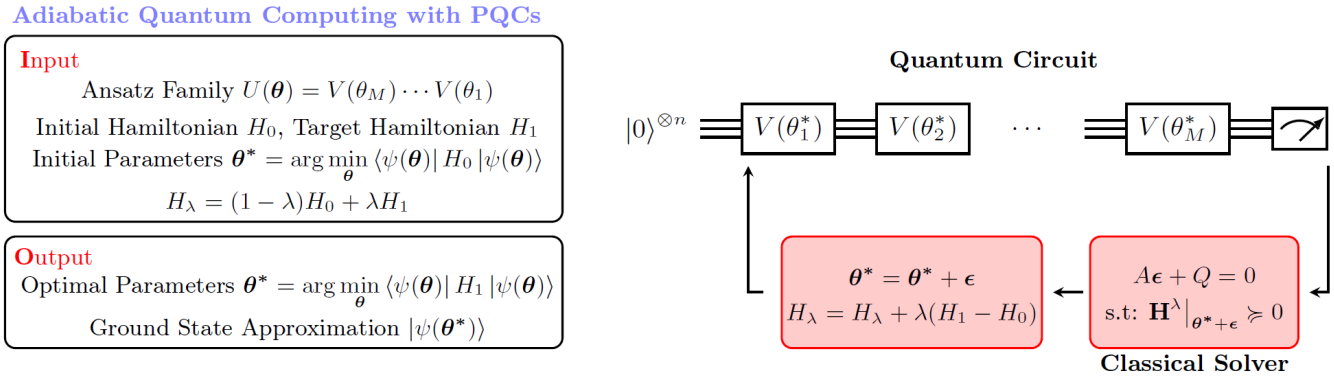


Figure 1: Adiabatic Quantum Computing with Parameterized Quantum Circuits (AQC-PQC).

ical problem:

$$\begin{aligned} &min\ ||\boldsymbol{\epsilon}|| \\ &subject\ to:\ A\boldsymbol{\epsilon} + \boldsymbol{Q} = 0, \\ &\boldsymbol{H}^\lambda\big|_{\boldsymbol{\theta}^*+\boldsymbol{\epsilon}} \succcurlyeq 0, \end{aligned} \tag{1}$$

where $\boldsymbol{H}^\lambda\big|_{\boldsymbol{\theta}^*+\boldsymbol{\epsilon}}$ *is the Hessian evaluated at the shifted point,* $\boldsymbol{Q} = \sum_i Q_i \hat{\boldsymbol{e}}_{\boldsymbol{i}}$ *is a vector and* A *is a matrix that are defined via their elements*

$$\begin{aligned} Q_i &= \lambda \frac{\partial}{\partial\theta_i} \left(\langle\psi(\boldsymbol{\theta})| H_2 |\psi(\boldsymbol{\theta})\rangle\right)\Big|_{\boldsymbol{\theta}^*} \\ A_{ij} &= \frac{\partial^2}{\partial\theta_i\partial\theta_j} \left(\langle\psi(\boldsymbol{\theta})| H_\lambda |\psi(\boldsymbol{\theta})\rangle\right)\Big|_{\boldsymbol{\theta}^*} \end{aligned}$$

Intuitively, we search for the smaller shift (first line), that results in a vanishing gradient (second line) that is a minimum (third line). The most important point is that the elements A_{ij}, Q_i and the Hessian correspond to observables calculated on the unperturbed state $|\psi(\boldsymbol{\theta}^*)\rangle$.

Inspired by AQC, we propose a novel algorithm that utilizes Theorem 1 and is significantly different than the standard framework of variational quantum algorithms. In AQC, the system of qubits is initialized in an easy-to-prepare ground state of a Hamiltonian H_0. Then, the system is allowed to interact under a time-dependent Hamiltonian $H(\tau) = (1-\tau)\,H_0 + \tau H_1$ (with $\tau \equiv t/t_f \in [0,1]$). If the evolution is slow enough, so that the system of qubits always remains in the instantaneous ground state, the system at time $t = t_f$ will find itself in the ground state of H_1. The total computational resources, i.e. the total time t_f needed to evolve to the desired ground state, depends on the spectral gap between the instantaneous first excited state and the ground state [8] which sometimes can become exponentially small.

The main idea of our algorithm is the following. We consider the discretized Hamiltonian

$$H_k = \left(1 - \frac{k}{K}\right) H_0 + \frac{k}{K} H_1. \tag{2}$$

Here the step subscript k has the role of the (discrete in our case) time. Let us set $\lambda := 1/K \ll 1$, and $H_2 := (H_1 - H_0)$. We can rewrite the step-dependent Hamiltonian as

$$H_k = H_0 + \lambda H_2 k. \tag{3}$$

We can easily see that $H_{k+1} - H_k = \lambda H_2$, and thus we can apply Theorem 1 for any consecutive pair of $\{H_k, H_{k+1}\}$. We start from H_0 and initialize the algorithm with the known ground state that corresponds to the initial parameters $\boldsymbol{\theta_0}$ such that $\boldsymbol{\theta_0} = \arg\min_{\boldsymbol{\theta}} \langle\psi(\boldsymbol{\theta})| H_0 |\psi(\boldsymbol{\theta})\rangle$. Then, for each step we compute the shift vector $\boldsymbol{\epsilon}$ and add it to the parameters corresponding to the ground state of the previous step to obtain the ground state of the next step. After K steps, the Hamiltonian becomes H_1 and the algorithm returns the desired ground state. You can visualize our algorithm in Figure 1 and see [9] for full details. The number of discrete steps K needed to always return the ground state of H_1 can be lower bounded by Theorem 2 stated below.

Theorem 2 *Consider a time-dependent Hamiltonian* $H(\tau) = (1-\tau)H_0 + \tau H_1$, $\tau \equiv t/t_f \in [0,1]$. *Let* $\Delta(\tau) \equiv E_1(\tau) - E_0(\tau)$ *be the instantaneous spectral gap and* $\delta_\tau(\lambda) \equiv E_0(\tau+\lambda) - E_0(\tau)$ *be the energy difference between the ground states at times* $\tau + \lambda$ *and* τ *respectively. Moreover, assume that the parameterized family of states contains the ground state for each* $\tau \in [0,1]$ *and* $|\delta_\tau(\lambda)| \ll \Delta(\tau+\lambda)$. *Then, AQC-PQC will always return the ground state of the target Hamiltonian* H_1 *at time* $\tau = 1$ *as long as we discretize the time-dependent Hamiltonian into* $K > K_0$ *steps where:*

$$K_0 \in \mathcal{O}\left(\frac{\text{poly}(n)}{\min_\tau \Delta(\tau)}\right) \tag{4}$$

Our algorithm offers a number of potential advantages. The classical part in our approach is a constrained linear solver (that can be performed very efficiently and with guarantees of finding the solution), while in variational approaches the classical part is an energy minimization. The vast majority of bottlenecks in VQAs come from the classical optimization part. First of all, a random initialization of parameters may lead to either bad performance or barren plateaux. Secondly, the emerging landscapes of VQAs are filled with a large amount of local minima that makes the parameters untrainable. Another key

Table 1: Probability of sampling the optimal solution.

MaxCut	Optimal Solution Overlap (%)					
	7 Qubits	**8 Qubits**	**9 Qubits**	**10 Qubits**	**11 Qubits**	**12 Qubits**
AQC-PQC	82.7	74.3	93.1	50	28.1	56.6
VQE	62.3	54.7	60.8	39.2	22.1	11.1
Number Partitioning	Optimal Solution Overlap (%)					
AQC-PQC	37.5	21.9	24.7	12.6	5	4.6
VQE	28.5	6.2	6.4	1.2	0.8	0.4

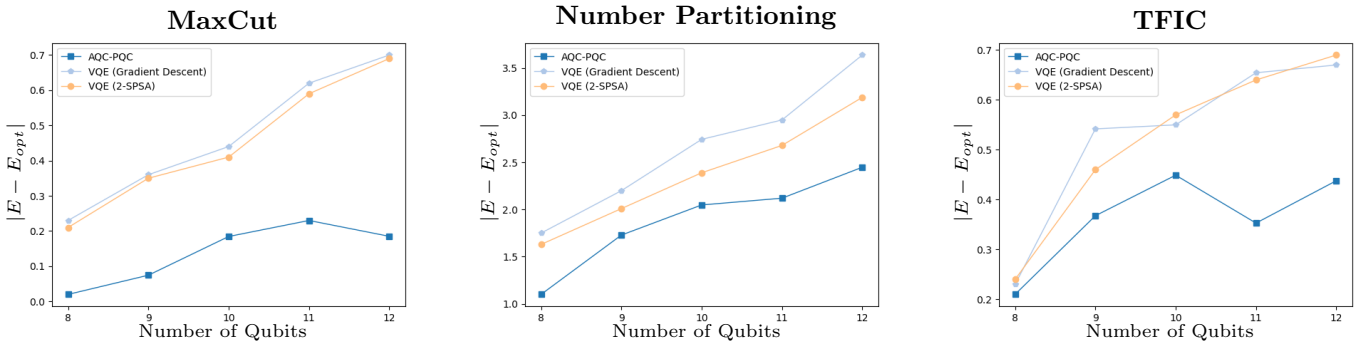


Figure 2: Performance of AQC-PQC algorithm compared to VQE (with 2-SPSA and Gradient Descent optimizers) for the MaxCut problem (left) and the Number Partitioning problem (right). The AQC-PQC algorithm with the dark blue line (square markers) is able to outperform both 2-SPSA and Gradient Descent on the quality of the output solution.

difference, and advantage of our approach, is that traditional variational quantum approaches require multiple quantum state preparations and number of shots that increases significantly as one approaches a minimum (as the gradients tend to zero). For each iteration, multiple quantum states need to be prepared (details depend on the classical optimizer used). However we neither know in advance how many iterations would be required to reach convergence, nor if the quantum state that we converge is the correct ground state. In contrast, in our strategy, we can mimic adiabatic quantum computing with only K steps that is typically much smaller than the iterations needed by a local classical optimization algorithm. Finally, our algorithm offers a direct advantage (over AQC) when the first excited state does not correspond to a minimum. In that case, the algorithm can choose larger steps as the Hessian of the first excited state would not be positive semidefinite and thus it would not correspond to a feasible solution of Eq. (1).

The performance of AQC-PQC was tested and compared with VQE on the Number Partitioning and the Max-Cut problem as well as on the TFIC problem. The first two problems correspond to classical combinatorial optimization problems with a diagonal Hamiltonian while the latter corresponds to a quantum problem with a non-diagonal Hamiltonian.

For the classical optimization problems, we chose to compare the methods on instance classes that we consider hard. Both of these problems have an intrinsic $\mathbb{Z}_2$ symmetry and so we chose instances with only two optimal solutions (one solution can be acquired from the other by flipping all qubits). Specifically, for the Max-Cut problem we sampled 100 *random weighted graphs* of sizes 8 to 12 while for the Number Partitioning problem, we sampled 100 instances of the same size as MaxCut with integers drawn from the interval $[0, 50]$.

Overall, we can see that AQC-PQC is able to outperform VQE in all instances, achieving overlap even five times larger in MaxCut and ten times larger in Number Partitioning (see Table 1). Moreover, as seen in Figure 2, the output states returned by AQC-PQC are significantly closer (in terms of energy) to the ground state compared to VQE. This is to be expected as the non-convexity of the cost landscape results in the classical optimization part of VQE to stuck in a local minimum. On the other hand, AQC-PQC provides a more robust strategy to navigate the (time-evolving) landscape. Provided that the number of steps is chosen accordingly and the ansatz family is expressive enough, the latter algorithm will always achieve a large overlap with the optimal solution.

For the last problem, we evaluated the performance of the two algorithms on TFIC for 100 random instances with the couplings (J_k, h) drawn uniformly at random from the uniform distribution. The results of the TFIC model are illustrated in Figure 2. Overall, we observe that AQC-PQC is able to return approximations of the ground state of the TFIC Hamiltonian that are closer compared to those returned by VQE. **Further details can be found at the full paper [9].**

References

[1] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[2] Tameem Albash and Daniel A Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1):015002, 2018.

[3] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.

[4] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S Kottmann, Tim Menke, et al. Noisy intermediate-scale quantum algorithms. *Reviews of Modern Physics*, 94(1):015004, 2022.

[5] Lennart Bittel and Martin Kliesch. Training variational quantum algorithms is np-hard. *Physical review letters*, 127(12):120502, 2021.

[6] Jarrod R McClean, Sergio Boixo, Vadim N Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature communications*, 9(1):1–6, 2018.

[7] Marco Cerezo, Akira Sone, Tyler Volkoff, Lukasz Cincio, and Patrick J Coles. Cost function dependent barren plateaus in shallow parametrized quantum circuits. *Nature communications*, 12(1):1791, 2021.

