Posters

August 29, 2023 (Tue.) [Poster Session II]

Engines for predictive work extraction from memoryful quantum stochastic processes

Ruo Cheng Huang[1 2 *] Paul M. Riechers[1 2 †] Mile Gu[1 2 ‡] Varun Narasimhachar[1 2 §]

[1] *School of Physical and Mathematical Sciences, Nanyang Technological University, 637371 Singapore, Singapore*
[2] *Quantum Hub, Nanyang Technological University, 637371 Singapore, Singapore*
[3]*Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Singapore 138632*

Abstract. Quantum information-processing techniques enable work extraction from a system's inherently quantum features, in addition to its classical free energy. Meanwhile, the science of computational mechanics affords tools for the predictive modelling of non-Markovian stochastic processes. We combine tools from these two sciences to develop a technique for predictive work extraction from non-Markovian stochastic processes with quantum outputs. We demonstrate that this technique can extract more work than non-predictive quantum work extraction protocols, on one hand, and predictive work extraction without quantum information processing, on the other. We discover a phase transition in the efficacy of memory for work extraction from quantum processes, which has no classical precedent. Our work opens the prospect of machines that harness environmental free energy in an essentially quantum time-varying form.

Keywords: Quantum Thermodynamics, Stochastic Processes, Computational Mechanics, Temporal Correlations

1 Introduction

In the earliest heat engines, a combustible fuel was burned to maintain a temperature gradient between hot and cold heat reservoirs. The second law of thermodynamics holds that no engine can sustainably function with a single reservoir [1, 2, 3, 4]. While thought experiments such as Maxwell's demon and Szilard's engine initially appear to defy this law [5], a more complete understanding of thermodynamics resolved the apparent paradox: the resource powering the engine need not be a temperature gradient, but may be *any* form of free energy—even information [6, 7, 8, 9, 10]. The emerging field of quantum thermodynamics has continued to expand the scope of "fuel" to increasingly general forms of free energy. There has been both theoretical and experimental advancement in constructing engines that can harness the free energy locked up in quantum coherence, over and above classical free energy [11, 12, 13, 14, 15].

The story does not stop there—in addition to *static* fuel, there is also a *dynamical* fuel-like resource embodied by complex thermodynamic processes. The framework of *computational mechanics* in complexity science offers powerful techniques for the characterization and manipulation of stochastic processes. The future behaviour of such a process in general cannot be known perfectly using data from its past. Nevertheless, temporal correlations, i.e., patterns in a process's behaviour over time, enable prediction. These correlations may even be *non-Markovian*, whereby the future of a process depend not only on its present, but also on its distant past. Epsilon machines and their quantum extensions [16, 17, 18] perform memory-optimal predictive modelling of stochastic processes. Pattern extractors [14, 19] leverage prediction to extract useful work from the classical free energy present in temporal patterns, exchanging heat with a single bath. However, these predictive engines are not equipped to harness the free energy locked up in quantum degrees of freedom.

1.1 Our contributions

Here, we develop the theoretical prototype for a *predictive quantum engine*: a machine that charges a battery by feeding on a multipartite quantum system whose parts are temporally correlated via a classical stochastic process [20]. It can extract free energy beyond what is accessible to current quantum engines or classical predictive engines. We present a systematic construction of such an engine for arbitrary classical processes and quantum output states. We illustrate its application on example stochastic processes of correlated non-orthogonal qubits. (Fig. 1). We also use this test case to benchmark the performance of our engine against various alternatives, including one without coherent quantum information processing, and one without predictive functionality. Our predictive quantum engine outperforms these alternatives in terms of work output. We show that parametrized processes of correlated non-orthogonal quantum outputs exhibit phase boundaries between parametric regions where memory of past observations can and cannot enhance the work yield—despite the apparently smooth change of memoryful correlations in the process across this boundary. The sudden lack of memory advantage is thus fundamentally thermodynamic (since prediction per se has more freedom than during work extraction) and fundamentally quantum (since classical engines can exploit all the process's inherent memory). Finally, we generalize the Information Processing Second Law (IPSL) to the quantum regime and derive the

[*] ruocheng001@e.ntu.edu.sg
[†] pmriechers@gmail.com
[‡] gumile@ntu.edu.sg
[§] varun.achar@gmail.com

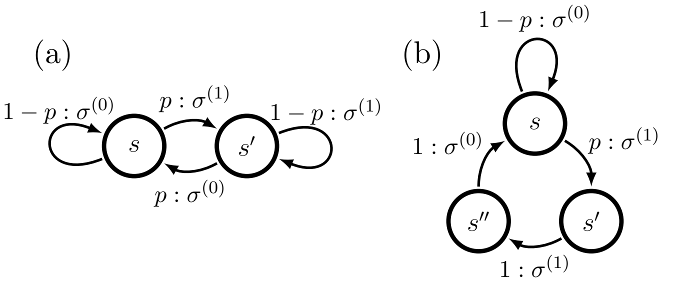


Figure 1: Latent-state sources of correlated quantum processes. Each arrow represents a transition between latent states; the label $p:\sigma^{(x)}$ indicates that the transition happens with probability p and produces a quantum state $\sigma^{(x)}$. (a) Perturbed-coin process. (b) 2-1 golden-mean process.

fundamental bounds on a quantum pattern engine's performance.

2 Framework

Rather than using memoryless quantum sources which produce independent and identically distributed (IID) quantum states at each discrete time [21, 22], we consider general *finite-state sources of quantum states* which can create nontrivial correlation across time. Some simple examples are depicted in Fig. 1. these memoryful sources of states can be represented by a hidden Markov Model (HMM). The states generated are separable but can have non-classical correlations in the form of discord [23]. These sources generalize the kindred 'classically controlled qubit sources' of Ref. [24].

We restrict the the engine to possess no quantum memory and limited classical memory. Each of the quantum states generated also has an immediate expiration date, hence the quantum states generated must be fed into the engine one at a time rather than storing everything for later time. We allow the HMM to be arbitrary in its alphabet, size, statistics as well as the quantum outputs' dimensionality and purity. Lastly, we assume that the source of the fuel tape is known exactly, which entails the complete knowledge of the HMM

The engine will operate relying on its internal memory which keeps track of "belief states", $\boldsymbol{\eta}_t$. The memory will allow the engine to predict what the next expected state, ξ_t, will be. The engine then attempts to extract work from the quantum states based on the identity of ξ_t. The battery storing the work will eventually be measured and the memory will be updated. The process proceeds cyclically as shown in Fig. 2.

3 Results

Here we provide a summary of our results.

1. We generalised the so-called "mixed-state presentation" into the quantum regime where the states are non-orthogonal [25, 26, 27, 28, 29].

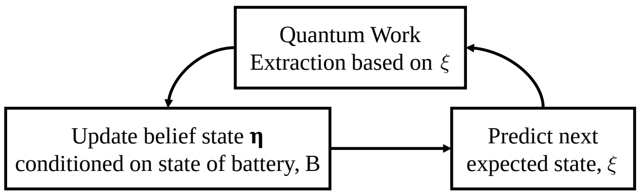

Quantum Work
Extraction based on ξ
Update belief state $\boldsymbol{\eta}$
conditioned on state of battery, B
Predict next
expected state, ξ

Figure 2: The protocol proceeds cyclically to fine-tune the belief state.

2. We demonstrated the superior performance of our engine by comparing it against other engines that lacked either memory or the ability to operate coherently on quantum states.
3. We discovered a phase transition in efficacy of knowledge in work extraction with respect to process parametrization. This along with the general performance of our engine can be found in Fig. 3
4. Finally, we provide a fundamental limit of this quantum pattern engine by invoking the quantum information processing second law to act as an upper bound.

4 Discussion

We developed the theoretical prototype for a *quantum pattern engine*: a machine that can adaptively extract useful work from quantum stochastic processes by exploiting knowledge of the temporal patterns they contain. We witnessed that, in the presence of coherence, the memory-assisted quantum approach will always outperform the memory-assisted classical approach. We also demonstrated its advantage over engines that can only harness static quantum resources. We found a phase transition marking the onset of memory advantage. It is an open question whether this phase transition coincides with the onset of quantum discord.

We found how to update the state of knowledge about any latent-state generator of a quantum process, given any POVM on the current quantum output. Furthermore, the fundamental thermodynamic limits of work extraction from correlated multi-partite quantum systems was found, setting the ultimate benchmark.

Despite the advances presented here, many open questions remain for future work.

Although designing the protocol guarantees maximal work extraction locally in time, it remains an interesting open question whether there is a superior steady-state approach that sacrifices short-term work extraction for greater knowledge and long-term returns. It may be possible to extend our method to more complex quantum processes, e.g., to those with entangled temporal correlations. This would, however, likely require a quantum memory. On the other hand, our method can immediately be adapted to applications where the pattern is spatial instead of temporal (e.g., states of many-body

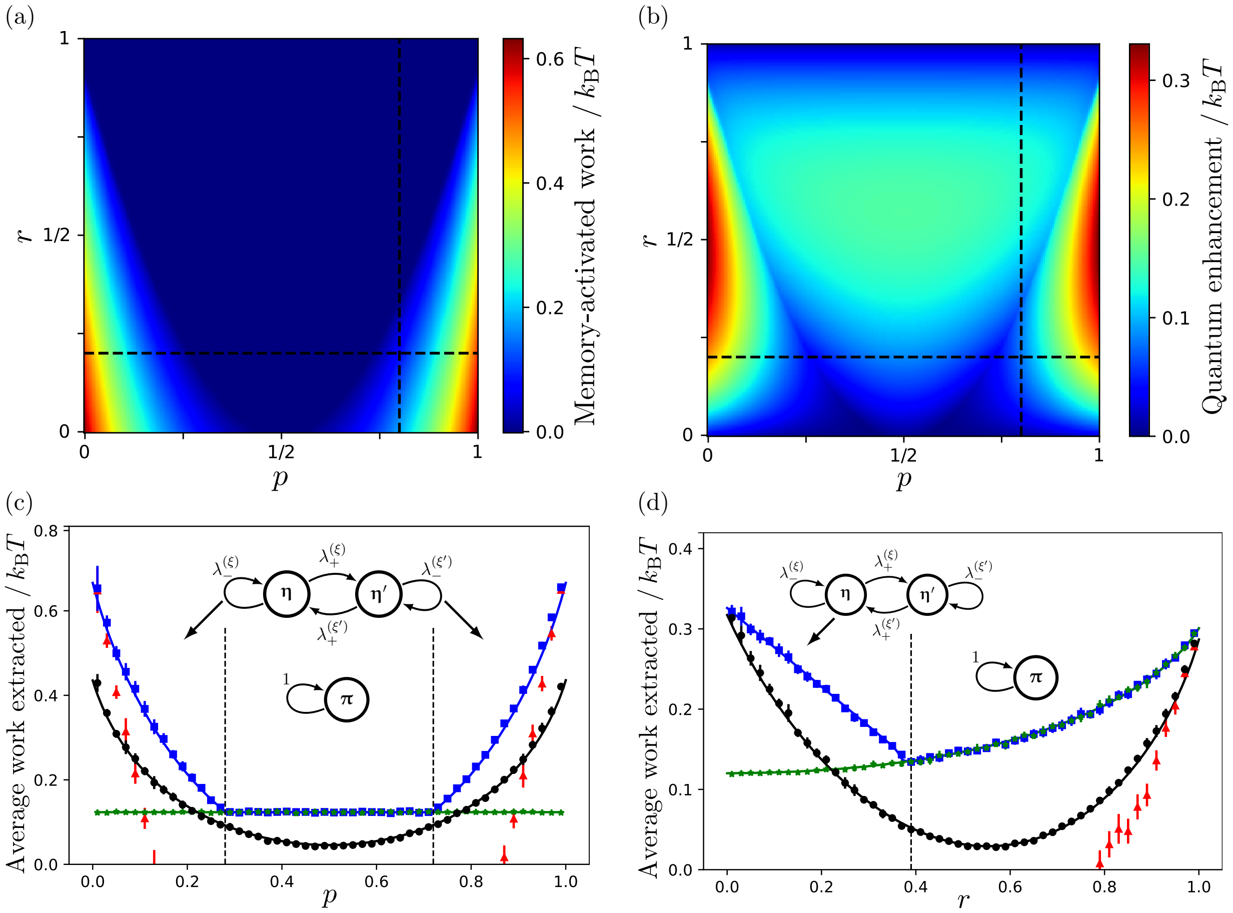


Figure 3: Comparison between average work-extraction rates of various approaches. (a) Memory enhancement of work extracted, compared to memoryless quantum approach. (b) Quantum enhancement of work extraction, compared to memoryful classical approach. Panels (c) and (d) reveal phase transitions in memory enhancement through cross-sections of parameter space. Analytic results (solid lines) and simulations (markers) are shown. Blue (squares) represents memory-assisted quantum approach; black (circles) represents memory-assisted classical approach; red (triangles) represents overcommitment approach; green (stars) represents memoryless quantum approach.

systems), and where the engine is constrained to operate locally on small regions at a time.

References

[1] Daniel V Schroeder. An introduction to thermal physics, 1999.

[2] Stephen J Blundell and Katherine M Blundell. *Concepts in thermal physics*. Oup Oxford, 2009.

[3] James Sethna. *Statistical mechanics: entropy, order parameters, and complexity*, volume 14. Oxford University Press, USA, 2021.

[4] Herbert B Callen. Thermodynamics and an introduction to thermostatistics, 1998.

[5] Leo Szilard. On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings. *Behavioral Science*, 9(4):301–310, 1964.

[6] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM journal of research and development*, 5(3):183–191, 1961.

[7] Charles H Bennett. The thermodynamics of computation—a review. *International Journal of Theoretical Physics*, 21(12):905–940, 1982.

[8] Dibyendu Mandal and Christopher Jarzynski. Work and information processing in a solvable model of Maxwell's demon. *Proceedings of the National Academy of Sciences*, 109(29):11641–11645, 2012.

[9] Dibyendu Mandal, HT Quan, and Christopher Jarzynski. Maxwell's refrigerator: an exactly solvable model. *Physical review letters*, 111(3):030602, 2013.

[10] Juan MR Parrondo, Jordan M Horowitz, and Takahiro Sagawa. Thermodynamics of information. *Nature physics*, 11(2):131–139, 2015.

[11] Paul Skrzypczyk, Anthony J Short, and Sandu Popescu. Work extraction and thermodynamics for

individual quantum systems. *Nature communications*, 5, 2014.

[12] Johan Åberg. Catalytic coherence. *Physical review letters*, 113(15):150402, 2014.

[13] Kamil Korzekwa, Matteo Lostaglio, Jonathan Oppenheim, and David Jennings. The extraction of work from quantum coherence. *New Journal of Physics*, 18(2):023045, 2016.

[14] Andrew JP Garner, Jayne Thompson, Vlatko Vedral, and Mile Gu. Thermodynamics of complexity and pattern manipulation. *Physical Review E*, 95(4):042140, 2017.

[15] Matteo Lostaglio. Thermodynamic laws for populations and quantum coherence: A self-contained introduction to the resource theory approach to thermodynamics. *arXiv preprint arXiv:1807.11549*, 2018.

[16] James P Crutchfield and Karl Young. Inferring statistical complexity. *Physical review letters*, 63(2):105, 1989.

[17] Cosma Rohilla Shalizi and James P Crutchfield. Computational mechanics: Pattern and prediction, structure and simplicity. *Journal of statistical physics*, 104(3):817–879, 2001.

[18] James P Crutchfield. Between order and chaos. *Nature Physics*, 8(1):17–24, 2012.

[19] Alexander B. Boyd, Dibyendu Mandal, and James P. Crutchfield. Correlation-powered information engines and the thermodynamics of self-correction. *Phys. Rev. E*, 95:012152, Jan 2017.

[20] Ruo Cheng Huang, Paul M Riechers, Mile Gu, and Varun Narasimhachar. Engines for predictive work extraction from memoryfull quantum stochastic processes. *arXiv preprint arXiv:2207.03480*, 2022.

[21] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2010.

[22] Philipp Strasberg, Gernot Schaller, Tobias Brandes, and Massimiliano Esposito. Quantum and information thermodynamics: A unifying framework based on repeated interactions. *Phys. Rev. X*, 7:021003, Apr 2017.

[23] Kavan Modi, Tomasz Paterek, Wonmin Son, Vlatko Vedral, and Mark Williamson. Unified view of quantum and classical correlations. *Phys. Rev. Lett.*, 104:080501, Feb 2010.

[24] Ariadna E Venegas-Li, Alexandra M Jurgens, and James P Crutchfield. Measurement-induced randomness and structure in controlled qubit processes. *Physical Review E*, 102(4):040102, 2020.

[25] James P Crutchfield. The calculi of emergence: computation, dynamics and induction. *Physica D: Nonlinear Phenomena*, 75(1-3):11–54, 1994.

[26] Christopher J Ellison, John R Mahoney, and James P Crutchfield. Prediction, retrodiction, and the amount of information stored in the present. *Journal of Statistical Physics*, 136(6):1005–1034, 2009.

[27] Sarah E Marzen and James P Crutchfield. Nearly maximally predictive features and their dimensions. *Physical Review E*, 95(5):051301, 2017.

[28] Paul M Riechers and James P Crutchfield. Spectral simplicity of apparent complexity. I. The nondiagonalizable metadynamics of prediction. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 28(3):033115, 2018.

[29] Alexandra M Jurgens and James P Crutchfield. Shannon entropy rate of hidden Markov processes. *Journal of Statistical Physics*, 183(2):1–18, 2021.

Swarm Intelligence for Routing on Quantum Repeater Networks

Hyensoo Choi[1] * Shigetora Miyashita[2] † Takahiko Satoh[3] ‡

[1] *Faculty of Environment and Information Studies, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-0882, Japan*

[2] *Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

[3] *Faculty of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama 223-8522, Japan*

Abstract. Ant colony optimization, a classical optimization heuristic, allows each node in the network to generate a routing table while none of each node has the knowledge of the entire path. It has been proposed with a quantum computer by Jannu *et al.* and extended to network routing protocols called AntNet by Di Caro *et al.*. This work applies the AntNet routing algorithm to generate and update routing tables of the quantum repeater network. Moreover, we also demonstrate it for a single routing table during local quantum repeater network operations. Our algorithm, namely quantum walking AntNet, makes the most of quantum networks by implementing edge selection based on quantum walks on a graph. The results pave the way for routing quantum repeater networks using a quantum computer.

Keywords: Ant colony optimization, quantum computing, quantum ant, quantum network, swarm intelligence

1 Introduction

Quantum networks have attracted considerable attention as they can teleport and process data based on quantum mechanical laws. Unlike classical networks that rely on bits, quantum networks use qubits that can simultaneously exist in an exponential number of states owing to the superposition. This theoretical basis leads to the development of several key ideas in practical implementation, such as the quantum internet architecture [Koz+23] that enable quantum communication among different nodes and quantum clock synchronization (QCS) [Tro+22] to achieve highly accurate and secure synchronization of clocks. Various researchers have demonstrated the protocols, such as China's Micius satellite experiment, but without considering quantum repeaters. Therefore, this work develops a path selection based on evaluating quantum links [Van+13] to build and maintain repeater-based quantum networks.

Here, we present an analysis of a novel routing algorithm *AntNet*, equipped with a quantum walk for easier discoverability of each node, especially over a mesh network topology. While the AntNet protocol broadcasts a massive amount of lightweight set of data and lets them hop around the network until the specific condition (e.g., the target node) is met to perform a standard random walk, we improve this by applying the rule of quantum walk in the moment of random path selection.

Like classical networks, it requires various technologies and research to be complete, covering diverse layers. One way to efficiently compute a blueprint of a new quantum networking protocol or algorithm is using a quantum network simulator, such as QuISP [Sat+22], SeQUeNCe [Wu+21] or NetSquid [Coo+21].

This algorithm can be applied to a quantum network. Still, we need to prepare a method to encode the concept of Ant into the quantum information to take advantage of quantum mechanics.

ACO relies on the random walk concept, and we see the potential for performance improvement by applying its quantum analogue, the quantum walk [Zha18]. We want to design an ACO algorithm variant by generating the entity of quantum walking ants in the algorithm and investigate the differences in terms of performance between our variant and classical ACO.

Many prior projects have shown the usage of classical AntNet in the physically extreme classical network environment, and the same rule also applies to the quantum network. One of the most widely-used networks that find its use in a hostile environment (e.g., under the ocean [Tro+22]) is the sensor network, a quantum sensor network in our case. Sensor nodes are often installed in extreme environments, or the connection goes through such locations. If then, we can think of several important physical locations requiring the installation of a sensor node, which is far from the gateway location but relatively close to the neighboring node. We need a quantum mesh network to easily deploy sensors wherever the network end node can reach [MZ21].

Lastly, one of the most significant advantages of ACO with quantum network routing is that it can leverage the power of the quantum walk algorithm. For instance, [Miy+23] demonstrated the effectiveness of quantum walk in exploring the relativistic spacetime. However, quantum walk can achieve the expansion of its functionality with the application for digraphs, and we can map into the existing routing problems.

[Zha18] showed us that Quantum Walk is performable on graphs and digraphs, which means it is applicable on the standard quantum network topology, just if we can run the quantum evolutionary algorithms on the physical network while hopping through edges. To achieve this, teleportation on the edges of quantum states is necessary, and we cannot perform in classical networks.

Our manuscript examines quantum repeaters within

*collodi@sfc.wide.ad.jp

†miyashita@atto.t.u-tokyo.ac.jp

‡satoh@ics.keio.ac.jp

the context of heterogeneous links. These systems involve distinct segments between repeaters, each possessing unique properties. These properties may pertain to the physical medium of the link (e.g., fiber optic versus free space), the length of the link, or the level of loss or noise inherent in the link. The design and implementation of quantum repeaters for such heterogeneous links pose significant challenges, as they necessitate handling diverse errors and inefficiencies across various links. To address these complexities, we introduce QAntNet—a novel algorithm developed specifically for quantum repeaters and premised on the principles of ant colony optimization. Within the QAntNet framework, we distribute 'ants' to identify the most optimal path. A distinctive attribute of this method is the entanglement of the ants with the (entangled) resources. We conceptualize this model as a quantum walk search operating on a quantum repeater network graph, enabling the optimal path selection.

2 Background

2.1 Routing in Classical Networks

Routing in the conventional network relies on the OSPF (Open Shortest Path First) protocol. The protocol is based on the distributive Dijkstra algorithm to generate a routing table regarding the link cost information. When each user performs a communication, the generated routing table is used to send the data packet to the destination using the shortest path possible.

Currently, the quantum network is actively adopting the qDijkstra algorithm to its routing protocol, which is precisely the same as the normal OSPF but includes the quantum link's unique cost calculation method to its evaluation factors.

2.2 Ant Colony Algorithm and AntNet

Ant colony algorithm or ant colony optimization algorithm (ACO) is mostly focused on finding the optimal path in the given space/network. Motivated by ants and their route-finding method, ACO introduces an abstract object called an ant and uses the state transition model to emulate the real ecology of ants on the graph [DBS06]. There are several algorithm variations, and they are called the ACO family.

Both ACO and Dijkstra are developed considering more generalized problems such as abstract graphs, but like OSPF, ACO also has its networking counterpart, AntNet [DD98]. This member of the ACO family is known to perform faster than OSPF on several network topologies [DD+98].

Following is the general description of the AntNet algorithm [DD98; Jia+12]:

The good quantum link is evaluated by the bell pair's fidelity and the cost of a link is determined by the average coherence time of the bell pair measured by each nodes over a time. Uptime can be measured by recording the begin/end time of the classical routine's quantum resource access.

Here, we explain how the conventional AntNet algorithm generates a routing table by passing a query function to its neighbours:

Walker Initialization. A starting node S generates a binary function f based on the query of the node user. The function f is a composition of several *commonly* predefined, primitive binary functions. (**e.g.** function $g(n)$ returns 1 if the node has more than n qubit. If not, $g(n)$ returns 0.)

First Evaluation. Node S runs f in itself. **If** 1**,** the procedure ends right there because the node S itself is the target. **If** 0**,** the node performs a pheromone database lookup, and picks a connection with highest pheromone value, which leads to the next node N. Now the starting node S passes function f to N with the list of visited nodes, concatenating itself to the list.

Next Node. When the arbitrary neighbouring node N receives a function f with additional data, it automatically executes the function part and repeats the procedure of node S, selecting the next node (by pheromone or random) and passing f with a list of received visited nodes, concatenated with N itself.

Peak. When the function f finally reaches the target node T, it will return $True$, and send the function+data back the way it came to the starting node, adding the pheromone value of every visiting node's edge-pheromone database.

Note: Every node owns its own key-value database table, where the key is a socket number of connection and value is a pheromone amount. With this method, for every single connection there are two key-value representing the pheromone of that connection, because every connections has two endpoints.

The above algorithm is classical. There are enough demonstrations of AntNet's unique functionalities, such as easy extensibility of the network without directly connecting every node to the gateway, which leads to less node installation cost (considering the extremely high cost of underground optical cable installation, even the slightest difference in cable usage can make a huge difference).

2.3 Routing in Quantum Networks

A quantum network is built for various purposes, and those are achieved by teleporting the quantum state without losing information. Entanglement Swapping, also known as two-qubit Bell state measurement (BSM), is how we teleport the quantum state from one location to another physically distant location. With this, the quantum network equips the ability to move the quantum information we want without violating the no-cloning theorem. Quantum routing protocol is a pre-defined set of promises between the network nodes in order to lead the message to the correct destination. [Pan+19; HPE19].

A method has been proposed for routing quantum repeater networks that uses the rate at which bell states are generated between neighboring repeaters. By defining the inverse of this rate as the link cost, we can generate the link table using the OSPF technique [Van+13].

3 Routing for Quantum Networks based on AntNet

This section briefly starts by explaining the advantage of quantum walks over classical random walks. We then discuss how to replace the classical AntNet with quantum walks.

3.1 Advantage of Quantum Walks over Random Walks

In this subsection, we briefly explain the advantage of quantum walks over classical random walks by reviewing the difference. Random walks are stochastic processes where the position of a walker moves through a graph probabilistically. The process is defined by a Markov chain of the form following according to Ref [Chi09]:

$$\frac{d}{dt}p(t) = (M - I)p(t). \tag{1}$$

This is nothing but a diffusion equation known as the lazy random walk. On the other hand, quantum walks are defined as a counterpart of random walks, and we can quantize the dynamics according to the Schrödinger equation.

$$i\frac{d}{dt}q(t) = Hq(t). \tag{2}$$

Here, the Hamiltonian H for a continuous-time quantum walk is defined by

$$H = -\gamma L - |w\rangle \langle w| \tag{3}$$

with a Laplacian matrix L. The Laplacian matrix comprises two sub-matrices: negative off-diagonal entries from the adjacency matrix A and positive diagonal elements from the degree matrix D such that $L = D - A$. These matrices allow us to search various types of graphs efficiently.

Quantum walks have been shown to offer exponential speedup in certain search problems compared to classical random walks. For example, a quantum walk-based search algorithm on an unsorted database can locate the desired item exponentially faster than classical random walks [CG04]. This is mainly because continuous-time quantum walks have a quadratic dispersion due to the Hamiltonian defined in Eq. (2), which makes the propagation faster than classical random walks. Furthermore, the propagation becomes even faster in discrete-time [Miy+23].

3.2 Proposal of AntNet with Quantum Walks

To utilize the advantage of the quantum walk mentioned above, we replace the random walking part of the conventional AntNet algorithm with the quantum walk.

The random walk used in ACO/AntNet is performing relatively independent actions in each node, and that means the edge selection process only cares about not visiting the previously visited nodes (and the pheromone value of every connected edge). However, in the case of Quantum Walk, the selection is determined by the unitary quantum operation on the arbitrary quantum state *passed* from the previous node.

From the quantum register's perspective, every node is required to have at least two quantum registers (e.g. quantum register A and B) for communication, to perform quantum teleportation with a previous node. For instance, when quantum register A receives a quantum state from a previous node, the current node will apply the quantum walk's unitary coin operator to the received quantum state and measure it to use the result to decide the next edge.

When the next destination is decided, the node will use another quantum register, B, to send the post-measurement quantum state to the next node on the quantum channel. The query function and history of visited nodes will be sent on the classical channel.

4 Conclusion

In conclusion, our manuscript presented an approach to finding the optimal path for quantum repeaters within heterogeneous link systems. We address the challenges inherent in such systems by proposing QAntNet, a quantum repeater-specific algorithm inspired by ant colony optimization. QAntNet leverages the optimal path by distributing ants through a quantum walk search. Since QAntNet can identify the optimal paths, it can be demonstrated to concrete graph problems for quantum repeaters across heterogeneous links. We anticipate that our findings will stimulate further advancements in the field of quantum communications, fostering the development of more efficient and robust quantum networks.

5 Acknowledgement

T.S. is supported by MEXT KAKENHI Grant Number 22K19781.

References

[Koz+23] Wojciech Kozlowski et al. *Architectural Principles for a Quantum Internet.* RFC 9340. Mar. 2023. DOI: 10.17487/RFC9340. URL: https://www.rfc-editor.org/info/rfc9340.

[Tro+22] James Troupe et al. “Quantum clock synchronization for future NASA deep space quantum links and fundamental science”. In: *arXiv preprint arXiv:2209.15122* (2022).

[Van+13] Rodney Van Meter et al. “Path selection for quantum repeater networks”. In: *Networking Science* 3 (2013), pp. 82–95.

[Sat+22] Ryosuke Satoh et al. “Quisp: a quantum internet simulation package”. In: *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE. 2022, pp. 353–364.

[Wu+21] Xiaoliang Wu et al. “SeQUeNCe: a customizable discrete-event simulator of quantum networks”. In: *Quantum Science and Technology* 6.4 (2021), p. 045027.

[Coo+21] Tim Coopmans et al. “Netsquid, a network simulator for quantum information using discrete events”. In: *Communications Physics* 4.1 (2021), p. 164.

[Zha18] Hanmeng Zhan. “Discrete quantum walks on graphs and digraphs”. In: (2018).

[MZ21] Adrià Mallorquí and Agustín Zaballos. “A heterogeneous layer-based trustworthiness model for long backhaul nvis challenging networks and an iot telemetry service for antarctica”. In: *Sensors* 21.10 (2021), p. 3446.

[Miy+23] Shigetora Miyashita et al. “Digital quantum simulator for the time-dependent Dirac equation using discrete-time quantum walks”. In: *arXiv preprint arXiv:2305.19568* (2023).

[DBS06] Marco Dorigo, Mauro Birattari, and Thomas Stutzle. “Ant colony optimization”. In: *IEEE computational intelligence magazine* 1.4 (2006), pp. 28–39.

[DD98] Gianni Di Caro and Marco Dorigo. “AntNet: Distributed stigmergetic control for communications networks”. In: *Journal of Artificial Intelligence Research* 9 (1998), pp. 317–365.

[DD+98] Gianni Di Caro, Marco Dorigo, et al. “An adaptive multi-agent routing algorithm inspired by ants behavior”. In: *Proceedings of PART98-5th Annual Australasian Conference on Parallel and Real-Time Systems*. 1998, pp. 261–272.

[Jia+12] Hao Jiang et al. “A quantum-inspired ant-based routing algorithm for WSNs”. In: *Proceedings of the 2012 IEEE 16th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE. 2012, pp. 609–615.

[Pan+19] Mihir Pant et al. “Routing entanglement in the quantum internet”. In: *npj Quantum Information* 5.1 (2019), p. 25.

[HPE19] Frederik Hahn, A Pappa, and Jens Eisert. “Quantum network routing and local complementation”. In: *npj Quantum Information* 5.1 (2019), p. 76.

[Chi09] Andrew M. Childs. “On the Relationship Between Continuous- and Discrete-Time Quantum Walk”. In: *Communications in Mathematical Physics* 294.2 (Oct. 2009), pp. 581–603. DOI: `10.1007/s00220-009-0930-1`. URL: `https://doi.org/10.1007%2Fs00220-009-0930-1`.

[CG04] Andrew M. Childs and Jeffrey Goldstone. “Spatial search by quantum walk”. In: *Physical Review A* 70.2 (Aug. 2004). DOI: `10.1103/physreva.70.022314`. URL: `https://doi.org/10.1103%2Fphysreva.70.022314`.

[Gho+22] Mrityunjay Ghosh et al. “A novel quantum algorithm for ant colony optimisation”. In: *IET Quantum Communication* 3.1 (2022), pp. 13–29.

[Jan+22] Srikanth Jannu et al. “Energy Efficient Quantum-Informed Ant Colony Optimization Algorithms for Industrial Internet of Things”. In: *IEEE Transactions on Artificial Intelligence* (2022).

[LN22] Linh Le and Tu N Nguyen. “DQRA: Deep quantum routing agent for entanglement routing in quantum networks”. In: *IEEE Transactions on Quantum Engineering* 3 (2022), pp. 1–12.

[Cal17] Marcello Caleffi. “Optimal routing for quantum networks”. In: *Ieee Access* 5 (2017), pp. 22299–22312.

Quantum Guesswork: a combinatorial instance of quantum hypothesis testing

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. We consider a game-theoretical scenario involving two parties, say Alice and Bob. At each round, Alice chooses a quantum state from a given ensemble, known to both parties, and sends it to Bob. Bob is allowed to perform any quantum operation on the state and to query Alice multiple times, one state at a time, until he correctly guesses the state. The game is repeated many times, and Bob's aim is to minimize the average number of queries needed. This problem, known as quantum guesswork, can be reframed as an instance of quantum hypothesis testing, and has therefore long been conjectured not to admit analytical solutions except for the cases in which the hypothesis testing problem is solvable, that is, for binary and symmetric ensembles.

Here, we disprove such a belief by deriving conditions under which the guesswork problem can be recast as a combinatorial problem, that is, an optimization over a finite set, and therefore can be solved analytically by exhaustive search. We further show that such conditions are verified by any qubit ensemble, thus conclusively settling the problem in dimension two, and we show that in that case the guesswork is equivalent to a combinatorial problem known as quadratic assignment problem (QAP). Finally, we introduce the (infinite) class of so-called benevolent qubit ensembles, which includes symmetric, informationally complete (SIC) and mutually unbiased basis (MUBs) ensembles, and we explicitly solve the corresponding QAP for such a class.

This presentation is based on Refs. [4, 7].

Keywords: quantum guesswork, quantum hypothesis testing, quadratic assignment problem

The guesswork problem can be conveniently framed as a theoretical game involving two parties, say Alice and Bob. At each round, Alice and Bob are given a classical and a quantum state, respectively, the latter solely dependant on the former. Bob queries Alice one state at a time until he correctly guesses her classical state, at which point he pays a cost that solely depends on the number of queries he had to perform. The probability distribution according to which classical states are sampled, the mapping between classical and quantum states, as well as the cost function, are known in advance to the parties. The minimum guesswork problem consists of the minimization of Bob's average cost.

For almost a decade, obtaining closed-form solutions for this problem has been considered impossible. In this respect, while discussing quantum hypothesis testing in relation to the guesswork problem, the Authors of Ref. [2] write (verbatim):

> *Closed-form result or optimal measurement is only known for some special quantum systems, e.g., the case with exactly two states, equiprobable symmetric states, or multiply symmetric states. We believe that it is also the case for minimum guesswork discrimination, because of the analogy between these two problems.*

Following this line of thought, Ref. [3] resorts to numerical techniques to approximately compute the guesswork, even for ensembles as "simple" as the trine and square ensembles of a qubit. The Authors therein write

[*] michele.dallarno.mv@tut.jp

> *Moreover, the SDP in (87) can be solved numerically to obtain the same value, providing a matching numerical lower bound; see [20] for the code involved, including a high-precision demonstration using the SDP solver SDPA-GMP [37] showing agreement to 200 digits.*

As recently as November 2022, Ref. [6] appeared which still resorts to numerics to approximate the value of the guesswork for the square ensemble.

In this contribution, we disprove this belief by deriving the closed-form expression of the guesswork for a class of ensembles that includes, for instance, any qubit ensemble (not necessarily binary nor symmetric) with uniform probability distribution. To achieve this, we demonstrate and exploit a symmetry of the guesswork problem (technically, a Z_2-covariance) that has been so far overlooked. It is important to remark that such a symmetry is not inherited by the ensemble (which may as well be not symmetric at all), but instead it is "injected" by the structure of the guesswork problem itself.