[8] Mohammad HS Amin. Consistency of the adiabatic theorem. *Physical review letters*, 102(22):220401, 2009.

[9] Ioannis Kolotouros, Ioannis Petrongonas, Miloš Prokop, and Petros Wallden. Adiabatic quantum computing with parameterized quantum circuits, 2023.

Short-depth Quantum Circuits for Probing Quantum Phase Transitions: Stochastic Series Expansion, Berry's Phase and Quantum Coherence

K. C. Tan[1 2 *] D. Bowmick[2] D. Liu[2] P. Sengupta[2]

[1] *Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610051, China*

[2] *School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Republic of Singapore*

Abstract. Limited coherence times is a major bottleneck in near-term quantum computing and short-depth quantum circuits are needed to operate within such limitations. We introduce several short-depth quantum algorithms that performs tasks such as (i) quantum stochastic series expansion (QSSE), (ii) Quantum Coherence Estimation (QCE) and (ii) Berry Phase Estimation (BPE). QSSE efficiently estimates expectation values of quantum observables, while QCE and BPE efficiently estimates non-local properties of a many-body system. We show how these algorithms avoid bottlenecks posed by exponentially increasing dimensions of quantum systems and can be useful to study the physics of many-body systems in the thermodynamic limit.

Keywords: Quantum Algorithms, Many-body Physics, Geometric Phase, Quantum Coherence

1 Introduction

Quantum computing can potentially revolutionize the study of many-body systems by making it possible to the simulate ever more complex interactions for very large system sizes[1]. A potential obstacle to achieving this is the limited coherence times of current quantum processors. For this reason, it is important to develop efficient algorithms that perform tasks efficiently using relatively shallow quantum circuits that can operate within the limitations of short coherence times, and yet is able to extract critical information about the physics of the many-body system. In the following sections we will discuss three quantum algorithms that are designed to do this: (i) quantum stochastic series expansion (QSSE)[2], (ii) Quantum Coherence Estimation (QCE) and (ii) Berry Phase Estimation (BPE). QSSE is a technique of extracting information about a many-body system without requiring us to first produce a simulated copy of the system within the quantum processor. On the other hand, QCE and BPE assumes that we already have a simulated copy of the system that is readily accessible, and the goal is to extract non-local quantum properties from the prepared state. Classical methods are typically bottlenecked by the exponentially increasing dimensions of quantum many-body systems. These quantum algorithms can be applied to study several different types of quantum phase transitions.

2 Quantum Stochastic Series Expansion for general Hamiltonians

The basic premise of Stochastic Series Expansion is that the expectation value of an observable O for a system with local Hamiltonian $H = \sum_i H_i$ at inverse temperature β can be written in the following way using Taylor's expansion as:

$$\langle O \rangle = \sum_{n,b,\alpha} \frac{\beta^n}{n!} \langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle \langle \alpha | O | \alpha \rangle / Z.$$

Here, the summation is over all possible strings $b = b_n \dots b_1$, expansion power n, and basis state $|\alpha\rangle$. Note that $\langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle$, as well as its complex conjugate $\langle \alpha | H_{b_1} \dots H_{b_n} | \alpha \rangle$ appears in the summation. Since $\langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle + \langle \alpha | H_{b_1} \dots H_{b_n} | \alpha \rangle = 2\,\mathrm{Re}\{\langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle\}$, only the real part of each term contributes to the expectation value. Hence we can equivalently write

$$\langle O \rangle = \sum_{n,b,\alpha} \frac{\beta^n}{n!} \mathrm{Re}\{\langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle\} \langle \alpha | O | \alpha \rangle / Z. \quad (1)$$

In order to implement quantum SSE, we only need to sample the real portion of $\langle \alpha | H_{b_n} \dots H_{b_1} | \alpha \rangle$ and ensure that it is nonnegative. We show that this can be done by adding a sufficiently large constant to the Hamiltonian (see Supplementary Information).

Suppose $M \geq n$ is the cutoff in the expansion power. For a fixed M, let $H_b := |h_b| \left[\mathrm{sgn}(h_b) \bigotimes_{i=1}^{N} \sigma_{b^i}^{(A_i)} + 2M \mathbb{1}_A \right]$. We note that this is an unequal superposition of 2 unitary operations that depends on the cutoff value M.

We introduce the state

$$|\psi_{\mathrm{in}}\rangle := |\alpha_A\rangle \left|\phi_{B_1}\right\rangle \dots \left|\phi_{B_n}\right\rangle |+_C\rangle, \quad (2)$$

where

$$\left|\phi_{B_i}\right\rangle := \sqrt{(2M)/(2M+1)} \left|0_{B_i}\right\rangle + \sqrt{1/(2M+1)} \left|1_{B_i}\right\rangle$$

and $|+_C\rangle := \frac{1}{\sqrt{2}}(|0_C\rangle + |1_C\rangle)$.

It can be shown that for any given expansion power n, we can define a unitary $V_{AB,C}$, controlled by qubit C, that satisfiies:

$$\begin{aligned} &\langle \psi_{\mathrm{in}} | V_{AB,C} | \psi_{\mathrm{in}} \rangle \\ &= \frac{\langle \alpha_A | H_{b_1} \dots H_{b_n} | \alpha_A \rangle + \langle \alpha_A | H_{b_n} \dots H_{b_1} | \alpha_A \rangle}{2(2M+1)^n \left| h_{b_n} \dots h_{b_1} \right|} \\ &= \frac{\mathrm{Re}\{\langle \alpha_A | H_{b_n} \dots H_{b_1} | \alpha_A \rangle\}}{(2M+1)^n \left| h_{b_n} \dots h_{b_1} \right|}, \end{aligned} \quad (3)$$

* bbtankc@gmail.com

which allows us to define an estimator for the relative weight:

$$q(n,b,\alpha) \;\coloneqq\; \left|\langle\psi_{\text{in}}|\, V_{AB,C}\, |\psi_{\text{in}}\rangle\right|^2$$

$$\equiv \left|\frac{\text{Re}\left\{\langle\alpha_A|\, H_{b_n}\ldots H_{b_1}\, |\alpha_A\rangle\right\}}{(2M+1)^n\left|h_{b_n}\ldots h_{b_1}\right|}\right|^2 \tag{4}$$

Note that the spectrum of $H_{b_i}/\left|h_{b_i}\right|$ is in the range $[0, 2M{+}1]$ so the absolute value of $\text{Re}\left\{\langle\alpha_A|\, H_{b_n}\ldots H_{b_1}\, |\alpha_A\rangle\right\}/\left|h_{b_n}\ldots h_{b_1}\right|$ is within the range $[0,(2M+1)^n]$. The configuration weights, and hence $\text{Re}\left\{\langle\alpha_A|\, H_{b_n}\ldots H_{b_1}\, |\alpha_A\rangle\right\}$, are always nonnegative.

From here, we sample the probability $q(n,b,\alpha)$ and perform a standard Metropolis algorithm according to the ratio $\sqrt{q(n',b',\alpha')/q(n,b,\alpha)}$. It turns out that this ratio will yield, on average, exactly the correct expression for the expectation value of an observable O in Stochastic Series Expansion.

3 Quantum algorithm for sampling quantum coherence

We now describe a quantum algorithm to efficiently sample the total quantum coherence of a many-body system. Briefly speaking, quantum coherence is a measure of the the amount of quantum superposition that exists between the orthogonal basis of a given quantum system[3]. There are many possible quantifiers for quantum coherence, but the one we will consider is called the l_2-norm of coherence. Recall that for any matrix M and basis $\{|i\rangle\}$, the l_2-norm is $\|M\|_{l_2} \coloneqq \sqrt{\sum_{i,j}|\langle i|M|j\rangle|^2}$. Correspondingly, the l_2-norm of coherence is the l_2-norm induced distance between a state ρ and the closest incoherent state σ:

$$C_{l_2}(\rho) \coloneqq \min_{\sigma\in\mathcal{I}} \|\rho-\sigma\|_{l_2}^2, \tag{5}$$

where $\mathcal{I}$ is the set of incoherent states i.e. the set of quantum states σ which has a diagonal density matrix in the basis $\{|i\rangle\}$. One can also write the l_2-norm of coherence in terms of the completely dephased state $\Delta\rho \coloneqq \sum_i \langle\rho|i|\rho\rangle\, |i\rangle\langle i|$, which gives

$$\begin{aligned} C_{l_2}(\rho) &\coloneqq \|\rho-\Delta\rho\|_{l_2}^2 \\ &= \sum_{i\neq j} |\langle i|\rho|j\rangle|^2. \end{aligned} \tag{6}$$

We now describe quantum circuits that estimate the l_2-norm of coherence. We write ρ in matrix form as

$$\rho = \sum_{ij} \rho_{ij}\, |i\rangle\langle j|\,. \tag{7}$$

For simplicity, we will assume that $\{|i\rangle\}$ is the computational basis. Consider the purity of the state:

$$\begin{aligned} \text{Tr}(\rho^2) &= \text{Tr}\Big(\sum_{ij}\rho_{ij}\,|i\rangle\langle j|\sum_{kl}\rho_{kl}\,|k\rangle\langle l|\Big) \\ &= \sum_i |\rho_{ii}|^2 + \sum_{i\neq j}\left|\rho_{ij}\right|^2. \end{aligned} \tag{8}$$

Note that in Eq. (8), the second term $C_{l_2}(\rho) \coloneqq \sum_{i\neq j}\left|\rho_{ij}\right|^2$ is exactly the l_2 norm of coherence.

Suppose we have two independent preparations of ρ,and let P_2 denote cyclic permutation of the two input states. We verify that:

$$\begin{aligned} \langle P_2\rangle &= \text{Tr}(P_2\rho\otimes\rho) \\ &= \text{Tr}(\rho^2), \end{aligned} \tag{9}$$

so we can easily measure the purity by measuring the expectation value of P_2.

Given the purity, we now need $\sum_i |\rho_{ii}|^2$, which is the purity of the completely dephased state $\Delta\rho$. For any n qubit state ρ, dephasing can be performed by copying the classical information from each qubit gate to ancillae qubits via a series of CNOT gates. We then measure the expectation value of P_2 on the state $\Delta\rho\otimes\Delta\rho$ to measure $\sum_i |\rho_{ii}|^2$. The final quantum coherence of the system is then given by

$$C_{l_2}(\rho) = \text{Tr}(\rho^2) - \text{Tr}(\Delta\rho^2). \tag{10}$$

We initially assumed that the basis $\{|i\rangle\}$ is the computational basis, but this is not a strict requirement. Suppose we want to measure the coherence with respect to some other set of basis vectors $\{|i'\rangle\}$ instead. In general, the two bases are related by some unitary operation $|i'\rangle = U\,|i\rangle$, so we can always measure the coherence with respect to $\{|i'\rangle\}$ by performing the previously described procedure on the state $U^\dagger\rho U$ instead of ρ.

4 Quantum algorithm for sampling discrete Berry phases

We now describe how to estimate the Berry phase of a many-body system using quantum circuits. We will use the definition of definition of the discrete Berry phase[4]. Suppose we have a collection of k quantum states $\{|\psi_i\rangle\}_{i=1}^k$ forming the closed loop $|\psi_1\rangle \to |\psi_2\rangle \to \ldots |\psi_k\rangle \to |\psi_1\rangle$. The discrete Berry Phase is:

$$\begin{aligned} \gamma_B^k &\coloneqq \text{Im}\{\ln(\langle\psi_1|\psi_2\rangle\ldots\langle\psi_k|\psi_1\rangle)\} \\ &= \sum_{j=1}^{k}\text{Im}\{\ln\left\langle\psi_j|\psi_{j+1 \bmod k}\right\rangle\}. \end{aligned} \tag{11}$$

This definition generalizes the standard definition of the Berry phase, in the limit of a large number of discrete points k. For a collection of k n-qubit pure states $\{|\psi_i\rangle\}_{i=1}^k$ we apply the cyclic permutation operation:

$$P_k\,|\psi_1\rangle\,|\psi_2\rangle\ldots|\psi_k\rangle = |\psi_k\rangle\,|\psi_1\rangle\ldots|\psi_{k-1}\rangle\,. \tag{12}$$

Observe that the expectation value of P_k is related to the discrete Berry phase:

$$\langle P_k\rangle = \langle\psi_1|\psi_k\rangle\,\langle\psi_2|\psi_1\rangle\ldots\langle\psi_k|\psi_{k-1}\rangle \tag{13}$$

$$= (\langle\psi_1|\psi_2\rangle\ldots\langle\psi_{k-1}|\psi_k\rangle\,\langle\psi_k|\psi_1\rangle)^* \tag{14}$$

To do this, we first perform a controlled operation with an ancilla qubit initialized in the state $|+\rangle$ such that

$$|\psi_1\rangle\ldots|\psi_k\rangle\,|+\rangle \tag{15}$$

$$\to (|\psi_1\rangle\ldots|\psi_k\rangle\,|0\rangle + P_k\,|\psi_1\rangle\ldots|\psi_k\rangle\,|1\rangle)/\sqrt{2}. \tag{16}$$

Measuring the ancilla in the $\{|+\rangle, |-\rangle\}$ basis yields the probabilities

$$P_\pm = (1 \pm \mathrm{Re}\{\langle P_k \rangle\})/2. \tag{17}$$

which allows us to find the real portion of $\langle P_k \rangle$. To find the imaginary portion, we apply an additional phase gate on the ancilla such that

$$|\psi_1\rangle \ldots |\psi_k\rangle |+\rangle \tag{18}$$

$$\rightarrow (|\psi_1\rangle \ldots |\psi_k\rangle |0\rangle + iP_k |\psi_1\rangle \ldots |\psi_k\rangle |1\rangle)/\sqrt{2}. \tag{19}$$

Measuring the ancilla in the $\{|+\rangle, |-\rangle\}$ basis as before, we obtain the probabilities

$$Q_\pm = (1 \mp \mathrm{Im}\{\langle P_k \rangle\})/2, \tag{20}$$

so $Q_- - Q_+$ gives us the imaginary portion of $\langle P \rangle$. Given both real and imaginary portions of $\langle P \rangle$, the discrete Berry phase is estimated by

$$\gamma_B^k = \mathrm{Im}\{\ln \langle P_k \rangle^*\} = -\arg\{\langle P_k \rangle\} \tag{21}$$

5 Conclusion

We proposed several short-depth quantum algorithms that performs tasks such as (i) quantum stochastic series expansion (QSSE), (ii) Quantum Coherence Estimation (QCE) and (ii) Berry Phase Estimation (BPE) that can be applied to study complex many-body system. In all cases the gate cost of implementing the algorithms scales linearly with the size of the system, while the depth cost scales only polylogarithmically. As a demonstration, we apply QSSE to estimate the energy of a frustrated spin-$\frac{1}{2}$ configuration and use QSSE and QCE to study topological phase transitions in spin-1 chains.

References

[1] J. Preskill. Quantum Computing in the NISQ era and beyond. *Quantum* **2**, 79 (2018).

[2] K. C. Tan, D. Bowmick, and P. Sengupta. Sign-problem free quantum stochastic series expansion algorithm on a quantum computer. *npj Quantum Inf.* **8**, 44 (2022).

[3] A. Streltsov, G. Adesso, and M. B. Plenio. Colloquium: Quantum coherence as a resource. *Rev. Mod. Phys.* **89**, 041003 (2017).

[4] K. Asbóth, L. Oroszlány, and A. Pályi. Berry phase, chern number. In *A Short Course on Topological Insulators: Band Structure and Edge States in One and Two Dimensions.* Springer International Publishing, pages 23–44, 2016.

Expectation values from any quantum measurements

Dominik Šafránek[1] * Dario Rosa[1 2] †

[1] *Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon - 34126, Korea*
[2] *Basic Science Program, Korea University of Science and Technology (UST), Daejeon - 34113, Korea*

Abstract. I will show that quite unintuitively, one can estimate an expectation value of a desired observable, by measuring any other observable. Also the entire probability distribution of outcomes can be estimated. This shows that quantum measurements carry much more information than was previously known. This result is useful experimentally, in any system in which there are limitations on which measurements can be performed. This is especially true for estimating expectation values of global observables, such as energy, in many-body quantum simulators, in which only local measurements can be performed.

Keywords: expectation values, measurement, many-body quantum systems, quantum simulators, quantum annealers

1 Summary

Measuring expectation values is one of the most important tasks in many fields of quantum physics. In many-body systems, it is very important to measure the energy, because it encodes many properties of the many-body systems, and informs us how close we are to the ground state. The latter is especially important in quantum annealers, in which the ground state of a Hamiltonian encodes a solution to a problem that could be very difficult to solve numerically, in a process called adiabatic quantum computing. However, without knowledge of the energy of the system, one cannot be sure whether the ground state was reached. Additionally, it is very difficult to measure this quantity, because at most two-local measurements can be performed. I will show that quite unintuitively, one can estimate an expectation value of a desired observable, by measuring any other observable [1]. Also the entire probability distribution of outcomes can be estimated. This shows that quantum measurements carry much more information than was previously known. This result is completely general and can be applied to any quantum systems, and especially to many-body systems. We perform simulations of four experimentally realized models and show that using this method, in some situations one can exclude over 95

2 Motivation and overlap with fields of quantum

Measuring observables is one of the central tasks in quantum physics. They appear in quantum information (entanglement witness), and its subfields quantum thermodynamics (entropy, work extraction), quantum speed limits (characteristic time scale of quantum systems), quantum metrology (homodyne measurement, Heisenberg limit in scaling, estimation of time and temperature). On a broader scale, expectation values also appear in Heisenberg's uncertainty principle, quantum field theory (vacuum expectation values), and nuclear physics (mean properties of atomic nuclei). However, in experimental setups (such as quantum simulators) only a very restricted set of measurements can be performed. Hence, we developed a theoretical framework to estimate any observable of interest by measuring any other observable. This is very important, for example, in adiabatic quantum computing and quantum annealing, in which case one wants to estimate the energy of the state, and see whether it is close to the ground state. Since these results are completely general, they also go beyond this specific application.

3 Methods

We provide rigorous proofs for four main theorems, regarding the bounds on the energy probability distribution and on the bound for the mean value of energy. We provide numerical simulations in small and large systems, for the Heisenberg model in the manuscript, and additionally the XY, Ising, and PXP models, all of which have been experimentally realized.

4 Interest to the community, and potential impact

- As mentioned above, expectation values appear in many fields of physics. Methods invented here can therefore be applied very broadly.
- Experimentally, in ion-trap quantum simulators, adiabatic quantum computers, and in quantum annealers (such as D-Wave), it is very important to measure expectation values. For example, measuring the expectation value of energy might confirm that one is close to the ground state, which represents the computational output. Yet, there are significant experimental limitations on the locality of the gates one can perform. Our method allows the estimation of global observables such as energy, by measuring only local observables, circumventing the experimental limitations.
- Related to that, we demonstrate in numerical simulations presented in our paper that the method

*dsafranekibs@gmail.com
†dario_rosa@ibs.re.kr

is highly efficient in situations relevant to state-of-the-art experimental capabilities. For example, two-qubit gates allowed the estimation of ground state energy below 4% of error, for the Heisenberg and Ising models. The Ising model is the model often used both in quantum simulators and quantum annealers.

- This is a very active field of research. The innovative method we developed provides a completely independent alternative to a highly impactful method of estimating expectation values using randomized measurements [2], Compared to their method, which requires measuring in a highly entangled basis, our method is significantly easier to implement in the current experimental platforms. This is because we do not put any experimental requirement on the types of measurements that experimenters can do. Thus, this method can be applied with the experimental capabilities currently at their disposal.

- At the same time, we delineate several new directions of investigation, which will spur additional theoretical research.

References

[1] Dominik Šafránek, Dario Rosa Expectation values from any quantum measurement arXiv: arXiv:2301.10428 [quant-ph], 2023.

[2] H.Y. Huang, R. Kueng, J. Preskill Predicting Many Properties of a Quantum System from Very Few Measurements J. Nat. Phys. 16, 1050–1057, 2023

Quantum-Relaxation Based Optimization Algorithms: Theoretical Extensions

Kosei Teramoto *† Rudy Raymond[2 1 3 ‡] Eyuri Wakakuwa[4 §] Hiroshi Imai[1 ¶]

[1] *Department of Computer Science, The University of Tokyo*
[2] *IBM Quantum, IBM Japan*
[3] *Quantum Computing Center, Keio University*
[4] *Department of Mathematical Informatics, Nagoya University*

Abstract. Quantum Random Access Optimizer (QRAO), proposed by Fuller et al., utilizes Quantum Random Access Code (QRAC) to encode binary optimization variables in a single qubit. Our research extends quantum-relaxation with a different QRAC that encodes three classical bits into two qubits, achieving an improved approximation ratio of 0.722 for the maximum cut problem, albeit with a reduced bit-to-qubit compression ratio, illustrating the inherent trade-off. Furthermore, we introduce a novel quantum relaxation that consistently maintains a 2x bit-to-qubit compression ratio, unlike the original work by Fuller et al. These findings provide new insights into quantum-relaxation based optimization.

Keywords: Quantum Relaxation, Quantum Random Access Codes, Quantum State Rounding, Maximum Cut Problem, Quantum Approximability

1 Backgrounds

Solving optimization problems is one of the most important tasks for which quantum computation is expected to be useful. Various quantum algorithms have been devised for NP-hard optimization problems such as QAOA (Quantum Approximate Optimization Algorithms) [3] proposed by Farhi, Goldstone, and Gutmann, and VQE (Variational Quantum Eigensolver) [13] proposed by Peruzzo et al. Although QAOA and VQE are classical-quantum hybrid algorithms designed for near-term devices capable of running only noisy shallow circuits, there are some significant challenges. The first issue is scalability. Because QAOA and VQE encode one classical bit into one qubit and the number of qubits of near-term quantum devices is at most several hundred qubits, the problem instance sizes are highly limited. The second issue is that it is not yet clear if the *quantumness* (i.e., quantum entanglement) of constant-depth QAOA and VQE can yield better results than classical optimization algorithms, as indicated in [12]. In other words, for combinatorial optimization, QAOA and VQE may have limited appeal for execution on a quantum computer from the outset.

Recently, a new classical-quantum hybrid optimization algorithm, QRAO (Quantum Random Access Optimization) [4], was proposed by Fuller et al. to address the aforementioned issues. Specifically, the QRAO encodes multiple classical bits (up to three) into one qubit using the $(3,1)$-QRAC (Quantum Random Access Code) [1, 6]. Here, (m,n)-QRAC means the quantum random access codes which encodes m classical bits into n qubits. This constant-factor improvement in scalability enabled Fuller et al. to conduct experiments with QRAO on superconducting quantum devices to solve the largest instances of a maximum cut problem (up to 40 nodes using only 15 qubits). Moreover, since QRAO searches for quantum states that correspond to solutions to the relaxation problem, which expands the solution space and potentially includes infeasible solutions, the quantum state that is eventually discovered is an entangled state that cannot be directly interpreted as a classical solution. These methods, known as *quantum-relaxation*, have been extended for more general quadratic programs [16]. To obtain the classical solution, *quantum state rounding* of the relaxed solution must be performed. Therefore, compared to standard VQE methods, QRAO may benefit from quantum entanglement if the entangled states result in better relaxed values. In other words, QRAO is inherently different from standard quantum-classical hybrid algorithms like QAOA and may benefit from quantum mechanical properties. There is experimental evidence that entanglement can help QRAO find optimal solutions in some instances [14].

The quantum state rounding algorithm, known as *magic state rounding*, employed in QRAO is inspired by the approximation algorithm for the maximum cut problem proposed by Goemans and Williamson, which has an approximation ratio of 0.879 [5]. This algorithm randomly selects pairs of two-bit-inverted relationships and decodes the encoded bits into one of two candidate outcomes by executing the corresponding quantum measurement. Through quantum information theoretic analysis, it has been proven that the approximation ratio of quantum relaxation using $(3,1)$-QRAC is 0.555, while that using $(2,1)$-QRAC is 0.625 [4]. While the optimality of standard QAOA or VQE is often assumed when the ground state is achieved, the approximation ratios of QRAO are determined irrespective of whether the ground state is reachable. In other words, these ratios are guaranteed as long as the relaxed value of the obtained quan-

†kogupi93@gmail.com
‡rudyhar@jp.ibm.com
§e.wakakuwa@gmail.com
¶imai@is.s.u-tokyo.ac.jp

[1]This research is based on the work conducted while K.T. was affiliated with the Dept. of Computer Science, The Univ. of Tokyo until March 2023.

tum state surpasses that of the classical optimal value. This aspect is critical, as finding the exact ground state can be QMA-hard [8].

The first consideration is approximation ratio bounds. The approximation ratios of quantum relaxations using $(3,1)$- and $(2,1)$-QRAC suggest an inherent trade-off between space efficiency and approximability: the higher the space compression ratio, the lower the approximation ratio. Furthermore, the approximation ratio bound of QRAO is significantly below that of Goemans and Williamson's 0.879 [5], which has been proven to be optimal under the Unique Game Conjecture (UGC) [9]. This discrepancy arises because the success probability of decoding each bit of the QRACs used in QRAO is relatively low. Specifically, the success probability of decoding each encoded bit is approximately 0.85 for $(2,1)$-QRAC and roughly 0.79 for $(3,1)$-QRAC [1, 6]. We note that even if the approximation ratio of quantum relaxations is less than that of the Goemans-Williamson algorithms [5], it is worth noting that quantum relaxations tend to perform better for instances with a small *gain* (see Appendix Ⅲ of the original QRAO paper [4]). The 'gain' refers to a parameter indicating how much the optimal MaxCut value deviates from the trivial lower bound (half the number of edges). The challenge of evaluating this gain is known as the MaxCutGain problem [2], which is also hard to approximate under the UGC [10]. This, along with the existence of already successful classical approximation algorithms for the MaxCut problem such as Goemans-Williamson [5], underscores the intrigue and potential of quantum-relaxation based optimizers.