In the qubit case, our solution recasts the guesswork, by definition a continuous optimization problem on Bob's quantum strategy, as a finite combinatorial problem known as the quadratic assignment problem (QAP) [8, 9, 10], whose solution can therefore be found by exhaustive search.

As such, the QAP by itself represents a closed-form solution of the guesswork problem, but it can further be solved for the class of so-called benevolent ensembles, which includes regular polygons, SICs and MUBs. For instance, from our result the following closed-form expression for the guesswork $G(\boldsymbol{\rho})$ of any regular N-polygon qubit ensemble $\boldsymbol{\rho}$ follows:

$$G\left(\boldsymbol{\rho}\right) = \frac{N+1}{2} - \frac{1}{2}\begin{cases} \frac{2\sqrt{3\cos\left(\frac{\pi}{N}\right)^2+1}}{N\sin\left(\frac{\pi}{N}\right)^2}, & \text{if } N \text{ even},\\ \frac{\cos\left(\frac{\pi}{2N}\right)}{N\sin\left(\frac{\pi}{2N}\right)^2}, & \text{if } N \text{ odd}.\end{cases}$$

whose generality and exactness can be compared with the aforementioned numerical approximated approaches.

References

[1] J. Massey, *Guessing and entropy*, Proceedings of 1994 IEEE International Symposium on Information Theory, 204 (1994).

[2] W. Chen, Y. Cao, H. Wang, Y. Feng, *Minimum guesswork discrimination between quantum states*, Quantum Information & Computation **15**, 0737 (2015).

[3] E. P. Hanson, V. Katariya, N. Datta, and M. M. Wilde, *Guesswork with Quantum Side Information*, IEEE Trans. Inform. Theory **68**, 322 (2022).

[4] M. Dall'Arno, F. Buscemi, and T. Koshiba, *Guesswork of a quantum ensemble*, IEEE Trans. Inform. Theory **68**, 3193 (2022).

[5] M. Dall'Arno, F. Buscemi, and T. Koshiba, *Classical computation of quantum guesswork*, arXiv:2112.01666.

[6] V. Katariya, N. Bhusal, C. You, *Experimental Guesswork with Quantum Side Information using Twisted Light*, arXiv:2211.08601.

[7] M. Dall'Arno, *Quantum guesswork*, arXiv:2302.06783.

[8] T. C. Koopmans, M. Beckmann, *Assignment problems and the location of economic activities*, Econometrica **25**, 53 (1957).

[9] S. Sahni, T. Gonzalez, *P-Complete Approximation Problems*, Journal of the ACM **23**, 555 (1976).

[10] E. Dragoti-Çela, *The quadratic assignment problem: Theory and algorithms* (Kluwer Academic Publishers, 1998, Dordrecht, The Netherlands).

Entanglement swapping via lossy channel

Wan Zo[1] [2] [*] Bohdan Bilash[1] Kyunghwan Oh[2] Yong-Su Kim[1] [†]

[1] *Center for Quantum Information, Korea Institute of Science and Technology (KIST), Seoul, 02792, Republic of Korea*

[2] *Department of Physics, Yonsei University, Seoul, 03722, Republic of Korea*

Abstract. Sharing quantum entanglement between long distances is a key resource of quantum communication network. Optical fiber channel loss is the main challenge of sharing photonic entanglement. Quantum repeater and spin-photon interface is promising candidate to implement quantum network. In this paper, we propose modified entanglement swapping scheme of Duan-Lukin-Cirac-Zoller (DLCZ) protocol using path encoded qubits for spin-photon entanglement scenario. we show the entanglement property by removing which-way information and verify channel loss robustness of our scheme. This proof-of-principle experiment paves the way towards overcoming optical channel loss problem.

Keywords: Quantum Entanglement, Entanglement swapping, Quantum optics

1 Introduction

Distributing quantum entanglement between long distance nodes is a key challenge of quantum communication network. Optical fiber channel loss is main problem of photonic quantum network and many approaches are proposed to overcome optical fiber channel loss. A potential solution to this problem is quantum repeater architecture[1] and spin-photon interface[2]. In spin-photon interface, particle spins are coupled to vacuum and single photon state. Single-photon entangled state is robust in terms of channel loss because there is only one photon that actually passes through the optical lossy channel[Fig.1].

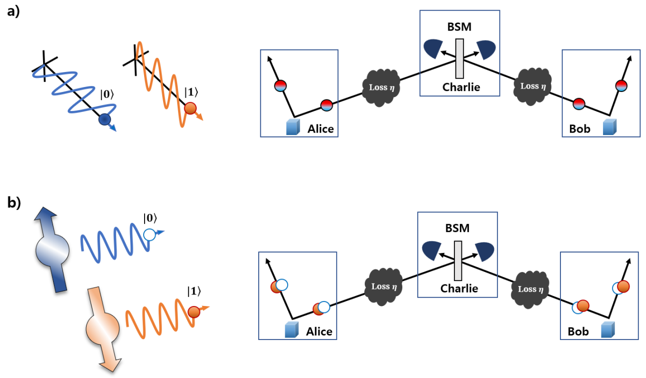


Figure 1: Entanglement swapping protocols based on polarization entanglement and spin-photon entanglement. **a)** Polarization entanglement swapping. Two photons traverse lossy channels, resulting in the entanglement distribution probability decreasing proportionally to η^2. **b)** Spin-photon entanglement swapping. In this scenario, only one photon traverses the lossy channel. Thus, the entanglement distribution probability decreases at a rate proportional to η.

In this work, we propose modified DLCZ scheme and experimentally demonstrate entanglement property of distributed state by showing which-way information visibility. We verify channel loss robustness of path encoded entanglement among two parties by manipulating channel loss for spin-photon interface scenario. For asymmetric channel losses, we restore state visibility by preparing asymmetric input states.

2 Theory

We have used two Spontaneous Parametric Down Conversion (SPDC) crystals and built Mach-Zehnder Interferometer (MZI)[Fig.2]. Beam displacer have separated pump laser. Thus, initial state as follows:

$$a\,|LVDH\rangle + a'\,|RVUH\rangle \tag{1}$$

H, V denote polarization and L,R,U,D denote left, right, up, down path. a and a' are related to input powers of crystals. L and R are connected to entanglement verification part (blue region), and U and D are connected to Bell state projection part (green region). waveplates and polarization beam splitter implement X, Y, and Z basis projection.

After the X basis projection, the state collapses onto single-photon path-encoded state and this state has entanglement property[3]. We show visibility curve by removing which-way information at the entanglement verification part. For pure state, this visibility is equal to concurrence[4][5].

3 Experiment

We have used 405nm CW laser and two Periodically-Poled KTP crystal for 405-810 down conversion process. beam separation distance is 4mm. input powers before crystals are 0.5mW for each. To avoid double pair generation in SPDC process, we have used low power. X basis projection visibilities are 0.9471±0.0065 and 0.9300±0.0070. Z basis projection visibilities are 0.06 and 0.07[Fig.3a]. To get entanglement distribution rate, we have measured coincidence counts of Alice's detector (D_A) and Bob's detector (D_B) . The HWP before PBS in entanglement verification part have been set to 0

[*] zowan@kist.re.kr

[†] yong-su.kim@kist.re.kr

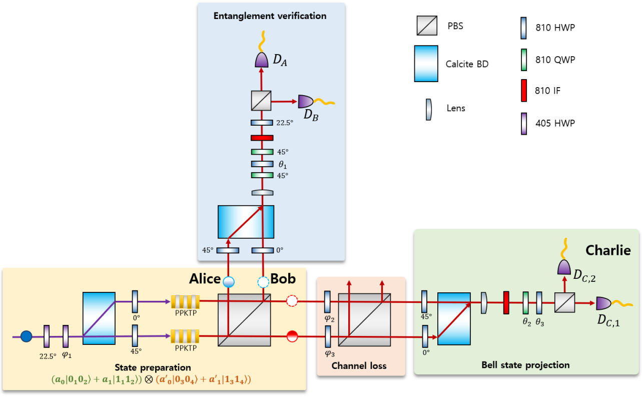


Figure 2: Experiment setup. We used Mach-Zehnder type interferometer with calcite type beam displacers. A type-2 SPDC crystal generates horizontal and vertical polarization photon. The set of beam displacer and polarization projector can removes which-way information. The Charlie part performs Bell state projection onto X, Y, and Z basis by changing angles of the HWP and QWP. We implemented channel loss by using a PBS and two HWPs before the PBS. BD: Beam Displacer, PBS: Polarization Beam Splitter, HWP: Half-Wave Plate, QWP: Quarter-Wave Plate, IF: Bandpass Interference Filter.

and Bell state projection part have been set to X basis projection. We have verified that coincidence counts linearly decrease as function of channel loss η. Coincidence counts are integrated up to 5 seconds[Fig.3b]. We have shown that X basis projection visibilities are maintained when we have changed channel loss. Two optical channel losses are same[Fig.3c].

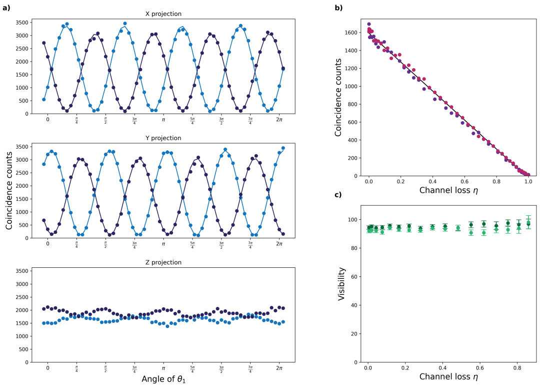


Figure 3: Experiment results. **a)** Visibilities for X, Y, and Z projections. The X-axis denotes the HWP angle within the Q-H-Q waveplates set. **b)** The entanglement distribution rate in terms of channel loss η. The HWP before PBS in entanglement verification part was set to 0 and Bell state projection part was set to X projection. Two types of dots denote coincidence counts between Alice's detector and Charlie's, and between Bob's and Charlie's. **c)** X projection visibilities in terms of channel loss η. Two types of dots denote visibilities of each coincidence set.

We have tested asymmetric channel loss, For instance, one optical channel has no loss and only one channel has loss. By changing HWP before first BD, we have manipulated input power ratio for two crystals and it have changed input state coefficient a, a'. This asymmetric initial state coefficients have compensated asymmetric optical channel loss. We have checked that degradation of visibilities by asymmetric channel loss (green dots) and restoration of visibility by compensation (violet and red dots)[Fig.4].

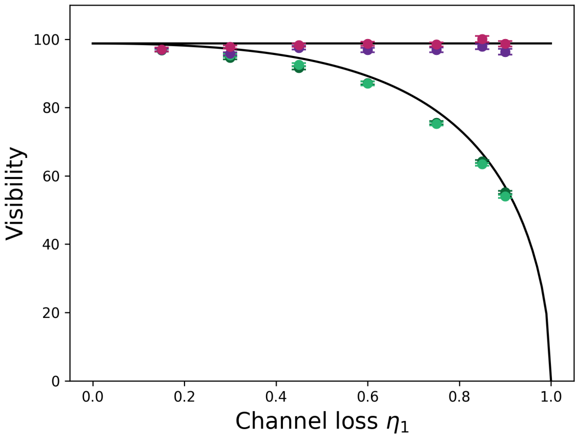


Figure 4: Visibility restorations for asymmetric channel losses. Two green dots denote visibilities for asymmetric channel losses and symmetric input states. Violet and red dots denote visibilities for asymmetric channel losses and asymmetirc input states.

4 Discussion

We have experimentally demonstrated entanglement property of distributed state by showing which-way information visibility. We have verified channel loss robustness of path encoded entanglement among two parties by manipulating channel loss for spin-photon interface scenario.

References

[1] Duan, LM. et al. Long-distance quantum communication with atomic ensembles and linear optics. *Nature 414, 413–418*, 2001.

[2] Vasconcelos, R. et al. Scalable spin–photon entanglement by time-to-polarization conversion. *npj Quantum Inf 6, 9*, 2020

[3] Caspar, P. et al. Heralded Distribution of Single-Photon Path Entanglement. *PhysRevLett.125.110506*, 2020

[4] Lee, Sang Min et al. Proposal for direct measurement of concurrence via visibility in a cavity QED system. *PhysRevA.77.040301*, 2008

[5] Florian Kaiser et al. Entanglement-Enabled Delayed-Choice Experiment. *Science338,637-640*, 2012

Deep Ising Born Machine (published in https://onlinelibrary.wiley.com/doi/10.1002/qute.202300033)

Zhu Cao[1] *　　Linlin Wang[†2 ‡]

[1] *East China University of Science and Technology*
[2] *East China Normal University*

Abstract. A quantum neural network (QNN) is a method to find patterns in quantum data and has a wide range of applications including quantum chemistry, quantum computation, quantum metrology, and quantum simulation. Efficiency and universality are two desirable properties of a QNN but are unfortunately contradictory. In this work, we examine a *deep Ising Born machine* (DIBoM), and show it has a good balance between efficiency and universality. More precisely, the DIBoM has a flexible number of parameters to be efficient, and achieves provable universality with sufficient parameters. The architecture of the DIBoM is based on generalized controlled-Z gates, conditional gates, and some other ingredients. To compare the universality of the DIBoM with other QNNs, we propose a fidelity-based expressivity measure, which may be of independent interest. Extensive empirical evaluations corroborate that the DIBoM is both efficient and expressive.

Keywords: Quantum Machine learning, Quantum Neural Network, Efficiency, Universality, Expressivity

1 Introduction

Machine learning (ML) has emerged as one of the most revolutionary techniques in recent years [Goodfellow *et al.*(2016)Goodfellow, Bengio, and Courville]. Despite its significance, ML necessitates a tremendous amount of computational power. However, with the waning effectiveness of Moore's law on the speed of classical processors [Waldrop(2016)], and the ever-increasing computational demands of state-of-the-art ML models, the future development of ML may face significant hindrances due to the shortage of adequate computational resources. Quantum computing [Nielsen and Chuang(2010)], a novel computing paradigm, holds the potential to sustainably advance ML. At present, quantum machine learning (QML) [Biamonte *et al.*(2017)Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, and Lloyd], which refers to the use of quantum computers for machine learning, is still in its nascent stage. Depending on whether the data and the learning algorithm are classical or quantum, QML can be categorized into four types: classical learning of classical data, quantum learning of classical data, classical learning of quantum data, and quantum learning of quantum data.

Among the four categories of QML, quantum learning of quantum data is arguably the most promising type to achieve a demonstrable exponential speedup over classical machine learning methods. Furthermore, quantum learning of quantum data has a diverse array of applications, including quantum chemistry [Arute *et al.*(2020)Arute, Arya, Babbush, Bacon, Bardin, Barends, Boixo, Broughton, Buckley, Buell *et al.*], quantum data compression [Romero *et al.*(2017)Romero, Olson, and Aspuru-Guzik], quantum error correction [Johnson *et al.*(2017)Johnson, Romero, Olson, Cao, and Aspuru-Guzik], quantum metrology [Koczor *et al.*(2020)Koczor, Endo, Jones, Matsuzaki, and Benjamin], quantum compiling [Sharma *et al.*(2020a)Sharma, Khatri, Cerezo, and Coles], quantum state diagonalization [LaRose *et al.*(2019)LaRose, Tikku, O'Neel-Judy, Cincio, and Coles], quantum simulation [Li and Benjamin(2017)], quantum fidelity estimation [Cerezo *et al.*(2020)Cerezo, Poremba, Cincio, and Coles], and consistent histories [Arrasmith *et al.*(2019)Arrasmith, Cincio, Sornborger, Zurek, and Coles]. It is worth noting that these quantum applications generate a substantial amount of quantum data, which in turn fuels the development of quantum learning of quantum data, analogous to how the vast amount of classical information drives the advancement of classical machine learning.

Methods of quantum learning of quantum data can be classified into two categories: those that belong to quantum neural networks (QNNs) and those that do not. Examples of methods that do not fall into the QNN category are the Harrow-Hassidim-Lloyd algorithm [Harrow *et al.*(2009)Harrow, Hassidim, and Lloyd], quantum principal component analysis [Lloyd *et al.*(2014)Lloyd, Mohseni, and Rebentrost], and quantum support vector machines [Rebentrost *et al.*(2014)Rebentrost, Mohseni, and Lloyd]. Various proposals of QNNs have been put forward in the literature [Schuld *et al.*(2015)Schuld, Sinayskiy, and Petruccione, Lewenstein(1994), Wan *et al.*(2017)Wan, Dahlsten, Kristjánsson, Gardner, and Kim, da Silva *et al.*(2016)da Silva, Ludermir, and de Oliveira, Gonsalves(2017), Kouda *et al.*(2005)Kouda, Matsui, Nishimura, and Peper, Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf], among which the current state-of-the-art is arguably given by Ref. [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf].

*caozhu55@gmail.com
†Corresponding author
‡llwang@cs.ecnu.edu.cn

In addition, specialized QNNs tailored to specific data inputs have also been developed, including quantum convolutional neural networks [Cong *et al.*(2019)Cong, Choi, and Lukin], quantum recurrent neural networks [Bausch(2020)], quantum generative adversarial networks [Lloyd and Weedbrook(2018)], quantum autoencoders [Bondarenko and Feldmann(2020)], quantum reservoir networks [Ghosh *et al.*(2021)Ghosh, Nakajima, Krisnanda, Fujii, and Liew], and quantum residual networks [Killoran *et al.*(2019)Killoran, Bromley, Arrazola, Schuld, Quesada, and Lloyd]. However, these specialized QNNs cannot perform universal quantum computation, which is essential for the general quantum learning task of learning the hidden mapping between a set of quantum input and label pairs. Consequently, we will focus our attention on general QNNs and the general quantum learning task hereafter.

Efficiency and universality are two desirable properties of a general QNN. While efficiency is measured in terms of the number of parameters in the model, which should be as small as possible, universality refers to the ability of a QNN to approximate an arbitrary unitary on n qubits. These two properties are often in conflict with each other. For instance, consider a basic QNN that applies a parametrized unitary U to the quantum input and produces an output that approximates the label. Here, the parametrized unitary U for n qubits is represented as

$$U = \exp\left[i\left(\sum_{j_1=0}^{3}\cdots\sum_{j_n=0}^{3}\alpha_{j_1,j_2,\ldots,j_n}\left(\sigma_{j_1}\otimes\cdots\otimes\sigma_{j_n}\right)\right)\right], \tag{1}$$

where σ_0 is the identity matrix, σ_1, σ_2, σ_3 are Pauli matrices, and $\alpha_{j_1,j_2,\ldots,j_n}$ are real parameters that are learned during training. Although this basic QNN is universal, as it can approximate any unitary by adjusting its parameters, it is not efficient since its number of parameters is 4^n. Ideally, the number of parameters should be polynomial in n for the model to be efficient.

This work investigates a *deep Ising Born machine* (DIBoM), which has a flexible number of parameters to mitigate the efficiency issue while retaining universality with sufficient parameters. The DIBoM consists of a generalized controlled-Z (CZ) gate, a conditional gate, a global or local cost function, and some other ingredients. By replacing the normal CZ gate with a generalized CZ gate in a hardware-efficient QNN [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven, Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz], we demonstrate that the expressivity can be increased through numerical evaluations. Along the way, we develop an expressivity measure, called fidelity-based expressivity, to characterize the expressivity of different QNN architectures. This measure may be of independent interest. Moreover, we theoretically prove that hardware-efficient QNN with generalized CZ gates can achieve universal quantum computation with sufficient parameters. The conditional gate is used to solve the problem of different input and output dimensions, and this approach can save space by a constant factor compared to dissipative QNNs [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Saleimann, Scheiermann, and Wolf]. In addition, the ablation study shows that this ingredient improves the expressivity of the DIBoM. We examine two variants of the DIBoM, one with a global cost function and the other with a local cost function, and show that the global cost function version has a wider range of applicability, while the local cost function is more trainable and can mitigate the barren plateau issue. We perform extensive experiments to compare the DIBoM with other QNN architectures, evaluate its different components, analyze the sensitivity of its performance to various parameters, and examine its robustness to noise. Our work invites further research on the design of QNNs with multiple desirable properties, and we hope it will stimulate further development of the architecture design of QNNs in general.

The roadmap for the rest of the paper is as follows. First, in Sec. 2, we review related works. Next, in Sec. 3, we present the DIBoM model and its training method. In Sec. 4, we analyze theoretically the properties of the DIBoM. We then turn to the empirical evaluation of the model, with Sec. 5 describing the simulation setup and Sec. 6 presenting the results. Finally, we conclude the paper in Sec. 7 and give a few outlooks.

2 Related works

In this section, we provide a review of the related works in the field, including hardware-efficient QNNs, dissipative QNNs, Ising Born machines, and Hamiltonian learning.

2.1 Hardware-efficient QNNs

We start by reviewing hardware-efficient QNNs [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven, Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz]. Hardware-efficient QNNs were proposed to reduce the exponential training cost of the basic QNN, and require only a polynomial number of resources. They are composed of alternating layers of single-qubit rotations and entangling gates such as CZ gates. The connectivity of the entangling gates can be either linear [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven] or pairwise [Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz], as shown in Figs. 1(a) and (b). The layer number of a hardware-efficient QNN can vary and so are the parameters of its single-qubit rotations. From now on, we will refer to the architecture in Ref. [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven] as *the* hardware-efficient QNN.

There are two drawbacks to the hardware-efficient QNN. First, to our knowledge, there is to date no proof that the hardware-efficient QNN presented in [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven] is capable of universal quantum computation even

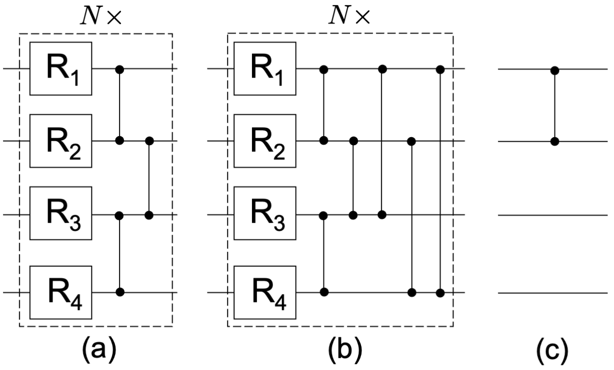


Figure 1: (a) A series of blocks where each block consists of single-qubit rotations with nearest-neighbor CZ gates; (b) A series of blocks where each block consists of single-qubit rotations with all-to-all CZ gates; (c) Unitary with a single CZ gate connecting the first two qubits.

with an exponential number of layers, see Sec. 4.2 for more details. In particular, it is not known whether it can be used to simulate a circuit with a single CZ gate, which is illustrated in Fig. 1(c). Secondly, it falls short of varying the relative number of qubits between the input and the output. The DIBoM resolves these shortcomings of the hardware-efficient QNN [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven] while retaining its merits.

Note that the definition of the hardware-efficient QNN varies in the literature, and in the broadest sense can include any QNN that can be implemented efficiently on some quantum hardware. This in particular includes the DIBoM as a special case, which is different from its usage in our work.

2.2 Dissipative QNN

We next review dissipative QNNs, a different type of QNNs that deal with quantum inputs and outputs of unmatched dimensions [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf, Sharma *et al.*(2020b)Sharma, Cerezo, Cincio, and Coles]. Initially, all hidden and output qubits are in the state $|0\rangle$. Dissipative QNNs apply a unitary on all input, hidden, and output qubits and subsequently trace out the input and hidden qubits,

$$\rho_{out} \equiv \mathrm{tr}_{in,hid}\left(U(\rho_{in} \otimes |0\cdots 0\rangle_{hid,out}\langle 0\cdots 0|)U^{\dagger}\right). \quad (2)$$

This process is illustrated in Fig. 2. Here, it can be easily seen that the dimensions of the quantum input ρ_{in} and the quantum output ρ_{out} need not be the same. Since this network architecture discards lots of qubits, it was given the name a *dissipative QNN* [Sharma *et al.*(2020b)Sharma, Cerezo, Cincio, and Coles].

However, a dissipative QNN also has several disadvantages. First, it has a larger space complexity due to its dissipative nature. When the dimensions of the quantum input and output are equal, a dissipative QNN has an overhead of 2 over a basic QNN in terms of the qubits used. Second, the resource that a dissipative QNN requires is still exponential, due to the same reason that an exponential number of parameters are needed to parametrize a unitary transformation. A DIBoM maintains the merit of a dissipative QNN that can handle unequal quantum input and output dimensions and in the meantime has the additional merits of having small space complexity and being resource efficient.

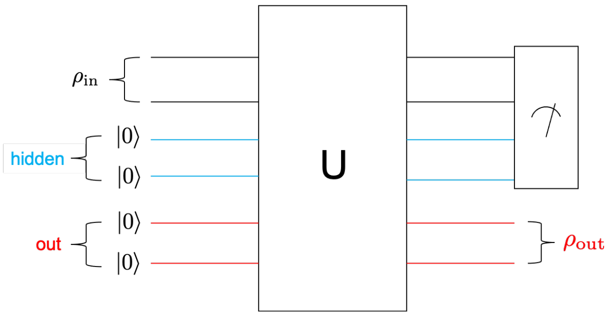


Figure 2: An illustration of a dissipative QNN. Here, ρ_{in} and ρ_{out} are its quantum input and quantum output respectively, and U is a unitary. The hidden and output qubits are all in the quantum state $|0\rangle$ initially.

2.3 Ising Born machine

We then review the Ising Born machine [Coyle *et al.*(2020)Coyle, Mills, Danos, and Kashefi], which bears a close resemblance to the DIBoM. We begin by defining a quantum Born machine [Cheng *et al.*(2018)Cheng, Chen, and Wang, Liu and Wang(2018), Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz], as illustrated in Fig. 3(a). A quantum Born machine is a class of models that consists of a parametrized quantum circuit followed by a quantum measurement. As the measurement outcome is determined by Born's rule, these models are named "quantum Born machines" [Cheng *et al.*(2018)Cheng, Chen, and Wang, Liu and Wang(2018), Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz]. Figure 3(b) provides a specific choice of the parametrized quantum circuit proposed by Ref. [Coyle *et al.*(2020)Coyle, Mills, Danos, and Kashefi], containing one fixed layer (Hadamard gates on all qubits) and two tunable layers (one with tunable two-qubit gates on all pairs of qubits and the other with tunable one-qubit gates on all qubits). Since tunable two-qubit gates mimic the Ising model, this model was named the "Ising Born machine" [Coyle *et al.*(2020)Coyle, Mills, Danos, and Kashefi]. Although the original Ising Born machine was designed for generative modeling of classical data, it can be adapted to the general quantum learning task by removing the final quantum measurement layer in Fig. 3(b). The main difference between the DIBoM and Ising Born machine lies in the number of tunable layers; the DIBoM can contain more than two tunable layers, thereby achieving higher expressive power than the Ising Born machine.

In particular, the Ising Born machine is not capable of universal quantum computation, while the DIBoM is.

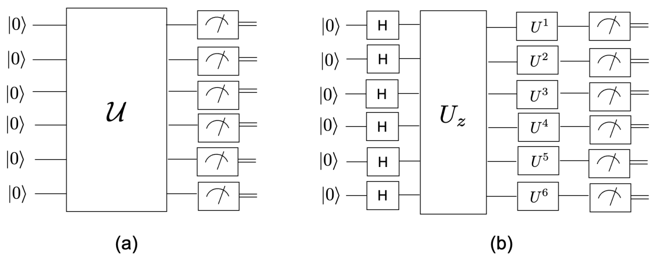


Figure 3: (a) The schematics of a quantum Born machine. Here, $\mathcal{U}$ is a parametrized quantum circuit. (b) The schematics of an Ising Born machine. Here, H is a Hadamard gate, U_z consists of tunable two-qubit gates on all pairs of qubits, and U^k ($k \in \mathbb{N}$) is a tunable single-qubit gate on qubit k.

2.4 Hamiltonian learning

Finally, the general quantum learning task is closely related to Hamiltonian learning [Cirstoiu *et al.*(2020)Cirstoiu, Holmes, Iosue, Cincio, Coles, and Sornborger, Barison *et al.*(2021)Barison, Vicentini, and Carleo], a topic that has attracted a tremendous amount of interest recently. The correspondence is as follows: the quantum input $|\phi_{in}\rangle$ corresponds to the initial quantum state of a system, and the quantum label $|\phi_{out}\rangle$ corresponds to the quantum state of the system after evolving for a time δt. The hidden mapping V can be associated with the time-evolution operator $e^{-iH\delta t}$ where H is the Hamiltonian of the system. By approximating V using QNNs, the Hamiltonian H is learnt.

3 Model

After reviewing related works, we now describe the DIBoM architecture. We first mathematically formulate the problem of quantum learning of quantum data in Sec. 3.1. Then in Sec. 3.2, we describe the DIBoM model which is targeted to this learning problem. Finally, we describe the training procedure of the DIBoM in Sec. 3.3.

3.1 Learning problem setup

We begin by introducing the quantum learning problem addressed by the DIBoM. Let $\mathcal{D}$ be an underlying distribution, and suppose we have N pairs of training samples and labels $(|\psi_i\rangle, |\phi_i\rangle) \in \mathcal{D}$, where $1 \leq i \leq N$. For each pair, we are given K copies of the input $\otimes_{i=1}^{N}(|\psi_i\rangle^{\otimes K}, |\phi_i\rangle^{\otimes K})$, as well as M copies of test samples denoted by $\rho^{\otimes M}$. The goal is to generate model outputs for each test sample that closely approximate the corresponding test label. We use the infidelity to measure the similarity between two quantum states, and assume that the training and test data are independently sampled from $\mathcal{D}$. To illustrate, consider an example where $|\psi_i\rangle$ is a randomly generated n-qubit pure state and its corresponding label is $|\phi_i\rangle = V|\psi_i\rangle$. Here V is a hidden n-qubit unitary that is independent of i and unknown to the model. This example will also be used in the evaluation section. It should be noted that $|\psi_i\rangle$ and $|\phi_i\rangle$ may not be of the same dimension for general $\mathcal{D}$.

3.2 Model architecture

After defining the learning problem, we now turn to the model of the DIBoM. The DIBoM takes a quantum state as an input and outputs a quantum state as an output which may have different dimensions. It is based on a basic quantum structure which is illustrated in Fig. 4. This basic quantum structure has three steps. First, the quantum input ρ_{in} together with a k-qubit ancilla $|0\rangle^{\otimes k}$ undergo a unitary transformation U that produces an intermediate state

$$\rho_{\text{inter1}} = U(\rho_{\text{in}} \otimes |0\rangle^{\otimes k})U^\dagger. \tag{3}$$

Second, part of the joint quantum state is measured, resulting in outcome j. Let ρ^j_{inter2} denote the post-measurement state conditioned on that the outcome is j. Third, another unitary V_j, that can depend on the outcome j in the second step, is applied to ρ^j_{inter2} to produce the output quantum state

$$\rho_{\text{out}} = V_j \rho^j_{\text{inter2}} V_j^\dagger. \tag{4}$$

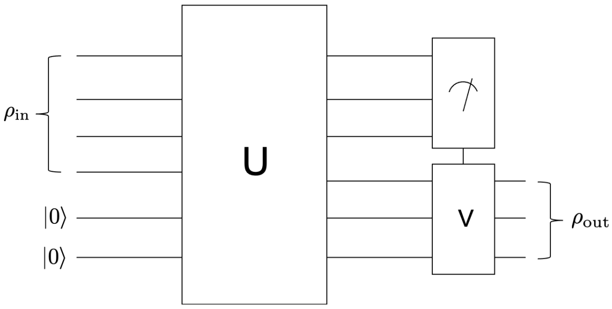


Figure 4: The basic quantum structure. Here, ρ_{in} and ρ_{out} are its quantum input and quantum output respectively, U is a unitary, and V is a controlled unitary by the result of the measurement outcome.

With the basic quantum structure at hand, we are ready to define the DIBoM. Instead of applying the unitaries U and V_j which would consume exponential time to train, the DIBoM uses a stack of layers as a substitute for U and V_j. The layers have two types. The first type consists of single-qubit rotations that are parametrized by α_j as follows:

$$U_{\text{SG}} = \exp\left[i(\sum_{j=1}^{3} \alpha_j \sigma_j)\right], \tag{5}$$

where σ_1, σ_2, σ_3 are Pauli matrices. The second type of layer applies a generalized CZ gate to all pairs of qubits:

$$U_{\text{CZ}} = \exp\left[-i\pi(\sum_{1\leq j<k\leq n} \beta_{jk} |11\rangle_{jk}\langle 11|_{jk})\right], \tag{6}$$

where β_{jk} is an arbitrary real number that interpolates smoothly between a CZ gate and an identity gate. It is worth noting that if $\beta_{jk}=1$, a CZ gate is applied on qubits j and k, and if $\beta_{jk}=0$, an identity gate is applied. The DIBoM is constructed as $\prod_{j=L/2}^{1}(U_{\mathrm{CZ}}^{j}U_{\mathrm{SG}}^{j})$ for L even and $U_{\mathrm{SG}}^{(L+1)/2}\prod_{j=(L-1)/2}^{1}(U_{\mathrm{CZ}}^{j}U_{\mathrm{SG}}^{j})$ for L odd, where L is the total number of layers and $\prod_{j=1}^{L}U^{j}$ is short for $U^{1}\cdots U^{L}$. An illustration of the DIBoM is shown in Fig. 5.

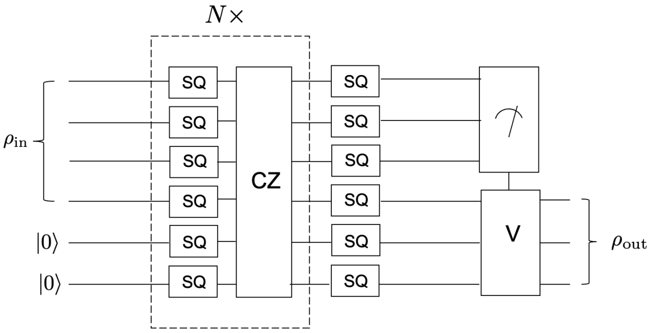


Figure 5: The schematics of a deep Ising Born machine. Here, SQ denotes a tunable single-qubit gate; CZ denotes generalized CZ gates between all pairs of qubits. The rest symbols have the same meanings as the basic quantum structure.

3.3 Training procedure

After presenting the DIBoM model, the next step is to discuss the training process. In this regard, two variants of the loss functions are considered. The first loss function, called the global loss function, has the form

$$\mathcal{L}_G = 1-\mathbb{E}_x\langle\phi_{\mathrm{out}}^{x}|\rho_{\mathrm{out}}^{x}|\phi_{\mathrm{out}}^{x}\rangle, \tag{7}$$

where $|\phi_{\mathrm{out}}^{x}\rangle$ is the correct label, ρ_{out}^{x} is the output of the DIBoM, and $\mathbb{E}_x$ stands for the expectation over the random variable x. The intuition behind this loss function can be seen through some special cases. If the correct label is identical to the model output, the loss is 0; otherwise, the loss is positive. Hence, by minimizing the loss function, the model output converges to the correct label. If the correct label is a mixed state $\sigma_{\mathrm{out}}^{x}$, the loss function can be easily generalized to $\mathcal{L}=1-\mathbb{E}_xF(\rho_{\mathrm{out}}^{x},\sigma_{\mathrm{out}}^{x})$, where $F(\rho,\sigma):=\left[\mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}\right]^2$.

The second loss function, called the local loss function, is given by

$$\mathcal{L}_L = 1-\frac{1}{nN}\sum_{x=1}^{N}\sum_{y=1}^{n} tr((|\phi_x^0\rangle_y\langle\phi_x^0|_y\otimes I_{\bar{y}})\rho_x^0(s)), \tag{8}$$

where $|\phi_x^0\rangle_y$ is the y-th qubit of the input state $|\phi_x^0\rangle$, $I_{\bar{y}}$ denotes completely mixed states for all qubits except the y-th qubit, N is the number of samples, and n is the number of qubits. Here, $\rho_x^0(s)$ represents the effective input which, when applied the unitary given by the current model, generates the correct quantum label. When applying the local loss function, it is assumed that the input quantum state $|\phi_x^0\rangle$ is a product state and has the form $|\phi_x^0\rangle=|\phi_x^0\rangle_1\otimes\cdots\otimes|\phi_x^0\rangle_n$, where n is the number of qubits. Note that no such assumption is made when applying the global loss function.