Additionally, another challenge that emerges is that the bit-to-qubit compression ratio in Fuller et al.'s QRAO diminishes below the anticipated value (that is, 3x when employing $(3,1)$-QRACs or 2x when employing $(2,1)$-QRACs) as the density of the graph instance escalates. This indicates a limitation in leveraging the space advantage of quantum relaxations in certain applications.

2 Our Results

In this study, we aim to address the above-mentioned issues regarding the approximability and the preservation of the bit-to-qubit compression ratio by enhancing quantum relaxation in two ways. To tackle the first problem, we incorporate the use of $(3,2)$-QRAC, which boasts a higher decoding success probability than $(3,1)$- or $(2,1)$-QRACs, enabling us to achieve a superior approximation ratio for the MaxCut problem, albeit with a slightly reduced bit-to-qubit compression ratio. As for the second challenge, we have developed a new quantum relaxation methodology that consistently ensures a 2x bit-to-qubit compression ratio, unlike the original quantum relaxation proposed by Fuller et al. In the following sections, we delve deeper into the outcomes associated with these measures. For an in-depth exploration of the technical details, we invite the readers to refer to [15]. We hope that our results lead to the analysis of the quantum approximability and practical efficiency of the quantum-relaxation based approaches.

2.1 Quantum Relaxation with Better Approximation Ratio Using $(3,2)$-QRACs

Firstly, we will show the formulation of the $(3,2)$-QRAC which encodes three classical bits into two qubits obtained by numerical calculation [7]. The success probability of decoding each encoded bit is $\frac{1}{2}+\frac{1}{\sqrt{6}} \approx 0.908$, and it is optimal among all $(3,2)$-QRACs based on the bound by Manvčinska and Storgaard [11]. Also, we extended the quantum relaxation by using this $(3,2)$-QRAC. The instance of the problem is encoded into the problem Hamiltonian, and the maximum eigenstate of the Hamiltonian is explored. By performing the quantum state rounding algorithm, we obtain the classical binary solution to the problem. Furthermore, we proved the approximation ratio bound of the above quantum-relaxation based optimization algorithm for the MaxCut problem as $\frac{13}{18} \approx 0.722$. The only assumption of the proof of the approximation ratio is the same as the one when using $(3,1)$- or $(2,1)$-QRACs, that is, the energy of the found candidate quantum state for the maximum eigenstate of the problem Hamiltonian exceeds the optimum value of the original problem instance. Although the space compression ratio of our quantum relaxation is $\frac{3}{2} = 1.5$ and is lower than the one using $(3,1)$- or $(2,1)$-QRACs, the approximation ratio bound is better. Our result is consistent with the trade-off between the space compression ratio and the approximability of the maximum cut problem. Though the obtained approximation ratio bound 0.722 is lower than that of Goemans and Williamson, the practical feasibility of quantum-relaxation based approaches is enhanced.

2.2 Space Compression Ratio Preserving Quantum Relaxation

To always guarantee the bit-to-qubit compression ratio of QRAO using $(3,1)$-QRAC is essential as in the original QRAO the ratio becomes lower as the density of the graph instance increases. This is because there is a constraint that the endpoints of each edge must be associated with different qubits. For example, if the graph instance is the complete graph, then the number of qubits needed to run QRAO is the same as the number of vertices. In such cases, the quantum-relaxation based optimizer has no space advantage against standard QAOA and VQE algorithms. In this research, we propose new types of encoding which encode up to two classical bits into a single-qubit by using the $(3,1)$-QRAC. The third encoded bit's position in $(3,1)$-QRAC corresponds to the parity of the two bits. This modification allows us to remove the constraint that the endpoints of each edge have to be assigned to different qubits. The space compression ratio of the algorithm is always 2x which is independent of the density of the graph instances. Unfortunately, non-trivial approximation ratio bound $\left(>\frac{1}{2}\right)$ does not exist generally. We calculate the approximation ratio of this new algorithm by using two parameters ϵ and λ as $\max\left\{\frac{81-14\sqrt{3}+14\sqrt{3}\lambda+8\epsilon}{81+162\epsilon}, \frac{27-14\lambda+12\epsilon}{27+54\epsilon}\right\}$. The parameter ϵ is defined by the equation $\mathrm{OPT} = \left(\frac{1}{2}+\epsilon\right)|E|$ where

OPT is the optimal cut value, and therefore ϵ quantifies the so-called MaxCutGain [2]. The parameter λ is the ratio of the edges whose endpoints are assigned to different qubits. By using the approximation ratio bound, we analyze the condition of the graph instance that our algorithm gives a non-obvious approximation ratio bound for the maximum cut problem.

References

[1] Andris Ambainis, Ashwin Nayak, Amnon Ta-Shma, and Umesh Vazirani. Dense quantum coding and quantum finite automata. *Journal of the ACM (JACM)*, 49(4):496–511, 2002.

[2] M. Charikar and A. Wirth. Maximizing quadratic programs: extending Grothendieck's inequality. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 54–60, 2004.

[3] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[4] Bryce Fuller, Charles Hadfield, Jennifer R Glick, Takashi Imamichi, Toshinari Itoko, Richard J Thompson, Yang Jiao, Marna M Kagele, Adriana W Blom-Schieber, Rudy Raymond, et al. Approximate solutions of combinatorial problems via quantum relaxations. *arXiv preprint arXiv:2111.03167*, 2021.

[5] Michel X Goemans and David P Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM (JACM)*, 42(6):1115–1145, 1995.

[6] Masahito Hayashi, Kazuo Iwama, Harumichi Nishimura, Rudy Raymond, and Shigeru Yamashita. (4, 1)-quantum random access coding does not exist—one qubit is not enough to recover one of four bits. *New Journal of Physics*, 8(8):129, 2006.

[7] Takashi Imamichi and Rudy Raymond. Constructions of quantum random access codes. In *Asian Quantum Information Symposium (AQIS)*, volume 66, 2018.

[8] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local Hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[9] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[10] Subhash Khot and Ryan O'Donnell. Sdp gaps and ugc-hardness for max-cut-gain. *Theory of Computing*, 5(4):83–117, 2009.

[11] Laura Mančinska and Sigurd AL Storgaard. The geometry of Bloch space in the context of quantum random access codes. *Quantum Information Processing*, 21(4):1–16, 2022.

[12] Giacomo Nannicini. Performance of hybrid quantum-classical variational heuristics for combinatorial optimization. *Physical Review E*, 99(1):013304, 2019.

[13] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):1–7, 2014.

[14] Kosei Teramoto, Rudy Raymond, and Hiroshi Imai. The role of entanglement in quantum-relaxation based optimization algorithms. *arXiv preprint arXiv:2302.00429*, 2023.

[15] Kosei Teramoto, Rudy Raymond, Eyuri Wakakuwa, and Hiroshi Imai. Quantum-relaxation based optimization algorithms: Theoretical extensions. *arXiv preprint arXiv:2302.09481*, 2023.

[16] Andrew Zhao and Nicholas C. Rubin. Quantum relaxation for quadratic programs over orthogonal matrices, 2023.

Non-Local and Quantum Advantages in Network Coding for Multiple Access Channels

Jiyoung Yun[1] * Ashutosh Rai[1] † Joonwoo Bae[1] ‡

[1] *School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST),* 291 *Daehak-ro, Yuseong-gu, Daejeon* 34141*, Republic of Korea*

Abstract. Devising efficient communication in a network consisting of multiple transmitters and receivers is a problem of immense importance in communication theory. Interestingly, resources in the quantum world have been shown to be very effective in enhancing the performance of communication networks. In this work, we study entanglement-assisted communication over classical network channels. We consider multiple access channels, an essential building block for many complex networks, and develop an extensive framework for n-senders and 1-receiver multiple access channels based on nonlocal games. We obtain generic results for computing correlation assisted sum-capacities of these channels. The considered channels introduce less noise on winning and more noise on losing the game, and the correlation assistance is classified as local (L), quantum (Q), or no-signaling (NS). Furthermore, we consider a broad class of multiple access channels such as depolarizing ones that admix a uniform noise with some probability and prove general results on their sum-capacities. Finally, we apply our analysis to three specific depolarizing multiple access channels based on Clauser-Horne-Shimony-Holt, magic square, and Mermin-GHZ nonlocal games. In all three cases we find significant enhancements in sum-capacities on using nonlocal correlations. We obtain either exact expressions for sum-capacities or suitable upper and lower bounds on them.

Keywords: Entanglement, Nonlocal games, Multiple Access Channels, Nonlocal Advantage

The quantum advantages in communication, for instance, quantum cryptographic protocols that establish the highest level of security without computational assumptions, have been recently extended to network information theory by showing that higher channel capacities beyond what the Shannon information theory dictates are possible by exploiting quantum and non-local correlations in channel coding. Channel capacities beyond local strategies are successfully shown for two-sender and two-receiver interference channels (ICs) [1, 2] and two-sender and one-receiver multiple access channels (MACs) [3, 4, 5].

The scenario of network communication is shared in common in the aforementioned network protocols where multiple senders, by applying non-signaling correlations, cooperate in channel coding and a noisy channel under the control of a malicious party is defined by a *nonlocal game*. We call a game *quantum pseudo-telepathy* if a quantum strategy wins it with certainty. Noise occurs whenever senders lose a game against the party. For instance, the Clauser-Horne-Shimony-Holt (CHSH) game can be applied, where a quantum strategy wins the game with probability $\cos^2(\pi/8) = 1/2+1/(2\sqrt{2}) \approx 85\%$ and a local strategy with probability $3/4 = 75\%$. We recall that the Popescu-Roherlich box wins the game with certainty. The magic square game is a psuedo-telepathy protocol, where quantum players win the game with certainty with the help of entangled states whereas classical ones do with probability $8/9$. Quantum senders with pre-shared entanglement encode messages such that they win a nonlocal game with the highest probability, possibly, beyond local strategies. Then, network channel capacities rely on both correlations shared by senders and their encoding to win the game with the best strategy.

There is a hierarchal structure among correlations: non-signaling, quantum, and local correlations. Consequently, channel coding with non-signaling correlations [1] or quantum correlations [2] against a noisy channel controlled by a malicious party with the CHSH game lead to a higher channel capacity over local strategies. When noise occurs by the magic square game, two quantum parties achieve a noise-free channel since entanglement suffices to allow them to win the game with certainty. In this sense, entanglement has been thus identified as a resource for quantum advantages [3]. Although quantum advantages in network information theory are thus evidenced, little is known about how to manip-

* jiyoungyun@kaist.ac.kr
† ashutosh.rai@kaist.ac.kr
‡ joonwoo.bae@kaist.ac.kr

ulate nonlocal encoding in order to achieve channel capacities. Moreover, from a communication-centric point of view it may not be realistic that a noise-free channel is defined when senders win a game; more generally, a noisy channel can be introduced even if senders win a game.

Summary of the results. *In the present contribution, we establish a generic framework for deriving channel capacities for n-sender multiple access channels. The framework applies to a general class of n-sender MACs, where a malicious party introduces distinct noise depending on whether senders win a game or not: any n-sender MAC falls into the consideration. Channel coding strategies with local (*L*), quantum (*Q*), and no-signaling (*NS*) resources are considered. We show the hierarchical structure of sum-capacities with local, quantum, and non-signaling encoding strategies. Remarkably, the bounds on sum-capacities are tight in most cases and also exact in some of them. The results may be compared to cases with particular channels studied in Ref. [3, 5], and significant improvements on sum-capacity separations are shown.*

Notations. Let $\mathcal{N}^{(A_1,...,A_n)\to B}$ denote a (discrete memoryless) n-sender MAC from n parties $\text{Alice}_1, ..., \text{Alice}_n$ to a single one, Bob. For $k \in \{1, ..., n\}$, we write by $x_k \in \mathcal{X}_k$ a message of Alice_k and by $y \in \mathcal{Y}$ of Bob. Then, an n-sender MAC corresponds to a mapping

$$\mathcal{N} : \mathcal{X}_1 \times \mathcal{X}_2 \times \cdots \times \mathcal{X}_n \to \mathcal{Y} \tag{1}$$

and is fully characterized by conditional probabilities $P_{\mathcal{N}}(y|x_1, ..., x_n)$. Then, let R_k denote an achievable rate for a party Alice_k and the sum rate satisfies that

$$R_1 + \cdots + R_n \leq I(A_1, A_2, \cdots, A_n : B) \tag{2}$$

In addition, channel coding $\text{E} : (m_1, ..., m_n) \to (x_1, ..., x_n)$, described by $P_{\text{E}}(x_1, ..., x_n|m_1, ..., m_n)$, that applies resources such as local, quantum, and non-signaling correlations may be incorporated so that a channel may be defined as, $\mathcal{N}^{\text{E}} := \mathcal{N} \circ \text{E}$. Then, the sum capacity for a channel $\mathcal{N}$ with channel coding assisted by correlations $\mathcal{R} \in \{\text{L}, \text{Q}, \text{NS}\}$ is given by

$$\mathcal{C}^{(\mathcal{R})}(\mathcal{N}^{\text{E}}) = \max_{\pi(m_1,...,m_n)} \left\{ \max_{\text{E}\in\mathcal{R}} I^{(\pi,\ \text{E})}(A_1, ..., A_n;\ B) \right\} \tag{3}$$

where the maximization runs over probabilities $\pi(m_1, ..., m_n) = \pi_1(m_1) \cdots \pi_n(m_n)$ and channel coding strategies E over a resource $\mathcal{R}$.

Let G denote a game. If encoded bits $(x_{11}, x_{12};\ ...\ ; x_{n1}, x_{n2}) \subset \mathcal{X}_{11} \times \mathcal{X}_{12} \times\ ...\ \times \mathcal{X}_{n1} \times \mathcal{X}_{n2}$ win the game, we write that $(x_{11}, x_{12};\ ...\ ; x_{n1}, x_{n2}) \in \mathcal{W}_G$. Then, a family of depolarizing MACs based on a nonlocal game G is defined as follows

$$\mathcal{N}_G(y_1, ..., y_n|x_{11}, x_{12};\ ...\ ; x_{n1}, x_{n2}) = \begin{cases} \eta\left(\prod_{k=1}^{n} \delta_{(x_{k1},y_k)}\right) + \frac{1-\eta}{\Delta}, & \text{if } (x_{11}, x_{12};\ ...\ ; x_{n1}, x_{n2}) \in \mathcal{W}_G \\ \frac{1}{\Delta} & \text{otherwise,} \end{cases} \tag{4}$$

where $\Delta = d^n$, $0 \leq \eta < 1$, and δ is the delta function.

Now let us consider a n-party nonlocal game G. The k^{th} player receives a question $x_{k1} \in \mathcal{X}_{k1}$ and answers with $x_{k2} \in \mathcal{X}_{k2}$. The players win on satisfying some winning condition $(x_{11}, x_{12};\ ...\ ; x_{n1}, x_{n2}) \in \mathcal{W}_G \subset \mathcal{X}_{11} \times \mathcal{X}_{12} \times\ ...\ \times \mathcal{X}_{n1} \times \mathcal{X}_{n2}$. Let us consider that respective question and answer set to each player are $\{0, 1, ..., d-1\}$ and $\{0, 1, ..., D-1\}$.

Main results. In what follows, we summarize the key results for a family of depolarizing MACs based on nonlocal games. We outline how these results can be applied to channels where noise occurs by the CHSH nonlocal game [6], the magic square game [7], and the Mermin-GHZ game [7, 8, 9].

Result 1. For depolarizing channels based on nonlocal game G the mutual information between input and output to the channels is given by

$$I(X;Y) = H(Y) - \log_2 \Delta\ +\ \omega\ F(\Delta, \eta), \tag{5}$$

where a winning probability is denoted by $\omega = \sum_{x\in\mathcal{W}_G} p(x)$ and

$$\begin{aligned} F(\Delta, \eta) \ &\equiv\ \log_2 \Delta + \frac{1+(\Delta-1)\eta}{\Delta} \log_2 \frac{1+(\Delta-1)\eta}{\Delta} \\ &+\ (\Delta-1)\frac{1-\eta}{\Delta} \log_2 \frac{1-\eta}{\Delta}. \end{aligned} \tag{6}$$

Result 2. For depolarizing channels based on nonlocal games G, for any encoding E and any probability distribution $\pi(m)$ for the message set, an absolute upper bound on sum-rate is give by

$$I(M;Y) \leq F(\Delta,\eta)$$

and the bound is achieved if $H(Y) = \log_2 \Delta$, $\omega = 1$, and $I(M;Y) = I(X;Y)$.

Given a game G, we write $r_{max} = \omega_{\mathrm{L}}^* \Delta$ for a highest winning probability ω_{L}^* with classical strategies for uniformly distributed questions.

Result 3. For depolarizing channels based on a game G and probability distribution $\pi(m)$ over message and encodings $\mathrm{E} \in \mathrm{L}$, a sum capacity assisted by local correlations is given by

$$\mathcal{C}^{(\mathrm{L})}(\mathcal{N}_G) \leq \max_{\pi(m)} \left\{ H(M) + F(\Delta,\eta) \max_{\mathcal{M}_{r_{max}}} \left\{ \sum_{m \in \mathcal{M}_{r_{max}}} \pi(m) \right\} \right\} - \log_2 \Delta \tag{7}$$

where $\mathcal{M} = \mathcal{M}_{r_{max}} \cup (\mathcal{M}_{r_{max}})^c$, of the set of all messages $\mathcal{M}$, such that $|\mathcal{M}_{r_{max}}| = r_{max}$.

In particular, let $\widetilde{G}$ denote a quantum pseudo-telepathy game. For some cases, the upper bound $F(\Delta,\eta)$ in Result 2 is tight and thus corresponds to the sum capacity.

Result 4. Consider a pseudo-telepathy-game $\widetilde{G}$ in which a resource set $\widetilde{\mathcal{R}}$ allows parties to win the game. Suppose that the local outputs of the winning strategy are uniformly random for every input. Then, for a depolarizing channel $\mathcal{N}_{\widetilde{G}}$, there exists an encoding assisted by the resource $\widetilde{R}$ that achieves the upper bound $F(\Delta,\eta)$. Thus, the sum-capacity of such channels is given by $\mathcal{C}(\mathcal{N}_{\tilde{G}}) = F(\Delta,\eta)$.

Let us apply the general results to specific MACs with the CHSH, the magic square, and the Mermin-GHZ games.

(i) The CHSH game. Consider two senders with message set $\{0,1\}$ and one receiver [6, 10] against a noisy channel with the CHSH game. Since non-signaling encoding wins the game with certainty, we have that from Result 4, the analytical expression for the no-signaling sum-capacities as $\mathcal{C}^{(\mathrm{NS})}(\mathcal{N}_{CHSH}) = F(4,\eta)$. From Result 3 an upper bound to $\mathcal{C}^{(\mathrm{L})}(\mathcal{N}_{CHSH})$ may be found, which is however not tight: we obtain the capacity with a different technique, see the full paper.

(ii) The magic square (MS) game. There are two senders with message set $\{0,1,2\}$ and one receiver [7, 10]. Quantum senders win the magic square game with certainty. It follows from Result 4 to have the analytical expression for quantum sum-capacities as $\mathcal{C}^{(\mathrm{Q})}(\mathcal{N}_{MS}) = F(9,\eta)$. The computation of the classical sum-capacities $\mathcal{C}^{(\mathrm{L})}(\mathcal{N}_{MS})$ becomes unfeasible. One can apply Result 3 to find upper bounds only.

(iii) The Mermin-GHZ (M-GHZ) game. There are n-senders with message set $\{0,1\}$ and one receiver [8, 9, 10]. Quantum senders win the Mermin-GHZ game with certainty and thus Result 4 applies to this case. An analytical expression for the quantum sum-capacities is computed as $\mathcal{C}^{(\mathrm{Q})}(\mathcal{N}_{M-GHZ}) = F(2^n,\eta)$. The computation of the classical sum-capacities $\mathcal{C}^{(\mathrm{L})}(\mathcal{N}_{M-GHZ})$ becomes unfeasible. One can have upper bounds to local capacities from Result 3.

To conclude, in the presented work [10] (the full technical version appended), we have developed a framework for analyzing sum-capacities of n-sender one-receiver multiple access channels based on nonlocal games and obtained generic results applicable to any such channels. We then modeled depolarizing channels in our framework and derived sum-capacities by considering three types of non-signaling encoding resource $\mathcal{R} \in \{L,\ Q,\ NS\}$. Finally, we showed sum-capacity separation results for some specific examples of nonlocal game based depolarizing channels. The general propositions derived in this work are applicable to a broad class of multiple access channels based on nonlocal games. Moreover, the method developed in this work can be applied to any multiple access channel based on

nonlocal games which introduces less (more) noise when its input satisfy winning (losing) condition in a corresponding nonlocal game. Our generic approach can be applied to study the sum-capacities, with resource sets $\mathcal{R} \in \{L, Q, NS\}$, of any bi-asymmetric n-senders and 1-receiver multiple accesses channels which introduces different degree of noise in the two branches of the channel.

References

[1] Y. Quek and P. W. Shor. *Quantum and superquantum enhancements to two-sender, two-receiver channels.* Phys. Rev. A **95**, 052329 (2017).

[2] J. Yun, A. Rai, and J. Bae. *Nonlocal Network Coding in Interference Channels.* Phys. Rev. Lett. **125**, 150502 (2020).

[3] F. Leditzky, M. A. Alhejji, J. Levin, and G. Smith. *Playing games with multiple access channels.* Nat. Commun. **11**, 1497 (2020).

[4] J. Notzel. *Entanglement enabled communication.* IEEE Journal on Selected Areas in Information Theory **1.2**, 401 (2020).

[5] A. Seshadri, F. Leditzky, V. Siddhu, and G. Smith. *On the separation of correlation-assisted sum capacities of multiple access channels.* IEEE International Symposium on Information Theory (ISIT), pp.2756-2761, (2022); see also arXiv:2205.13538 [cs.IT].

[6] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner *Bell nonlocality.* Rev. Mod. Phys. 86, 419 (2014).

[7] G. Brassard, A. Broadbent, and A. Tapp. *Quantum Pseudo-Telepathy.* Foundations of Physics **35**, 1877 (2005), (also on arXiv: quant-ph/0407221).

[8] D. M. Greenberger, M. A. Horne, and A. Zeilinger *Going beyond Bell's theorem, In Bell's Theorem.* Quantum Theory and Conceptions of the Universe (M. Kafatos, editor), pages 69–72, 1988.

[9] N. D. Mermin. *Quantum mysteries revisited.* American Journal of Physics **58(8)**, 731 (1990).

[10] J. Yun, A. Rai, and J. Bae. *Non-Local and Quantum Advantages in Network Coding for Multiple Access Channels.* arXiv: 2304.10792 [quant-ph] (2023).

Quantum Optimized Centroid Initialization (QOCI)

Nicholas R. Allgood[1] * Ajinkya Borle[1] † Charles K. Nicholas[1] ‡

[1] *University of Maryland Baltimore County*

Abstract. One of the major benefits of quantum computing is the potential to resolve complex computational problems faster than can be done by classical methods. There are many prototype-based clustering methods in use today, and the selection of the starting nodes for the center points is often done randomly. Clustering often suffers from accepting a local minima as a valid solution when there are possibly better solutions. We will present the results of a study to leverage the benefits of quantum computing for finding better starting centroids for prototype-based clustering.

Keywords: quantum, clustering, annealing, k-means, adiabatic

1 Introduction

Clustering of data must start from somewhere, and the well-known k-means and k-medoids are no exception. These methods of clustering select a center point to start grouping related data around them to form their clusters. In much of the literature on the subject, we find that the center points to start with are selected at random. To get the best and most accurate groups of related data, we must have accurate center points to formulate a cluster, whether it be a computed mean or a sample from the data. Many clustering methods employ an iterative verification of the selected center point to confirm it is the best candidate to be the centroid. [1]

We could perhaps save some computational overhead if we begin the clustering algorithm with the *best* possible center nodes computed prior to forming the cluster. In classical cases, this could be quite computationally expensive, significantly more so than using a random process to select the starting nodes. With the advantages of quantum computing, we could theoretically offload this process to a quantum computer, which would then calculate the best possible center nodes. A major benefit to doing this, other than potential performance increases, is even if a performance increase isn't found, we immediately know that we have the best possible set of center nodes and we simply have to continue forming the clusters by grouping pieces of related data. A final point is that we can take some comfort in knowing that with a quantum process, we are getting a true global minimum based on the entire sample data and avoid the well-known *local minima*.