With the loss function defined (either $\mathcal{L}_G$ or $\mathcal{L}_L$, for simplicity denoted as $\mathcal{L}$), we describe the procedure to train the network with quantum computers in two steps: (i) calculate the loss function with quantum computers; (ii) update the parameters by performing gradient descent on the loss function. For the first step, we first compute the quantity $\langle\phi_{\mathrm{out}}^{x}|\rho_{\mathrm{out}}^{x}|\phi_{\mathrm{out}}^{x}\rangle$. To this end, we exploit the quantum circuit plotted in Fig. 6 [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf]. Through straightforward calculation, one can verify that this circuit takes $|\phi_{\mathrm{out}}^{x}\rangle$ and ρ_{out}^{x} as inputs and outputs $(1+\langle\phi_{\mathrm{out}}^{x}|\rho_{\mathrm{out}}^{x}|\phi_{\mathrm{out}}^{x}\rangle)/2$. By a linear transformation and some classical computation, the loss function $\mathcal{L}$ is obtained.

For the second step, we calculate the derivative of the loss function for all parameters and update the parameters accordingly by performing gradient descent. For a parameter y^{μ} (either α_j or β_{jk}) in any layer, we calculate its derivative by running the loss function calculation twice as follows,

$$\frac{\delta\mathcal{L}}{\delta y^{\mu}}=\frac{\mathcal{L}(y^{\mu})-\mathcal{L}(y^{\mu}-\epsilon)}{\epsilon}, \tag{9}$$

where ϵ is a small value. Note that this method assumes the availability of a high-precision quantum computer, as a noisy quantum computer may yield a derivative that is far from the true value. An alternative way would be calculating an analytic derivative directly with quantum computers, but this requires further investigation and is left as future work.

Then we update each parameter y^{μ} in the k-th iteration to minimize the loss function by the rule

$$y_{k+1}^{\mu}=y_k^{\mu}-\eta\frac{\delta\mathcal{L}}{\delta y_k^{\mu}}, \tag{10}$$

where η is the learning rate. When η is small enough, the loss function always decreases by the parameter update. We note that the DIBoM is efficiently trainable as it has only a polynomial number of parameters.

The training strategy presented here does not aim to optimize the efficiency or computation cost of the training algorithm. Instead, we chose this strategy to evaluate the model's performance in terms of its converged loss. If the model can converge to a low test loss with some strategies, it is highly probable that the training strategy presented here will also result in a low test loss. This makes it quite ideal for testing the performance of the model in terms of its converged loss. Notably, there are training strategies such as the parameter-shift rule [Wierichs *et al.*(2022)Wierichs, Izaac, Wang, and Lin] that can significantly reduce the number of quantum circuit evaluations, and gradient descent methods that offer faster convergence. For example, one may utilize Nesterov acceleration [Nesterov()] which is also a

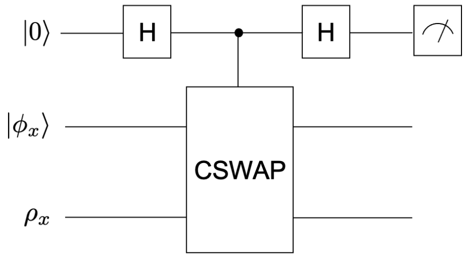


Figure 6: Circuit to calculate the loss function. Here, H is the Hadamard gate, $CSWAP$ is the controlled-SWAP gate, $|\phi_x\rangle$ is the ground-truth label, and ρ_x is the model output.

first-order optimization method (utilizing only first-order derivatives) to speed up the convergence. The k-th iteration of the parameter y^μ in Nesterov acceleration has the form,

$$\begin{aligned} x^\mu_{k+1} &= y^\mu_k + \frac{k-1}{k+2}(y^\mu_k - y^\mu_{k-1}), \\ y^\mu_{k+1} &= x^\mu_{k+1} - \eta \frac{\delta \mathcal{L}}{\delta x^\mu_{k+1}}. \end{aligned} \tag{11}$$

It can be shown that when $\eta = 1/\mathsf{L}$ where L is the Lipschitz constant of $\mathcal{L}$, the convergence rate by the above iteration rule is $O(1/k^2)$, quadratically better than $O(1/k)$ of Eq. (10).

Second-order or higher-order optimization methods can offer further improvements in convergence by utilizing second-order derivatives $\partial^2\mathcal{L}/\partial y^\mu \partial y^\nu$. However, the computational cost of each iteration step in second-order optimization methods is $O(N^2)$, in contrast to $O(N)$ of first-order optimization methods, where N is the number of parameters. In classical neural networks, N is usually on the order of 10^8, making second-order optimization methods computationally too expensive. As a result, first-order optimization methods are generally preferred. Similarly, in QNNs, second-order optimization methods were considered inferior to first-order methods in cases where the learning problem required a large N to solve. However, recent advances have shown that second-order methods can be substantially sped up [Gacon *et al.*(2021)Gacon, Zoufal, Carleo, and Woerner], making them a competitive alternative.

The training procedures presented above update all parameters simultaneously, and we refer to them as *simultaneous training*. Another training method, known as *layer-by-layer training* [Skolik *et al.*(2021)Skolik, McClean, Mohseni, van der Smagt, and Leib], offers an alternative approach. In each training step, the parameters of one layer are updated using gradient descent, while the parameters of all other layers are fixed. The layer to be trained can be selected in a round-robin manner, from layer 1 to layer L and then repeated, where L is the number of layers. Alternatively, choosing the layer randomly is also a plausible approach. For the remainder of this paper, we will use the round-robin approach for layer-by-layer training, unless otherwise specified.

4 Theoretical Analysis

In this section, we conduct a theoretical analysis of the DIBoM architecture from three perspectives to gain insight into its properties. Specifically, we examine its flexibility with unequal input and output dimensions in Sec. 4.1, its balance between expressive power and efficiency in Sec. 4.2, and compare it with other models in Sec. 4.3.

4.1 Input-output dimension

We start by showing that the DIBoM can support unequal input and output dimensions. This is due to its underlying structure, which was illustrated in Figure 4. The DIBoM can accommodate a larger or smaller number of input qubits m than output qubits n by adjusting the number of ancilla qubits and the qubits to be measured. There are two cases to consider. First, if $m < n$, an ancilla $|0\rangle^{\otimes(m-n)}$ can be used, and no measurement is required after the unitary U. Second, if $m > n$, no ancilla is needed, and a measurement can be performed on $m - n$ qubits after the unitary U.

As a result, quantum teleportation can be instantiated by the DIBoM as follows. We begin with a single-qubit quantum input ρ_{in} and an ancilla in the state $|00\rangle$. A unitary operator U is next applied to the system, which leaves the quantum input unchanged and entangles the ancilla into an EPR pair. A measurement is subsequently performed on both the quantum input and one of the qubits in the EPR pair. Based on the measurement outcome, an appropriate unitary operation is applied to the other qubit of the EPR pair. The result is the original quantum state being teleported to the quantum output ρ_{out}.

4.2 Expressive power

Next we show another theoretical property of the DIBoM, namely its ability to perform universal quantum computation. This is achieved by a reduction from a well-known result that $2^{O(n)}$ layers of single-qubit gates and CZ gates suffice for universal quantum computation, where n is the number of qubits [Nielsen and Chuang(2010)]. Note that a general circuit with $2^{O(n)}$ layers of single-qubit gates and CZ gates is inequivalent to the hardware-efficient QNN [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven]. For example, Fig. 1(c), which belongs to the class of circuits with single-qubit gates and CZ gates, cannot be converted into the form of the hardware-efficient QNN [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven], while it can be turned into the form of DIBoM as we will see shortly.

Given a circuit $\mathcal{C}$ with $2^{O(n)}$ layers of single-qubit gates and CZ gates that approximates the desired unitary U within an error of ϵ, we convert it to the structure of a DIBoM in two steps.

1. In the first step, we split each layer of $\mathcal{C}$ into two layers, with the first layer containing only single-qubit gates and the second layer containing only

CZ gates. The resulting circuit is called $\mathcal{C}_2$.

2. In the second step, we fill missing single-qubit gates in the odd layer of $\mathcal{C}_2$ with identity single-qubit gates, and fill missing generalized CZ gates in the even layer of $\mathcal{C}_2$ with identity two-qubit gates.

An illustration of this reduction is shown in Fig. 7. Hence a DIBoM with $2 \times 2^{O(n)} = 2^{O(n)}$ layers is capable of universal quantum computation.

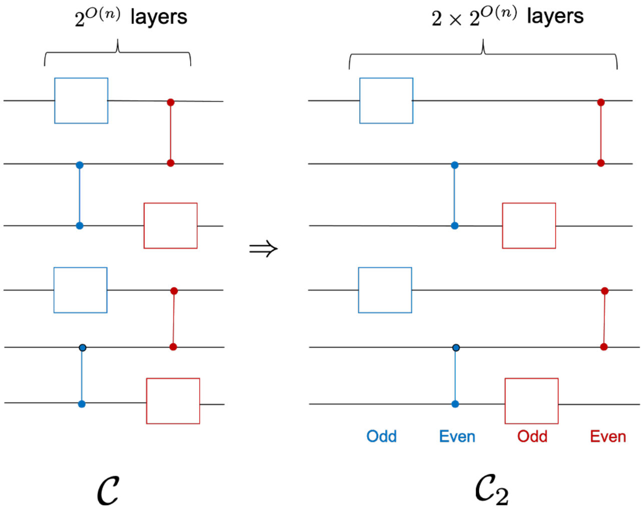


Figure 7: A reduction from $\mathcal{C}$ to $\mathcal{C}_2$. Each layer of $\mathcal{C}$ consists of a mixture of one-qubit gates and CZ gates. In the reduction, every layer of $\mathcal{C}$ is split into two layers, the first involving only single-qubit gates and the second involving only CZ gates.

The representation power of the DIBoM forms a hierarchy that varies with the number of layers. At one end of the spectrum, when the DIBoM has a polynomial depth, it possesses a limited number of parameters, making it highly efficient. Conversely, at the other end of the spectrum, when the DIBoM has an exponential depth, it has the ability to approximate universal quantum computation with high precision. Hence the DIBoM balances the efficiency and the expressive power quite well.

To quantitively compare the expressivity of the DIBoM and other QNN architectures, we propose an expressivity measure

$$E(\mathsf{A}) = \min_{U,|\phi\rangle} \max_{\theta} \left|\langle\phi|\mathsf{A}(\theta)^\dagger U|\phi\rangle\right|, \qquad (12)$$

where $\mathsf{A}(\theta)$ and θ are the parametrized circuit and its parameters, U is an arbitrary unitary, and $|\phi\rangle$ is an arbitrary pure quantum state.

To understand this measure, let us consider two special cases. First, when $\mathsf{A}(\theta)$ can recover any unitary, we can choose θ such that $\mathsf{A}(\theta) = U$ and hence $E(\mathsf{A}) = 1$. Second, when $\mathsf{A}(\theta)$ is a fixed unitary, i.e. θ is empty, for an arbitrary $|\phi\rangle$, we can select a unitary U such that $\mathsf{A}(\theta)|\phi\rangle$ and $U|\phi\rangle$ are orthogonal quantum states, hence $E(\mathsf{A}) = 0$. Due to its close relationship with fidelity, we call this expressivity measure, *fidelity-based expressivity* (FBE). Compared with other expressivity measures of QNNs, such as covering-number-based expressivity (CNBE) [Du *et al.*(2022)Du, Tu, Yuan, and Tao], FBE has the advantages of having no extra parameters (CNBE has a parameter ϵ), and having the range $[0, 1]$ (CNBE is not upper bounded by a constant). In addition, FBE can be computed through continuous optimization, while some measures such as CNBE require discrete optimization which is usually harder to compute.

We now compare the expressive power of the DIBoM and hardware-efficient QNN through FBE. Specifically, we consider a three-qubit learning task and plot the FBE of the DIBoM and hardware-efficient QNN with L layers as a function of L. (More details of the plot can be found in Appendix A.) Recall that the hardware-efficient QNN consists of alternating layers of single-qubit rotations and fixed CZ gates, while the DIBoM consists of alternating layers of single-qubit rotations and generalized CZ gates. Our results, as shown in Fig. 8, indicate that the DIBoM outperforms the hardware-efficient QNN by a substantial margin when the layer number is the same. For instance, with 21 layers, the DIBoM achieves an FBE exceeding 0.77 (where higher values indicate superior performance), whereas the hardware-efficient QNN achieves only around 0.57. (Note however that DIBoM has more parameters than the hardware-efficient QNN with the same number of layers and hence this does not imply that DIBoM is strictly superior to the hardware-efficient QNN.) Finally, the Ising Born machine, which corresponds to a DIBoM with $L = 3$, exhibits significantly lower expressivity than the DIBoM with 21 layers, as shown in the figure.

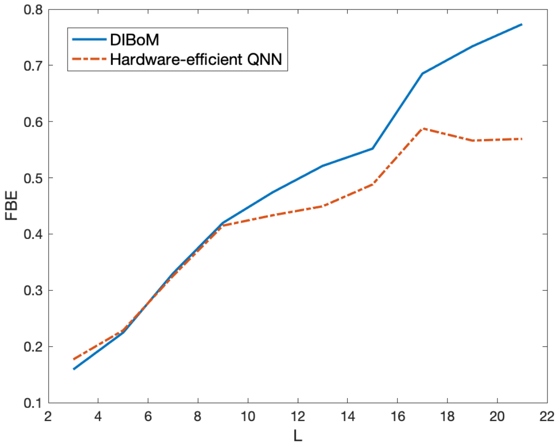


Figure 8: Fidelity-based expressivity (FBE) of the DIBoM and hardware-efficient QNN as a function of the layer number L.

The use of FBE also allows for a quantitative evaluation of the balance between efficiency and expressivity in DIBoMs. In Fig. 9, we present such an evaluation for a 3-qubit DIBoM. Efficiency is quantified as the logarithm of the number of parameters, while expressivity is measured using FBE. The endpoints of the curve are obtained through theoretical analysis, where a 0-parameter DIBoM corresponds to an FBE value of 0 and a 13449-parameter DIBoM corresponds to an FBE value of 1. (The proof for the latter fact can be found in Appendix B.) The remaining data points are computed numerically. The quality of the balance is characterized by the area of the purple region, with a smaller area indicating a better balance. It is worth noting that for some QNN architectures such as the hardware-efficient QNN, this area is not even guaranteed to be finite. In the figure, we show that through DIBoM, this area can be made finite. An interesting question for future work is how to achieve the smallest possible area of the purple region by optimizing over different QNN architectures.

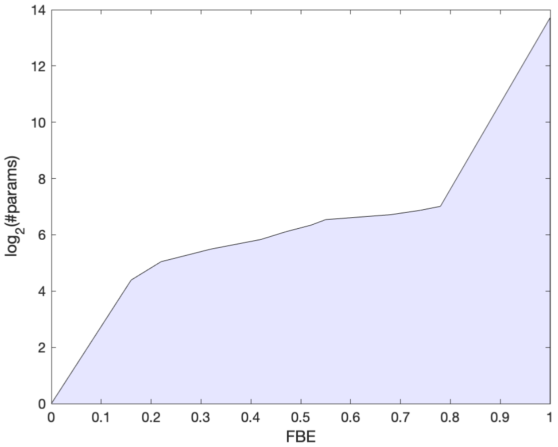


Figure 9: Quantitative characterization of the balance between efficiency and expressivity for a 3-qubit DIBoM. Here, the y axis stands for efficiency which is measured as the logarithm of the number of parameters, and the x axis stands for expressivity which is measured using fidelity-based expressivity (FBE).

4.3 Comparisons with other models

With the theoretical properties of the DIBoM at hand, we are ready to compare the DIBoM with other QNNs theoretically. First, we compare it to a basic QNN, as defined by Eq. (1). The DIBoM has the advantage that its number of parameters is quadratic while a basic QNN has an exponential number of parameters.

Second, we compare it to a dissipative QNN [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf]. In the case that the input and output have the same quantum dimension, the DIBoM uses only half the number of qubits required by a dissipative QNN. In addition, the number of parameters that a DIBoM uses is exponentially smaller than that of a dissipative QNN.

5 Empirical evaluation setup

To further investigate the properties of the DIBoM architecture, we conduct an extensive empirical evaluation of the DIBoM in this and the following sections. In this section, we present the setup of the evaluation, while the results of the evaluation are presented in the next section. The setup consists of two parts. In the first part, we describe the synthetic dataset that is used in the evaluation of the DIBoM. In the second part, we give the classical simulation for the training of the DIBoM.

5.1 Dataset

We start with the construction of the synthetic datasets used in the empirical evaluation. The samples in each dataset are of the form $\left|\phi_x^{in}\right\rangle$ where $x = 1, \ldots, N$ and N is the number of samples. Each sample is associated with a corresponding label $|\phi_x^{out}\rangle = V\left|\phi_x^{in}\right\rangle$ where $x = 1, \ldots, N$. The unitary V is referred to as the *intrinsic unitary* of the data and is hidden from the training models. If not otherwise specified, the samples are 2-qubit states and chosen randomly. In our main experiment, we generate a total of 20 samples, which are randomly divided into equal-sized training and test sets (50:50).

5.2 Classical simulation of training

With the dataset in place, we next describe how to evaluate the performance of a DIBoM on the dataset. Due to the lack of a quantum computer, we classically simulate the training procedure of the DIBoM and examine the result. In the following, we describe the classical simulation of the network training for a DIBoM. Let $L+1$ be the number of layers in the network, where layer 0 is the input layer and layer L the output layer. The transition from layer $l-1$ to layer l is given by

$$\rho_x^l(s) \quad = \quad U^l(s)\rho_x^{l-1}(s){U^l}^\dagger(s), \tag{13}$$

where s is any parameter of the model (such as α_j or β_{jk}) and $U^l(s)$ is the unitary in layer l. The loss function is computed as $\mathcal{L}(s) = 1 - C(s)$, where

$$C(s) = \frac{1}{N}\sum_{x=1}^{N} \langle \phi_x^L | \rho_x^L(s) | \phi_x^L \rangle, \tag{14}$$

with N being the number of data points, $\rho_x^L(s)$ denoting the output state of the network, and $\left|\phi_x^L\right\rangle$ denoting the label.

In each iteration, the unitaries in the network are updated by $U^l(s+\epsilon) = e^{i\epsilon K^l(s)}U^l(s)$. Therefore, the network training is equivalent to obtaining $K^l(s)$ in each iteration. To this end, we first calculate the derivative of C with respect to the parameter s, which is

$$\frac{dC}{ds} = \lim_{\epsilon\to 0}\frac{C(s+\epsilon) - C(s)}{\epsilon}, \tag{15}$$

where ϵ is a small positive number. To evaluate this derivative, we first obtain the expression of $C(s+\epsilon)$. For the parameter $s+\epsilon$, the input quantum state stays unchanged as $\rho_x^0(s+\epsilon) = \rho_x^0 = \left|\phi_x^0\right\rangle\left\langle\phi_x^0\right|$. The quantum output, by the composition of layers, is however changed and can be expressed as

$$\rho_x^L(s+\epsilon) \quad = \quad \prod_{l=L}^{1} e^{i\epsilon K^l(s)}U^l(s)\rho_x^0 \prod_{l=1}^{L} {U^l}^\dagger(s)e^{-i\epsilon K^l(s)}.$$

We then substitute the updated quantum output corresponding to the parameter $s+\epsilon$ into the derivative of C, obtaining

$$\frac{dC}{ds} = \frac{i}{N}\sum_x \mathrm{tr}(\sum_{l=L}^{1} M^l(s)K^l(s)), \qquad (16)$$

where N is the number of samples, tr denotes the trace operation, and $M^l(s)$ is defined as

$$\begin{aligned} M^l(s) &= [\prod_{j=l}^{1} U^j(s)\rho_x^0 \prod_{j=1}^{l} {U^j}^\dagger(s), \\ &\qquad \prod_{j=l+1}^{L} {U^j}^\dagger(s)|\phi_x^L\rangle\langle\phi_x^L| \prod_{j=L}^{l+1} U^j(s)]. \end{aligned}$$

Here, $[\cdot,\cdot]$ denotes the commutator operation. The derivation of Eq. (16) is given in Appendix **??**.

To maximize the increase of C, $K^l(s)$ should be chosen such that dC/ds is maximized. For this purpose, we consider $K^l(s)$ which corresponds to a general n-qubit unitary $U^l(s)$. To avoid overfitting, we impose regularization on the parameters. Specifically, we regularize the parameters $K^l_{\alpha_1,\cdots,\alpha_n}(s)$ which are defined as

$$K^l(s) = \sum_{\alpha_1,\cdots,\alpha_n} K^l_{\alpha_1,\cdots,\alpha_n}(s)(\otimes_{k=1}^n \sigma^{\alpha_k}). \qquad (17)$$

Hence, the combined objective (which both maximizes the derivative of C and minimizes the change of network parameters) to be maximized is

$$\begin{aligned} C_2 &= \frac{dC}{ds} - \lambda\sum_{\alpha_i} K^l_{\alpha_1,\cdots,\alpha_n}(s)^2 \\ &= \frac{i}{N}\sum_x \mathrm{tr}(\sum_{l=L}^{1} M^l(s)K^l(s)) - \lambda\sum_{\alpha_j} K^l_{\alpha_1,\cdots,\alpha_n}(s)^2 \\ &= \frac{i}{N}\sum_x \mathrm{tr}_{\alpha_1,\ldots,\alpha_n}(\mathrm{tr}_{rest}(\sum_{l=L}^{1} M^l(s)K^l(s))) \\ &\quad -\lambda\sum_{\alpha_j} K^l_{\alpha_1,\cdots,\alpha_n}(s)^2. \end{aligned} \qquad (18)$$

To maximize C_2, we calculate its derivative with respect to $K^l_{\alpha_1,\cdots,\alpha_n}(s)$ as $i\sum_x \mathrm{tr}_{\alpha_1,\cdots,\alpha_n}(\mathrm{tr}_{rest}(M^l(s))(\otimes_{k=1}^n\sigma^{\alpha_k}))/N - 2\lambda K^l_{\alpha_1,\cdots,\alpha_n}(s)$. By setting it to 0 and solving for $K^l_{\alpha_1,\cdots,\alpha_n}(s)$, we obtain

$$K^l_{\alpha_1,\cdots,\alpha_n}(s) = \frac{i}{2N\lambda}\sum_x \mathrm{tr}_{\alpha_1,\cdots,\alpha_n}(\mathrm{tr}_{rest}(M^l(s))(\otimes_{k=1}^n\sigma^{\alpha_k}).$$

There is a caveat that C_2 might be always negative, in which case we should not update s. To ensure that the solution of $dC_2/dK^l_{\alpha_1,\ldots,\alpha_n} = 0$ results in an increase in C, we explicitly check the value of C before updating s. We substitute the obtained value of $K^l_{\alpha_1,\ldots,\alpha_n}(s)$ back to $K^l(s)$ and obtain

$$\begin{aligned} K^l(s) &= \frac{i}{2N\lambda}\sum_{\alpha_1,\cdots,\alpha_n}\sum_x \mathrm{tr}_{\alpha_1,\cdots,\alpha_n}(\mathrm{tr}_{rest}(M^l(s)) \\ &\quad (\otimes_{k=1}^n\sigma^{\alpha_k}))(\otimes_{k=1}^n\sigma^{\alpha_k}) \\ &= i2^n\sum_x \mathrm{tr}_{rest}(M^l(s))/(N\lambda). \end{aligned} \qquad (19)$$

Finally the unitary U^l is updated by

$$U^l(s+\epsilon) = \exp(-\epsilon 2^n \sum_x \mathrm{tr}_{rest}(M^l(s))/(N\lambda))U^l(s).$$

Now we consider the specific unitaries used by the DIBoM which can be categorized into three cases:

1. The first case is U^j_{SG}, which is a single-qubit unitary on the qubit j. The corresponding K for this unitary is $K^l(s) = \sum_{\alpha=0}^3 K^l_\alpha(s)\sigma^\alpha$. To obtain the optimal $K^l_\alpha(s)$, we set $dC_2/dK^l_\alpha(s) = 0$ and obtain

$$K^l_\alpha(s) = \frac{i}{2N\lambda}\sum_x \mathrm{tr}_j(\mathrm{tr}_{rest}(M^l(s))\sigma^\alpha), \qquad (20)$$

where "rest" denotes qubits other than qubit j. Substituting the expression back into $K^l(s)$, we have

$$\begin{aligned} K^l(s) &= \frac{i}{2N\lambda}\sum_\alpha\sum_x \mathrm{tr}_j(\mathrm{tr}_{rest}(M^l(s))\sigma^\alpha)\sigma^\alpha \\ &= i\sum_x \mathrm{tr}_{rest}(M^l(s))/(N\lambda). \end{aligned} \qquad (21)$$

Therefore the unitary is updated as $U^l(s+\epsilon) = \exp(i\epsilon K^l(s))U^l(s)$.

2. The second case is a product of single-qubit gates $U^{\otimes}_{SG}$, the corresponding K of which has the form

$$K^l(s) = \sum_{j=1}^n\sum_{\alpha=0}^3 K^l_{j,\alpha}(s)\sigma^\alpha_j. \qquad (22)$$

By letting $dC_2/dK^l_{j,\alpha}(s) = 0$, we obtain

$$K^l_{j,\alpha}(s) = \frac{i}{2N\lambda}\sum_x \mathrm{tr}_j(\mathrm{tr}_{[n]\backslash\{j\}}(M^l(s))\sigma^\alpha_j), \qquad (23)$$

where $[n]\backslash\{j\}$ refers to all qubits except qubit j. Substituting this expression back into $K^l(s)$, we have

$$\begin{aligned} K^l(s) &= \frac{i}{2N\lambda}\sum_{j=1}^n\sum_\alpha\sum_x \mathrm{tr}_j(\mathrm{tr}_{[n]\backslash\{j\}}(M^l(s))\sigma^\alpha_j)\sigma^\alpha_j \\ &= i\sum_{j=1}^n\sum_x \mathrm{tr}_{[n]\backslash\{j\}}(M^l(s))/(N\lambda). \end{aligned} \qquad (24)$$

Therefore the unitary is updated as $U^l(s+\epsilon) = \exp(i\epsilon K^l(s))U^l(s)$.

3. The third case is the collection of generalized CZ gates on all pairs of qubits U_{CZ}. Its corresponding K is $K^l(s)=\sum_{1\le j<k\le n}K^l_{jk}(s)|11\rangle_{jk}\langle 11|$. By setting $dC_2/dK^l_{jk}(s)=0$, we obtain

$$K^l_{jk}(s)=\frac{i}{2N\lambda}\sum_x \mathrm{tr}_{j,k}(\mathrm{tr}_{[n]\backslash\{j,k\}}M^l(s))|11\rangle_{jk}\langle 11|, \tag{25}$$

where $[n]\backslash\{j,k\}$ refers to the set of qubits excluding qubits j and k. Substituting this expression back into $K^l(s)$, we have

$$\begin{aligned} K^l(s) &= \frac{i}{2N\lambda}\sum_{j,k}\sum_x \mathrm{tr}_{j,k}(\mathrm{tr}_{[n]\backslash\{j,k\}}(M^l(s))|11\rangle_{jk} \\ &\quad \langle 11|)|11\rangle_{jk}\langle 11| \\ &= i[\sum_x\sum_{j,k}\mathrm{tr}_{j,k}(\mathrm{tr}_{[n]\backslash\{j,k\}}(M^l(s))|11\rangle_{jk}\langle 11|) \\ &\quad |11\rangle_{jk}\langle 11|]/(2N\lambda). \end{aligned}$$

Therefore the unitary is updated as $U^l(s+\epsilon)=\exp(i\epsilon K^l(s))U^l(s)$.

Three final remarks are in order. First, the hyperparameter λ is set to 0.5 in the simulation unless otherwise stated. Second, the classical simulation of the gradient descent for each parameter s in layer-by-layer training is identical to the simultaneous training method. Third, the classical simulation for training controlled unitaries V_j is similar, and the specifics are deferred to Appendix C.

6 Empirical evaluation results

Having presenting the simulation setup, this section proceeds to present the empirical evaluation results. We first empirically compare the performance of DIBoM with other QNNs in Sec. 6.1. Next, we conduct an ablation study on the DIBoM in Sec. 6.2 to investigate the individual components of the model. Then in Sec. 6.3, we assess the sensitivity of the performance of the DIBoM to its various parameters. Additionally, in Sec. 6.4, we analyze the robustness of the DIBoM to noise. Finally, in Sec. 6.5, we mitigate the barren plateau issue in the training of the DIBoM. Some auxiliary details pertaining to the construction of the datasets used in this section are presented in Appendix D.

6.1 Performance comparison

To begin, we analyze the training performance of the DIBoM and compare it to that of other models, considering both the converged loss and the model complexity. Additionally, we explore two different training methods and evaluate the gap between training and test performance. Moreover, we plot the optimization landscape of the DIBoM and dissipative QNN to gain a deeper understanding of their respective training processes.

We first test the simultaneous training and layer-by-layer training methods and display four training results with different datasets in Fig. 10. The results indicate that the model's loss converges to 0. Comparing the two training methods, we observe that layer-by-layer training consistently performs worse than simultaneous training. Thus, we will solely utilize simultaneous training in future simulations.

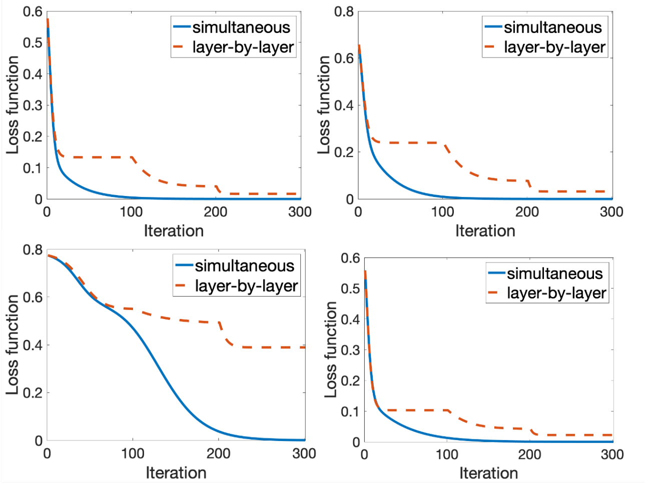


Figure 10: Four learning curves for a DIBoM with two training methods.

We next examine the DIBoM's prediction accuracy on the test set and assess its gap with the training accuracy, as depicted in Fig. 11. Notably, the training loss and test loss are nearly identical, with the test loss occasionally being smaller than the training loss. This suggests that statistical fluctuations rather than generalization errors may cause the deviation between the training and test losses. Given the close proximity of the two losses, we will exclusively evaluate the test loss in subsequent simulations.

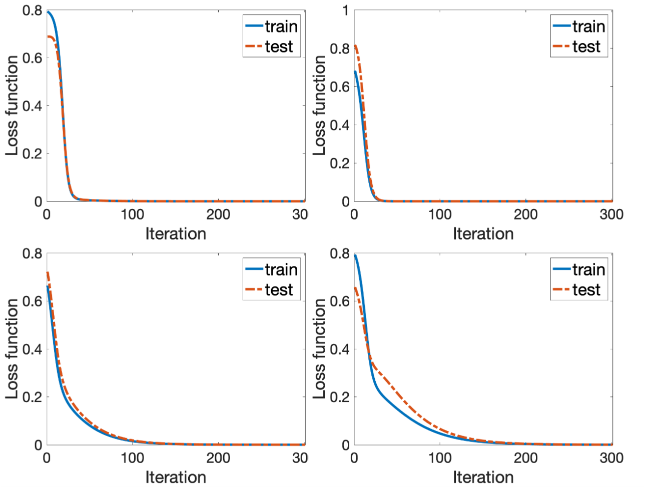


Figure 11: Four training and test learning curves for a DIBoM.

We then compare the DIBoM with three other QNNs: a hardware-efficient QNN [McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven], a dissipative QNN [Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, Scheiermann, and Wolf] and an Ising

Born machine [Coyle *et al.*(2020)Coyle, Mills, Danos, and Kashefi]. In the simulation, we set the number of qubits to be 2, the DIBoM to contain five alternating layers of tunable single-qubit gates and tunable two-qubit gates, the hardware-efficient QNN to have the same structure as the DIBoM but with all tunable two-qubit gates replaced by CZ gates, and the Ising Born machine to contain a layer of tunable two-qubit gates followed by a layer of tunable single-qubit gates. The intrinsic unitary V of this simulation is restricted to have the same structure as the DIBoM but with unknown parameters. We compare the models based on two criteria. First, we compare them in terms of their performance, as shown in Fig. 12. We observe that both the dissipative QNN and the DIBoM reach zero loss while the hardware-efficient QNN and the Ising Born machine do not. This could be attributed to the limited expressive power of the hardware-efficient QNN and the Ising Born machine. We also observe that the dissipative QNN converges faster than the DIBoM. Second, we compare them in terms of the number of model parameters in Fig. 13. We observe that the dissipative QNN requires significantly more parameters than the other three models. As the number of qubits increases, the ratio of the number of parameters of the DIBoM to that of the dissipative QNN tends to 0. Hence, there is a tradeoff between performance and the number of parameters.

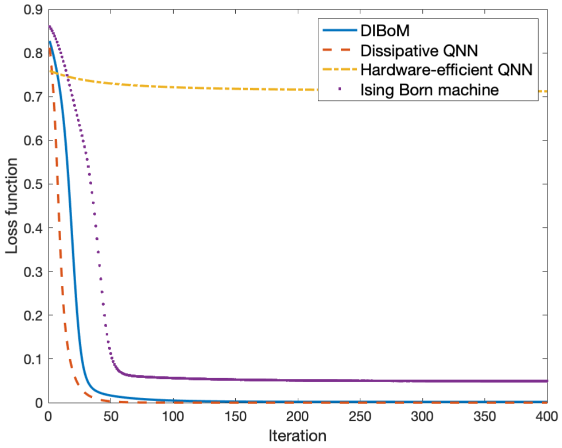


Figure 12: Comparison of a DIBoM with a dissipative QNN, a hardware-efficient QNN and an Ising Born machine in terms of the performance for a structured unitary V.

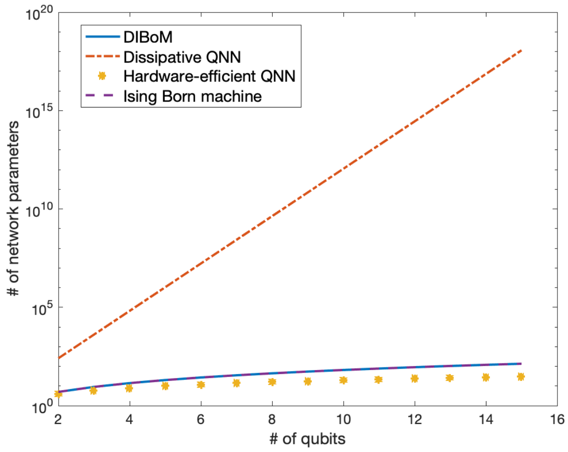


Figure 13: Comparison of a DIBoM with a dissipative QNN, a hardware-efficient QNN and an Ising Born machine in terms of the number of model parameters.

To investigate why the DIBoM and dissipative QNN can achieve zero loss, we plot the optimization landscapes of them with two parameters varying and other parameters fixed. The result of the DIBoM is shown in Fig. 14, where one parameter to be changed is from the single-qubit unitary gate (parameter 1) and the other parameter to be changed is from the generalized CZ gate (parameter 2). Despite the highly non-convex landscape, all local minima are global minima, explaining why the DIBoM can always converge to zero loss. The result of the dissipative QNN is shown in Fig. 15. It can be seen that there is no flat region in the landscape and this explains the fast convergence of a dissipative network. However, there exists a local minimum in the middle of the figure which does not coincide with the global minimum. Hence, whether dissipative networks can always be trained to reach a global minimum requires further investigation.

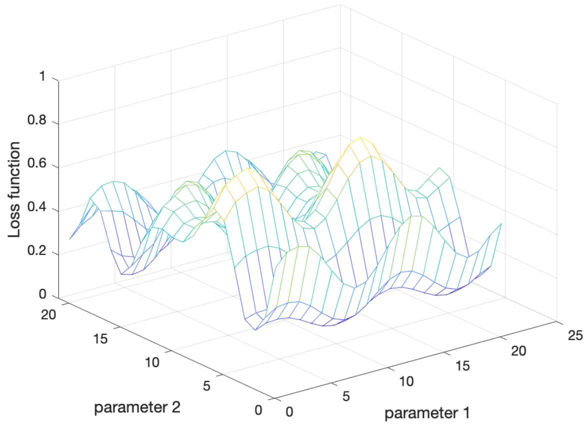


Figure 14: Optimization landscape of a DIBoM with two parameters varying and other parameters fixed.

To facilitate a performance comparison between the DIBoM and other models, we have assumed that the hidden unitary V has the same structure as the DIBoM, but with unknown parameters. To further evaluate the DIBoM's performance, we conduct a test where V is a random unitary, with $n = 3$ qubits and 10 layers for both the DIBoM and hardware-efficient QNN. The results are shown in Fig. 16. The DIBoM outperforms the hardware-efficient QNN and the Ising Born machine but has lower accuracy than the dissipative QNN. As compensation, the number of parameters in a DIBoM scales only quadratically with n, whereas the dissipative QNN

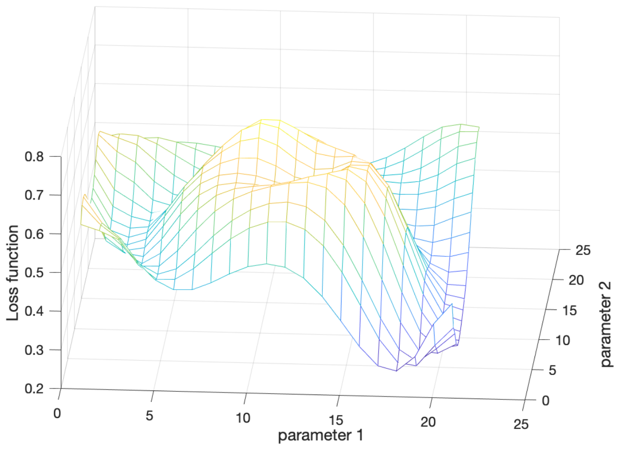


Figure 15: Optimization landscape of a dissipative QNN with two parameters varying and other parameters fixed.

has exponential scaling.

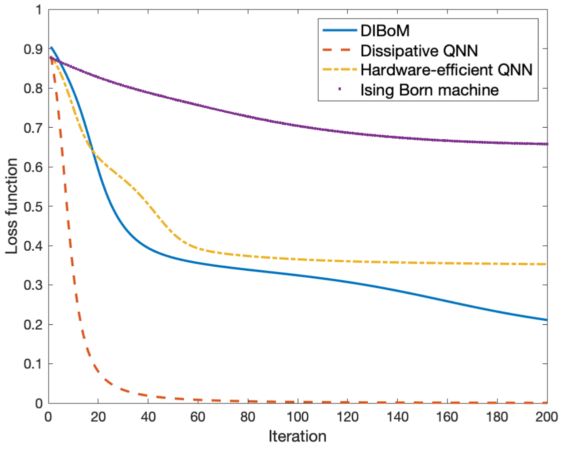


Figure 16: Comparison of a DIBoM with a dissipative QNN, a hardware-efficient QNN and an Ising Born machine in terms of the performance for a random unitary V.

6.2 Ablation study

After analyzing the overall performance of the DIBoM architecture, we next alter some components of the DIBoM to get a better understanding of the contribution of each component of the DIBoM to its performance.

First, we investigate the case that the DIBoM consists of a layer of one single-qubit gate (U_{SG}^2), which acts exclusively on the second qubit. The training result is plotted in Fig. 17, where it is evident that the fidelity of the model output reaches 1 after sufficient training. Moreover, it converges rapidly, reaching zero loss by the 14th iteration. To delve deeper into the optimization process and understand why the optimization of the single-qubit unitary does not trap at a local minimum, we plot the loss function as a function of two parameters while keeping all other parameters fixed. Both parameters are from the single-qubit unitary, and the resulting plot is depicted in Fig. 18. Despite the highly non-convex optimization landscape, it is noticeable that all local minima have approximately the same value, explaining the achievement of zero loss during training of a single-qubit unitary.

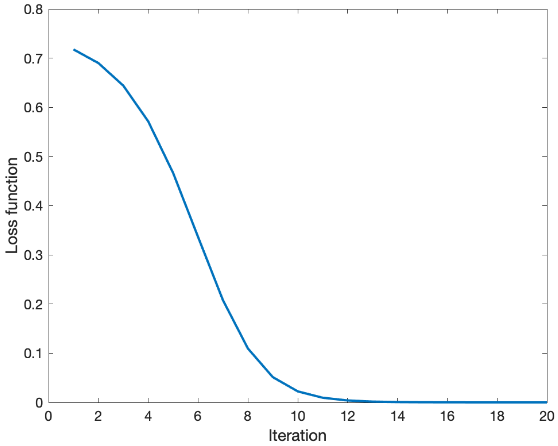


Figure 17: Learning curve of a variant of the DIBoM (U_{SG}^2).

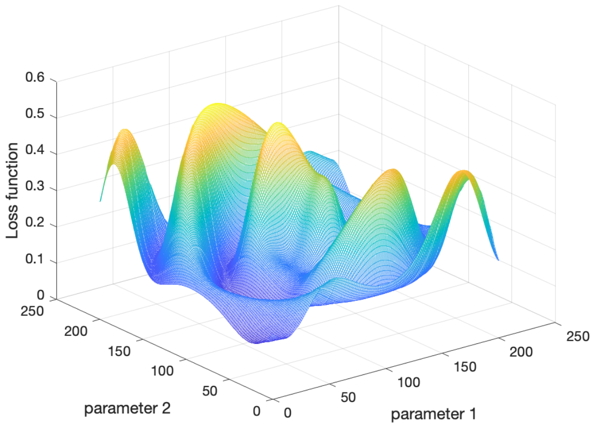


Figure 18: Optimization landscape of a variant of the DIBoM (U_{SG}^2) with two parameters varying and other parameters fixed.

Next we examine the case that the DIBoM consists of one layer of generalized CZ gates (U_{CZ}). As displayed in Fig. 19, the fidelity of the model output approaches unity after sufficient training, similar to the case of single-qubit gates. Note however that the initial loss for the generalized CZ gate case is lower than that of the single-qubit gate case. This observation suggests that a generalized CZ gate is more rigid than a single-qubit gate, implying a narrower range of variation in the quantum output induced by the former. Note also that convergence for the generalized CZ gate case occurs around the 50th iteration, which is slower than the convergence rate for the single-qubit gate case, indicating the former case is comparatively harder to train than the latter case.

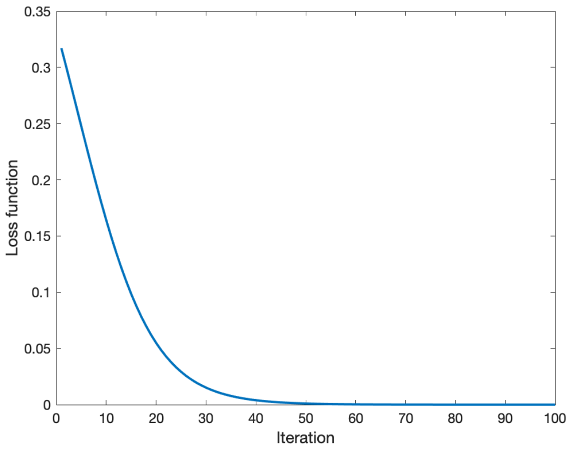


Figure 19: Learning curve for a variant of the DIBoM (U_{CZ}).

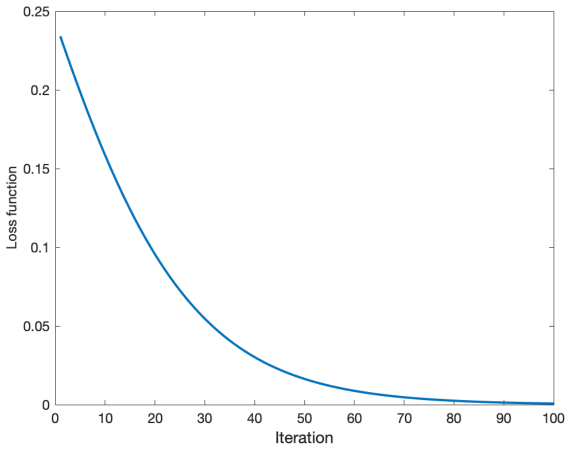


Figure 21: Learning curve of a DIBoM ($U_{CZ}U^2_{SG}$) with the second layer trained and the first layer fixed.

Then we investigate the DIBoM ($U_{CZ}U^2_{SG}$) where one of the layers is fixed. For the case that the second layer is fixed, the training result is shown in Fig. 20. As illustrated, the training converges to zero loss and the convergence is quite fast, reaching the optimal loss at the 40th iteration. Conversely, for the case that the first layer of the DIBoM is fixed while the second layer is to be trained, the convergence is slower, as evidenced by the training results shown in Fig. 21. This slower convergence could be attributed to the flatter optimization landscape of the generalized CZ gate. Nonetheless, the training also eventually converges to zero loss.

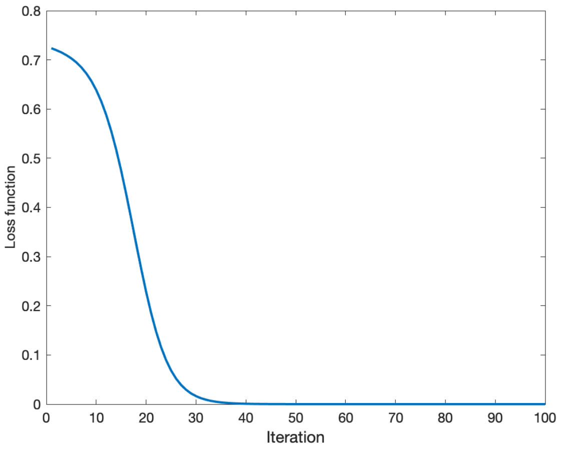


Figure 20: Learning curve of a DIBoM ($U_{CZ}U^2_{SG}$) with the first layer trained and the second layer fixed.

Next we examine a network comprising a layer of single-qubit gates on all qubits followed by a layer of generalized CZ gates ($U_{CZ}U^{\otimes}_{SG}$). Here, $U^{\otimes}_{SG}$ is a n-qubit gate that can be decomposed as a tensor product of parameterized single-qubit gates, potentially featuring varying parameters. We refer to the former layer as a *product gate* layer. Figure 22 compares the performance of the product gate case to the single gate setting. It can be seen that the product gate case converges to a zero loss while the single gate case does not, implying that the product gate case has more expressive power.

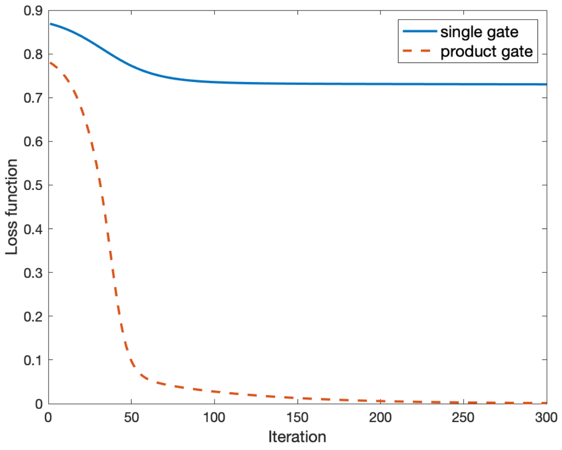


Figure 22: The performance comparison between the product-gate variant of the DIBoM ($U_{CZ}U^{\otimes}_{SG}$) and a DIBoM ($U_{CZ}U^2_{SG}$).

An ablation study of generalized CZ gates is then conducted by comparing the following two models. The first model, denoted as "generalized CZ", utilizes DIBoM with 3 layers ($U^{\otimes}_{SG}U_{CZ}U^{\otimes}_{SG}$). The second model, denoted as "normal CZ", is almost identical to the first model, except that all the generalized CZ gates in the second layer are replaced with normal CZ gates. The study's results are presented in Fig. 23, which indicates that the first model achieves a much lower loss than the second model. This finding suggests the strong effectiveness of generalized CZ gates.

To evaluate the effect of the controlled unitary V_j, we conduct an ablation study. For this purpose, we modify the training dataset to have different input and output dimensions. Specifically, we construct the training dataset as $(|\psi_i\rangle \otimes |0\rangle \otimes |0\rangle, |\psi_i\rangle)_{1\leq i\leq N}$, where $|\psi_i\rangle$ represents a

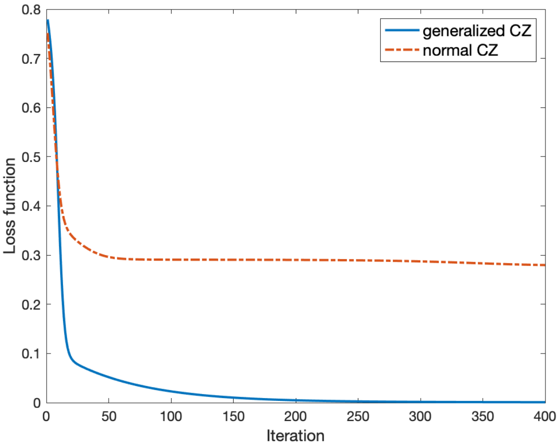


Figure 23: Performance comparison between a DIBoM and an ablated DIBoM where generalized CZ gates are replaced by normal CZ gates.

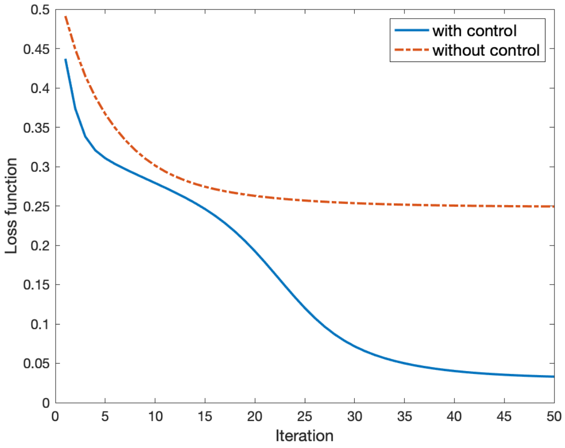


Figure 24: The performance comparison between a full DIBoM with controlled gates and an ablated DIBoM without controlled gates.

randomly generated pure qubit using the Haar measure. The objective of the model is to transform the quantum sample such that the third qubit approximates the quantum label as closely as possible. We denote the complete DIBoM model as *with control*. The model first applies a three-qubit unitary U, measures the first two qubits in the computational basis, and then uses the measurement result j to perform the controlled unitary V_j on the third qubit. On the other hand, the ablated version, denoted as *without control*, includes the same first two components but lacks the final component. More precisely, the unitary U is composed of three layers $U_{SG}U_{CZ}U_{SG}$, and there are four single-qubit unitaries V_j $(0 \leq j \leq 3)$ to be optimized. The results are illustrated in Fig. 24, which indicates that the complete DIBoM model achieves significantly lower loss than the ablated version. This suggests that the complete DIBoM model can effectively learn the quantum teleportation protocol from scratch, while the ablated version lacks this capability.

6.3 Parameter sensitivity

So far, we have investigated the performance of the DIBoM and the contributions of its individual components to this performance. For a more comprehensive evaluation of the DIBoM, we conduct experiments to test the effect of different parameters, including the data size, the number of layers, the number of qubits per layer, and the regularization constant.

We begin by examining the relation between the number of training samples N and the performance characterized by the training loss. The result is shown in Fig. 25, which clearly demonstrates that the larger the sample size, the faster the convergence. This is because more samples help to estimate the intrinsic unitary V better during training. We next test the relation between the number of samples and the gap between the training performance and the test performance, as shown in Fig. 26. It can be seen that as the number of samples increases, the gap gradually becomes smaller. This is expected as a larger number of samples results in a smaller variance and thus, improved generalization performance.

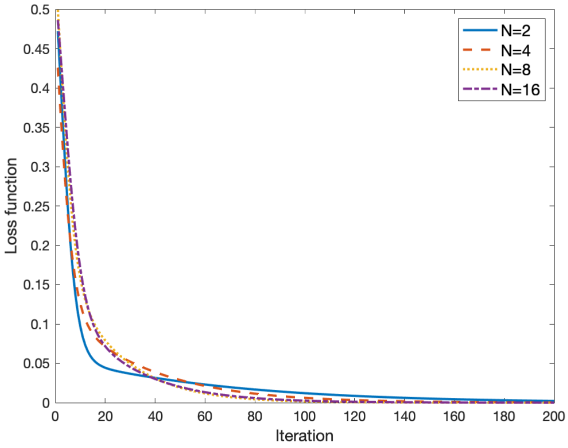


Figure 25: The training performance of a DIBoM with a varying number N of training samples.

Next, we test the effects of the layer number L on the performance of the product gate variant of the DIBoM $(\cdots U_{SG}^{\otimes}U_{CZ}U_{SG}^{\otimes})$, as shown in Fig. 27. It can be seen that the losses of all cases converge to zero loss, indicating that the network can scale to many layers without adversely affecting trainablity.

Then we test the number of qubits n that the input and output contain, also on the product gate variant of the DIBoM $(U_{CZ}U_{SG}^{\otimes})$. As shown in Fig. 28, for all cases from 2 qubits to 4 qubits, the DIBoM can train well, with zero loss convergence. However, the case of $n = 5$ did not perform as well, resulting in non-zero loss. Therefore, we perform four additional simulations for $n = 5$, which are displayed in Fig. 29. It is evident that although the loss function eventually converges to zero loss, this requires a larger number of iterations. Moreover, the training curve displays both slow-varying and fast-varying re-

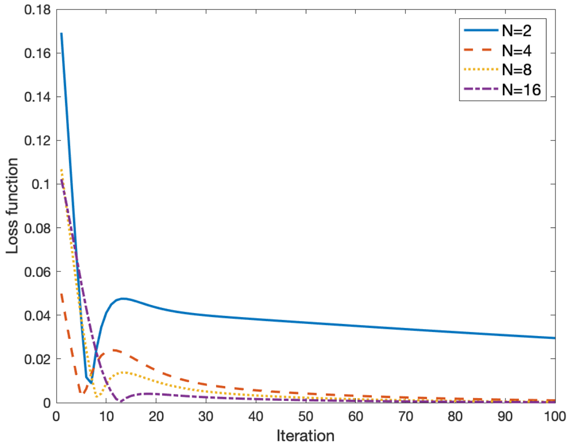


Figure 26: The gap between the training and test performance of a DIBoM with N training samples.

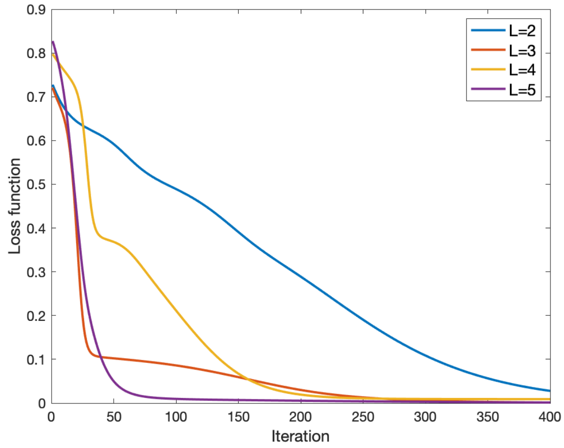


Figure 27: The performance of a DIBoM with a different number L of layers.

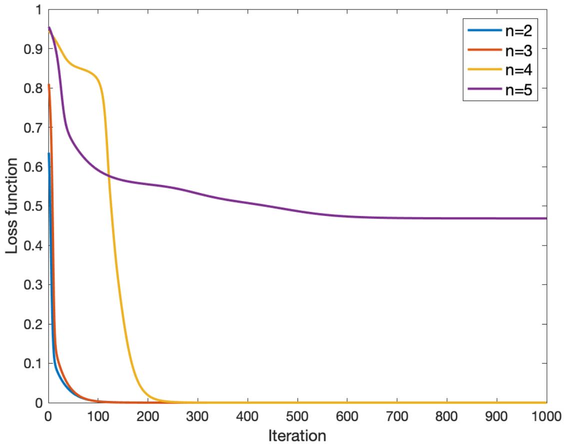


Figure 28: The performance of a DIBoM with a different number of qubits n per layer.

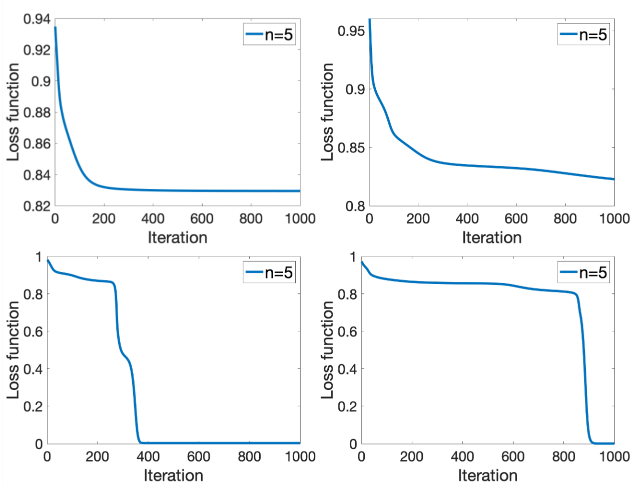


Figure 29: Four learning curves of a DIBoM with 5 qubits per layer.

gions, a phenomenon that is already observable in the case of $n = 4$, but is more pronounced when $n = 5$.

In addition, we test the effect of the regularity constraint parameter λ on the performance of the DIBoM. The results, shown in Fig. 30, indicate an optimal value of λ at 0.5. Deviating from this value leads to worse convergence.

Finally, we assess the ability of a DIBoM $(\cdots U_{SG}^{\otimes} U_{CZ} U_{SG}^{\otimes})$ to approximate an arbitrary unitary using 2 to 5 layers, alternating between product gate and generalized CZ gate layers. The results, presented in Fig. 31, demonstrate that the converged loss decreases as the number of layers increases. Notably, with only 5 layers, the DIBoM already achieves a low loss.

6.4 Robustness to noise

We further assess the robustness of a DIBoM to noise in the data since real-world data inevitably contains noise. To this end, we manually corrupt some of the training data and evaluate its effect on the test loss. The corruption ratio of the training data varies from 10% to 100%, with an uncorrupted dataset, denoted by "original", acting as the control. Given a corruption ratio, say 30%, we randomly select 30% of real quantum data and replace it with fake data $\left(\left|\phi_x^{in}\right\rangle, \left|\phi_x^{out}\right\rangle\right)$, where $\left|\phi_x^{in}\right\rangle$ and $\left|\phi_x^{out}\right\rangle$ are Haar random n-qubit pure states. The results are visualized in Fig. 32. On the negative side, with a gradual increase of the corrupted ratio, the performance gradually degrades, as evidenced by the comparison of solid and dashed lines. On the positive side, even with up to 60% of corrupted data, the DIBoM remained effective, which indicates a high level of robustness against noise.

We also investigate the noise robustness of the model with respect to varying numbers of layers. To perform this investigation, we set a fixed corruption ratio of 20% and vary the number of layers, while monitoring the loss at the 300 iterations. The results are presented in Fig. 33. Notably, our analysis reveals that the noise robustness of the model remains consistent across different numbers of layers. This finding highlights the scalability of the

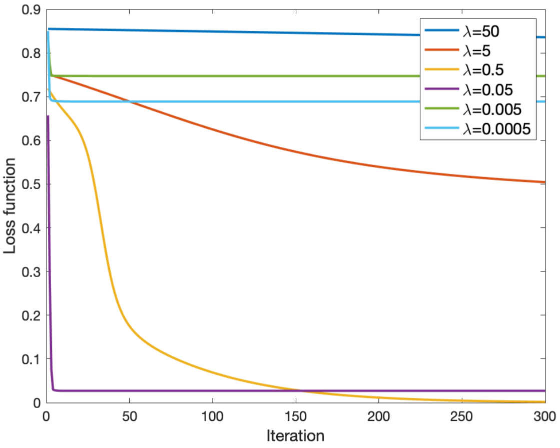


Figure 30: The performance of a DIBoM with different parameters of the regularity constraint λ.

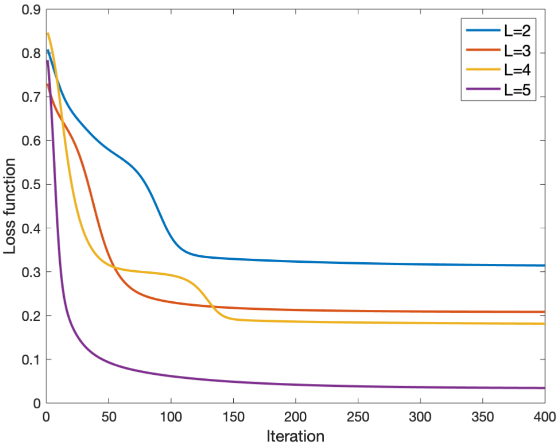


Figure 31: The performance of a DIBoM with L layers on a generic quantum learning dataset.

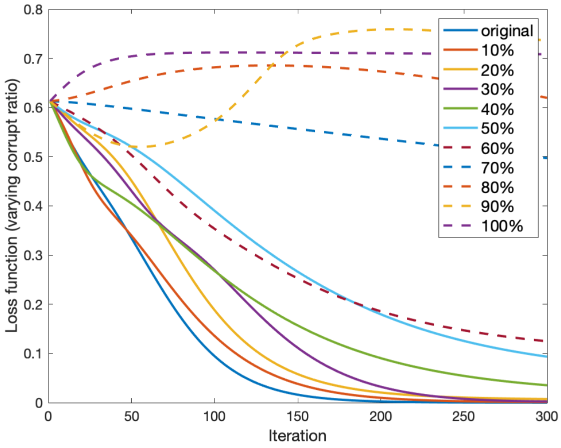


Figure 32: The performance of a DIBoM with various proportions of the training data corrupted.

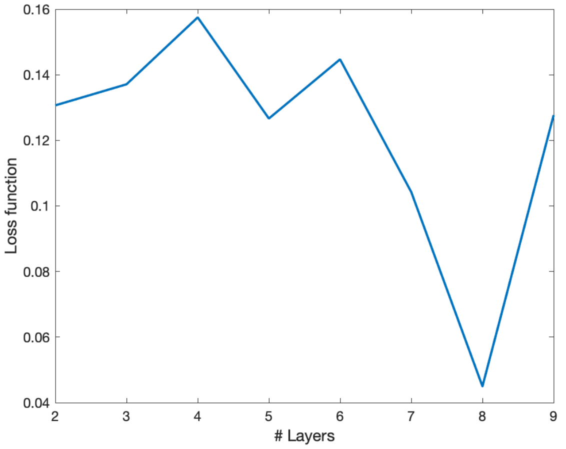


Figure 33: Noise robustness of the DIBoM with different numbers of layers.

DIBoM and suggests that increasing the number of layers does not negatively impact the noise robustness of the model.

6.5 Barren plateau

In Sec. 6.3, we have observed that the model with the global loss function Eq. (7) already suffers from the barren plateau (slow-varying region) issues for $n = 5$ qubits. Previous work by Cerezo et al. [Cerezo *et al.*(2021)Cerezo, Sone, Volkoff, Cincio, and Coles] demonstrated that local cost functions can mitigate this issue. Hence we incorporate the local cost function described in Eq. (8). We defer the details of the classical simulation of the model under the local loss function to Appendix E. To accommodate the local cost function, we design the training data as a product state $|\phi_{in}\rangle = |\phi_{in}\rangle_1 \otimes \cdots \otimes |\phi_{in}\rangle_n$ for n qubits, which we refer to as *product-form training data.* Notably, a zero local cost function for this data implies a zero global cost function. We examine the performance of the local cost function by training the model with 2 to 5 qubits, as shown in Fig. 34. We observe that all learning curves converge to zero loss quickly, suggesting that the barren plateau issue is mitigated. This stands in stark contrast to the global cost function, which exhibits the barren plateau phenomenon when $n = 5$.

We have also plotted the comparison between local and global cost functions in Fig. 35. The local cost function always reaches zero loss faster than the global cost function. More importantly, the curve of the local cost function lacks a flat region, suggesting that the barren plateau phenomenon is mitigated for the DIBoM with a local cost function. Notably, we observe that the barren plateau issue is also mitigated for a global cost function with product-form training data. This observation suggests that product-form training data may be easier to train than entangled training data.

Finally, we discuss the computation time required for our simulations. For the case $n = 8$, which is the largest simulation we performed, the computation takes about 1 hour on an 8-core 3.2GHz computer. The number of parameters for this case is $4n + n(n-1)/2 = 60$. It is

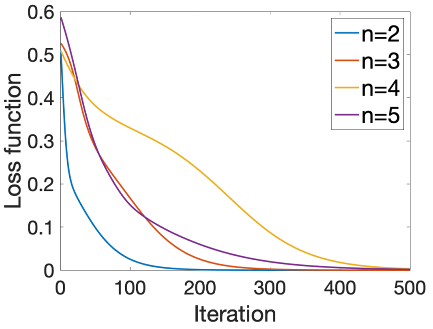


Figure 34: The performance of a DIBoM with a local cost function. Here, n is the number of qubits.

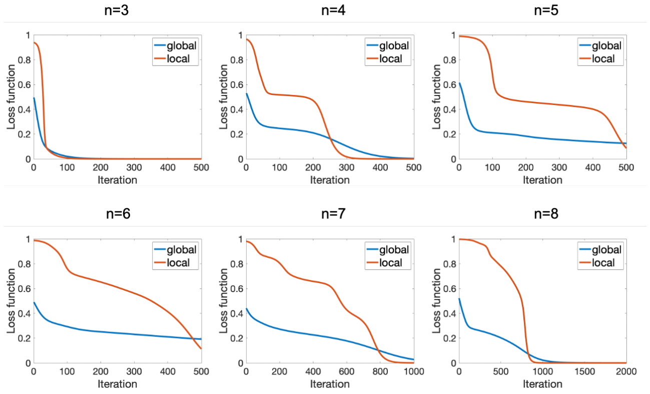


Figure 35: Comparisons between global and local cost functions with different numbers of qubits n. Here, n ranges from 3 to 8. The redline depicts the local cost function while the blue line depicts the global cost function.

worth noting that the computation time is exponential in n, which is a consequence of the classical computation requiring multiple multiplications on density matrices of size $2^n \times 2^n$ in each iteration. Each of these multiplications takes time 2^{3n}, resulting in a cost of 256^3 per multiplication when $n = 8$. Furthermore, the number of iterations is polynomially related to n, with approximately 2000 iterations required for $n = 8$. Hence, the computation burden on a classical computer is substantial. However, it is important to note that intermediate-scale quantum computers have the potential to substantially decrease computation time to a polynomial of n, as the classical manipulation of $2^n \times 2^n$ density matrices is no longer necessary. As such, we expect that current noisy intermediate-scale quantum (NISQ) devices will boost the trainable size to a few hundred qubits.

7 Discussion

In summary, we examined a deep Ising Born machine (DIBoM) and showed it has a good balance between efficiency and universality. Specifically, we described its model architecture and its training procedure. Additionally, through theoretical analysis, we demonstrated that the DIBoM has the capability of universal quantum computation. Apart from the theoretical analysis, we empirically evaluated the performance of the DIBoM and compared it with other QNNs. Our evaluations revealed that the DIBoM has a moderate number of parameters while being quite expressive. Along the way, we introduced a new expressivity measure called fidelity-based expressivity, which may be of independent interest.

There are two potential limitations of the DIBoM: trainability and generalizability. Trainability refers to the ability to find the global minimum in a polynomial number of iterations with respect to the number of qubits. In our simulations, we have shown that the DIBoM is trainable for a moderate number of qubits, but it is unclear whether it remains trainable for a large number of qubits, which is beyond the capability of our simulation. Moreover, there is no theoretical guarantee that the DIBoM is trainable, and recent negative results [Anschuetz and Kiani(2022)] suggest that most shallow and local QNN architectures are not trainable.

Generalizability refers to the ability to achieve low test error given a small training error. In our simulations, we have empirically observed that the DIBoM has low test error, but we have no theoretical guarantee for this fact. When the number of parameters of a QNN is much larger than the number of training data, over-parameterization can cause overfitting, making it difficult to achieve theoretical guarantees of generalizability. This is a challenging problem even for classical neural models.

Therefore, it is crucial to develop QNN architectures that achieve efficiency, universality, provable trainability, and provable generalizability simultaneously. The DIBoM only addresses the first two goals, leaving much room for improvement.

There are a few other promising avenues for future research. First, applying the DIBoM to downstream quantum learning tasks is likely to be both fruitful and interesting. Second, an experimental demonstration of the DIBoM on quantum hardware would be interesting. Third, due to the exponential cost of the classical simulation, a NISQ device may show a speed advantage in training the DIBoM, which makes it an ideal target for showing quantum supremacy on practical problems.

The tradeoff between efficiency and universality is also worth further investigation. To achieve universality as defined in our work, an ansatz needs to have exponentially many parameters because the ansatz cannot express all unitaries if the dimension of its Hilbert space is smaller than that of $\mathsf{SU}(n)$. This implies full universality and efficiency cannot be achieved simultaneously for any quantum learning model. There are several directions to further explore the tradeoff between universality and efficiency and the associated design of quantum learning models.

First, by relaxing the definition of universality, there may exist more interesting tradeoff between universality and efficiency. However, this makes the research landscape more complex since there are a lot of ways to weaken the notion of universality. Previous research has considered weakening the universality to the class of functions that maps 0 to the ground state of a Hamiltonian which is the sum of poly(n) Pauli bases [Biamonte(2021)], that is real and continuous [Goto *et al.*(2021)Goto, Tran, and Nakajima], that can be described by a quantum circuit with a polynomial number of gates where each gate acts on a constant number of qubits [Cai *et al.*(2022)Cai, Ye, and Deng], and that is boolean [Herman *et al.*(2022)Herman, Raymond, Li, Robles, Mezzacapo, and Pistoia]. Besides these choices, there are many other choices available, potentially infinitely many. For example, one can consider the class of function that maps 0 to the thermal state of a Hamiltonian which is the sum of poly(n) Pauli bases, that is complex and meromorphic, that can be described by a quantum circuit with a logarithmic number of gates where each gate acts on a logarithmic number of qubits, to just name a few. How to achieve these different types of universality while maintaining efficiency in a strict sense is an interesting research question.

Another direction is to replace universality by restricting the ansatz to contain the solution one is looking for. In this case, the ansatz becomes problem specific, which is not a universal ansatz that can deal with all learning problems. Following this line, after taking a learning problem, one should design a specific ansatz that suits this problem which requires additional manual work and expertise. How to reduce the manual efforts in designing a specific ansatz for a given learning problem (such as combinatorial optimization problems [Zhou *et al.*(2020)Zhou, Wang, Choi, Pichler, and Lukin], learning the ground state of a Hamiltonian [Motta *et al.*(2020)Motta, Sun, Tan, O'Rourke, Ye, Minnich, Brandão, and Chan], simulating quantum dynamics [Sparrow *et al.*(2018)Sparrow, Martín-López, Maraviglia, Neville, Harrold, Carolan, Joglekar, Hashimoto, Matsuda, O'Brien *et al.*], or drug discovery for a specific disease [Cao *et al.*(2018)Cao, Romero, and Aspuru-Guzik]) is an interesting question on its own.