2 Problem Formulation

Quantum annealers such as those used by D-Wave [2], rely on a formulation of problem as an Ising Hamiltonian or a Quadratic Binary Optimization Problem (QUBO) both of which are equivalent expressions of a problem. We formulate our QUBO from the input data which is heavily inspired by non-negative matrix factorization [3],[4]. One drastic change is that our formulation will allow both positive and negative real numbers.[5] In non-negative matrix factorization, we assume $V = WH$ where V the product of the two matrices W and H and both V and H contain only non-negative real numbers. Non-negative matrix factorization has an inherent cluster property where it will automatically cluster columns of input data $V = (v_1, ..., v_n)$. The approximation of V via $V \cong WH$ is obtained by finding W and H that minimize the error function $||V - WH||_F$, subject to $W \geq 0, H \geq 0$ and F being the Frobenius or L^2-norm. To encode our problem as a QUBO, we first created a series of *substitution variables* that represent different combinations of unknown variables for our W and H matrices.

3 Solution Formulation

With our formulation defined, we submit our QUBO to a specified solver. The solver works using an adiabatic process over a period of time and during that time our unknown variables are replaced by computed values from the specific solver. One of the major benefits is that the solver is examining a much larger range of combinations and the most correct solution is the one that corresponds to the lowest energy value. When we get our result, we are returned a result that is series of binary values. These binary values read together are the result with the left most qubit being designated the *sign qubit* and the remaining qubits used to represent the value [6]. In a similar fashion to Borle and Lomonaco, we also use a radix-2 approximation of the binary value which results in only supporting integers for the coordinates of a centroid.

4 Results

Using the free developer account provided by D-Wave, we have a monthly time limit we are allowed as such many of our sample sizes are limited in scope. Our results are utilizing random Gaussian blobs utilizing *make_blobs* from scikit-learn. We also ran our experiments using the MOTIF data set, which is a fairly large data set consisting of malware metadata primarily used for machine learning. [7] We also use and compare between three different processes provided by the D-Wave Ocean SDK: TABU, Simulated Annealing, and D-Wave's Hy-

*allgood1@umbc.edu
†aborle1@umbc.edu
‡nicholas@umbc.edu

brid BQM solver. [1] MOTIF is a multi-dimensional data set so to reduce our data to two usable dimensions, we used principal component analysis (PCA). To measure our clustering performance, we used a variety of common metrics utilized in cluster analysis: inertia, silhouette score, homogeneity, completeness and v-measure. As an additional metric, we also recorded the number of overall iterations k-means took when using the centroids. For cluster inertia and iterations, the lower value is more optimal where for silhouette, homogeneity, completeness, and v-measure a higher value is more optimal. We had an upper bound of the number of iterations allowed by k-means set to 10000.

4.1 Gaussian Centroids

Using random Gaussian centroids, we chose a seed value of 0 and a cluster size k of 3. The figures below show a comparison between random centroids and those generated from QOCI utilizing the TABU, simulated annealing, and Hybrid BQM processes.

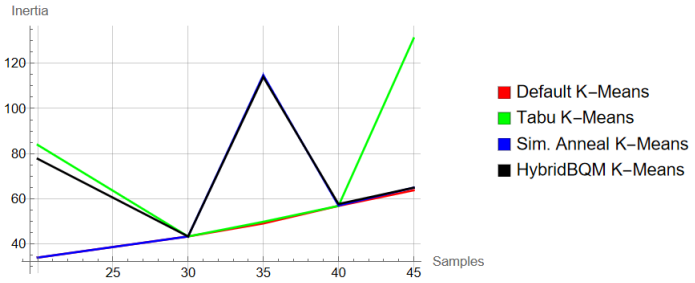


Figure 1: Gaussians: Inertia

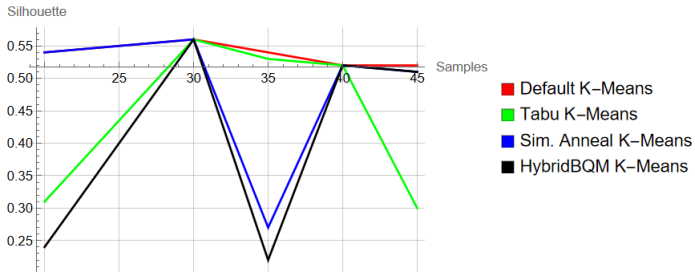


Figure 2: Gaussians: Silhouette Scores

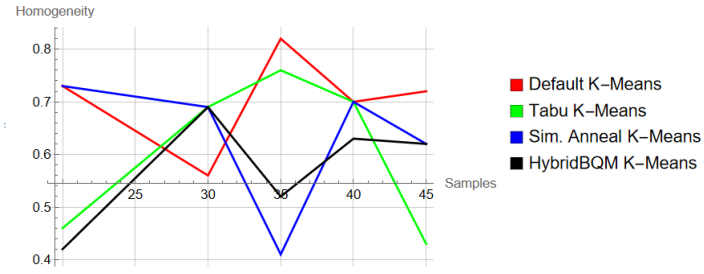


Figure 3: Gaussians: Homogeneity Scores

[1] https://www.dwavesys.com/resources/white-paper/hybrid-solvers-for-quadratic-optimization

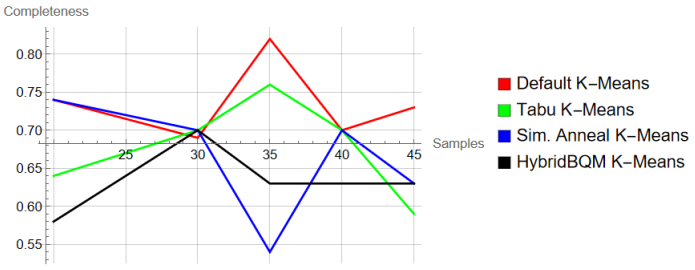


Figure 4: Gaussians: Completeness Scores

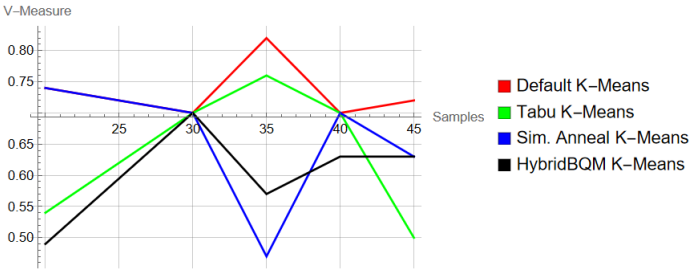


Figure 5: Gaussians: V-Measure Scores

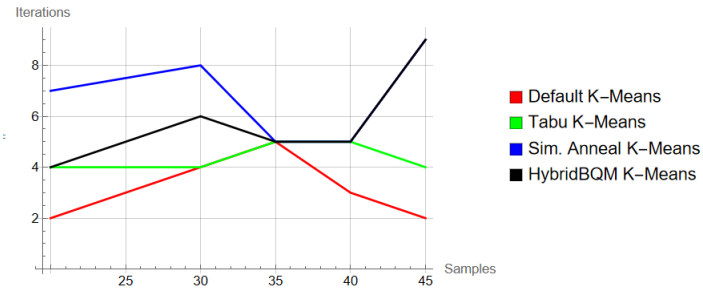


Figure 6: Gaussians: Iterations

4.2 MOTIF Centroids

Using randomly chosen data from the MOTIF data set, we chose a seed value of 0 and a cluster size k of 3. The figures below show a comparison between random centroids and those generated from QOCI utilizing the TABU, simulated annealing, and Hybrid BQM processes.

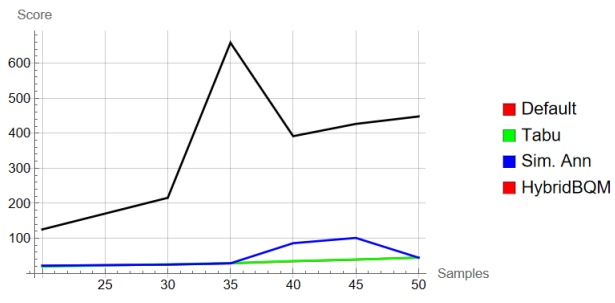


Figure 7: MOTIF: Inertia

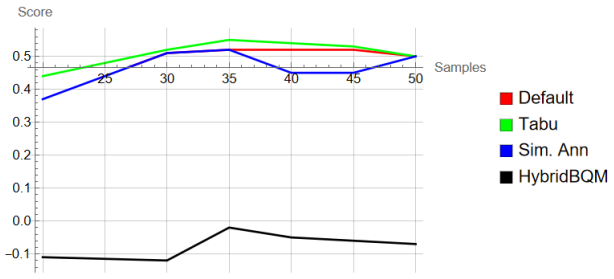


Figure 8: MOTIF: Silhouette Score

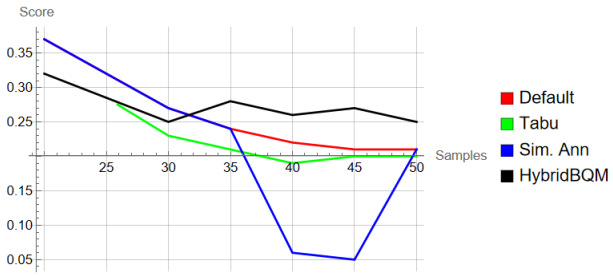


Figure 9: MOTIF: Homogeneity Score

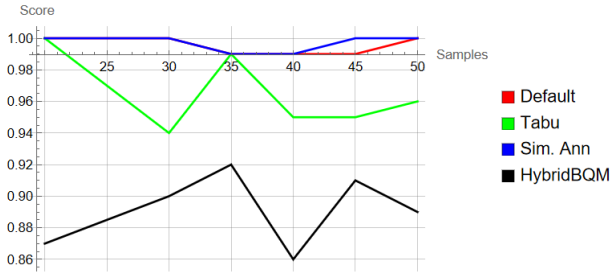


Figure 10: MOTIF: Completeness Score

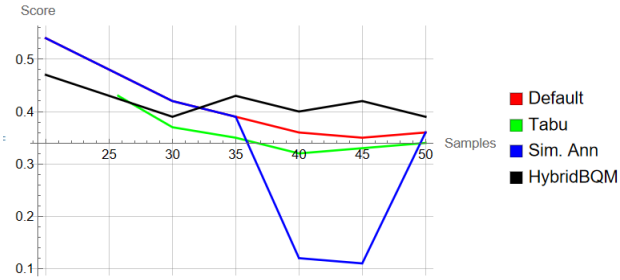


Figure 11: MOTIF: V-Measure

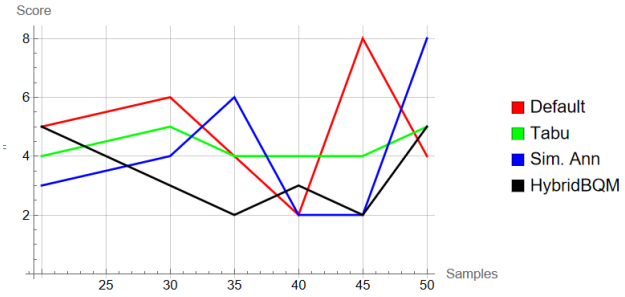


Figure 12: MOTIF: Iterations

5 Conclusion

To summarize our findings for the Gaussian data, all processes for cluster inertia were nearly the same with the exception of the TABU solver which proved to be the least optimal. For the silhouette score, the random centroids along with Hybrid BQM were both similar and the more optimal processes. For the remaining scores of homogeneity, completeness, v-measure and iterations, the random centroids were more optimal.

Summarizing our findings for the MOTIF data set, the inertia was similar on all processes except the Hybrid BQM which was the least optimal. This was also true when comparing the silhouette score. For the homogeneity score the Hybrid BQM process was more optimal. The completeness score was quite close only between simulated annealing and Hybrid BQM. With v-measure, the Hybrid BQM again was the more optimal process. In terms of number of iterations taken by k-means, random, TABU, and Hybrid BQM all were similar in terms of number of iterations while simulated annealing had the highest number of iterations.

References

[1] Dubes R. C. Jain A. K. *Algorithms for Clustering Data.* Prentice-Hall, 1988.

[2] D-Wave. D-wave. `https://dwavesys.com`, 2020.

[3] Jingu Kim and Haesun Park. Sparse nonnegative matrix factorization for clustering. Technical Report GT-CSE-08-01, Georgia Institute of Technology, 2008.

[4] Christian Bauckhage. k-means clustering is matrix factorization, 2015. `https://arxiv.org/abs/1512.07548`.

[5] Fred W. Glover and Gary A. Kochenberger. A tutorial on formulating QUBO models. *CoRR*, abs/1811.11538, 2018.

[6] Ajinkya Borle and Samuel J. Lomonaco. Analyzing the quantum annealing approach for solving linear least squares problems. *arXiv 1809.07649*, 2018.

[7] Robert J. Joyce, Dev Amlani, Charles Nicholas, and Edward Raff. MOTIF: A large malware reference dataset with ground truth family labels. *Computers and Security*, 124, Issue C, January 2023.