This work was supported by the National Natural Science Foundation of China (12105105), the Natural Science Foundation of Shanghai (21ZR1415800), the Shanghai Sailing Program (21YF1409800), and the startup fund from East China University of Science and Technology (JKH01221665 and YH0142206).

A Evaluation details of the Fidelity-based Measure

Here, we detail how we numerically evaluate an upper bound of the fidelity-based measure (FBE) for any QNN architecture in Fig. 8.

For a QNN architecture $\mathsf{A}(\theta)$, we first select $k = 100$ random unitaries U_i $(1 \leq i \leq k)$ and $m = 10$ random pure states ϕ_j $(1 \leq j \leq m)$. Then for any U_i, we optimize θ to minimize

$$s_i = \frac{1}{m}\sum_{j=1}^{m} |\langle \phi_j | \mathsf{A}(\theta)^\dagger U_i | \phi_j \rangle|. \tag{26}$$

The bound $\min_i s_i$ is then taken as the upper bound of the FBE.

B Concrete number of parameters for a universal 3-qubit DIBoM

In this section, we provide the precise number of parameters needed for a 3-qubit DIBoM to be universal.

To begin with, according to Section 4.5.1 of Ref. [Nielsen and Chuang(2010)], a 3-qubit unitary can be decomposed as a product of $2^{3-1}(2^3-1) = 28$ two-level unitaries. Our next goal is to decompose a two-level 3-qubit unitary further.

As per Section 4.5.2 of Ref. [Nielsen and Chuang(2010)], a two-level 3-qubit unitary can be decomposed as a product of at most 5 controlled-controlled single-qubit unitaries. Our next goal is to decompose a controlled-controlled single-qubit unitary further.

Figure 4.18 of Ref. [Nielsen and Chuang(2010)] reveals that a controlled-controlled single-qubit unitary can be decomposed as a product of 2 CZ gates and 3 controlled single-qubit unitary gates. Our next goal is to decompose a controlled single-qubit unitary further.

Figure 4.6 of Ref. [Nielsen and Chuang(2010)] states that a controlled single-qubit unitary can be decomposed as a product of two CZ gates and single-qubit gates. Consequently, a controlled-controlled single-qubit unitary can be decomposed as a product of $2+3\times 2 = 8$ CZ gates and single-qubit gates, a two-level 3-qubit unitary as a product of $8\times 5 = 40$ CZ gates and single-qubit gates, and a 3-qubit unitary as a product of $40 \times 28 = 1120$ CZ gates and single-qubit gates. This implies that a $(1120 \times 2 + 1)$-layer 3-qubit DIBoM is sufficient to realize any 3-qubit unitary.

A $(1120 \times 2 + 1)$-layer 3-qubit DIBoM contains 1121 single-qubit gate layers and 1120 generalized CZ gate layers. This translates to a total of $1121 \times 9 + 1120 \times 3 = 13449$ parameters.

%endwidetext

C Simulation of the controlled unitary V_j

In this section, we present the classical simulation that involves the controlled unitaries V_j. Before any measurements, the initial quantum state ρ^0 is evolved to the following quantum state:

$$\rho^1 = U^k \cdots U^1 \rho^0 {U^1}^\dagger \cdots {U^k}^\dagger. \tag{27}$$

After measuring with outcome i, the post-measurement state is given by

$$\rho_i^2 = \mathrm{tr}_A(\rho^1(|i\rangle\langle i|_A \otimes I_B)). \tag{28}$$

The post-measurement states then undergo another series of unitaries and become

$$\rho^3 = \sum_i V_i^l \cdots V_i^1 \rho_i^2 {V_i^1}^\dagger \cdots {V_i^l}^\dagger \triangleq \sum_i \mathcal{E}^i(\rho^1), \tag{29}$$

where $\mathcal{E}^i$ is a quantum operation that acts on the state ρ^1.

After getting the output of the quantum circuit ρ^3, the cost function can be written as

$$C(s) = \frac{1}{N}\sum_{x=1}^{N}\langle\psi_x|\rho_x^3|\psi_x\rangle = \frac{1}{N}\mathrm{tr}(|\psi_x\rangle\langle\psi_x|\rho_x^3). \tag{30}$$

Since

$$\begin{aligned}\rho^3(s+\epsilon) = &\sum_i e^{i\epsilon K_{2,i}^l}V_i^l\cdots e^{i\epsilon K_{2,i}^1}V_i^1\mathrm{tr}_A[e^{i\epsilon K_1^k}U^k\cdots e^{i\epsilon K_1^1}\\ &U^1\rho_x^0{U^1}^\dagger e^{-i\epsilon K_1^1}\cdots {U^k}^\dagger e^{-i\epsilon K_1^k}(|i\rangle\langle i|_A\otimes I_B)]\\ &{V_i^1}^\dagger e^{-i\epsilon K_{2,i}^1}\cdots {V_i^l}^\dagger e^{-i\epsilon K_{2,i}^l},\end{aligned} \tag{31}$$

we can evaluate the derivative of the cost function as

$$\frac{dC}{ds} = \lim_{\epsilon>0}\frac{C(s+\epsilon)-C(s)}{\epsilon} = \frac{1}{N}\mathrm{tr}(|\psi_x\rangle\langle\psi_x|X), \tag{32}$$

where

$$\begin{aligned}X = &\sum_{i=1}^{4}\{[iK_{2.i}^l, V_i^l\cdots V_i^1\rho_i^2{V_i^1}^\dagger\cdots{V_i^l}^\dagger]+\cdots+\\ &V_i^l\cdots V_i^2[iK_{2.i}^1, V_i^1\rho_i^2{V_i^1}^\dagger]{V_i^2}^\dagger\cdots{V_i^l}^\dagger+\\ &+\mathcal{E}^i([iK_1^k(s), U^k\cdots U^1\rho^0{U^1}^\dagger\cdots{U^k}^\dagger])+\cdots+\\ &+\mathcal{E}^i(U^k\cdots U^2[iK_1^1(s), U^1\rho^0{U^1}^\dagger]{U^2}^\dagger\cdots{U^k}^\dagger)\}.\end{aligned} \tag{33}$$

To update the parameter in the network, we minimize the function

$$\frac{dC}{ds} - \lambda\sum_{\alpha_1,\cdots,\alpha_n}K^2_{\alpha_1,\cdots,\alpha_n}(s)^2, \tag{34}$$

where $K_{\alpha_1,\cdots,\alpha_n}(s)$ is related to $K_1(s)$ by

$$K_1(s) = \sum K_{\alpha_1,\cdots,\alpha_n}(s)\otimes_{k=1}^n\sigma^{\alpha_k}. \tag{35}$$

We will focus on two specific cases of $K_1(s)$: $K_1(s)\to\sigma_j^\alpha$ and $K_1(s)\to|11\rangle_{jk}$. The former case involves one qubit, while the latter case only involves two qubits. If $K_1^j(s)$ only acts on qubit j, we let

$$K_1(s) = K_1^j(s)\otimes I_{\bar{j}}. \tag{36}$$

If $K_1^{j,k}(s)$ acts on qubits j and k, we let

$$K_1(s) = K_1^{j,k}(s)\otimes I_{\bar{j,k}}. \tag{37}$$

This ends the classical simulation of the controlled unitaries.

D Simulation details

This section presents additional simulation setups for the figures in Sec. 6.

To know beforehand that the global optimal loss of the DIBoM can reach 0 with a suitable tuning of its parameters, we make the following restrictions on the intrinsic unitary V. For Figs. 10, 11, and 12, the intrinsic unitary V is restricted to a single-qubit unitary multiplied by a generalized CZ gate, with the single-qubit unitary acting on the second qubit. For Fig. 17, the intrinsic unitary V is restricted to a single-qubit unitary acting on the second qubit. For Fig. 19, the intrinsic unitary V is restricted to a layer of generalized CZ gates. For Fig. 20, the intrinsic unitary V is chosen in such a way that it is obtained by a single-qubit gate multiplied by a generalized CZ gate, where the generalized CZ gate of V is set to be identical to the fixed second layer of the DIBoM. For Fig. 21, the intrinsic unitary V is also obtained by a single-qubit gate multiplied by a generalized CZ gate, with the single-qubit gate matching the first layer of the DIBoM. For Figs. 22, 28, 34, and 35, the intrinsic unitary V is restricted to a product gate layer followed by a generalized CZ gate layer, with the product gate acting on all qubits of the quantum input. For Fig. 23, the intrinsic unitary V is restricted to have the same circuit structure as the DIBoM but with different parameters. For Fig. 27, the intrinsic unitary V is restricted to an alternating product of a product gate layer and a generalized CZ gate layer with a total of L layers, where L is the given layer number.

Some auxiliary setups are as follows. For Fig. 11, both the training and test samples are drawn from the same distribution, meaning that the intrinsic unitary V that transforms the input into the output is identical for both sets. The number of samples in the test set is fixed at 10. For Fig. 31, the intrinsic unitary V is randomly selected from all 2-qubit unitaries, with no additional constraints.

E Simulation of local cost function

When simulating the local cost function, there are two changes compared to simulating the global cost function. Firstly, the input and output are reversed. Secondly, the input $|\phi\rangle\langle\phi|$ is substituted by $|\phi\rangle_i\langle\phi|\otimes I_{\bar{i}}$.

Let us start with the first change. The ground truth unitary is V, hence the ground truth output is $|\phi_x^L\rangle = V|\phi_x^0\rangle$, where $|\phi_x^0\rangle$ is the quantum input. In the reverse setup, we will compare the "model" input $\rho_x^0 = U^\dagger|\phi_x^L\rangle\langle\phi_x^L|U$ with the actual input $|\phi_x^0\rangle\langle\phi_x^0|$ in the cost function:

$$C_{reverse}(s) = \frac{1}{N}\sum_{x=1}^{N}\langle\phi_x^0|\rho_x^0(s)|\phi_x^0\rangle. \tag{38}$$

A crucial observation is

$$\frac{dC_{reverse}}{ds} = \frac{i}{N}\sum_x\mathrm{tr}(\sum_{l=L}^{1}M^l(s)K^l(s)), \tag{39}$$

where $M^l(s)$ and $K^l(s)$ are exactly the same as the ones in Appendix **??**. The optimization of the reversed cost function $C_{reverse}$ is hence equivalent to that of the original cost function C, thus completing the first change.

Moving on to the second change, we note that the input $|\phi_x^0\rangle$ is a product state, which allows us to express it as $|\phi_x^0\rangle = |\phi_x^0\rangle_1\otimes\cdots\otimes|\phi_x^0\rangle_n$. As a result, the local cost function takes the form

$$C_{local}(s) = \frac{1}{nN}\sum_{x=1}^{N}\sum_{y=1}^{n}\mathrm{tr}((|\phi_x^0\rangle_y\langle\phi_x^0|_y\otimes I_{\bar{y}})\rho_x^0(s)), \tag{40}$$

where $I_{\bar{y}}$ is the identity operator acting on all subsystems except the y-th one. Accordingly, the derivative of C_{local} with respect to the parameter s can be expressed as

$$\frac{dC_{local}}{ds} = \frac{i}{Nn}\sum_{x} \mathrm{tr}(\sum_{l=1}^{L} M^{l}_{local}(s)K^{l}(s)), \qquad (41)$$

where

$$\begin{aligned} M^{l}_{local}(s) \;=\; & [\prod_{j=l}^{1} U^{j}(s)(\sum_{y} |\phi^{0}_{x}\rangle_{y}\langle\phi^{0}_{x}|_{y} \otimes I_{\bar{y}}) \prod_{j=1}^{l} {U^{j}}^{\dagger}(s), \\ & \prod_{j=l+1}^{L} {U^{j}}^{\dagger}(s)\rho^{0}_{x} \prod_{j=L}^{l+1} U^{j}(s)]. \end{aligned}$$

The remaining procedure is identical to that of the global cost function.

References

[Goodfellow *et al.*(2016)Goodfellow, Bengio, and Courville] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016) http://www.deeplearningbook.org.

[Waldrop(2016)] M. M. Waldrop, Nature News **530**, 144 (2016).

[Nielsen and Chuang(2010)] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2010).

[Biamonte *et al.*(2017)Biamonte, Wittek, Pancotti, Rebentrost, Wiebe, and Lloyd] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, Nature **549**, 195 (2017).

[Arute *et al.*(2020)Arute, Arya, Babbush, Bacon, Bardin, Barends, Boixo, Broughton, Buckley, Buell *et al.*] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, *et al.*, Science **369**, 1084 (2020).

[Romero *et al.*(2017)Romero, Olson, and Aspuru-Guzik] J. Romero, J. P. Olson, and A. Aspuru-Guzik, Quantum Science and Technology **2**, 045001 (2017).

[Johnson *et al.*(2017)Johnson, Romero, Olson, Cao, and Aspuru-Guzik] P. D. Johnson, J. Romero, J. Olson, Y. Cao, and A. Aspuru-Guzik, arXiv preprint arXiv:1711.02249 (2017).

[Koczor *et al.*(2020)Koczor, Endo, Jones, Matsuzaki, and Benjamin] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, New Journal of Physics **22**, 083038 (2020).

[Sharma *et al.*(2020a)Sharma, Khatri, Cerezo, and Coles] K. Sharma, S. Khatri, M. Cerezo, and P. J. Coles, New Journal of Physics **22**, 043006 (2020a).

[LaRose *et al.*(2019)LaRose, Tikku, O'Neel-Judy, Cincio, and Coles] R. LaRose, A. Tikku, É. O'Neel-Judy, L. Cincio, and P. J. Coles, npj Quantum Information **5**, 1 (2019).

[Li and Benjamin(2017)] Y. Li and S. C. Benjamin, Physical Review X **7**, 021050 (2017).

[Cerezo *et al.*(2020)Cerezo, Poremba, Cincio, and Coles] M. Cerezo, A. Poremba, L. Cincio, and P. J. Coles, Quantum **4**, 248 (2020).

[Arrasmith *et al.*(2019)Arrasmith, Cincio, Sornborger, Zurek, and A. Arrasmith, L. Cincio, A. T. Sornborger, W. H. Zurek, and P. J. Coles, Nature communications **10**, 1 (2019).

[Harrow *et al.*(2009)Harrow, Hassidim, and Lloyd] A. W. Harrow, A. Hassidim, and S. Lloyd, Physical Review Letters **103**, 150502 (2009).

[Lloyd *et al.*(2014)Lloyd, Mohseni, and Rebentrost] S. Lloyd, M. Mohseni, and P. Rebentrost, Nature Physics **10**, 631 (2014).

[Rebentrost *et al.*(2014)Rebentrost, Mohseni, and Lloyd] P. Rebentrost, M. Mohseni, and S. Lloyd, Physical Review Letters **113**, 130503 (2014).

[Schuld *et al.*(2015)Schuld, Sinayskiy, and Petruccione] M. Schuld, I. Sinayskiy, and F. Petruccione, Physics Letters A **379**, 660 (2015).

[Lewenstein(1994)] M. Lewenstein, Journal of Modern Optics **41**, 2491 (1994).

[Wan *et al.*(2017)Wan, Dahlsten, Kristjánsson, Gardner, and Kim K. H. Wan, O. Dahlsten, H. Kristjánsson, R. Gardner, and M. Kim, npj Quantum information **3**, 1 (2017).

[da Silva *et al.*(2016)da Silva, Ludermir, and de Oliveira] A. J. da Silva, T. B. Ludermir, and W. R. de Oliveira, Neural Networks **76**, 55 (2016).

[Gonsalves(2017)] C. P. Gonsalves, NeuroQuantology **15**, 22 (2017).

[Kouda *et al.*(2005)Kouda, Matsui, Nishimura, and Peper] N. Kouda, N. Matsui, H. Nishimura, and F. Peper, Neural Computing & Applications **14**, 114 (2005).

[Beer *et al.*(2020)Beer, Bondarenko, Farrelly, Osborne, Salzmann, K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann, D. Scheiermann, and R. Wolf, Nature Communications **11**, 1 (2020).

[Cong *et al.*(2019)Cong, Choi, and Lukin] I. Cong, S. Choi, and M. D. Lukin, Nature Physics **15**, 1273 (2019).

[Bausch(2020)] J. Bausch, Advances in Neural Information Processing Systems **33** (2020).

[Lloyd and Weedbrook(2018)] S. Lloyd and C. Weedbrook, Physical Review Letters **121**, 040502 (2018).

[Bondarenko and Feldmann(2020)] D. Bondarenko and P. Feldmann, Physical Review Letters **124**, 130502 (2020).

[Ghosh *et al.*(2021)Ghosh, Nakajima, Krisnanda, Fujii, and Liew] S. Ghosh, K. Nakajima, T. Krisnanda, K. Fujii, and T. C. Liew, Advanced Quantum Technologies **4**, 2100053 (2021).

[Killoran *et al.*(2019)Killoran, Bromley, Arrazola, Schuld, Quesada, and Lloyd] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, Physical Review Research **1**, 033063 (2019).

[McClean *et al.*(2018)McClean, Boixo, Smelyanskiy, Babbush, and Neven] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, Nature communications **9**, 1 (2018).

[Benedetti *et al.*(2019)Benedetti, Garcia-Pintos, Perdomo, Leyton-Ortega, Nam, and Perdomo-Ortiz] M. Benedetti, D. Garcia-Pintos, O. Perdomo, V. Leyton-Ortega, Y. Nam, and A. Perdomo-Ortiz, npj Quantum Information **5**, 1 (2019).

[Sharma *et al.*(2020b)Sharma, Cerezo, Cincio, and Coles] K. Sharma, M. Cerezo, L. Cincio, and P. J. Coles, arXiv preprint arXiv:2005.12458 (2020b).

[Coyle *et al.*(2020)Coyle, Mills, Danos, and Kashefi] B. Coyle, D. Mills, V. Danos, and E. Kashefi, npj Quantum Information **6**, 1 (2020).

[Cheng *et al.*(2018)Cheng, Chen, and Wang] S. Cheng, J. Chen, and L. Wang, Entropy **20**, 583 (2018).

[Liu and Wang(2018)] J.-G. Liu and L. Wang, Physical Review A **98**, 062324 (2018).

[Cirstoiu *et al.*(2020)Cirstoiu, Holmes, Iosue, Cincio, Coles, and Sornborger] C. Cirstoiu, Z. Holmes, J. Iosue, L. Cincio, P. J. Coles, and A. Sornborger, npj Quantum Information **6**, 1 (2020).

[Barison *et al.*(2021)Barison, Vicentini, and Carleo] S. Barison, F. Vicentini, and G. Carleo, Quantum **5**, 512 (2021).

[Wierichs *et al.*(2022)Wierichs, Izaac, Wang, and Lin] D. Wierichs, J. Izaac, C. Wang, and C. Y.-Y. Lin, Quantum **6**, 677 (2022).

[Nesterov()] Y. Nesterov, in *Sov. Math. Dokl*, Vol. 27.

[Gacon *et al.*(2021)Gacon, Zoufal, Carleo, and Woerner] J. Gacon, C. Zoufal, G. Carleo, and S. Woerner, Quantum **5**, 567 (2021).

[Skolik *et al.*(2021)Skolik, McClean, Mohseni, van der Smagt, and Leib] A. Skolik, J. R. McClean, M. Mohseni, P. van der Smagt, and M. Leib, Quantum Machine Intelligence **3**, 1 (2021).

[Du *et al.*(2022)Du, Tu, Yuan, and Tao] Y. Du, Z. Tu, X. Yuan, and D. Tao, Physical Review Letters **128**, 080506 (2022).

[Cerezo *et al.*(2021)Cerezo, Sone, Volkoff, Cincio, and Coles] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Nature communications **12**, 1 (2021).

[Anschuetz and Kiani(2022)] E. R. Anschuetz and B. T. Kiani, Nature Communications **13**, 7760 (2022).

[Biamonte(2021)] J. Biamonte, Physical Review A **103**, L030401 (2021).

[Goto *et al.*(2021)Goto, Tran, and Nakajima] T. Goto, Q. H. Tran, and K. Nakajima, Physical Review Letters **127**, 090506 (2021).

[Cai *et al.*(2022)Cai, Ye, and Deng] H. Cai, Q. Ye, and D.-L. Deng, Quantum Science and Technology **7**, 025014 (2022).

[Herman *et al.*(2022)Herman, Raymond, Li, Robles, Mezzacapo, a D. Herman, R. Raymond, M. Li, N. Robles, A. Mezzacapo, and M. Pistoia, arXiv preprint arXiv:2204.05286 (2022).

[Zhou *et al.*(2020)Zhou, Wang, Choi, Pichler, and Lukin] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, Physical Review X **10**, 021067 (2020).

[Motta *et al.*(2020)Motta, Sun, Tan, O'Rourke, Ye, Minnich, Bran M. Motta, C. Sun, A. T. Tan, M. J. O'Rourke, E. Ye, A. J. Minnich, F. G. Brandão, and G. K.-L. Chan, Nature Physics **16**, 205 (2020).

[Sparrow *et al.*(2018)Sparrow, Martín-López, Maraviglia, Neville, C. Sparrow, E. Martín-López, N. Maraviglia, A. Neville, C. Harrold, J. Carolan, Y. N. Joglekar, T. Hashimoto, N. Matsuda, J. L. O'Brien, *et al.*, Nature **557**, 660 (2018).

[Cao *et al.*(2018)Cao, Romero, and Aspuru-Guzik] Y. Cao, J. Romero, and A. Aspuru-Guzik, IBM Journal of Research and Development **62**, 6 (2018).

1

1

1, 4.2, B

Device-independent self-testing of unsharp measurements

Prabuddha Roy[1] * A.K.Pan[1 2] †

[1] *National Institute of Technology Patna, Ashok Rajpath, Patna, Bihar 800005, India*
[2] *Department of Physics, Indian Institute of Technology Hyderabad, Telengana-502284, India*

Abstract. In this work [1], we provide device-independent (DI) self-testing of the unsharp instrument through the quantum violation of two Bell inequalities where the devices are uncharacterized and the dimension of the system remains unspecified. We introduce an elegant sum-of-squares approach to derive the dimension-independent optimal quantum violation of Bell inequalities which plays a crucial role. Note that the standard Bell test cannot self-test the post-measurement states and consequently cannot self-test unsharp instrument. The sequential Bell test possess the potential to self-test an unsharp instrument. We demonstrate that there exists a trade-off between the maximum sequential quantum violations of the Clauser-Horne-Shimony-Holt inequality, and they form an optimal pair that enables the DI self-testing of the entangled state, the observables, and the unsharpness parameter. Further, we extend our study to the case of elegant Bell inequality. Since an actual experimental scenario involves losses and imperfection, we demonstrate robustness of our certification to noise.

Keywords: sharing of nonlocality, device-independent self-testing, unsharp measurement

Device-independent (DI) self-testing [2, 3, 4, 5, 6] is the strongest certification protocol based on the Bell test. In DI scenario, the devices remain uncharacterized (black-boxes) and the dimension of the quantum system is considered to be finite but unbounded. Essentially, the optimal quantum violation of a Bell's inequality self-tests the entangled state and the observables in any arbitrary dimension. In this work, we provide the DI self-testing of unsharp measurements based on the two Bell's inequalities, viz., the CHSH inequality [7], and the elegant Bell inequality [8]. By unsharp measurements, here we mean the noisy variants of the sharp projective measurements so that the number of measurement operators equals the number of projectors.

Note that, both the above mentioned inequalities are optimized in quantum theory for sharp measurement and any unsharp measurement naturally gives rise to sub-optimal quantum values. Thus, to DI self-test unsharp instrument one has to introduce a protocol based on a sub-optimal quantum violation of Bell's inequality. However, in an ideal scenario, the sub-optimal quantum value in a Bell test may come from different sources, viz., inappropriate choices of entangled state or observables than that are required for optimal quantum violation, or from unsharp measurements. If one can ensure that the entangled state and observables remains same that are required for optimal quantum violation then it is indeed the unsharp instrument that is responsible for providing the sub-optimal quantum value. This, in turn, self-tests the unsharpness parameter of the quantum instrument.

We demonstrated that the sequential Bell test has the potential to provide DI self-testing. Specifically, we first consider the CHSH inequality, and by introducing an elegant sum-of-squares (SOS) approach we derive the optimal quantum violation without assuming the dimension of the system. Further, we consider the sequential sharing of nonlocality [9, 10, 11, 12] when multiple observers on one side perform unsharp measurements. Note that, the sharing of non-locality without assuming the dimension has not hitherto been demonstrated. Using our SOS approach, we optimize the quantum violations for two sequential observers and show that there is a trade-off between those two quantum violations. Moreover, we show that the trade-off in turn enables us to certify the state, observables and unsharpness parameter of the instrument in a DI way. Note however that, in the practical implementation of our protocol, there will be inevitable noise and imperfection which forbids obtaining the maximum quantum values. For that case, we provide a range within which the unsharpness parameter should belong. The more perfect the experimental realization more accurate one can self-test the unsharpness parameter. We extend our protocol for the case of elegant Bell inequality, where we demonstrate that at most three observers can sequentially share the preparation contextuality and demonstrate the self-testing of two unsharpness parameters.

1 Results

In this paper [1], we have demonstrated the following results:

i) First, we derive the optimal quantum value of CHSH inequality without assuming the dimension of the system by introducing sum-of-squares (SOS) approach. This in turn uniquely certifies the state and observables.

ii) Next, we consider the sequential sharing scenario where on one side of the shared entangled state, Alice performs sharp measurement and on the other side multiple independent Bobs performs unsharp measurement. By optimizing the quantum violations for sequential Bobs in the Bell-CHSH scenario and from the trade-off between the quantum violations, we first showcase that at most two independent Bobs can showcase quantum advantage over preparation non-contextual (local) model in the DI way.

iii)We then prove that the sub-optimal sequential

*prabuddharoy.94@gmail.com
†akp@phy.iith.ac.in

quantum values are maximized for the same state and observables as required for the optimal violation. Hence, two maximized sub-optimal quantum violations form an optimal pair enabling the DI self-testing of the unsharp instrument.

iv) According to Naimark's theorem, any non-projective measurement can be modelled as sharp measurements in a higher dimension. Since in DI scenario there is no bound on the dimension one may argue that sub-optimal quantum violation arises due to the inappropriate choices of observables in higher dimension and not from the unsharp measurement. Our dimension-independent optimization of the sequential quantum violations of Bell's inequality bypasses the constraints that would arise from Naimark's theorem and enables the DI self-testing of unsharp instruments.

v) Moreover, We provide an argument to demonstrate how our certification protocol is robust to the noise by providing a range within which the unsharpness parameter can belong.

vi) Following the similar approach, we have additionally demonstrated the sequential scenario of nonlocality for elegant Bell inequality in the DI manner and certify the states, local observables and unsharpness parameters which provide such advantage.

2 Figures and Tables

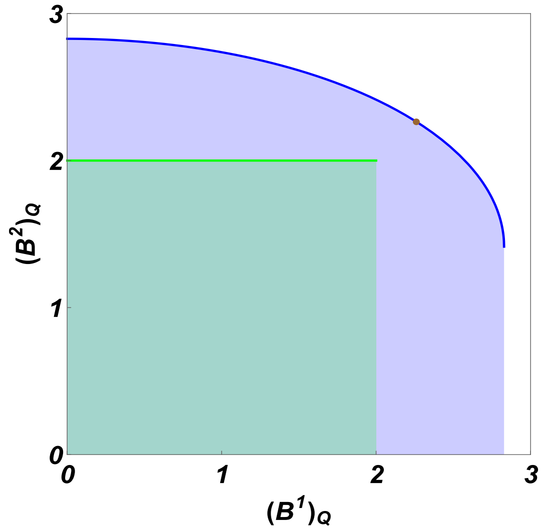


Figure 1: Optimal trade-off between quantum bound of CHSH inequality of Bob$_1$ and Bob$_2$ is shown by the solid blue curve while the shaded portion gives the suboptimal range. The solid green line is for classical bound of CHSH inequality for the same two Bobs.

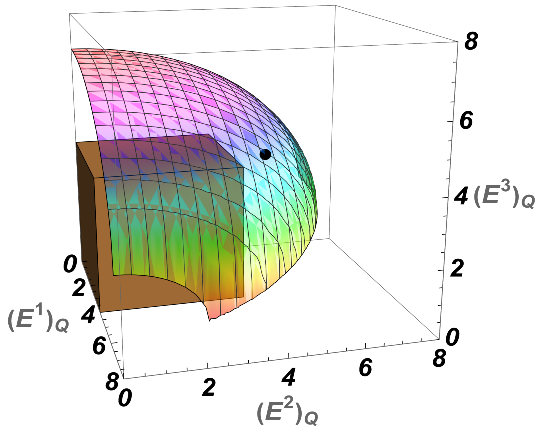


Figure 2: Optimal trade-off between quantum bound of elegant Bell inequality of Bob$_1$, Bob$_2$ and Bob$_3$. The black point on the three-dimensional graph indicates the point which certifies the unsharpness parameters λ_1 and λ_2 when quantum values of all three sequential Bobs are considered to be equal.

References

[1] P. Roy and A K Pan, Device-independent self-testing of unsharp measurements, New J. Phys. 25 013040 (2023).

[2] D. Mayers and A. Yao, Self testing quantum apparatus, 10.48550/ARXIV.QUANT-PH/0307205 (2003).

[3] M. McKague, Self-testing in parallel, New Journal of Physics 18, 045013 (2016).

[4] M. McKague, Self-testing in parallel with CHSH, Quantum 1, 1 (2017).

[5] I. Supíc and J. Bowles, Self-testing of quantum systems: a review , Quantum 4, 337 (2020).

[6] I. Supíc, D. Cavalcanti, and J. Bowles, Device-independent certification of tensor products of quantum states using single-copy self-testing protocols,Quantum 5, 418 (2021).

[7] J.S. Bell, On the Einstein Podolsky Rosen paradox, Physics, **1**, 195 (1964).

[8] N. Gisin, Bell inequalities: many questions, a few answers, arXiv:quant-ph/0702021.

[9] R. Silva, N. Gisin, Y. Guryanova, and S. Popescu, Multiple Observers Can Share the Nonlocality of Half of an Entangled Pair by Using Optimal Weak Measurements, Phys. Rev. Lett. 114, 250401 (2015)

[10] A. Kumari and A.K. Pan, Sharing nonlocality and nontrivial preparation contextuality using the same family of Bell expressions, Phys. Rev. A 100, 062130 (2019).

[11] P. J. Brown and R. Colbeck, Arbitrarily Many Independent Observers Can Share the Nonlocality of a Single Maximally Entangled Qubit Pair, Phys. Rev. Lett. 125, 090401 (2020)

[12] T. Zhang and S-M. Fei, Sharing quantum nonlocality and genuine nonlocality with independent observables, Phys. Rev. A 103, 032216 (2021).

Generation of Braiding Operators for Topological Quantum Computing

Abdellah Tounsi[1] Nacer Eddine Belaloui[1] Mohamed Messaoud Louamri[1] Amani Mimoun[1] Achour Benslama[1] Mohamed Taha Rouabah[1] *

[1] *Constantine Quantum Technologies,*
Laboratoire de Physique Mathématique et Subatomique,
Frères Mentouri University Constantine 1, Ain El Bey Road, Constantine, 25017, Algeria.

Abstract. Topological quantum computing is a theoretical framework that enables unitary operations to be performed without introducing errors due to the system's dynamics. This framework relies on the statistical properties of anyon particles, which are governed by the braid group. However, computing the anyonic braid matrices for topological quantum computing is considered challenging. In this study, we propose a systematic method for computing such braid matrices for quantum circuit-based anyonic states. This method can serve as the foundation for a general topological quantum computing simulator, facilitating the study of complex topological quantum circuits in the context of any anyon model. In this presentation, we aim to provide a comprehensive review of the foundational principles of anyon model theory. Specifically, we describe the methodology for constructing qubits/qudits based on anyonic states. Furthermore, we present a generalized formula for systematically computing the matrix components of the braid operations, which we validate through algebraic techniques. Moreover, we reproduce well-known quantum gates from previous studies to demonstrate the validity of our method.

Keywords: Fault-tolerant quantum computing, topological quantum computing, anyons, quantum circuit compilation

Experimental implementations of quantum computers are susceptible to errors that arise from the interaction of qubits with their environment. To tackle this challenge, multiple approaches are being investigated, including cooling and fabrication technologies, as well as quantum error mitigation techniques. However, for achieving fault-tolerant quantum computing, quantum error correction is necessary. Nonetheless, the practical application of quantum error correction is limited by the threshold theorem, which stipulates that a sufficiently low error rate is required for reliable quantum computation.

In recent decades, topological quantum systems have emerged as a promising approach for storing and processing quantum information in a robust manner, as they are immune to local noise [1]. Topological quantum computation (TQC) focuses on two-dimensional quantum systems that support excitations with fractional statistics, known as anyons. These anyons exhibit statistics that differ from those of fermions and bosons, making them unique. A system of anyons possesses a topologically protected Hilbert space that grows exponentially with the number of anyons, and quantum information can be processed through the braiding of anyons. Furthermore, topological quantum computers, based on anyon models, have been shown to be Turing-complete i.e universal. Experimental evidence has confirmed that quasi-particles produced by the Fractional Quantum Hall Effect (FQHE) exhibit anyonic statistics [2]. Other potential systems for constructing a topological quantum computer include lattice models such as the Kitaev toric code, Kitaev honeycomb model, and Levin-Wen model.

Anyon models serve as the frameworks for topological quantum computation (TQC), providing a description of the fusion rules among different types of anyons with distinct topological charges. In the context of anyon models, quantum information is encoded in the permissible fusion states of the anyons and manipulated through the exchange of anyons, which is known as braiding. The fusion state of a pair of anyons is determined by a non-local collective property of that pair, and it is not accessible to either individual anyon. Therefore, quantum information encoded in anyons must be resilient to local perturbations, as it is protected by the non-local properties of the anyons in the fusion states.

To account for the various ways of fusing more than two anyons, fusion matrices F are introduced, which allow for basis transformations. On the other hand, the exchange of two isolated anyons is described by rotation matrices R. By solving the hexagon and pentagon identities, one can determine the explicit values of the components of F and R. These identities ensure the algebraic consistency of any anyon model. Methods for solving such equations are already established in the literature on quantum groups, providing a thorough explanation of the general analytical solution of the hexagon and pentagon equations for any $SU(2)_k$ anyon model.

The process of exchanging two adjacent anyons is commonly referred to as a braid operation. The braid group is generated by the set of all possible braid operators for a given number of anyons. These operators must satisfy the Yang-Baxter equation, which is also a consistency relation that gives the algebraic structure of anyon models. In addition to the hexagon and

*m.taha.rouabah@umc.edu.dz

pentagon equations, the Yang-Baxter equation also has practical applications in the implementation of unit-testing for numerical packages used to compute the braid generators. It serves as a crucial tool for validating the accuracy and reliability of numerical computations in the context of topological quantum computing.

The explicit matrix representation of a braid generator can be obtained by selecting the basis of the fusion space and applying the relevant F and R moves. In this work, we investigated the action of braid generators on a group of identical anyons that are grouped in sets of a specific number. Each set is dedicated to representing a qubit/qudit. Representing qubits/qudits with a definite number of anyons preserves the circuit model of quantum computing. The formula introduced in this work ensures a systematic method for computing braid generators with the assistance of computational units. However, this method does not guarantee the reduction of computational complexity since the size of the braid generator grows exponentially with the number of anyons.

To demonstrate the reliability of our approach, we utilized the algebraic relations of the braid group. Additionally, we have shown how to reproduce well-known topological quantum gates from previous works.

This approach is highly valuable for studying the compilation of quantum circuits on topological systems that utilize anyonic states [3]. The method has already been implemented in the `TQSim` library [4] for the Fibonacci model and can be extended to all $SU(2)_k$ anyon models, as general solutions for F and R matrices are available [5]. Simulating topological quantum circuits is extremely useful for testing the compilation of quantum gates and developing valuable quantum circuits.

References

[1] A. Kitaev. *Fault-tolerant quantum computation by anyons.* Annals of Physics, 303(1):2 – 30, January 2003. doi:https://doi.org/10.1016/S0003-4916(02)00018-0.

[2] H. Bartolomei, et al. *Fractional statistics in anyon collisions.* Science, 368(6487):173–177, 2020. doi: 10.1126/science.aaz5601.

[3] M. T. Rouabah, et al. *Compiling single-qubit braiding gate for fibonacci anyons topological quantum computation.* Journal of Physics: Conference Series, 1766(1):012029, January 2021. doi:10.1088/1742-6596/1766/1/012029.

[4] A. Tounsi, et al. *TQSim, a topological quantum computing simulator based on anyon models.* https://github.com/Constantine-Quantum-Tech/tqsim, 2022.

[5] E. Génetay Johansen et al. *Fibonacci anyons versus majorana fermions: A monte carlo approach to the compilation of braid circuits in* $\mathrm{SU}(2)_k$ *anyon models.* PRX Quantum, 2:010334, Mar 2021. doi: 10.1103/PRXQuantum.2.010334.

Extended Abstract: Thermodynamic Criteria of Entanglement for Multi-Qubit Systems and their Experimental Verification

Jitendra Joshi* and T S Mahesh

Department of Physics and NMR Research Center,
Indian Institute of Science Education and Research, Pune 411008, India

Mir Alimuddin* and Manik Banik

Department of Physics of Complex Systems, S.N. Bose National Center for Basic Sciences,
Block JD, Sector III, Salt Lake, Kolkata 700106, India.

The phenomenon of quantum entanglement underlies several important protocols that enable emerging quantum technologies. Being an extremely delicate resource entangled states easily get perturbed by their external environment, and thus makes the question of entanglement certification immensely crucial for successful implementation of the protocols involving entanglement. In this work, we propose a set of entanglement criteria for multi-qubit systems that can be easily verified by measuring certain thermodynamic quantities. In particular, the criteria depend on the difference in optimal works extractable from an isolated quantum system under global and local interactions respectively. As a proof of principle, we demonstrate the proposed thermodynamic criteria on nuclear spin registers of up to 10 qubits using Nuclear Magnetic Resonance architecture. We prepare noisy Greenberger–Horne–Zeilinger class of states in star-topology systems and certify their entanglement through our proposed criteria. We also provide elegant means of entanglement certification in multi-qubit systems when only partial or even no knowledge about the state is available.

Introduction.– Quantum entanglement, a fascinating characteristic of multipartite quantum systems, is crucial in the advancement of quantum information theory, communication protocols, quantum computation, and quantum sensing. While entanglement between two subsystems can be separable or entangled, multipartite systems exhibit various forms of entanglement. This type of entanglement has proven useful in distributed protocols and the potential development of the quantum internet. Verifying whether a state is entangled or not is essential for successful implementation of these protocols. Existing methods involve complex calculations or impractical tomographic knowledge. However, recent research has introduced thermodynamic quantities that can quantify entanglement in multipartite pure quantum states. We propose utilizing ergotropic work, a thermodynamic quantity, as an entanglement certifier for N-qubit systems. By considering different information about the state's spectral properties, we present several entanglement certifiers. We demonstrate the effectiveness of our method by implementing it on nuclear spin registers of up to 10 qubits using Nuclear Magnetic Resonance architecture. Our approach not only verifies entanglement experimentally but also allows certification of multipartite states when complete state knowledge is unavailable.

Theory.– A generic state of an N-qubit system is described by a density operator $\rho_{A_1\cdots A_N} \in \mathcal{D}\left((\mathbb{C}^2)^{\otimes N}\right)$. A state is fully separable if it is a probabilistic mixture of fully product states. States lying outside the set of fully separable states are entangled. Different kinds of entanglement are possible for multi-qubit systems, characterized by the set of states separable across a bipartite cut denoted as $\mathcal{S}[X|X^c]$.

Quantum entanglement has implications for work extraction from a quantum system. The ergotropy measures the optimally extractable work from a state, which is the difference between the energy of the state and the minimum energy achievable under unitary transformations [1]. During optimal work extraction, the system evolves to the passive state, which is the lowest energy state with the same spectral properties as the initial state.

In the multipartite scenario, different parts of the system can be probed separately for work extraction. The thermodynamic quantity $\Delta_{X|X^c}$ captures the difference in work extraction between the global system and the subsystem X in the X vs X^c configuration. A thermodynamic entanglement criterion is derived based on $\Delta_{X|X^c}$, involving the spectral properties of the states and Hamiltonians (see Theorem 1 in the main manuscript attached below).

The derived entanglement criterion states that for an N-qubit state separable across X vs X^c cut, $\Delta_{X|X^c}$ is upper-bounded by a specific function involving the spectral properties of the global state and the subsystem. Three versions of the criterion are presented in this work: one requiring knowledge of the state's spectral properties ($\delta^{GL}_{X|X^c}$) (resembling Nielsen-Kempe criterion [2]) and one based solely on the global state's spectral properties ($\delta^{G}_{X|X^c}$), and the third one does not require any information about the state ($\delta^{I}_{X|X^c}$). The latter criterion is experimentally less demanding.

Experimental tests of these thermodynamic entanglement criteria are conducted using specific classes of entangled states (noisy GHZ) in an NMR setup [3–5]. The results confirm the effectiveness of the criteria in detecting entanglement in the tested states. Further details and analysis are provided in the appendix.

Experimental results.– Figure 1 (a) The ^{1}H spectrum corresponds to the singlet or Werner state of pair of hydrogen spins BRTP. (b) The NMR pulse sequence to produce Werner class states with controlled purity and then to certify the presence or absence of entanglement. Here rectangles with $\Theta_\beta = e^{-i\Theta(I_{1\beta}+I_{2\beta})}$ represent RF rotations, delays represent free-evolutions, and PFG is the Pulsed-Field Gradient [6, 7]. (c) Plot of $\Delta_{1|2}$ in units of $\hbar\omega_H$ versus purity of the Werner class. The vertical dotted lines indicate the purity threshold for entanglement: the left one marks purity $\lambda = 1/3$, above which $\Delta_{1|2}$ surpasses $\delta^G_{1|2}$ and the state becomes entangled; the right one marks $\lambda = 1/2$, corresponding to the state-independent bound $\delta^I_{1|2}$.

Figure 2, (a) The NMR pulse sequence to prepare GHZ / MSSM class states with controlled purity on

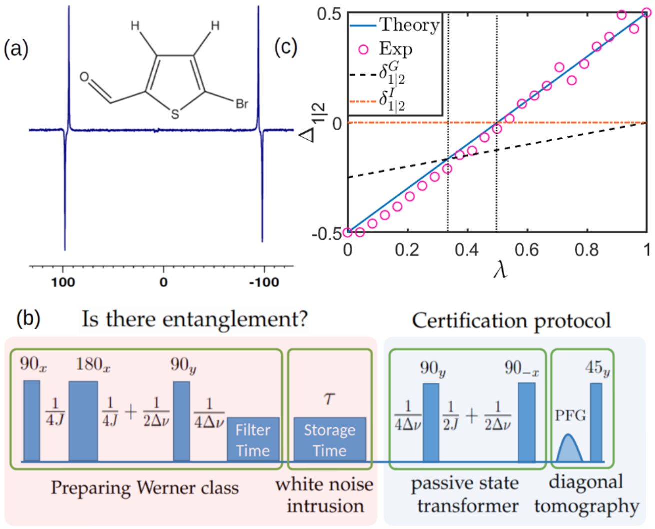


FIG. 1

an STR and then certify the presence or absence of entanglement. The vertical line shows the instant when the passive state is created. The dashed pulses cancel each other. (b) ^{19}F spectra of FAN corresponding to one-pulse experiment (front) and to the three-qubit GHZ state (back). (c) Plot $\Delta_{1|23}$ (in the unit of $\hbar\omega_H$) vs purity λ for 3-qubit noisy GHZ states. Comparing the values of $\Delta_{1|23}$ and $\delta^G_{1|23}$, $\delta^I_{1|23}$, we identify the threshold values marked by the dotted lines: $\lambda = 3/7$ and $\lambda = 0.68$, respectively. Above these thresholds, the state exhibits entanglement. (d,f) ^{31}P spectra of TMP corresponding to one pulse experiment (front), and to GHZ (d, back) and MSSM class (f, back). (e,g) $\Delta_{1|1^c}$ vs purity λ for the 10-qubit GHZ and MSSM classes. Here $\lambda = 0.499$ and $\lambda = 0.957$ for $\delta^G_{1|1^c}$ and $\delta^I_{1|1^c}$ marks the entanglement threshold boundary.

Discussions.– Efficient manipulation of entanglement in multipartite systems is crucial for various quantum technologies and protocols involving distributed quantum information. Different quantum architectures, such as linear optical devices, ion-trap systems, superconducting qubits, NMR systems, and hybrid quantum circuits, are being explored for this purpose. Certifying entanglement in the state being used is a crucial step in these efforts. While Bell tests provide device-independent certification schemes for bipartite and multipartite systems, they often require spatial separation among the entangled parts. Witness-based methods, on the other hand, require complete tomographic information about the state.

In contrast, our proposed thermodynamic criteria offer a less demanding approach to certify entanglement by measuring global and local ergotropic works. When the spectral information of the state is known, our criteria are equivalent to the Nielsen-Kempe majorization criteria. We experimentally validate these thermodynamic entanglement criteria in an NMR architecture using specific classes of noisy entangled states with 2, 3, and 10 qubits. Furthermore, our thermodynamic approach enables entanglement certification even when the knowledge about the state is limited. This opens up avenues for certifying entanglement

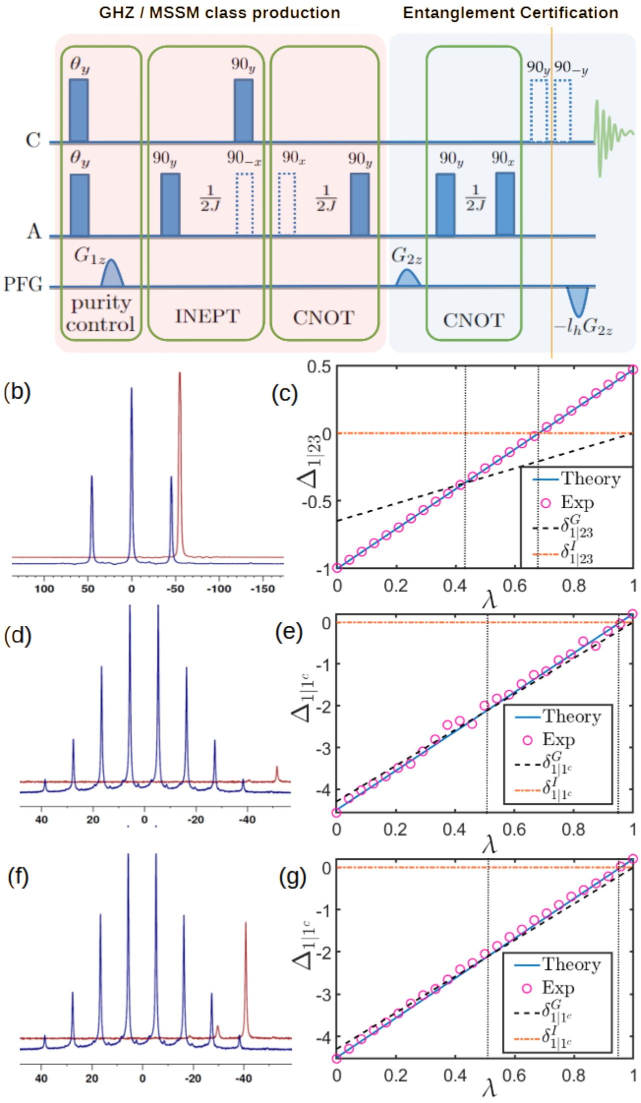


FIG. 2

in various quantum architectures with different degrees of state purity.

In terms of future research, several questions arise from our study. Generalizing the thermodynamic criteria for systems with arbitrary local dimensions and capturing more exotic forms of entanglement, such as genuine multipartite entanglement, would be important. Testing the criteria for other important classes of states, such as Bell diagonal states and X-states, would be interesting. Implementing the state-independent criteria experimentally is another intriguing direction, and recent work in this area may provide valuable insights. Finally, testing these thermodynamic criteria in other quantum architectures would be of great interest.

* These authors contribute equally in this project and share the equal credit of first authorship.

[1] A. E. Allahverdyan, R. Balian, and T. M. Nieuwenhuizen, Europhysics Letters (EPL) **67**, 565 (2004).

[2] M. A. Nielsen and J. Kempe, Phys. Rev. Lett. **86**, 5184 (2001).

[3] J. A. Jones, S. D. Karlen, J. Fitzsimons, A. Ardavan, S. C. Benjamin, G. A. D. Briggs, and J. J. L. Morton, Science **324**, 1166 (2009).

[4] T. S. Mahesh, D. Khurana, V. R. Krithika, G. J. Sreejith, and C. S. S. Kumar, J. Phys. Condens. Matter **33**, 383002 (2021).

[5] A. Shukla, M. Sharma, and T. Mahesh, Chem. Phys. Lett. **592**, 227 (2014).

[6] "$90_{x/y}$ represents a $\pi/2$ rotation with rotation axis along x/y, $(1/2j)$ and $(1/4j)$ delays represent the free evolution under the scalar coupling, $(1/2\delta\nu) \equiv 90^1_z 90^2_{-z}$ and $(1/4\delta\nu) \equiv 45^1_z 45^2_{-z}$ represent the evolution under the chemical shift and pfg is the crusher gradient used to dephasing coherences,".

[7] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, and N. J. Skelton, *Protein NMR spectroscopy: principles and practice* (Academic press, 1996).

On the optimal quantum state of BPSK-type asymmetric quantum communication in an attenuated environment

Suguru Sameshima[1] * Tiancheng Wang[2] † Souichi Takahira[1] ‡ Shogo Usami[1] §
Tsuyoshi Sasaki Usuda[3] ¶

[1] *Graduate School of Science and Technology, Meijo University, Aichi, 468-8502 Japan.*
[2] *Faculty of Informatics, Kanagawa University, Kanagawa, 221-8686, Japan.*
[3] *Graduate School of Information Science and Technology, Aichi Prefectural University, Aichi, 480-1198 Japan.*

Abstract. In this paper, we study the optimal sending quantum state for BPSK-type asymmetric quantum communication in attenuated environments. We are improving the Nair state. Nair state is the optimal sending quantum state in a noiseless environment. We study the optimal sending quantum state in an attenuated environment under constrained conditions for restricted the photon number state used. We also study the error probability and the sending quantum state, with the aim of discovering the optimal sending quantum state in an attenuated environment when an any quantum state can be used.

Keywords: asymmetric quantum communication, entanglement, phase shift keying

1 Introduction

We consider quantum communication with asymmetric settings (e.g., [1, 2]). The asymmetry here means that the sender/receiver in the communication system exhibits a difference in ability, cf. a master unit and a slave unit. The master unit side has an abundance of energy and computational power as a base station. In contrast, the slave unit side has small-capacity batteries like those in mobile communication devices, various Internet-of-Things (IoT) devices and sensor tags in IoT and sensor networks, and satellites in space communication. The assumption is that, with such devices on the slave unit side, it is more difficult to perform complicated calculation processing compared with that on the master unit side. Such forms of communication have been put into practical use in current technology, but in this research, we shall consider whether it is possible to bring in quantum technology and pursue performances uniquely available to mechanisms utilizing quantum properties. For this reason, the chief policy is to exploit actively quantum entanglement, which is one such property.

Asymmetric quantum communication involves a master unit and a slave unit, as described below. We will consider the case where the slave unit is the sender and the master unit is the receiver. The slave unit does not need to generate the sending light used for communication, but only modulates it. We assume M-ary phase-shift keying (PSK) as the modulation scheme. For scheme using amplitude-shift keying (ASK), the reliability of communications already has been studied [4]. And for scheme using PSK, the reliability and security of communications already has been studied [5, 6]. This paper considers the optimal sending quantum state in an attenuated environment, which was considered as a future work in [5].

2 Asymmetric quantum communication

We proposed the asymmetric quantum communication system by reference to quantum reading [3, 5]. As the basic form of an asymmetric communication, we consider the setup of a sender and a receiver (see Fig.1) [5]. Here, the sender corresponds to the weaker unit (hereinafter referred to as the slave unit), and the receiver corresponds to the stronger unit (hereinafter referred to as the master unit). The communication from the master unit to the slave unit is taken to be a usual quantum communication. In the case that information is sent from the slave unit to the master unit, the communication is performed in the manner illustrated in Fig.1. First, the entangled light beam $|\Psi\rangle_{\mathrm{SR}}$ of modes S and R is generated by the master unit (receiver); in mode S, the light illuminates a slave unit (Sender). If the sending information bit is $m \in \{0, 1, \ldots, M-1\}$, the mode S light undergoes a unitary transformation denoted by $\hat{U}_m^{(\mathrm{S})}$. The output $\hat{U}_m^{(\mathrm{S})} \otimes \hat{I}^{(\mathrm{R})} |\Psi\rangle_{\mathrm{SR}}$ is the input to the detector. A joint measurement is then performed on the state of the composite system SR. Here, $\hat{I}^{(\mathrm{R})}$ is the identity operator of the mode R, indicating that no action is being made on the state of mode R.

3 Sending quantum state

3.1 Nair state

Nair and his colleagues compose a MPSK-based system based on the setup of quantum reading and considered what kind of light source would be optimal, here meaning minimizing the error probability under the energy constraint. Here, the energy constraint is to constrain the average number of photons $\langle N_{\mathrm{S}}\rangle$ from the light source to be less than or equal to a given value N_{S}. And, the Nair state, a quantum state constructed with reference to the results of Nair's optimal state is follows [5]

$$|\Psi\rangle_{\mathrm{SR}}^{(\mathrm{Nair})} = \sum_{n=0}^{M-1} \sqrt{p_n}\, |n\rangle_{\mathrm{S}} |n\rangle_{\mathrm{R}}\,, \qquad (1)$$

*223426010@ccmailg.meijo-u.ac.jp
†wang@kanagawa-u.ac.jp
‡takahira@meijo-u.ac.jp
§susami@meijo-u.ac.jp
¶usuda@ist.aichi-pu.ac.jp

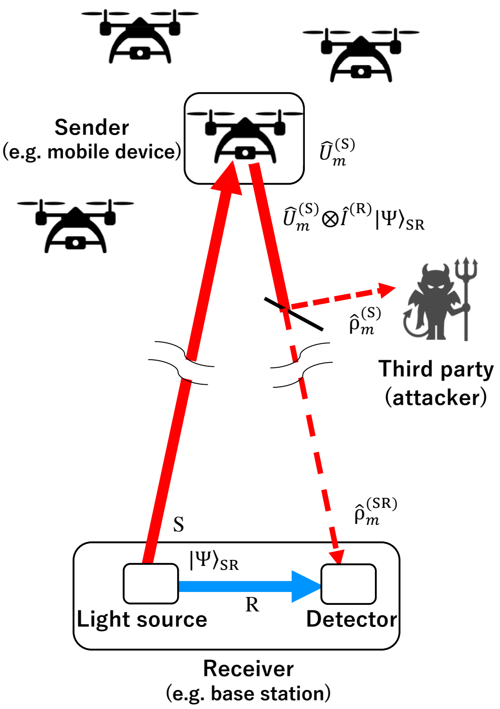

Sender
(e.g. mobile device)
$\hat{U}_m^{(\mathrm{S})}$
$\hat{U}_m^{(\mathrm{S})}\otimes\hat{I}^{(\mathrm{R})}|\Psi\rangle_{\mathrm{SR}}$
$\hat{\rho}_m^{(\mathrm{S})}$
Third party
(attacker)
$\hat{\rho}_m^{(\mathrm{SR})}$
S
$|\Psi\rangle_{\mathrm{SR}}$
R
Light source
Detector
Receiver
(e.g. base station)

Figure 1: Asymmetric quantum communication.

where, $|n\rangle$ is the photon number state and p_n is set to minimize the error probability of the legitimate receiver [7]. Now, it is also necessary to consider the error probability of the third party as well, since there may be a third party who eavesdrops on the information. The error probability of third parties should be as large as possible. In PSK-type asymmetric quantum communication in a noiseless environment, the Nair state is optimal. Optimal means that the error probability of the legitimate receiver is minimum (best for the legitimate receiver) and that of the third party is maximum (worst for the third party). However, the Nair state is not optimal in an attenuated environment. The error probability of the legitimate receiver is not minimum [5]. In fact, the Nair state showed a worse error probability than the quasi-Bell state and the TMSV state [5].

3.2 Improved Nair state

From the above, we need to search for optimal quantum states in an attenuated environment. However, the search is difficult because it requires to consider an infinite dimensional space. Therefore, we focus on the fact that the Nair state uses photon number states from $|0\rangle$ to $|M-1\rangle$. That is, under the condition that only $|0\rangle$ to $|M-1\rangle$ photon number states are used, we search for the set of coefficients p_n that shows the minimum error probability in an attenuated environment. And, we research the changes when that condition is relaxed to $|M\rangle$. From the above, we improve (1) as follows

$$|\Psi\rangle_{\mathrm{SR}} = \sum_{n=0}^{N_{max}} \sqrt{p_n}\, |n\rangle_{\mathrm{S}}\, |n\rangle_{\mathrm{R}}\,. \tag{2}$$

(2) use photon number states from $|0\rangle$ to $|N_{max}\rangle$ ($N_{max} \geq M-1$). If $N_{max} = M-1$, (2) is similar to the Nair state. Also, if $N_{max} > M-1$, the degree of freedom is higher compared to the Nair state, which can only be used up to $|M-1\rangle$.

4 Error performance

We use the Kraus representation to represent an attenuated channel.

4.1 Kraus representation

Let $\hat{\rho}^{(\mathrm{in})}$ be the input to an attenuated channel and $\hat{\rho}^{(\mathrm{out})}$ the output. Then,

$$\hat{\rho}^{(\mathrm{out})} = \sum_{k=0}^{\infty} \hat{E}_k \hat{\rho}^{(\mathrm{in})} \hat{E}_k^{\dagger}, \tag{3}$$

where

$$\hat{E}_k = \sum_{n=0}^{\infty} \sqrt{\begin{pmatrix} n \\ k \end{pmatrix}} \sqrt{\eta^{n-k}(1-\eta)^k}\, |n-k\rangle \langle n| \tag{4}$$

is the Kraus operator representing an attenuated channel.

4.2 Phase shift operator

When PSK is used as the sender's modulation scheme $\hat{U}_m^{(\mathrm{S})}$ in Fig. 1 is

$$\hat{U}_m^{(\mathrm{S})} = \hat{R}^{(\mathrm{S})}\left(\theta = \frac{2m\pi}{M}\right) = \mathrm{e}^{\mathbf{i}\frac{2m\pi}{M}\hat{N}^{(\mathrm{S})}}, \tag{5}$$

where $\hat{R}(\cdot)$ is the phase shift operator, $\hat{N}$ is the number operator. In this manner, the slave unit shifts the phase $\frac{2m\pi}{M}$ of the mode S depending on the classical information m.

4.3 Reliability

We describe the error probability of the master unit. If the received quantum state in BPSK communication is $\{\rho_0^{(\mathrm{out})}, \rho_1^{(\mathrm{out})}\}$, the minimum error probability with a quantum optimal receiver is:

$$P_{\mathrm{e}}^{min} = \frac{1}{2}\left(1 - \left|\left|\rho_0^{(\mathrm{out})} - \rho_1^{(\mathrm{out})}\right|\right|\right), \tag{6}$$

where $||\cdot||$ is the trace norm.

4.4 Security

Next, we consider the error probability by a third party. For the Nair state, an error probability of a third party is maximal, even in a noiseless environment. Since an attenuation does not improve the error probability, the error probability of the third party remained at maximum [5]. The improved Nair state has the same form as the Nair state; the only difference is expansion coefficients. As a result, the error probability of the third party is the maximum in both noiseless and attenuated environments (worst for the third party): $1-\dfrac{1}{M}$.

4.5 Error performance

We investigate the error probability of BPSK asymmetric quantum communication at the legitimate receiver

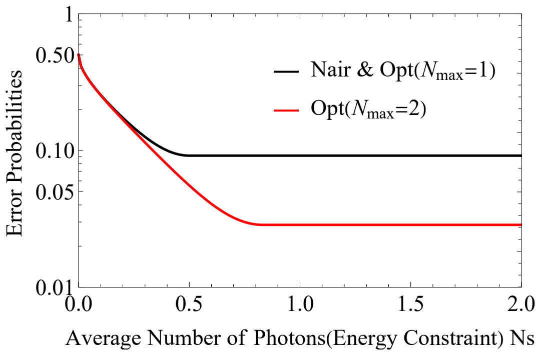


Figure 2: The error probabilities of the BPSK-type asymmetric quantum communication with Nair states and improved Nair state (Opt($N_{max}=1,2$)) in an attenuated environments with the transmissivity $\frac{2}{3}$.

side in an attenuated environment. The sending quantum states are the Nair state and the improved Nair state with $N_{max}=1,2$. The error probability of the Nair state was calculated in the previous paper [5]. The error probability of the improved Nair state is calculated using (6):

$$P_{\mathrm{e}}^{N_{max}=1} = \frac{1}{2}\left(1-2\sqrt{\eta p_0 p_1}\right), \tag{7}$$

$$P_{\mathrm{e}}^{N_{max}=2} = \frac{1}{2}\left\{1-\frac{1}{2}\left(A+B+\sqrt{A^2+B^2}\right)\right\}, \tag{8}$$

where

$$A = 2\sqrt{\eta p_1}\left(\sqrt{p_0}+\sqrt{2}(1-\eta)\sqrt{p_2}\right), B = 2\eta\sqrt{\eta p_1 p_2}.$$

If we set N_{S} a concrete value, p_0 and p_1 in (8) are uniquely determined. Therefore, it is sufficient to search for the minimum error probability within the energy constraint. On the other hand, for (8), if one of p_0, p_1, p_2 is determined, the other two are also uniquely determined. Therefore, after minimizing one variable, the minimum error probability is searched for in the energy constraint. In this paper, we used a numerical search to minimize.

The above minimization is calculated with the transmissivity $\eta = \frac{2}{3}$ and the result is shown in Fig. 2. First, we can see that the Nair state and the improved Nair state with $N_{max}=1$ coincide. The error probability decreases with N_{S} when $0 \leq N_{\mathrm{S}} \leq \frac{1}{2}$ and are constant when $N_{\mathrm{S}} > \frac{1}{2}$. Specifically, the value of p_n is $\{p_0, p_1\} = \{1-N_{\mathrm{S}}, N_{\mathrm{S}}\}$ for $0 \leq N_{\mathrm{S}} \leq \frac{1}{2}$ and $\{p_0, p_1\} = \{\frac{1}{2}, \frac{1}{2}\}$ for $N_{\mathrm{S}} > \frac{1}{2}$.

Next, we look at the improved Nair state when $N_{max} = 2$. The N_{S} is almost similar to the case $N_{max}=1$ up to about 0.1, and the difference with $N_{max}=1$ starts to widen from about 0.2. And the minimum error probability of 0.0286... is then achieved with the $N_{\mathrm{S}} \fallingdotseq 0.83$, $\{p_0, p_1 p_2\} = \{0.335, 0.500, 0.165\}$. The distribution of the set of p_n is Fig. 3. For $N_{\mathrm{S}}=0$, p_0 starts at 1 and p_1, p_2 at 0. As the N_{S} increases, p_0 decreases, p_1 increases and p_2 also increases slowly. The increase in p_1 becomes slower from $N_{\mathrm{S}} \approx 0.5$, and the minimum error probability is achieved at $N_{\mathrm{S}} \fallingdotseq 0.83$, where $p_1 = \frac{1}{2}$.

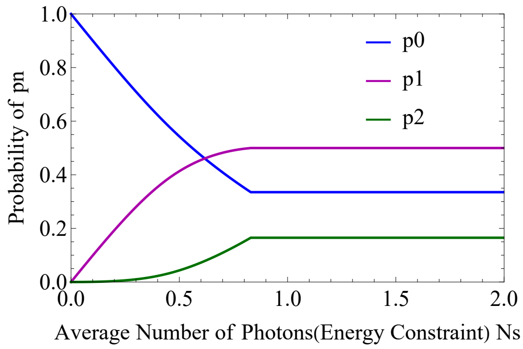


Figure 3: Coefficients p_n of the BPSK-type asymmetric quantum communication with improved Nair state (Opt($N_{max}=2$)) in an attenuated environments with the transmissivity $\frac{2}{3}$.

4.6 Discussion

For the condition we can only use a superposition of $|0\rangle$ and $|1\rangle$, the Nair state, which is an optimal state for a noiseless environment, and the optimal state for an attenuated environment coincide and $\{p_0, p_1\} = \{\frac{1}{2}, \frac{1}{2}\}$, the minimum error probability is achieved. One might think that more $|1\rangle$ should be used in the presence of attenuation, but it turns out that this is not the case.

Let us look at the improved Nair state when $N_{max} = 2$. It can be seen that before the N_{S} is about 0.1, it is similar to the case $N_{max}=1$. This is probably to be because the value of p_2 is close to zero, so that the set of p_n in $N_{max}=1$ and $N_{max}=2$ are close. This is also considered to be the case when increasing N_{max}. In other words, when the N_{S} is so small that a some p_n can be regarded as zero, the results when constrained to N_{max} and the results when $N_{max} \to \infty$ can be considered almost equivalent. It also looks at the N_{S} that achieve the minimum error probability. For $\eta = 1$, the Nair state is optimal, so $N_{\mathrm{S}} = 0.5$, for $\eta = \frac{2}{3}$, $N_{\mathrm{S}} \fallingdotseq 0.83$. It has also been confirmed that the N_{S} achieving the minimum error probability also increases when the transmissivity is further increased. This research is a future work.

5 Conclusion

In this paper, we have studied the asymmetric quantum communication using BPSK in an attenuated environment. In order to find the optimal sending quantum state in an attenuated environment, the optimal sending quantum state is derived numerically when the photon number states used are constrained from $|0\rangle$ to $|N_{max}\rangle$ and the error probability characteristics are analyzed. And we also investigate p_n. A future work is to investigate the error probability characteristics when the constraint on the number of photon states used is more relaxed. This will allow us to advance the research towards the discovery of the optimal sending quantum state under unconstrained conditions in an attenuated environment.

Acknowledgments: This work has been supported in part by JSPS KAKENHI Grant Number 20H00581, 20K20397, 21K04064, 22K20437, and The Nitto Foundation.

References

[1] S. Acharya, R. Alonso, M. Franklin, and S. Zdonik, "Broadcast disks: data management for asymmetric communication environments," ACM SIGMOD Record **24**, pp.199-210, (1995).

[2] M. Adler and B. M. Maggs, "Protocols for asymmetric communication channels," Proc. 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280), Palo Alto, CA, USA, (1998).

[3] S. Pirandola, "Quantum reading of a classical digital memory," Phys. Rev. Lett. **106**, 090504, (2011).

[4] T. Wang and T. S. Usuda, "Error performance of amplitude shift keying-type asymmetric quantum communication system," Entropy **24**, issue5, 708, (2022).

[5] 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. **J106-B**, no.3, pp. 112-125, (2023). (in Japanese)

[6] S. Sameshima, T. Wang, S. Usami, and T. S. Usuda, "PSK-type asymmetric quantum communication and its attenuation characteristics," Proc. 2022 International Symposium on Information Theory and Its Applications (ISITA2022), pp.241-245, Tsukuba, Japan, (2022).

[7] R. Nair, B. J. Yen, S Guha, J. H. Shapiro, and S. Pirandola, "Symmetric M-ary phase discrimination using quantum-optical probe states," Phys. Rev. **A86**, 022306, (2012).

[*]aliphy80@gmail.com [†]ananyaphys.c@gmail.com
[‡]govindlalsidhardh@gmail.com
[§]ramkrishnapatra033@gmail.com [¶]samrat9sen5@gmail.com
[#]srcphy825@gmail.com [♣]sahiln1112@gmail.com
[♢]manik11ju@gmail.com

Extended Abstract

Advantage of Hardy's Nonlocal Correlation in Reverse Zero-Error Channel Coding

Mir Alimuddin[*],[1] Ananya Chakraborty[†],[1] Govind Lal Sidhardh[‡],[1] Ram Krishna Patra[§],[1]
Samrat Sen[¶],[1] Snehasish Roy Chowdhury[#],[2] Sahil Gopalkrishna Naik[♣],[1] and Manik Banik[♢][1]

[1]*Department of Physics of Complex Systems, S. N. Bose National Center for Basic Sciences,*
Block JD, Sector III, Salt Lake, Kolkata 700106, India.
[2]*Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 BT Road, Kolkata, India.*

Hardy's argument constitutes an elegant proof of quantum nonlocality as established by the seminal Bell's theorem. In this Letter, we report an exotic application of Hardy's nonlocal correlations. We devise a simple communication task and show that the expected payoff of the task cannot be positive given that only 1-cbit communication is allowed from the sender to the receiver, who otherwise can share an unlimited amount of classical correlation. Interestingly, the same classical channel can ensure a positive payoff when assisted with correlations exhibiting Hardy's nonlocality. As it turns out, among all the 2-input-2-output no-signaling correlations, only Hardy's correlation can ensure a positive payoff while assisting the 1-cbit channel. This further prompts us to show that in the correlation assisted reverse zero-error channel coding scenario, where the aim is to simulate a noisy channel exactly by a noiseless one in assistance with correlations, assistance of non-maximally pure entangled states – even with vanishingly zero amount of entanglement – could be preferable over the maximal one.

Keywords: NS correlation, Hardy's correlation, Maximally and Non-maximally entangled states, Entanglement assisted simulation of noisy channel, channel capacity.

Arxiv link of the main manuscript. – arXiv:2303.06848

Introduction.– The pioneering work of J. S. Bell establishes one of the most striking departures of quantum theory from a classical worldview [1] (see also [2]). Violation of a Bell-type inequality, as demonstrated in several milestone experiments [3–8], endorses that certain correlations obtained from multipartite entangled states are not compatible with a *local-realistic* description [9]. Apart from Bell-type inequalities, another technique popularly known as 'nonlocality without inequality' proof is often used to establish the nonlocal behaviour of quantum theory. While the first proof of this kind for tripartite quantum systems is due to Greenberger-Horne-Zeilinger [18] (see also [19]), for bipartite systems, such a proof was first proposed by Hardy [20], which is considered to be "simpler and more compelling than the arguments that underlie the derivation of Bell-CH inequality" [21]. More recently, Hardy-type nonlocality proofs have shown to be useful in several practical tasks [22–27].

Here, we report a novel application of Hardy's nonlocal correlation in the simplest communication scenario. In particular, we show that Hardy's nonlocal correlation shared between two distant parties can empower the communication utility of a perfect classical channel. This is quite striking due to the following reasons. Firstly, nonlocal correlations are compatible with the no signalling (NS) principle, and hence they by themselves cannot be used for information transfer. Second, the advantage reported here is different and much more elementary than the nonlocal advantage demonstrated in the communication complexity scenario [13].

A two party guessing game.– We start by introducing a distributed guessing game played between two distant players (say) Alice and Bob. There is a Referee (say) Charlie who, in each run of the game, provides four closed boxes, numbered 1 to 4, to Bob, who has to open one of these boxes. Some of these boxes contain a bomb that will explode upon opening the box. Among the boxes that don't contain the bomb, some may be empty, some may contain a dollar bill, and others may even prompt Bob to pay a dollar bill to Charlie. In each run of the game, Charlie randomly picks one among four different arrangements of these boxes, as shown in Fig.1. Charlie then informs his choice to Alice, who then tries to help Bob in picking a box. However, only 1-bit of classical communication is allowed from Alice to Bob, which may be further assisted with preshared NS correlations shared between them. From now on, we will call this the distributed mine-hunting (DMH) game.