Quantum Capacities of Transducers

Chiao-Hsuan Wang[1 2 3 *] Fangxin Li[4] Liang Jiang[4]

[1]*Department of Physics and Center for Theoretical Physics, National Taiwan University, Taipei 10617, Taiwan*
[2]*Center for Quantum Science and Engineering, National Taiwan University, Taipei 10617, Taiwan*
[3]*Physics Division, National Center for Theoretical Sciences, Taipei, 10617, Taiwan*
[4]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

Abstract. High-performance quantum transducers, which faithfully convert quantum information between disparate physical carriers, are essential in quantum science and technology. Different figures of merit, including efficiency, bandwidth, and added noise, are typically used to characterize the transducers' ability to transfer quantum information. Here we utilize quantum capacity, the highest achievable qubit communication rate through a channel, to define a single metric that unifies various criteria of a desirable transducer. We investigate the optimal designs of quantum transduction schemes and show that the highest quantum capacity is achieved by transducers with a maximally flat conversion frequency response, analogous to Butterworth electric filters.

Keywords: Quantum Communications, Quantum Networks, Quantum Transductions, Quantum Channels, Quantum Frequency Conversions

1 Introduction

Classically, transducers are devices, such as antenna and microphones, that can convert signal from one physical platform to another. In quantum technology, transducers are essential elements that can faithfully convert quantum information between physical systems with disparate information carriers [1–3]. High-performance quantum transducers are the key to realize quantum networks [4–7] by interconnecting local quantum processors, such as microwave superconducting systems [8, 9], with long-range quantum communication carriers, such as optical fibers [10]. Tremendous progress has been made in a variety of coherent platforms for microwave-to-optical [11–23], microwave-to-microwave [24, 25], and optical-to-optical [26–29] frequency conversion.

Coherent conversion of quantum information between distinct devices is a challenging task. A functional quantum transducer has to satisfy demanding criteria simultaneously — high conversion efficiency, broad bandwidth, and low added noise — and its performance has been characterized by these three figures of merit [30]. On the other hand, a unified metric to assess the quantum communication capability of transducers is lacking. For example, one transducer may have a high conversion efficiency but operates within a narrow bandwidth, another may allow broadband conversion at a lower efficiency. It is hard to compare their transmission capability given separate criteria.

Quantum capacity, the highest achievable quantum communication rate through a channel [31–34], provides a natural metric to characterize the performance of quantum transducers. Consider a generic direct quantum transduction process by propagating external signals through a coupled bosonic chain [35]. After sending an input signal through the transducer, the output signal will be a mixture of the input signal and environmental noise. Assuming the environmental noise is thermal and that the transducer has no amplification effect, the action of the transducer can be described as a bosonic thermal-loss channel that attenuates the input state and combines it with a noisy thermal state [36]. We can thus model direct quantum transducers as bosonic thermal-loss channels and evaluate their quantum capacities.

In this work, we use quantum capacity to assess the intrinsic quantum communication capability of transducers. Using the continuous-time pure-loss quantum capacities of transducers as benchmarks, we discover that the optimal designs of transducers are those with maximally flat frequency response around the unity-efficiency conversion peak. Under the physical constraint of a bounded maximal coupling rate $g_{\max}$ between the bosonic modes, the maximal continuous-time quantum capacity $Q^{\max} \approx 31.4 g_{\max}$ is achieved by maximally flat transducers implemented by a long bosonic chain. We further include the effect of thermal noise from the environment by considering additive lower and upper bounds on quantum capacities of thermal-loss channels. Our methods provide a unified quantity to assess the performance of transducers across various physical platforms and suggest a fundamental limit on the quantum communication rate set by the physical coupling strength.

2 Capacity as a metric for transducers

We use the concept of quantum capacities of bosonic channels to assess the performance of direct quantum transducers. The quantum capacity quantifies the maximal achievable qubit communication rate through a quantum channel. Here we focus on direct quantum transduction achieved by directly converting quantum signals between bosonic modes via a coherent interface. At a given frequency ω in the appropriate rotating frame, assuming no intrinsic losses and no amplification gain, a direct quantum transducer with conversion efficiency $\eta[\omega]$ can be modeled as a Gaussian thermal-loss channel [36] described by the relation between the input and output

*chiaowang@phys.ntu.edu.tw

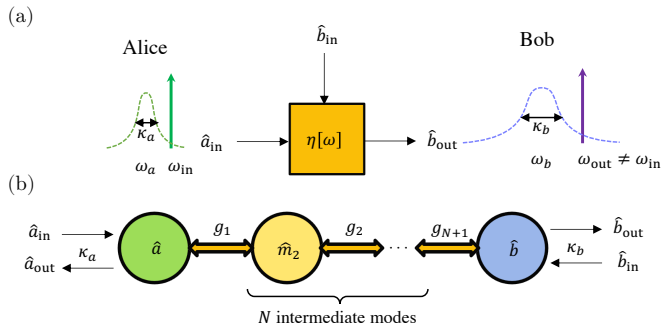


Figure 1: Generic model of quantum transducers. (a) A quantum transducer that can faithfully convert quantum states between different input and output frequencies ω_{in} and ω_{out} (in the lab frame), which is modeled as a thermal-loss channel with transmittance $\eta[\omega]$. (b) Schematic of a N-stage quantum transducer through a coupled bosonic chain connected to external input and output signals.

modes, up to phase shifts,

$$\hat{b}_{\text{out}}[\omega] = \sqrt{\eta[\omega]}\hat{a}_{\text{in}}[\omega] - \sqrt{1-\eta[\omega]}\hat{b}_{\text{in}}[\omega], \tag{1}$$

where $\hat{a}_{\text{in}}[\omega]$ is the input signal mode sent out by Alice, $\hat{b}_{\text{out}}[\omega]$ is the output signal mode received by Bob, and $\hat{b}_{\text{in}}[\omega]$ is the noisy input state from the environment with a mean thermal photon number $\bar{n}[\omega] = \left\langle \hat{b}^\dagger_{\text{in}}[\omega]\hat{b}_{\text{in}}[\omega] \right\rangle$ (see Fig. 1(a)). Note that we have no access to the reflective signal at Alice's side.

When the thermal photon number from the environment is negligible, $\bar{n} \approx 0$ for optical systems or via cooling [25, 37], this special case of thermal-loss channels is called the pure-loss channel. For pure-loss channels, their capacities are additive and can be analytically determined. Specifically, for one-way quantum communication (for example, from Alice to Bob only), for discrete-time signals at a given frequency ω with a fixed conversion efficiency $\eta[\omega]$, the one-way pure-loss capacity is given by [38]

$$q_1[\omega] = \max\left\{\log_2\left(\frac{\eta[\omega]}{1-\eta[\omega]}\right), 0\right\}, \tag{2}$$

which is the maximal amount of quantum information that can be reliably transmitted per channel use. This channel has infinite quantum capacity for ideal conversions, $\eta \to 1$, $q_1 \to \infty$, and has vanishing capacity when more than half of the signal is lost, $\eta \in [0, 1/2)$, $q_1 = 0$.

In reality, a quantum transducer has a finite conversion band and the conversion efficiency should be frequency-dependent. Treating different frequency modes within the conversion band as parallel quantum channels and taking the continuous limit in ω, here we define a continuous-time one-way pure-loss capacity of a quantum transducer,

$$Q_1 \equiv \int q_1[\omega]d\omega. \tag{3}$$

In contrast to the discrete-time one-way pure-loss capacity expression Eq. (2) that quantifies the maximal achievable quantum communication rate per channel use, the continuous-time quantum capacity defined in Eq. (3) is the maximal amount of quantum information that can be reliably transmitted through the transducer per unit time. This form of capacity is a direct analog to the Shannon capacity of classical continuous-time communication channels subject to frequency-dependent uncorrelated noises [39].

If the pure-loss channel is further assisted by two-way classical communication (between Alice and Bob) and local operations, the corresponding discrete-time two-way pure-loss capacity [40] is given by

$$q_2[\omega] = -\log_2\left(1-\eta[\omega]\right). \tag{4}$$

This channel again has infinite quantum capacity for ideal conversions, $\eta \to 1$, $q_2 \to \infty$, but has vanishing capacity only when the efficiency goes to zero, $\eta \to 0$, $q_2 = 0$. The corresponding continuous-time two-way pure-loss capacity is defined as

$$Q_2 \equiv \int q_2[\omega]d\omega. \tag{5}$$

The continuous-time pure-loss quantum capacities Q_1 and Q_2 defined above incorporate both concepts of efficiency and bandwidth and set the fundamental limit on the quantum communication rate based upon intrinsic transducer properties. To characterize these maximal achievable rates, we have assumed that infinite energy is available at the transducers. In practice, quantum capacities of transducers shall be lower in energy-constrained scenarios [41, 42]. We emphasize that Q_1 and Q_2 have the unit of qubits per second, and we will show in later text that these highest achievable communication rates are linked to the maximal coupling rates in the underling physical transducer system.

3 Physical limits on the capacities

The conversion efficiency of a transducer, $\eta[\omega]$, is determined by the parameters of its underlying physical implementation. We are interested in how the quantum capacities of transducers Q_1 and Q_2 are limited by the physical parameters of the transduction platform. Consider the generic model of direct quantum transducer [11–25, 27] implemented by a coupled bosonic chain with $N+2$ bosonic modes $\hat{m}_j$, where the two end modes, $\hat{m}_1 = \hat{a}$ and $\hat{m}_{N+2} = \hat{b}$, are coupled to external signal input and output ports at rates $\kappa_1 = \kappa_a$ and $\kappa_{N+2} = \kappa_b$ respectively (see Fig. 1(b)). Coherent quantum conversion can be realized by propagating bosonic signals from mode $\hat{a}$ (at frequency ω_a) to mode $\hat{b}$ (at frequency ω_b) through N intermediate stages, and we call this interface a N-stage quantum transducer. The conversion efficiency of a N-stage transducer is a frequency-dependent function determined by system parameters [12, 35],

$$\eta_N = \eta_N[\omega](\kappa_a, \kappa_b, \{\Delta_j\}, \{g_j\}), \tag{6}$$

where Δ_j is the detuning of mode $\hat{m}_j$ in the rotating frame of the laser drive(s) that bridges the up- and down-conversions between the input and output signals, and g_j is the coupling strength of the beam-splitter type interaction between the neighboring bosonic pair $\hat{m}_j$ and $\hat{m}_{j+1}$. Here we have assumed the system has no intrinsic losses and we will take g_j's to be real and positive without loss of generality.

For realistic physical implementations, the coherent coupling between neighboring modes is typically the most demanding resource. Therefore, under the physical constraint $\forall j, g_j \leq g_{\max}$, we look for the optimized choice of parameters κ_a, κ_b, Δ_j's, and g_j's to achieve the maximal possible Q_1 and Q_2 for N-stage quantum transducers.

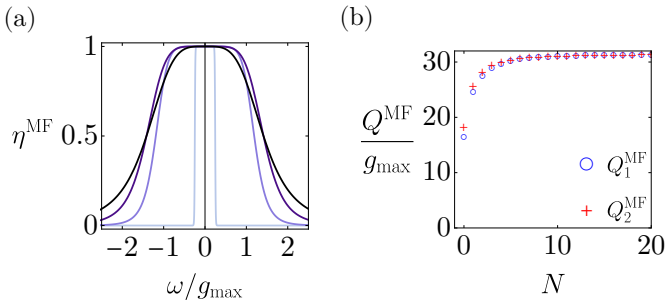


Figure 2: Diagrams for N-stage quantum transducers with maximally flat conversion efficiency. (a) Maximally flat efficiency function $\eta_N^{\mathrm{MF}}[\omega]$ for different N. (b) The continuous-time one-way pure-loss capacity, $Q_1^{N,\mathrm{MF}}$ (blue circle), and the continuous-time two-way pure-loss capacity, $Q_2^{N,\mathrm{MF}}$ (red cross), as a function of N.

Using the continuous-time pure-loss capacities as the benchmarks, we find that maximal values of Q_1 and Q_2 are achieved when the N-stage quantum transducer has a maximally flat (MF) efficiency (see Fig. 2(a)), a direct analog to a $(N+2)$-th order Butterworth low-pass electric filter. At large N, the continuous-time pure-loss quantum capacities saturate to the same value

$$\lim_{N\to\infty} Q_1^{N,\mathrm{MF}} = \lim_{N\to\infty} Q_2^{N,\mathrm{MF}} \equiv Q^{\max} = \frac{4\sqrt{3}\pi}{\log(2)} g_{\max}. \quad (7)$$

(See Fig. 2(b).) The above expression represents a physical limit on the maximal achievable quantum communication rate through a transducer, $Q^{\max} \approx 31.4 g_{\max}$ (qubit/sec). The quantum communication rate through a transducer is limited by the maximal available coupling strength within the bosonic chain.