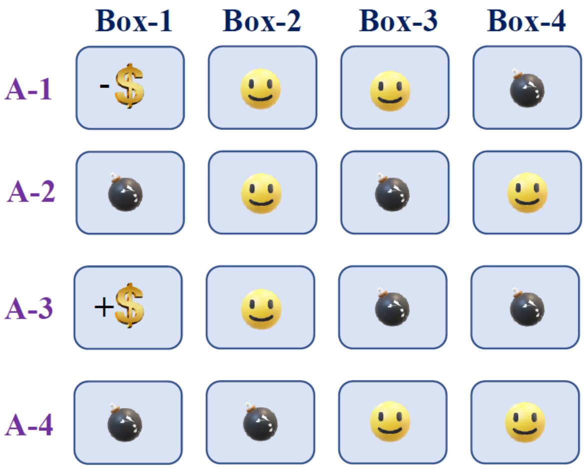


Figure 1. Distributed mine-hunting (DMH) game. Opening a box with $' + \$'$ assures dollar bill gain for Bob while opening a box with $' - \$'$ demands him to pay dollar bill to Charlie. A box with a 'smile' neither offers nor demands any dollar bill, whereas a box with a 'bomb' turns out to be fatal to Bob. Alice knows which of the arrangements {A-1,$\cdots$,A-4} Charlie chooses in a particular run and tries with a limited classical channel to help Bob to optimize his expected dollar gain.

General scenario.– The aforesaid game can be formally studied within a more generic set-up. Alice and Bob are given some collaborative payoff depending on the classical index $z \in \mathcal{Z}$ produced by Bob, given that Alice received some classical message $m \in \mathcal{M}$ sampled from a probability distribution $\{p(m) \mid m \in \mathcal{M}\}$. Such a game, in fact, is completely specified by the payoff matrix $\mathcal{G} \equiv (g_{mz})$, where $g_{mz} \in \overline{\mathbb{R}}$ is the reward/payoff given when Bob produced the index 'z' provided Alice received the message 'm'. For instance, the DMH game is specified by the following payoff matrix:

$$\mathcal{G}_{DMH} \equiv \begin{array}{c||c|c|c|c} \mathcal{M}\backslash\mathcal{Z} & 1 & 2 & 3 & 4 \\ \hline 1 & -1 & 0 & 0 & -\infty \\ 2 & -\infty & 0 & -\infty & 0 \\ 3 & +1 & 0 & -\infty & -\infty \\ 4 & -\infty & -\infty & 0 & 0 \end{array} \tag{1}$$

Payoffs in (1) quantitatively capture the scenario of DMH game. A reward of $-\infty$ for the box containing the bomb captures the notion that choosing such a box must be avoided at all costs [30]. The reward 0 corresponds to the event where the players survive but do not receive any reward. Events with reward $+1$ (-1) correspond to the scenario where the players receive (pay) some dollar bill from (to) Charlie. The game matrix and the sampling distribution of Alice's inputs are common knowledge to the players.

Alice and Bob are cooperative in nature and aim to maximize the payoff. Their collaborative strategy depends on the available resources, which can be broadly categorized into two types: (i) correlation shared between them before the game starts and (ii) communication from Alice to Bob which is allowed even after the game starts. While correlations can further be classified into classical, quantum, or the more general no signalling types, for direct communication, a classical or quantum channel can be used. The strategy employed by the players can be represented as a $|\mathcal{M}| \times |\mathcal{Z}|$ matrix $S \equiv (s_{mz})$, where s_{mz} denotes the probability of producing the output $z \in \mathcal{Z}$ by Bob given that Alice received the message $m \in \mathcal{M}$. Manifestly, entries of such a matrix are non-negative with row sum one. In this work, we refer to this matrix as a strategy matrix which, up to a transposition, is identical to the notion of 'channel matrix' used in [31]. Given such a strategy matrix S, the average payoff can be obtained as

$$\mathcal{P}(S) = \sum_{z,m} p(m)\, g_{mz}\, s_{mz}\ . \tag{2}$$

As it is evident, there will always be a perfect strategy for such a game if $\log_2 |\mathcal{M}|$ bits of communication are allowed from Alice to Bob. Interesting situations arise when communication is limited which might further be aided by preshared correlations of different types. In the next section, we analyze different such cases for DMH game and present some novel results.

Results.– We start with the scenario where only 1 bit of classical communication is allowed from Alice to Bob, and they can share an unlimited amount of share randomness, *i.e.*, classical correlation. By $\Omega_{n_c+SR}(|\mathcal{M}|,|\mathcal{Z}|)$ we denote the set of strategy matrices obtained when n-bits classical communication and unlimited amount of shared randomness are available. The set $\Omega_{n_c+SR}(|\mathcal{M}|,|\mathcal{Z}|)$ forms a polytope with extreme points S^e's representing strategy matrices obtained through deterministic encoding $\mathbb{E} : \mathcal{M} \to \{0,1\}^n$ at Alice's end and deterministic decoding $\mathbb{D} : \{0,1\}^n \to \mathcal{Z}$ at Bob's end [31] (see also [32]). Entries of such extreme strategy matrices can only be 0 or 1, and since only n-cbits are allowed, such a strategy can have at most 2^n non-zero columns. Our first technical result is to limit the optimal success probability of the game in Eq.(1) for classical strategies.

Theorem 1. *When players communicate using 1 bit of information assisted by a pre-shared unlimited classical correlation, the maximum possible average payoff of the DMH game is limited to zero.*

Theorem 2. *A strictly positive average payoff in DMH game can be ensured when a 2-input-2-output Hardy's nonlocal correlation is available to assist the 1-cbit communication channel from Alice to Bob.*

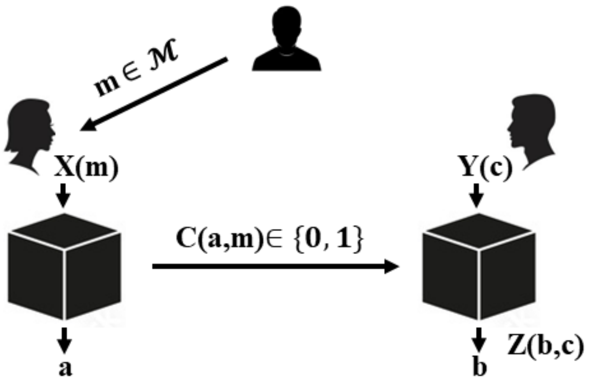


Figure 2. General strategy to play a game $\mathcal{G}$ when the 1-cbit communication channel is assisted with NS correlation. Alice computes the input $x = X(m)$ to her part of the nonlocal box based on the message $m \in \mathcal{M}$ received from Referee. The output a of the nonlocal box at her end and the message m determines the classical bit $c = C(a, m)$ sent to Bob. Bob then inputs $y = Y(c)$ into his end of the NS box and obtains the output b. Finally, he generates his guess as $z = Z(b, c)$.

Theorem 3. *Any* 2*-input-*2*-output NS correlation providing a strictly positive payoff in DMH game as an assistance to the* 1*-cbit channel must exhibit Hardy's nonlocality.*

Theorem 4. *Two qubit maximally entangled state together with* 1*-cbit channel from Alice to Bob does not result in a strategy ensuring strictly positive average payoff in DMH game.*

This theorem has an interesting implication. It shows that there exists a communication task wherein **a non-maximally pure entangled state can be preferable over the maximally entangled one even when the entanglement of the former is vanishingly zero.** More formally we can deduce the following corollary.

Corollary 1. *For every non maximally entangled state* $|\psi\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ *there exists a strategy matrix* S_ψ *such that* $S_\psi \in \Omega_{1_c+SR+|\psi\rangle}(4,4)$ *but* $S_\psi \notin \Omega_{1_c+SR+|\phi^+\rangle}(4,4)$.

Conclusion and outlook.– In conclusion, the present work establishes a novel use of quantum entanglement in zero-error information theory [49] (see also [50]) whose motivation arises from the fact that in many real-world critical applications no errors can be tolerated, and in practice, the communication channel can only be available for a finite number of times. In particular, we show that quantum correlations exhibiting Hardy's nonlocality can empower the communication *utility* of a perfect classical communication channel. In the supplemental part of the main manuscript [51], we have also shown that similar results can be obtained by considering generalization of Hardy's nonlocality argument as proposed by Cabello [29]. Our work also motivates many questions for future study. For instance, it would be interesting to see whether any nonlocal correlation can be made useful as a communication resource in the sense discussed here. It will also be interesting to see whether maximally entangled states of higher dimensions provide some advantage in the DMH game.

[1] J. S. Bell; On the Einstein Podolsky Rosen paradox, Physics Physique Fizika **1**, 195 (1964); On the Problem of Hidden Variables in Quantum Mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[2] N. D. Mermin; Hidden variables and the two theorems of John Bell, Rev. Mod. Phys. **65**, 803 (1993).

[3] 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 (1969).

[4] S. J. Freedman and J. F. Clauser; Experimental Test of Local Hidden-Variable Theories, Phys. Rev. Lett. **28**, 938 (1972)

[5] A. Aspect, P. Grangier, and G. Roger; Experimental Tests of Realistic Local Theories via Bell's Theorem, Phys. Rev. Lett. **47**, 460 (1981).

[6] A. Aspect, P. Grangier, and G. Roger; Experimental Realization of Einstein-Podolsky-Rosen-Bohm Gedankenexperiment: A New Violation of Bell's Inequalities, Phys. Rev. Lett. **49**, 91 (1982).

[7] 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).

[8] 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).

[9] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner; Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[10] A. K. Ekert; Quantum cryptography based on Bell's theorem, Phys. Rev. Lett. **67**, 661 (1991).

[11] J. Barrett, L. Hardy, and A. Kent; No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005).

[12] U. Vazirani and T. Vidick; Fully Device-Independent Quantum Key Distribution, Phys. Rev. Lett. **113**, 140501 (2014).

[13] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf; Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010).

[14] S. Pironio *et al.* Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[15] E. G. Cavalcanti and H. M. Wiseman; Bell Nonlocality, Signal Locality and Unpredictability (or What Bohr Could Have Told Einstein at Solvay Had He Known About Bell Experiments), Found Phys **42**, 1329 (2012).

[16] A. Chaturvedi and M. Banik; Measurement-device-independent randomness from local entangled states, EPL **112**, 30003 (2015).

[17] R. Colbeck and R. Renner; Free randomness can be amplified, Nature Phys **8**, 450 (2012).

[18] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going Beyond Bell's Theorem, In: Kafatos, M. (eds) Bell's theorem, quantum theory and conceptions of the universe, Fundamental Theories of Physics, vol 37 (1989).

[19] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger; Bell's theorem without inequalities, Am. J. Phys. **58**, 1131 (1990).

[20] L. Hardy; Quantum mechanics, local realistic theories, and Lorentz-invariant realistic theories, Phys. Rev. Lett. **68**, 2981 (1992).

[21] N. D. Mermin; Quantum mysteries refined, Am. J. Phys. **62**, 880 (1994).

[22] S. Das, M. Banik, A. Rai, MD R.Gazi, and S. Kunkri; Hardy's nonlocality argument as a witness for postquantum correlations, Phys. Rev. A **87**, 012112 (2013).

[23] A. Mukherjee, A. Roy, S. S. Bhattacharya, S. Das, Md. R. Gazi, and M. Banik; Hardy's test as a device-independent dimension witness, Phys. Rev. A **92**, 022302 (2015).

[24] H-W Li, M. Pawłowski, R. Rahaman, G-C Guo, and Z-F Han; Device- and semi–device-independent random numbers based on noninequality paradox, Phys. Rev. A **92**, 022327 (2015).

[25] R. Ramanathan, M. Horodecki, H. Anwer, S. Pironio, K. Horodecki, M. Grünfeld, S. Muhammad, M. Bourennane, and P. Horodecki; Practical No-Signalling proof Randomness Amplification using Hardy paradoxes and its experimental implementation, arXiv.1810.11648.

[26] A. Rai, M. Pivoluska, M. Plesch, S. Sasmal, M. Banik, and S. Ghosh; Device-independent bounds from Cabello's nonlocality argument, Phys. Rev. A **103**, 062219 (2021).

[27] A. Rai, M. Pivoluska, S. Sasmal, M. Banik, S. Ghosh, and M. Plesch; Self-testing quantum states via nonmaximal violation in Hardy's test of nonlocality, Phys. Rev. A **105**, 052227 (2022).

[28] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter; Zero-error channel capacity and simulation assisted by nonlocal correlations, IEEE Trans. Info. Theory **57**, 5509 (2011).

[29] A. Cabello; Bell's theorem with and without inequalities for the three-qubit Greenberger-Horne-Zeilinger and W states, Phys. Rev. A 65, **032108** (2002).

[30] Although uses of finite payoffs are common in game theory, the framework is quite flexible to incorporate wide range of situations including those with ±∞ payoffs. Such extreme payoffs generally appear in financial trading games, gambling games, and decision-making games. A classic example is the "Stag Hunt" game. If the hunters are part of a community that relies on hunting for survival then the failure to catch a stag leads to starvation and the payoff for each hunter could be considered $-\infty$.

[31] P. E. Frenkel and M. Weiner; Classical Information Storage in an n-Level Quantum System, Commun. Math. Phys. **340**, 563 (2015).

[32] M. Dall'Arno, S. Brandsen, A. Tosini, F. Buscemi, and V. Vedral; No-Hypersignaling Principle, Phys. Rev. Lett. **119**, 020401 (2017).

[33] For the set $\Omega_{n_c+SR}(|\mathcal{M}|,|\mathcal{Z}|)$, number of vertices N can be calculated using the formula $N = \sum_{k=1}^{2^n} k! \begin{pmatrix} |\mathcal{Z}| \\ k \end{pmatrix} \begin{Bmatrix} |\mathcal{M}| \\ k \end{Bmatrix}$ provided in [32]. Here $\begin{pmatrix} |\mathcal{Z}| \\ k \end{pmatrix} := \frac{|\mathcal{Z}|!}{k!(|\mathcal{Z}|-k)!}$ and $\begin{Bmatrix} |\mathcal{M}| \\ k \end{Bmatrix} := \sum_{j=0}^{k} \frac{1}{k!}(-1)^{k-j} \begin{pmatrix} k \\ j \end{pmatrix} j^{|\mathcal{M}|}$.

[34] 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 (1969).

[35] J. L. Cereceda; Quantum mechanical probabilities and general probabilistic constraints for Einstein-Podolsky-Rosen-Bohm experiments, Found. Phys. Lett. **13**, 427 (2000).

[36] R. Rabelo, L. Y. Zhi, and V. Scarani; Device-Independent Bounds for Hardy's Experiment, Phys. Rev. Lett. **109**, 180401 (2012).

[37] B. S. Cirel'son; Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[38] S. Goldstein; Nonlocality without inequalities for almost all entangled states for two particles, Phys. Rev. Lett. **72**, 1951 (1994).

[39] T.F. Jordan, Testing Einstein-Podolsky-Rosen assumptions without inequalities with two photons or particles with spin 1/2, Phys. Rev. A **50**, 62 (1994).

[40] Charles H. Bennett and Stephen J. Wiesner; Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states, Phys. Rev. Lett. **69**, 2881 (1992).

[41] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal; Entanglement-Assisted Classical Capacity of Noisy Quantum Channels, Phys. Rev. Lett. **83**, 3081 (1999).

[42] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter; Improving Zero-Error Classical Communication with Entanglement, Phys. Rev. Lett. **104**, 230503 (2010).

[43] M. B. Plenio and S. Virmani; An introduction to entanglement measures, Quant. Inf. Comput. **7**, 1 (2007).

[44] P. E. Frenkel and M. Weiner; On entanglement assistance to a noiseless classical channel, Quantum **6**, 662 (2022).

[45] R. K. Patra, S. G. Naik, E. P. Lobo, S. Sen, T. Guha, S. S. Bhattacharya, M. Alimuddin, and M. Banik; Classical analogue of quantum superdense coding and communication advantage of a single quantum, arXiv:2202.06796.

[46] M. A. Nielsen; Conditions for a Class of Entanglement Transformations, Phys. Rev. Lett. **83**, 436 (1999).

[47] C. E. Shannon; A mathematical theory of communication, Bell System Technical Journal **27**, 379 (1948).

[48] A. S. Holevo; Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel, Problems Inform. Transmission **9**, 177 (1973).

[49] C. Shannon; The zero error capacity of a noisy channel, IRE Trans. Inf. Theory **2**, 8 (1956).

[50] J. Korner and A. Orlitsky; Zero-error information theory, IEEE Tran. Inf. Theory **44**, 2207 (1998).

[51] Mir Alimuddin and Ananya Chakraborty and Govind Lal Sidhardh and Ram Krishna Patra and Samrat Sen and Snehasish Roy Chowdhury and Sahil Gopalkrishna Naik and Manik Banik; Advantage of Hardy's Nonlocal Correlation in Reverse Zero-Error Channel Coding arXiv,2303.06848.

Twin-field quantum key distribution with partial phase postselection

Yao Zhou[1 2 3 4] Zhen-Qiang Yin[1 2 3 4 *]

[1] *CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei, Anhui 230026, China*

[2] *CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei, Anhui 230026, China*

[3] *Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*

[4] *State Key Laboratory of Cryptology, P. O. Box 5159, Beijing 100878, China*

Abstract. In recent years, a revolutionary breakthrough called twin-field (TF) QKD has been developed to overcome the linear key-rate constraint and greatly increases the achievable distance. Later, no-phase-postselection TF-QKD was proposed and became a popular variant, since the removal of phase postselection leads to a higher key rate. However, the achievable distance is decreased compared to the original one. Here, we propose a TF-QKD protocol with partial phase postselection. Namely, its code mode is still free from global phase randomization and postselection to make sure the advantage of the high key rate remain. On other hand, phase postselection is introduced in the decoy mode to improve the performance. Applying an operator dominance condition, we prove universal security of the proposed protocol in the finite-key regime against coherent attacks, and numerical simulations confirm its potential advantages in terms of key rate and achievable distance.

Keywords: quantum cryptography, twin-field quantum key distribution, phase postselection

1 Introduction

Quantum key distribution (QKD) allows two remote parties to share information-theoretically secure keys. In 2018, M. Lucamarini et al. proposed the original twin-field (TF) QKD [1] to overcome the fundamental rate-distance limit [2, 3] of QKD without needing quantum repeaters. Once it was proposed, it has received extensive attention due to the advantages of measurement-device-independence and significant improvement on achievable distances. Soon after, the security of TF-QKD was strictly proved in [4].

Roughly speaking, the original TF-QKD [1, 4] consists of code-mode and decoy-mode. In the former one, the phase 0 or π of a weak coherent pulse is modulated to encode raw key bit 0 or 1 respectively. The latter one is similar but with different intensities, the function of which is to monitor security. Additionally, a continuous random phase θ_a (θ_b) from $[0, 2\pi)$ is applied in each optical pulse by Alice (Bob). After receiving the announcement from Eve, Alice and Bob postselect the cases satisfying $\theta_a \approx \theta_b$ to generate sifted key bits. This phase randomization and postselection play important role in security, but obviously reduce the key rate per trial. As an alternative, the variant called no-phase-postselection (NPP) TF-QKD [5, 6, 7] removes the phase randomization in code mode, thus its key rate will be free of reduction due to phase postselection. In its decoy mode, phase-randomization remains but postselection is also bypassed for simplicity. Intuitively, phase postselection inevitably reduces the key rate, but it seems helpful for Alice and Bob to monitor the security. Inspired by this idea, we propose a variant of the TF-type protocol with partial phase postselection. In our protocol, code mode can be the same as NPP-TF-QKD, thus still free from global phase randomization and postselection, which guarantees the advantage of high key rate remains. Conversely, the phase postselection is introduced in the decoy mode to help Alice and Bob to bound information leakage more accurately.

2 Protocol

Our protocol process is as follows.

Step 1: Alice and Bob randomly choose code mode and decoy mode with probabilities p_0 and $1-p_0$, respectively. In each mode, they send weak coherent pulses independently to the untrusted third party Charlie.

Step 2: In the code mode, Alice (Bob) randomly generates a key bit $\kappa_a(\kappa_b) \in \{0,1\}$ and prepares the corresponding weak coherent state $|e^{i\kappa_{a(b)}\pi}\sqrt{\mu}\rangle$ with intensity μ. Additionally, Bob introduces a randomized phase $\delta_b \in [-\frac{\Delta}{2}, \frac{\Delta}{2}]$ (typically $\Delta = \pi/8$). In the decoy mode, Alice (Bob) prepares nothing (vacuum state $\mu_0 = 0$) with probability (conditional probability in the decoy mode, the same below) $p_{10}/(1-p_0)$ or a weak coherent state with intensity randomly choosing from μ_1 and μ_2 with probabilities $p_{11}/(1-p_0)$ and $p_2/(1-p_0)$. For Alice and Bob, they modulate a phase $\theta_{a(b)}$ on their own weak coherent state for both intensities μ_1 and μ_2, where $\theta_{a(b)}$ is a randomized phase chosen uniformly from $[0, 2\pi)$.

Step 3: Charlie makes the two incoming states interfere on a beamsplitter (BS), two single-photon detectors L and R are located at its two distinct outputs. He records which detector clicks.

Step 4: Steps 1 to 3 will be repeated N_{tot} times. When the quantum communication is over, Charlie publicly announces all the information about the detection events. Only the trials where just one of the two detectors clicked are retained for further processing, all the other trials are discarded. Alice and Bob announce the intensities for each remaining trial. When Alice and Bob both choose

[*] yinzq@ustc.edu.cn

the code mode (both send the intensity μ), they record their key bits κ_a and κ_b sequentially to form the sifted key string. Note that if the click of the right detector R was announced, Bob should flip his corresponding key bit κ_b. When Alice and Bob both choose the decoy mode, they retain the trials in which they both send a vacuum state. For other cases in decoy mode, they announce their random phases and only the trials that meet both of the following conditions will be retained: (1) both of them choose the same intensity, (2) phase postselection condition: the random phase θ_a and θ_b satisfies $|\theta_a - \theta_b| \bmod \pi \le \frac{\Delta}{2}$, e.g., $|-\frac{17}{8}\pi| = |-\frac{\pi}{8}| = |\frac{17}{8}\pi| = \frac{\pi}{8}$. We denoted the length of the sifted key string as K_0. We also denote the numbers of the retained trials in which Alice and Bob both send the intensities μ_0, μ_1 or μ_2 as K_{10}, K_{11} and K_2, respectively, and $K_1 \equiv K_{10} + K_{11}$.

Step 5: According to K_0, K_1 and K_2, Alice and Bob can share a secret key string with length G from their sifted key string K_0 after error correction and privacy amplification [8] with a failure probability no larger than ϵ_{sec}.

3 Result

Based on the operator dominance method [9], we proof the security of our protocol in the finite-key region and get the final key length $G = K_0 - [K_0 h(f(K_1, K_2)/K_0)] - H_{\text{EC}} - \zeta - \zeta'$, where $f(K_1, K_2)$ is the analytic function of K_1, K_2 used to estimate the upper bound on the phase error. Here we define $H_{\text{EC}} = f_{\text{cor}} K_0 h(e_{\text{bit}})$ as the cost of error correction, where f_{cor} is the error correction inefficiency, e_{bit} is the bit error rate in code mode, $h(x) = -x\log_2(x) - (1-x)\log_2(1-x)$. Meanwhile, ζ' bits are consumed to ensure that the failure probability of error verification is up to $2^{-\zeta'}$. The protocol is ϵ_{sec}-secure with a small security parameter $\epsilon_{\text{sec}} = \sqrt{2}\sqrt{\epsilon + 2^{-\zeta}} + 2^{-\zeta'}$. Note that, in the numerical parts of our work, we fix $\epsilon_{\text{sec}} = 2^{-31}$ by assuming $\epsilon = 2^{-66}$, $\zeta = 66$ and $\zeta' = 32$.

We denote the distance between Alice and Bob as L (in km) and its loss is 0.2dB/km. We set the error correction inefficiency $f_{\text{cor}} = 1.1$ in our protocol. The detection efficiency of Charlie's apparatus is $\eta_d = 30\%$, so the channel transmittance from Alice (Bob) to Charlie is $\eta = 10^{\frac{-0.2L}{20}} \eta_d$. The dark count probability (also the count rate of the vacuum state) of each detector is $p_d = 10^{-8}$ per pulse and the intrinsic error rate due to imperfect optical interference visibility is $e_d = 0.03$. Using the same simulation model as [10], We optimize the group of parameters $(\mu, \mu_1, \mu_2, p_0, p_{10}, p_{11}, p_2, \Delta)$ to maximize the secret key rate G for any communication distance L.

We compare our results with the results [9] of the original NPP-TF-QKD protocol in fig.1 for the $N_{\text{tot}} = 10^{13}, 10^{15}$ and 10^{18} cases. The achievable distances of our protocol are 426 km, 443 km and 450 km in the three cases, which are 17 km, 26 km and 31 km longer than the original protocol, respectively. At short or medium distances, such as 340 km, the key rate bits per pulse of our protocol are doubled than original protocol. Additionally, our idea of partial phase postselection is applicable

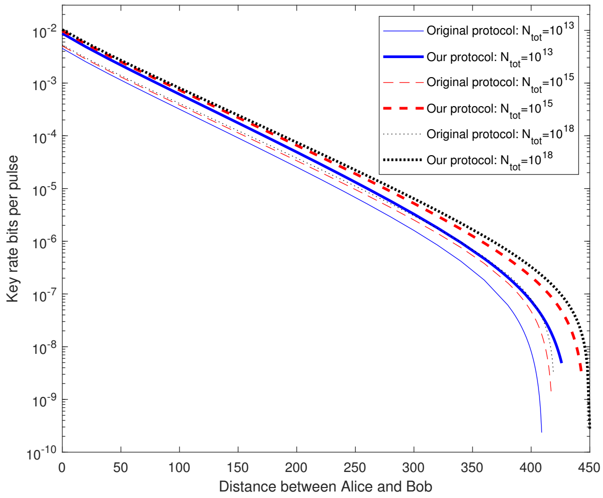


Figure 1: The key rate bits per pulse as a function of distance L between Alice and Bob. We show the result of our protocol and the original NPP-TF-QKD protocol for the $N_{\text{tot}} = 10^{13}, 10^{15}$ and 10^{18} cases.

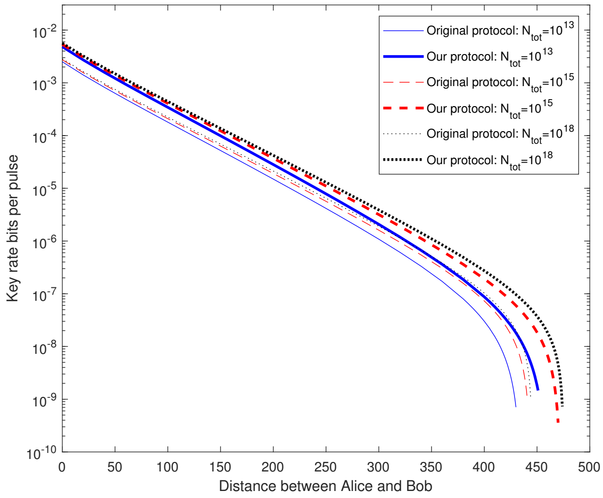


Figure 2: The key rate bits per pulse as a function of distance L between Alice and Bob in $N_{\text{tot}} = 10^{13}$, 10^{15} and 10^{18} cases for the original four-phase protocol [10] and our improved protocol.

in other variants of TF-QKD. Indeed, the four-phase TF-QKD proposed in [10] is very similar to NPP-TF-QKD, except that in code mode both Alice and Bob encode bit in phase from sets $\{0, \pi\}$ or $\{\pi/2, -\pi/2\}$ and then select the rounds with the same coding sets to generate sifted K_0 key bits. Evidently, we can introduce the proposed phase postselection to the decoy mode in this protocol. The improvement results are shown in fig.2. The achievable distances of improved four-phase protocol are 451 km, 470 km and 474 km when sending 10^{13}, 10^{15} and 10^{18} pulses, which are 21 km, 29 km and 30 km longer than the original four-phase protocol. At short or medium distances, such as 340 km, the key rate bits per pulse of the improved protocol are also doubled than the original protocol.

In summary, we propose a variation of TF-type QKD with partial phase postselection and give the security proof in the finite-key regime. Phase postselection being still free in code mode but introduced in decoy mode significantly improves the key rate and achievable distance. Our idea can also be used in the optimized four-phase twin-field protocol proposed by [10], which confirms its potential advantages in terms of key rate and achievable distance.

References

[1] Lucamarini et al. Overcoming the rate–distance limit of quantum key distribution without quantum repeaters. In *Nature 557.7705*, pages 400–403, 2018.

[2] Stefano Pirandola et al. Fundamental limits of repeaterless quantum communications. In *Nat. Commun. 8.1*, pages 15043, 2017.

[3] Masahiro Takeoka et al. Fundamental rate-loss tradeoff for optical quantum key distribution. In *Nat. Commun. 5.1*, pages 5235, 2014.

[4] Xiongfeng Ma et al. Phase-Matching Quantum Key Distribution. In *Phys. Rev. X 8 (3)*, pages 031043, 2018.

[5] Cui, Chaohan et al. Twin-Field Quantum Key Distribution without Phase Postselection. In *Phys. Rev. Applied 11 (3)*, pages 034053, 2019.

[6] Curty et al. Simple security proof of twin-field type quantum key distribution protocol. In *npj Quantum Information 5.1*, pages 64, 2019.

[7] Jie Lin et al. Simple security analysis of phase-matching measurement-device-independent quantum key distribution. In *Phys. Rev. A 98 (4)*, pages 042332, 2018.

[8] Renato Renner et al. Security of Quantum Key Distribution. In *PhD thesis. School Swiss Federal Institute of Technology Zurich*, 2005.

[9] Maeda et al. Repeaterless quantum key distribution with efficient finite-key analysis overcoming the rate-distance limit. In *Nat. Commun. 10.1*, pages 3140, 2019.

[10] Shuang Wang et al. Twin-field quantum key distribution over 830-km fibre. In *Nat. Photonics 16.2*, pages 154-161, 2022.

Extended abstract:
Negativity as a Resource for Memory Reduction in Stochastic Process Modeling

Kelvin Onggadinata[1 3] Andrew Tanggara[1 2 *] Dagomir Kaszlikowski[1 3]

[1]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543.*
[2]*Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673.*
[3]*Department of Physics, National University of Singapore, 3 Science Drive 2, Singapore 117543.*

Abstract. Finding the most efficient model that generates a stochastic process is an important and ubiquitous task in the field of quantitative science. It is known that for a given classical model with least amount of memory, as quantified by its statistical complexity, one can generally construct a quantum model with less memory. We first show that the presence of negativity in the quasi-probability representation (QPR) of a quantum machine is necessary for such memory advantage, and see that the amount of negativity grows as the memory advantage gets larger. From this intuition, we propose a quasi-probabilistic model from a class larger than those of quantum QPRs that saturates the absolute information-theoretic memory lower bound that for any model to perfectly generate a stochastic process, known as the excess entropy. We view and show that negativity acts as a resource to memory reduction in the same way it has been a resource in obtaining advantages in quantum information and quantum computation tasks.

Keywords: stochastic process modeling, hidden Markov model, quasi-probability representation, negative probabilities, Renyi entropies

Stochastic process modelling is an essential task which have found plenty of applications in the fields of quantitative science involving time-series data. One of the most ubiquitously used model is the Hidden Markov Model (HMM) [Vid11]; [Vid14]; [RJ86], which consists of an internal state that transitions over-time as it emulates the statistics of the given stochastic process that it models. The efficiency of such an HMM, which has been pervasively quantified using an entropic measure on its stationary distribution, called the *statistical complexity*, has been one of the primary focus of the field of computational mechanics [CY89]; [SC01] which has found many practical applications [Jel98]; [GJ95]; [FST98]; [HKS10]; [Yan+08] [CF97], as well as fundamentally interesting from both information-theoretical and physical standpoints.

Notably at the center of computational mechanics is the *ϵ-machine*, which is the unique HMM with least amount of memory among all *predictive* HMMs, namely those that perfectly emulates the future statistics of a stochastic process given only the past data. It has been shown more recently that one can obtain an HMM with lower statistical complexity for some stochastic processes by resorting to a construction with quantum-mechanical internal states and unitary dynamics, called the *q-machine* [Gu+12]; [MAC16]; [Yan+18]; [Rie+16]; [BTG18]; [Liu+19]; [LC19]; [Ell+20]; [Ell21] (also similar quantum HMM constructions in [WC08]; [MBW12]; [MW16]). Furthermore it was shown that a specific q-machine construction in [Liu+19] that involves a one-to-one mapping between the states of ϵ-machine and the states of the q-machine along with some optimized phase factor is optimal among all q-machines. Despite these fruitful excursions into the quantum regime, it remains the question of how one can construct an HMM that saturates the absolute information-theoretic memory lower bound necessary for any HMM that perfectly generates a stochastic process, known as the *excess entropy* (as it was posed as an open question in [Gu+12]). In this work, we address this long-standing open question using the framework of quasi-probabilities.

As a natural alternative representation to the Hilbert space representation of quantum mechanics, we first resort to the quasi-probability representation (QPR) [FE08]; [FE09]; [Fer11] of the q-machines. We first show that the presence of negativity in the QPR is a necessary condition for quantum memory advantage, despite not be-

[*] andrew.tanggara@gmail.com

ing sufficient, as we show in a construction of a non-negative discrete Wigner QPR for some q-machines that exhibits memory advantage. Nevertheless with evidences showing that the amount of negativity are larger for q-machines with smaller statistical complexity, we proceed to an analysis of a class of quasi-probabilities larger than that of quantum QPR. Here, we answer the long-standing open question of the existence of an HMM that saturates the excess entropy in a construction of a quasi-probabilistic HMM called the *n-machine*.

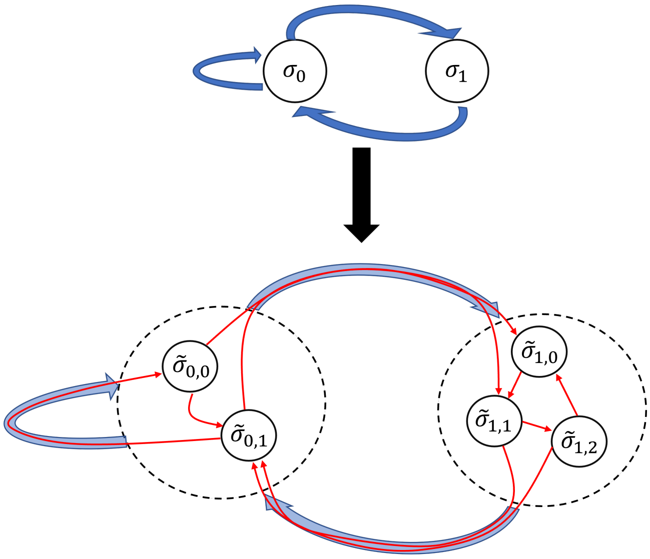


Figure 1: Illustration of n-machine construction from a classical HMM.

The n-machine construction takes an ϵ-machine or a q-machine that generates a certain stochastic process and "splits" one or more of its internal states into multiple states for which internal transitions between them can be with negative probability (hence still normalized) such that it still enables any n-machine to faithfully generates the statistics of any physically realizable stochastic process. We show that it is possible to construct such machine with statistical complexity saturating the excess entropy for stochastic processes realizable by a finite-state ϵ-machine or a q-machine. Another notable feature of this construction is that the resulting n-machine can be thought of as a generalization of the discrete Wigner quasi-probability representation (QPR) of a q-machine, where some features of the n-machine manifests in the QPR stationary distribution and transition matrix. Furthermore, we have found evidence showing that the amount of negativity present in the stationary state of the n-machine is proportional to its memory advantage over the ϵ-machine, as can be shown from constructions for some stochastic processes.