4 Thermal Noise

For realistic transduction schemes within a noisy environment, the quantum capacity will decrease due to the effect of thermal noise. The quantum capacities of Gaussian thermal-loss channels have yet to be analytically determined, but we can approach their values using additive upper and lower bound expressions. We find that the quantum capacities of maximally flat transducers are less susceptible to thermal loss at large N, and the difference between the upper bound, lower bound, and $Q^{\max}$ also vanishes at large N. Based on the above property and numerical evidence, it is highly likely that maximally flat transducers are still optimal under the effect of thermal loss.

5 Discussion

We have used the continuous-time quantum capacities to characterize the performance of direct quantum transducers. By considering the generic physical model of an externally connected bosonic chain with a bounded coupling rate $g_{\max}$, we showed that the maximal qubit communication rate of a transducer is given by $Q^{\max} \approx 31.4 g_{\max}$. Such maximal capacity is achieved by maximally flat N-stage quantum transducers with $N \to \infty$. Note that our result has no contradiction to the Lieb-Robinson bound [43] — after signals arrive at a delayed time, increasing with N as predicted by Lieb and Robinson, the qubit communication rate is upper-bounded by the quantum capacity of the transducer that saturates to a finite value $Q^{\max}$ at large N in the optimal scenario.

This work provides a fundamental limit of transducer capacities in terms of coupling strength, and offers a quantitative comparison for direct transducers across platforms that consolidates distinct metrics of efficiency, bandwidth, and added thermal noise. Our method can be directly extended to transducers with intrinsic losses by considering the dependence of the conversion efficiency η_N on the intrinsic dissipation rates [12, 35]. Intriguing future works include exploring bosonic encodings, such as GKP codes [44], to approach the quantum capacity bound and investigating superadditivity of general quantum capacities.

Acknowledgements

We acknowledge support from the ARO (W911NF-18-1-0020, W911NF-18-1-0212), ARO MURI (W911NF-16-1-0349, W911NF-21-1-0325), AFOSR MURI (FA9550-19-1-0399, FA9550-21-1-0209), AFRL (FA8649-21-P-0781), DoE Q-NEXT, NSF (OMA-1936118, EEC-1941583, OMA-2137642), NTT Research, Packard Foundation (2020-71479), and the NSTC of Taiwan (111-2112-M-002-049-MY3).

Technical Version

For technical details, please refer to our published article in the appendix: Chiao-Hsuan Wang, Fangxin Li, and Liang Jiang. "Quantum capacities of transducers." Nature Communications 13(1): 6698, 2022. (https://doi.org/10.1038/s41467-022-34373-8)

References

[1] Nikolai Lauk, Neil Sinclair, Shabir Barzanjeh, Jacob P Covey, Mark Saffman, Maria Spiropulu, and

Christoph Simon. Perspectives on quantum transduction. *Quantum Sci. Technol.*, 5(2):020501, 2020.

[2] Nicholas J. Lambert, Alfredo Rueda, Florian Sedlmeir, and Harald G.L. Schwefel. Coherent Conversion Between Microwave and Optical Photons—An Overview of Physical Implementations. *Adv. Quantum Technol.*, 3(1):1900077, 2020.

[3] Xu Han, Wei Fu, Chang-Ling Zou, Liang Jiang, and Hong X. Tang. Microwave-optical quantum frequency conversion. *Optica*, 8(8):1050, 2021.

[4] Chip Elliott. Building the quantum network. *New J. Phys.*, 4(1):46, 2002.

[5] H. J. Kimble. The quantum internet. *Nature (London)*, 453(7198):1023–1030, 2008.

[6] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin. A quantum network of clocks. *Nat. Phys.*, 10(8):582–587, 2014.

[7] Christoph Simon. Towards a global quantum network. *Nat. Photonics*, 11(11):678–680, 2017.

[8] Alexandre Blais, Arne L. Grimsmo, S. M. Girvin, and Andreas Wallraff. Circuit quantum electrodynamics. *Rev. Mod. Phys.*, 93(2):025005, 2021.

[9] Atharv Joshi, Kyungjoo Noh, and Yvonne Y Gao. Quantum information processing with bosonic qubits in circuit QED. *Quantum Sci. Technol.*, 6(3): 033001, 2021.

[10] Hiroki Takesue, Shellee D. Dyer, Martin J. Stevens, Varun Verma, Richard P. Mirin, and Sae Woo Nam. Quantum teleportation over 100 km of fiber using highly efficient superconducting nanowire single-photon detectors. *Optica*, 2:832–835, 2015.

[11] Linran Fan, Chang Ling Zou, Risheng Cheng, Xiang Guo, Xu Han, Zheng Gong, Sihao Wang, and Hong X. Tang. Superconducting cavity electrooptics: A platform for coherent photon conversion between superconducting and photonic circuits. *Sci. Adv.*, 4(8):eaar4994, 2018.

[12] Yuntao Xu, Ayed Al Sayem, Linran Fan, Chang Ling Zou, Sihao Wang, Risheng Cheng, Wei Fu, Likai Yang, Mingrui Xu, and Hong X. Tang. Bidirectional interconversion of microwave and light with thin-film lithium niobate. *Nat. Commun.*, 12(1):1–7, 2021.

[13] Amir H Safavi-Naeini and Oskar Painter. Proposal for an optomechanical traveling wave phonon-photon translator. *New J. Phys.*, 13(1):013017, 2011.

[14] R. W. Andrews, R. W. Peterson, T. P. Purdy, K. Cicak, R. W. Simmonds, C. A. Regal, and K. W. Lehnert. Bidirectional and efficient conversion between microwave and optical light. *Nat. Phys.*, 10(4):321–326, 2014.

[15] A P Higginbotham, P S Burns, M D Urmey, R W Peterson, N S Kampel, B M Brubaker, G Smith, K W Lehnert, and C A Regal. Harnessing electro-optic correlations in an efficient mechanical converter. *Nat. Phys.*, 14(10):1038–1042, 2018.

[16] R. Hisatomi, A. Osada, Y. Tabuchi, T. Ishikawa, A. Noguchi, R. Yamazaki, K. Usami, and Y. Nakamura. Bidirectional conversion between microwave and light via ferromagnetic magnons. *Phys. Rev. B*, 93(17):174427, 2016.

[17] Na Zhu, Xufeng Zhang, Xu Han, Chang-Ling Zou, Changchun Zhong, Chiao-Hsuan Wang, Liang Jiang, and Hong X. Tang. Waveguide cavity optomagnonics for microwave-to-optics conversion. *Optica*, 7 (10):1291, 2020.

[18] Xu Han, Wei Fu, Changchun Zhong, Chang Ling Zou, Yuntao Xu, Ayed Al Sayem, Mingrui Xu, Sihao Wang, Risheng Cheng, Liang Jiang, and Hong X. Tang. Cavity piezo-mechanics for superconducting-nanophotonic quantum interface. *Nat. Commun.*, 11 (1):1–8, 2020.

[19] Mohammad Mirhosseini, Alp Sipahigil, Mahmoud Kalaee, and Oskar Painter. Superconducting qubit to optical photon transduction. *Nature (London)*, 588(7839):599–603, 2020.

[20] Jingshan Han, Thibault Vogt, Christian Gross, Dieter Jaksch, Martin Kiffner, and Wenhui Li. Coherent Microwave-to-Optical Conversion via Six-Wave Mixing in Rydberg Atoms. *Phys. Rev. Lett.*, 120(9): 93201, 2018.

[21] Jonathan R. Everts, Matthew C. Berrington, Rose L. Ahlefeldt, and Jevon J. Longdell. Microwave to optical photon conversion via fully concentrated rare-earth-ion crystals. *Phys. Rev. A*, 99(6):063830, 2019.

[22] John G. Bartholomew, Jake Rochman, Tian Xie, Jonathan M. Kindem, Andrei Ruskuc, Ioana Craiciu, Mi Lei, and Andrei Faraon. On-chip coherent microwave-to-optical transduction mediated by ytterbium in YVO4. *Nat. Commun.*, 11(1):1–6, 2020.

[23] Yuta Tsuchimoto, Zhe Sun, Emre Togan, Stefan Fält, Werner Wegscheider, Andreas Wallraff, Klaus Ensslin, Ataç İmamoğlu, and Martin Kroner. Large-bandwidth Transduction Between an Optical Single Quantum Dot Molecule and a Superconducting Resonator. *PRX Quantum*, 3(3), 2022.

[24] Baleegh Abdo, Katrina Sliwa, Flavius Schackert, Nicolas Bergeal, Michael Hatridge, Luigi Frunzio, A Douglas Stone, and Michel Devoret. Full coherent frequency conversion between two propagating microwave modes. *Phys. Rev. Lett.*, 110(17):173902, 2013.

[25] F Lecocq, J B Clark, R W Simmonds, J Aumentado, and J D Teufel. Mechanically Mediated Microwave Frequency Conversion in the Quantum Regime. *Phys. Rev. Lett.*, 116(4):043601, 2016.

[26] Francesco Morichetti, Antonio Canciamilla, Carlo Ferrari, Antonio Samarelli, Marc Sorel, and Andrea Melloni. Travelling-wave resonant four-wave mixing breaks the limits of cavity-enhanced all-optical wavelength conversion. *Nat. Commun. 2011 21*, 2 (1):1–8, 2011.

[27] Jeff T. Hill, Amir H. Safavi-Naeini, Jasper Chan, and Oskar Painter. Coherent optical wavelength conversion via cavity optomechanics. *Nat. Commun.*, 3(1):1–7, 2012.

[28] Kristiaan De Greve, Leo Yu, Peter L. McMahon, Jason S. Pelc, Chandra M. Natarajan, Na Young Kim, Eisuke Abe, Sebastian Maier, Christian Schneider, Martin Kamp, Sven Höfling, Robert H. Hadfield, Alfred Forchel, M. M. Fejer, and Yoshihisa Yamamoto. Quantum-dot spin–photon entanglement via frequency downconversion to telecom wavelength. *Nature (London)*, 491(7424):421–425, 2012.

[29] Peter Lodahl, Sahand Mahmoodian, and Soren Stobbe. Interfacing single photons and single quantum dots with photonic nanostructures. *Rev. Mod. Phys.*, 87(2):347–400, 2015.

[30] Emil Zeuthen, Albert Schliesser, Anders S. S rensen, and Jacob M Taylor. Figures of merit for quantum transducers. *Quantum Sci. Technol.*, 5(3):34009, 2020.

[31] Benjamin Schumacher and M. A. Nielsen. Quantum data processing and error correction. *Phys. Rev. A*, 54(4):2629–2635, 1996.

[32] Seth Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55(3):1613–1622, 1997.

[33] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Theory*, 51(1):44–55, 2005.

[34] Michael M Wolf, David Pérez-García, and Geza Giedke. Quantum capacities of bosonic channels. *Phys. Rev. Lett.*, 98(13):130501, 2007.

[35] Chiao-Hsuan Wang, Mengzhen Zhang, and Liang Jiang. Generalized matching condition for unity efficiency quantum transduction. *Phys. Rev. Res.*, 4 (4):L042023, 2022.

[36] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J Cerf, Timothy C Ralph, Jeffrey H Shapiro, and Seth Lloyd. Gaussian quantum information. *Rev. Mod. Phys.*, 84(2):621–669, 2012.

[37] Mingrui Xu, Xu Han, Chang Ling Zou, Wei Fu, Yuntao Xu, Changchun Zhong, Liang Jiang, and Hong X Tang. Radiative Cooling of a Superconducting Resonator. *Phys. Rev. Lett.*, 124(3):033602, 2020.

[38] A. S. Holevo and R. F. Werner. Evaluating capacities of bosonic Gaussian channels. *Phys. Rev. A*, 63 (3):1–14, 2001.

[39] Robert G. Gallager. *Information theory and reliable communication*. Wiley, New York, 1968. ISBN 978-0-471-29048-3.

[40] Stefano Pirandola, Riccardo Laurenza, Carlo Ottaviani, and Leonardo Banchi. Fundamental limits of repeaterless quantum communications. *Nat. Commun. 2017 81*, 8(1):1–15, 2017.

[41] Kunal Sharma, Mark M. Wilde, Sushovit Adhikari, and Masahiro Takeoka. Bounding the energy-constrained quantum and private capacities of phase-insensitive bosonic Gaussian channels. *New J. Phys.*, 20(6):063025, 2018.

[42] Kyungjoo Noh, Stefano Pirandola, and Liang Jiang. Enhanced energy-constrained quantum communication over bosonic Gaussian channels. *Nat. Commun. 2020 111*, 11(1):1–10, 2020.

[43] Elliott H Lieb and Derek W Robinson. The finite group velocity of quantum spin systems. *Commun. Math. Phys.*, 28(3):251–257, 1972.

[44] Daniel Gottesman, Alexei Kitaev, and John Preskill. Encoding a qubit in an oscillator. *Phys. Rev. A*, 64 (1):012310, 2001.