Moreover for some stochastic processes, we show that the n-machine can be more efficient than the most general classical and quantum HMMs, called the *generative* HMMs, which internal dynamics are *oracular*, i.e. containing information about both the future and the past [LA09a]; [LA09b]; [Löh12]; [Rue+18]; [Ell21]. It is to be noted that a consequence of oracular nature of generative models is that it introduces non-determinism in the internal transition probability of the HMM, which makes the problem of minimizing the statistical complexity of classical or quantum generative HMMs for a given stochastic process a persistently hard problem, let alone determining whether one that saturates the excess entropy bound exists. Our n-machine construction hence circumvents this notoriously elusive problem for the more general class of quasi-probabilistic HMMs by saturating the excess entropy lower bound. Combined with the aforementioned n-machine feature that generalizes the discrete Wigner QPR, we see the n-machine construction as a promising first step towards a further, more-restricted construction for optimal generative classical and quantum HMMs. The class of n-machines can be also viewed as a restricted subclass of the class of all HMMs operating within the regime of generalized probabilistic theories (GPTs), a general physical theory that contains quantum theory and classical theory, which is known as an HMM *quasi-realization* [Vid14]; [FLW22]. Our n-machine construction can be seen as part of the long-standing efforts of finding the set of principles that single out quantum theory from any other GPTs [Bar+10]; [Lam18]; [Mül21]; [Plá16]; [Sch+21]; [Sha21]; [GM20]; [CY14]; [Chi+20].

In light of this quasi-probabilistic HMM construction, we use a Rényi entropic definition of statistical complexity and excess entropy with established operational interpretations and have found many applications in cryptography [Ben+95]; [Ren08]; [Buh+08], coding theory [Gal65]; [Csi72]; [Ari73]; [Csi95]; [PV10] and quantum theory [BPP12]; [DFW14]; [Dup+14]; [BCW14]. We show that they still enjoy most of the properties satisfied by their traditional counterparts for any classical and quantum HMMs, hence validates their utility as a measure of information. The use of these Rényi entropic measure may be of separate interests due to its features, such as a correspondence that we that found to the overlap between internal states of the q-machine (which their memory advantage has been attributed to).

For the n-machines, we have showed that these Rényi excess entropy and Rényi statistical complexity are well-defined, despite negative values in the n-machine stationary distribution. With the increasing amount of interests in the use of entropic measures on quasi-probabilities [MF00]; [WS07]; [BL19]; [BLZ22]; [OKK22]; [KJ22], we also argue that this is a reasonable measure of information for quasi-probabilities. In particular, we refer to the studies of classical simulation of quasi-probability sampling [AB14]; [PWB15] which has found applications in the study of classical simulations of quantum computation [PWB15]; [RRC16]; [KJ22]; [Kou+22], quantum error-correction [TBG17]; [Tak21]; [Tak+22], classical simulation of quantum memory channels [Yua+21], and local simulation of non-local quantum channels [MF21]. We made a simple observation by taking the Rényi entropy of the classical probability distribution P that simulates sampling of a given quasi-probability distribution Q and found that the entropy of P is simply the sum of entropy of Q plus a logarithmic factor of the amount of negativity present in Q. The latter quantity is precisely what has been identified as a resource for quantum computational advantage known as *mana* [Vei+14], whereas the amount of negativity has been shown as the overhead of classical simulation of quasi-probability sampling [PWB15]; [KJ22]. We argue that this observation shows that the entropy of a quasi-probability Q is simply part of how much information contained in its classical simulation P subtracted with the amount of overhead cost of running this simulation, hence giving us the "true" information content of Q. In addition, our use of entropic measure for quasi-probabilistic stochastic processes can also may also further the study of the information-theoretic principles of physical theories from an operational perspective of stochastic processes generation, which may be tied to uses of entropy of quasi-probabilities to derive quantum theory [WS07]; [OKK22]; [BLZ22].

We view that negativity acts as a resource in obtaining memory advantage in stochastic process modelling, analogous to how it acts as a resource in obtaining advantages in numerous computational [ME12]; [Vei+12]; [Vei+14]; [How+14]; [KK21], communication [AB11]; [AB14]; [OBC14], and metrology [Arv+20] tasks. It is interesting to ask a further question on identifying information-theoretic principles, such as the data-processing inequality, that one can impose on a stochastic process for it to have a realizable model within some physical theory and identify the most efficient model among all such theories. One way that can directly follow from this work is to identify such principle(s) to obtain the QPR and frame representations of quantum theory [FE08]; [FE09]; [Fer11] from the n-machine construction which puts further restriction on the quasi-probabilistic HMM construction.

References

[AB11] Samson Abramsky and Adam Brandenburger. "The sheaf-theoretic structure of non-locality and contextuality". In: *New Journal of Physics* 13.11 (2011), p. 113036 (cit. on p. 3).

[AB14] Samson Abramsky and Adam Brandenburger. "An Operational Interpretation of Negative Probabilities and No-Signalling Models". In: *Horizons of the Mind. A Tribute to Prakash Panangaden: Essays Dedicated to Prakash Panangaden on the Occasion of His 60th Birthday.* Ed. by Franck van Breugel et al. Cham: Springer International Publishing, 2014, pp. 59–75. DOI: 10.1007/978-3-319-06880-0_3. URL: https://doi.org/10.1007/978-3-319-06880-0_3 (cit. on p. 3).

[Ari73] Suguru Arimoto. "On the converse to the coding theorem for discrete memoryless channels (corresp.)" In: *IEEE Transactions on Information Theory* 19.3 (1973), pp. 357–359 (cit. on p. 2).

[Arv+20] David RM Arvidsson-Shukur et al. "Quantum advantage in postselected metrology". In: *Nature communications* 11.1 (2020), p. 3775 (cit. on p. 3).

[Bar+10] Howard Barnum et al. "Entropy and information causality in general probabilistic theories". In: *New Journal of Physics* 12.3 (2010), p. 033024 (cit. on p. 2).

[BCW14] Mario Berta, Patrick J Coles, and Stephanie Wehner. "Entanglement-assisted guessing of complementary measurement outcomes". In: *Physical Review A* 90.6 (2014), p. 062127 (cit. on p. 2).

[Ben+95] Charles H Bennett et al. "Generalized privacy amplification". In: *IEEE Transactions on Information theory* 41.6 (1995), pp. 1915–1923 (cit. on p. 2).

[BL19] Adam Brandenburger and Pierfrancesco La Mura. *Axioms for Rényi Entropy with Signed Measures.* Available online: www.adambranburger.com (accessed on 27 February 2023). 2019 (cit. on p. 3).

[BLZ22] Adam Brandenburger, Pierfrancesco La Mura, and Stuart Zoble. "Rényi Entropy, Signed Probabilities, and the Qubit". In: *Entropy* 24.10 (2022), p. 1412 (cit. on p. 3).

[BPP12] Gustavo Martın Bosyk, M Portesi, and A Plastino. "Collision entropy and optimal uncertainty". In: *Physical Review A* 85.1 (2012), p. 012108 (cit. on p. 2).

[BTG18] Felix C. Binder, Jayne Thompson, and Mile Gu. "Practical Unitary Simulator for Non-Markovian Complex Processes". In: *Phys. Rev. Lett.* 120 (24 June 2018), p. 240502. DOI: 10.1103/PhysRevLett.120.240502. URL: https://link.aps.org/doi/10.1103/PhysRevLett.120.240502 (cit. on p. 1).

[Buh+08] Harry Buhrman et al. "Possibility, impossibility, and cheat sensitivity of quantum-bit string commitment". In: *Physical Review A* 78.2 (2008), p. 022316 (cit. on p. 2).

[CF97] James P. Crutchfield and David P. Feldman. "Statistical complexity of simple one-dimensional spin systems". In: *Phys. Rev. E* 55 (2 Feb. 1997), R1239–R1242. DOI: 10.1103/PhysRevE.55.R1239. URL: https://link.aps.org/doi/10.1103/PhysRevE.55.R1239 (cit. on p. 1).

[Chi+20] Giulio Chiribella et al. "General Bayesian theories and the emergence of the exclusivity principle". In: *Physical Review Research* 2.4 (2020), p. 042001 (cit. on p. 2).

[Csi72] Imre Csiszár. "A class of measures of informativity of observation channels". In: *Periodica Mathematica Hungarica* 2.1-4 (1972), pp. 191–213. DOI: 10.1007/BF02018661. URL: https://doi.org/10.1007/BF02018661 (cit. on p. 2).

[Csi95] Imre Csiszár. "Generalized cutoff rates and Rényi's information measures". In: *IEEE Transactions on information theory* 41.1 (1995), pp. 26–34. DOI: 10.1109/18.370121. URL: https://ieeexplore.ieee.org/document/370121 (cit. on p. 2).

[CY14] Giulio Chiribella and Xiao Yuan. "Measurement sharpness cuts nonlocality and contextuality in every physical theory". In: *arXiv preprint arXiv:1404.3348* (2014) (cit. on p. 2).

[CY89] James P. Crutchfield and Karl Young. "Inferring statistical complexity". In: *Phys. Rev. Lett.* 63 (2 July 1989), pp. 105–108. DOI: 10.1103/PhysRevLett.63.105. URL: https://link.aps.org/doi/10.1103/PhysRevLett.63.105 (cit. on p. 1).

[DFW14] Frederic Dupuis, Omar Fawzi, and Stephanie Wehner. "Entanglement sampling and applications". In: *IEEE Transactions on Information Theory* 61.2 (2014), pp. 1093–1112 (cit. on p. 2).

[Dup+14] Frédéric Dupuis et al. "One-shot decoupling". In: *Communications in Mathematical Physics* 328 (2014), pp. 251–284 (cit. on p. 2).

[Ell+20] Thomas J. Elliott et al. "Extreme Dimensionality Reduction with Quantum Modeling". In: *Phys. Rev. Lett.* 125 (26 Dec. 2020), p. 260501. DOI: 10.1103/PhysRevLett.125.260501. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.260501 (cit. on p. 1).

[Ell21] Thomas J. Elliott. "Memory compression and thermal efficiency of quantum implementations of nondeterministic hidden Markov models". In: *Phys. Rev. A* 103 (5 May 2021), p. 052615. DOI: 10.1103/PhysRevA.103.052615. URL: https://link.aps.org/doi/10.1103/PhysRevA.103.052615 (cit. on pp. 1, 2).

[FE08] Christopher Ferrie and Joseph Emerson. "Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations". In: *Journal of Physics A: Mathematical and Theoretical* 41.35 (2008), p. 352001 (cit. on pp. 1, 3).

[FE09] Christopher Ferrie and Joseph Emerson. "Framed Hilbert space: hanging the quasi-probability pictures of quantum theory". In: *New Journal of Physics* 11.6 (2009), p. 063040 (cit. on pp. 1, 3).

[Fer11] Christopher Ferrie. "Quasi-probability representations of quantum theory with applications to quantum information science". In: *Reports on Progress in Physics* 74.11 (2011), p. 116001 (cit. on pp. 1, 3).

[FLW22] Marco Fanizza, Josep Lumbreras, and Andreas Winter. "Quantum theory in finite dimension cannot explain every general process with finite memory". In: *arXiv preprint arXiv:2209.11225* (2022) (cit. on p. 2).

[FST98] Shai Fine, Yoram Singer, and Naftali Tishby. "The hierarchical hidden Markov model: Analysis and applications". In: *Machine learning* 32 (1998), pp. 41–62. DOI: 10.1023/A:1007469218079. URL: https://doi.org/10.1023/A:1007469218079 (cit. on p. 1).

[Gal65] R Gallager. "A simple derivation of the coding theorem and some applications". In: *IEEE Transactions on Information Theory* 11.1 (1965), pp. 3–18 (cit. on p. 2).

[GJ95] Zoubin Ghahramani and Michael Jordan. "Factorial Hidden Markov Models". In: *Advances in Neural Information Processing Systems*. Ed. by D. Touretzky, M.C. Mozer, and M. Hasselmo. Vol. 8. MIT Press, 1995. URL: https://proceedings.neurips.cc/paper/1995/file/4588e674d3f0faf985047d4c3f13ed0d-Paper.pdf (cit. on p. 1).

[GM20] Andrew JP Garner and Markus P Mueller. "Characterization of the probabilistic models that can be embedded in quantum theory". In: *arXiv preprint arXiv:2004.06136* (2020) (cit. on p. 2).

[Gu+12] Mile Gu et al. "Quantum mechanics can reduce the complexity of classical models". In: *Nature Communications* 3.1 (2012). DOI: 10.1038/ncomms1761. URL: https://www.nature.com/articles/ncomms1761 (cit. on p. 1).

[HKS10] Robert Haslinger, Kristina Lisa Klinkner, and Cosma Rohilla Shalizi. "The Computational Structure of Spike Trains". In: *Neural Computation* 22.1 (Jan. 2010), pp. 121–157. ISSN: 0899-7667. DOI: 10.1162/neco.2009.12-07-678. URL: https://doi.org/10.1162/neco.2009.12-07-678 (cit. on p. 1).

[How+14] Mark Howard et al. "Contextuality supplies the 'magic' for quantum computation". In: *Nature* 510.7505 (2014), pp. 351–355. DOI: https://doi.org/10.1038/nature13460. URL: https://www.nature.com/articles/nature13460 (cit. on p. 3).

[Jel98] Frederick Jelinek. *Statistical methods for speech recognition*. MIT press, 1998 (cit. on p. 1).

[KJ22] Nikolaos Koukoulekidis and David Jennings. "Constraints on magic state protocols from the statistical mechanics of Wigner negativity". In: *npj Quantum Information* 8.1 (2022), p. 42. DOI: 10.1038/s41534-022-00551-1. URL: https://www.nature.com/articles/s41534-022-00551-1 (cit. on p. 3).

[KK21] Dagomir Kaszlikowski and Paweł Kurzyński. "A little bit of classical magic to achieve (super-) quantum speedup". In: *Foundations of Physics* 51.3 (2021), p. 55 (cit. on p. 3).

[Kou+22] Nikolaos Koukoulekidis et al. "Faster Born probability estimation via gate merging and frame optimisation". In: *Quantum* 6 (2022), p. 838 (cit. on p. 3).

[LA09a] Wolfgang Löhr and Nihat Ay. "Non-sufficient memories that are sufficient for prediction". In: *Complex Sciences: First International Conference, Complex 2009, Shanghai, China, February 23-25, 2009. Revised Papers, Part 1 1*. Springer. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 265–276. DOI: https://doi.org/10.1007/978-3-642-02466-5_25. URL: https://link.springer.com/chapter/10.1007/978-3-642-02466-5_25 (cit. on p. 2).

[LA09b] Wolfgang Löhr and Nihat Ay. "ON THE GENERATIVE NATURE OF PREDICTION". In: *Advances in Complex Systems* 12.02 (2009), pp. 169–194. DOI: 10.1142/S0219525909002143. URL: https://doi.org/10.1142/S0219525909002143 (cit. on p. 2).

[Lam18] Ludovico Lami. "Non-classical correlations in quantum mechanics and beyond". In: *arXiv preprint arXiv:1803.02902* (2018) (cit. on p. 2).

[LC19] Samuel P Loomis and James P Crutchfield. "Strong and weak optimizations in classical and quantum models of stochastic processes". In: *Journal of Statistical Physics* 176.6 (2019), pp. 1317–1342 (cit. on p. 1).

[Liu+19] Qing Liu et al. "Optimal stochastic modeling with unitary quantum dynamics". In: *Phys. Rev. A* 99 (6 June 2019), p. 062110. DOI: 10.1103/PhysRevA.99.062110. URL: https://link.aps.org/doi/10.1103/PhysRevA.99.062110 (cit. on p. 1).

[Löh12] Wolfgang Löhr. "Predictive models and generative complexity". In: *Journal of Systems Science and Complexity* 25 (2012), pp. 30–45. DOI: 10.1007/s11424-012-9173-x. URL: https://doi.org/10.1007/s11424-012-9173-x (cit. on p. 2).

[MAC16] John R Mahoney, Cina Aghamohammadi, and James P Crutchfield. "Occam's quantum strop: Synchronizing and compressing classical cryptic processes via a quantum channel". In: *Scientific reports* 6.1 (2016), p. 20495. DOI: 10.1038/srep20495. URL:

https://doi.org/10.1038/srep20495 (cit. on p. 1).

[MBW12] Alex Monras, Almut Beige, and Karoline Wiesner. *Hidden Quantum Markov Models and non-adaptive read-out of many-body states*. 2012. arXiv: 1002.2337 [quant-ph] (cit. on p. 1).

[ME12] Andrea Mari and Jens Eisert. "Positive Wigner functions render classical simulation of quantum computation efficient". In: *Physical review letters* 109.23 (2012), p. 230503 (cit. on p. 3).

[MF00] G. Manfredi and M. R. Feix. "Entropy and Wigner functions". In: *Phys. Rev. E* 62 (4 Oct. 2000), pp. 4665–4674. DOI: 10.1103/PhysRevE.62.4665. URL: https://link.aps.org/doi/10.1103/PhysRevE.62.4665 (cit. on p. 3).

[MF21] Kosuke Mitarai and Keisuke Fujii. "Overhead for simulating a non-local channel with local channels by quasiprobability sampling". In: *Quantum* 5 (2021), p. 388 (cit. on p. 3).

[Mül21] Markus Müller. "Probabilistic theories and reconstructions of quantum theory". In: *SciPost Physics Lecture Notes* (2021), p. 028 (cit. on p. 2).

[MW16] Alex Monras and Andreas Winter. "Quantum learning of classical stochastic processes: The completely positive realization problem". In: *Journal of Mathematical Physics* 57.1 (2016), p. 015219 (cit. on p. 1).

[OBC14] G Oas, J Acacio de Barros, and C Carvalhaes. "Exploring non-signalling polytopes with negative probability". In: *Physica Scripta* 2014.T163 (2014), p. 014034 (cit. on p. 3).

[OKK22] Kelvin Onggadinata, Pawel Kurzynski, and Dagomir Kaszlikowski. "Qubit from the classical collision entropy". In: *arXiv preprint arXiv:2205.00773* (2022) (cit. on p. 3).

[Plá16] Martin Plávala. "All measurements in a probabilistic theory are compatible if and only if the state space is a simplex". In: *Physical Review A* 94.4 (2016), p. 042108 (cit. on p. 2).

[PV10] Yury Polyanskiy and Sergio Verdú. "Arimoto channel coding converse and Rényi divergence". In: *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE. 2010, pp. 1327–1333 (cit. on p. 2).

[PWB15] Hakop Pashayan, Joel J Wallman, and Stephen D Bartlett. "Estimating outcome probabilities of quantum circuits using quasiprobabilities". In: *Physical review letters* 115.7 (2015), p. 070501 (cit. on p. 3).

[Ren08] Renato Renner. "Security of quantum key distribution". In: *International Journal of Quantum Information* 6.01 (2008), pp. 1–127 (cit. on p. 2).

[Rie+16] Paul M. Riechers et al. "Minimized state complexity of quantum-encoded cryptic processes". In: *Phys. Rev. A* 93 (5 May 2016), p. 052317. DOI: 10.1103/PhysRevA.93.052317. URL: https://link.aps.org/doi/10.1103/PhysRevA.93.052317 (cit. on p. 1).

[RJ86] L. Rabiner and B. Juang. "An introduction to hidden Markov models". In: *IEEE ASSP Magazine* 3.1 (1986), pp. 4–16. DOI: 10.1109/MASSP.1986.1165342. URL: https://ieeexplore.ieee.org/abstract/document/1165342 (cit. on p. 1).

[RRC16] Saleh Rahimi-Keshari, Timothy C Ralph, and Carlton M Caves. "Sufficient conditions for efficient classical simulation of quantum optics". In: *Physical Review X* 6.2 (2016), p. 021039 (cit. on p. 3).

[Rue+18] Joshua B. Ruebeck et al. "Prediction and generation of binary Markov processes: Can a finite-state fox catch a Markov mouse?" In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 28.1 (2018), p. 013109. DOI: 10.1063/1.5003041. URL: https://doi.org/10.1063/1.5003041 (cit. on p. 2).

[SC01] Cosma Rohilla Shalizi and James P Crutchfield. "Computational mechanics: Pattern and prediction, structure and simplicity". In: *Journal of statistical physics* 104 (2001), pp. 817–879. DOI: 10.1023/A:1010388907793. URL: https://doi.org/10.1023/A:1010388907793 (cit. on p. 1).

[Sch+21] David Schmid et al. "Characterization of noncontextuality in the framework of generalized probabilistic theories". In: *PRX Quantum* 2.1 (2021), p. 010331 (cit. on p. 2).

[Sha21] Farid Shahandeh. "Contextuality of general probabilistic theories". In: *PRX Quantum* 2.1 (2021), p. 010330 (cit. on p. 2).

[Tak+22] Ryuji Takagi et al. "Fundamental limits of quantum error mitigation". In: *npj Quantum Information* 8.1 (2022), p. 114 (cit. on p. 3).

[Tak21] Ryuji Takagi. “Optimal resource cost for error mitigation”. In: *Physical Review Research* 3.3 (2021), p. 033178 (cit. on p. 3).

[TBG17] Kristan Temme, Sergey Bravyi, and Jay M Gambetta. “Error mitigation for short-depth quantum circuits”. In: *Physical review letters* 119.18 (2017), p. 180509 (cit. on p. 3).

[Vei+12] Victor Veitch et al. “Negative quasi-probability as a resource for quantum computation”. In: *New Journal of Physics* 14.11 (2012), p. 113011. DOI: 10.1088/1367-2630/14/11/113011. URL: https://iopscience.iop.org/article/10.1088/1367-2630/14/11/113011 (cit. on p. 3).

[Vei+14] Victor Veitch et al. “The resource theory of stabilizer quantum computation”. In: *New Journal of Physics* 16.1 (2014), p. 013009 (cit. on p. 3).

[Vid11] Mathukumalli Vidyasagar. “The complete realization problem for hidden Markov models: a survey and some new results”. In: *Mathematics of Control, Signals, and Systems* 23.1-3 (2011), pp. 1–65 (cit. on p. 1).

[Vid14] M. Vidyasagar. *Theory and Applications to Biology*. Princeton: Princeton University Press, 2014. ISBN: 9781400850518. DOI: doi:10.1515/9781400850518. URL: https://doi.org/10.1515/9781400850518 (cit. on pp. 1, 2).

[WC08] Karoline Wiesner and James P Crutchfield. “Computation in finitary stochastic and quantum processes”. In: *Physica D: Nonlinear Phenomena* 237.9 (2008), pp. 1173–1195 (cit. on p. 1).

[WS07] William K Wootters and Daniel M Sussman. “Discrete phase space and minimum-uncertainty states”. In: *arXiv preprint arXiv:0704.1277* (2007) (cit. on p. 3).

[Yan+08] Jae-Suk Yang et al. “Increasing market efficiency in the stock markets”. In: *The European Physical Journal B* 61 (2008), pp. 241–246 (cit. on p. 1).

[Yan+18] Chengran Yang et al. “Matrix Product States for Quantum Stochastic Modeling”. In: *Phys. Rev. Lett.* 121 (26 Dec. 2018), p. 260602. DOI: 10.1103/PhysRevLett.121.260602. URL: https://link.aps.org/doi/10.1103/PhysRevLett.121.260602 (cit. on p. 1).

[Yua+21] Xiao Yuan et al. “Universal and operational benchmarking of quantum memories”. In: *npj Quantum Information* 7.1 (2021), p. 108 (cit. on p. 3).

Improved finite-size local randomness from CHSH games with simple constraints

Chun-Yu Chen[1 2 *] Kai-Siang Chen[2] Kai-Min Chung[1] Min-Hsiu Hsieh[3]
Yeong-Cherng Liang[2 4] Gelo Noel M. Tabia[2 4 †]

[1] *Institute of Information Science, Academia Sinica, Taipei 115, Taiwan*
[2] *Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan*
[3] *Hon Hai (Foxconn) Research Institute, Taipei, Taiwan*
[4] *Physics Division, National Center for Theoretical Sciences, Taipei 106, Taiwan*

Abstract. Device-independent randomness generation uses the violation of a Bell inequality to verify that the outputs of a nonlocal game are truly random. We focus on "local" randomness expansion protocols, where the extracted bits are random even to one of the involved parties. By incorporating zero-probability constraints into the Clauser-Horne-Shimony-Holt (CHSH) nonlocal game, we enhance the extractable rate in both asymptotic and finite-size regimes. Using the generalized entropy accumulation theorem and refining the second-order correction terms, we achieve a rate of up to 0.9 bit-per-round in our modified protocols, surpassing the standard CHSH game's maximum of 0.55 bits. Our results demonstrate some tolerance even without strictly enforcing the zero-probability constraints.

Keywords: Device-independent, local randomness, finite-size analysis, entropy accumulation, nonlocal game

In cryptography, randomness is an indispensable resource. Classically, there are no perfect classical sources of randomness. At best, one can exploit physical processes such as thermal or radio noise to produce unpredictable but partially biased and correlated bits. However, quantum physics allows us to obtain truly random outcomes from measurements on quantum systems [1], [2]. For cryptographic applications, the honest participants require perfect and secret randomness, which means a sequence of bits that is uniformly distributed from the perspective of some adversary who might hold some side information. So even with quantum devices, we need to make sure they are functioning properly, otherwise, an adversary could still exploit the noise in the implementation to gain some knowledge about the random bits.

To circumvent this possibility, we can adopt a device-independent (DI) approach based on nonlocal games such as the Clauser-Horne-Shimony-Holt (CHSH) game [3]. In a two-party nonlocal game, two cooperating but non-communicating players, Alice and Bob, receive inputs from a referee, and they each reply with an output to the referee. The referee decides if the players win the game or not by checking whether the combination of the inputs and outputs satisfies a winning condition. The winning condition is known by both players beforehand so they can choose a strategy that can achieve a maximal winning probability. A typical nonlocal game is designed in such a way that there is a gap between classical strategy, a strategy that can be described by the local hidden variable theory [4], and quantum strategy, a strategy that can be realized by quantum states and measurements.

Device-independent (DI) cryptography is based on the observation that for certain nonlocal games, players that employ a quantum strategy can achieve a score that is better than what can be reached by any classical strategy. Often we consider those nonlocal games with a property known as self-testing or rigidity [5]. If the players Alice and Bob achieve a winning probability close to the optimal quantum value, then the self-testing [6] property guarantees that the underlying state and measurements must be close to the optimal quantum strategy for that game, up to some inessential local isometries.

A nontrivial implication of the observed super-classical score is that genuine randomness can be extracted from the outputs of the nonlocal game, which lead to DI applications like quantum key distribution [7], [8] and random number generation [7], [9]–[11]. In the aforementioned examples, we usually consider "global" or non-blind randomness, where the unpredictability is from the perspective of some external third party to the nonlocal game. In contrast, as proposed by Miller and Shi [12], we may instead consider "local" or blind randomness, where the unpredictability is from the perspective of other players in the game. In [13], near-optimal bounds of local randomness for the CHSH game have been provided. Note that local randomness can also be thought of as the standard randomness with a higher level of security. Therefore, the local randomness extracted in this scenario can be used in any cryptographic task that required standard randomness.

Because the players in a nonlocal game cannot communicate, we expect Alice's (Bob's) statistics for output $a(b)$ to be independent of Bob's (Alice's) inputs $y(x)$. This means that the correlations $\vec{P}(a,b|x,y)$ must satisfy no-signaling (NS) constraints, i.e., $\sum_b P(a,b|x,y) = P(a|x)$ for all y and $\sum_a P(a,b|x,y) = P(b|y)$ for all x. Geometrically, the set $\mathcal{NS}$ of NS correlations for a fixed number of inputs and outputs forms a convex polytope. Since the only other constraint on any $\vec{P} \in \mathcal{NS}$ is that the conditional probabilities are nonnegative, then any $\vec{P}$ that belongs to the no-signaling boundary (NSB) should have some $P(a,b|x,y) = 0$.

Recently, Chen *et al.* [14] demonstrated that there are several classes of quantum correlations that lie on the NSB, each corresponding to violations of the CHSH inequality subjected to some extra zero-probability constraints. These correlations

[*] cchunyok@gmail.com
[†] gelonoel-tabia@gs.ncku.edu.tw

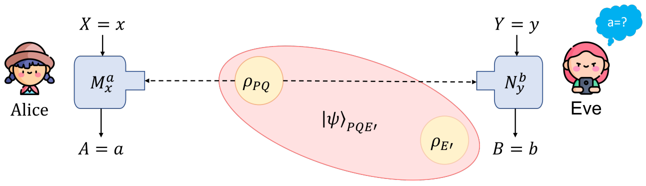


Figure 1: The scenario for a nonlocal game and Eve's strategy for guessing Alice's output.

can be classified according to the number and the relative positions of the vanishing probabilities [14, Table IV]. All these classes were proven to exhibit self-testing properties. In this work, we show that these NSB quantum correlations can provide more local randomness for a given winning probability in the CHSH game.

Let us describe the scenario involved in a protocol for local randomness. Here the adversary Eve is one of the players in the nonlocal game. Consider a CHSH game where Eve tries to guess Alice's outputs as shown in Fig 1. One of Eve's viable strategies can be described as follows (Fig. 1). Suppose in every round, Alice and Eve share a bipartite state $\rho_{P_iQ_i}$, where P_i and Q_i denote Alice and Eve's respective quantum systems in i-th round. We allow Eve to prepare this state so she can also create a purification of it, i.e., $|\psi\rangle_{P_iQ_iE'_i}$. Then, she sends system P_i to Alice and keeps systems Q_i and E'_i. Afterward, Alice and Eve perform the measurements for the nonlocal game according to their respective inputs $X_i = x, Y_i = y$ for the i-th round. Additionally, Alice generates an extra output $C_i = w(X_i, Y_i, A_i, B_i)$ to record the results of the nonlocal game in the i-th round, i.e., $C_i = 1(C_i = 0)$ for a win (loss). After n rounds, Alice can check the statistics of $\{C_i\}_i$ to determine whether Eve is honest or not. More precisely, Alice can set a better-than-classical threshold w_{th} that can be achieved with an honest implementation of the chosen nonlocal game. If the statistics of $\{C_i\}_i$ give a value that is lower than this threshold w_{th}, she aborts the game and discards all the outputs $\{A_i\}_i$. At the end of the protocol, Eve's quantum side information is $E'^n = E'_1E'_2...E'_n$. If we denote the classical information: $A^n = A_1A_2...A_n$, and similar for B^n, X^n, Y^n. The exact amount of local randomness that can be extracted from the outputs is given by the n-round conditional smooth min-entropy, i.e., $H^{\varepsilon}_{min}(A^n|B^n, X^n = x^n, Y^n = y^n, E'^n)$.

In Fu and Miller's work [13], they computed the local randomness by Eve's guessing probability of Alice's output using a second measurement that depends on Alice's input. From this guessing probability, we can compute the min-entropy—a relevant quantity for estimating the amount of extractable randomness in the one-shot setting, a setting that considers the worse case in a single shot. However, the one-shot setting can only promise the randomness we can extract from a source follows a certain probability distribution, which is still fairly unrealistic since with finite data, we can only try to estimate the probability distribution (and thus its associated winning probability) using the relative frequencies, but this is a good estimate only in the asymptotic limit. To this end, note that the entropy accumulation theorem (EAT) [15] and its generalized version (GEAT) [16] are powerful frameworks for bounding the conditional smooth min-entropy, a quantity that also counts the deviation of the estimated probability distribution, for a sequential quantum cryptographic protocol [17]–[19].

The original EAT applies to multi-round protocols with a model of side information consisting of a static quantum variable and a sequence of independent classical variables that satisfy Markov conditions [15, Chapter 4]. On the other hand, GEAT can be used in protocols where the full (quantum) side information is updated in every round, so long as no new side information of past outputs is created [16, Chapter 4]. Unlike in QKD protocols where either version is applicable due to the equivalence between prepare-and-measure and entanglement-based schemes, in local randomness expansion, the adversary acquires new quantum side information in each round as a player in the nonlocal game so the use of GEAT is imperative. Indeed, this is one of the examples considered in [16], however, they only presented a figure of single-round conditional von Neumann entropy bounds for the CHSH game.

In this work, we employ GEAT to compute the amount of extractable local randomness incorporating finite-size effects for the various classes of NSB correlations described in [14]. To apply GEAT, the single-round conditional von Neumann entropy is required. For this we follow Brown-Fawzi-Fawzi method [20], which uses an integral approximation of the logarithm combined with the Navascués-Pironio-Acín (NPA) hierarchy [21], [22] to obtain an entropy lower bound with semidefinite programming (SDP).[1] Finally, to obtain the finite-size extractable rates, we adopt techniques from [23], [24] to build the min-tradeoff function, which is essential for obtaining the second-order correction. The min-tradeoff function depends on the testing probability γ and some parameter $f_\perp$. For our construction, we set the testing probability at $\gamma = 10^{-2}$. Then for a fixed number of rounds N, we scan through the range of $f_\perp$, and the order α of the Renyi entropy to find the values that lead to the highest rates.

Our results indicate that by adding some simple constraints, namely zero-probability constraints (or even relaxing this by allowing small values), the extractable local randomness increases in some range of lower winning probabilities. In the photonic DI randomness expansion experiment by Liu *et al.* [23], they demonstrated that when the winning probability is around 0.7525, then the number of rounds required for achieving a nonzero extractable rate of non-blind randomness with the CHSH game is at least 8.951×10^{10}. In contrast, if we instead consider the local randomness from the CHSH game with some simple additional constraints, then as shown in Figure 2, we can achieve a nonzero rate with far fewer rounds, i.e., $N \approx 7.6 \times 10^8$. Furthermore, we can also relax the zero constraints to inequality constraints that tolerated a small value, e.g., we take $\delta_{\text{zero}} = 10^{-3}$, which has been shown to be attainable in practice [25]. With this small tolerance, we can still get a nonzero rate if we have 6.7×10^9 rounds. This illustrates that our results exhibit some robustness to implementation errors, making it quite experimentally feasible.

[1] We use Peter Wittek's python package, `ncpol2sdpa`, to construct NPA hierarchy. The latest version of the package is currently maintained by Peter Brown and can be found in https://github.com/peterjbrown519/ncpol2sdpa.

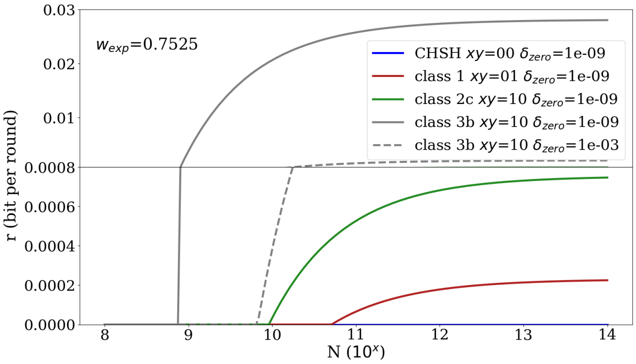


Figure 2: Finite local randomness per round for the standard CHSH, class 1, class 2c, and class 3b (The numbers of the classes indicate the number of the zero-probability constraints to be imposed in the protocol). The dashed line shows class 3b where we tolerate a comparably detectable value $\delta_{\text{zero}} = 10^{-3}$ for the zeros.

Moreover, if we consider the highest achievable winning probabilities for each class, we see a substantial improvement in the amount of local randomness compared to the standard CHSH. For instance, if we look at the asymptotic rates, for CHSH at the maximum winning probability of 0.8535, we can extract about 0.5523 random bits per round. But when we consider class 3b, which has a maximum winning probability of 0.7837, we can get a much higher rate of 0.8975 random bits per round. This is nontrivial and even potentially useful.

In general, we observe that the amount of extractable local randomness is more when the correlation has zeros. These can be roughly understood as follows: When adding extra constraints to an optimization problem, we are taking a subset of the feasible solutions that satisfy the original constraints. And for a minimization problem, the optimal value from the feasible set is always a valid lower bound on the optimal value from its subset. Moreover, in the case of the highest winning probability for each class, the self-testing property restricts Eve's attacks in the sense that they cannot deviate too much from a particular quantum strategy in the nonlocal game.

Furthermore, while classes with more zeros typically produce more local randomness (higher rates) for the same winning probability, we observe that the best possible rate does not necessarily trend with the number of zeros. This suggests that the amount of local randomness also depends on the structure of the correlation, or geometrically to the particular boundary it belongs to.

Because we have observed an improvement in the local randomness rate of CHSH by adding simple constraints, this suggests a potentially similar gain in the standard DI (global) randomness. In addition, it may be worth noting that when we consider the CHSH game with two zero-probability constraints, one possibility leads to an optimal quantum strategy where the parties have one identical measurement setting. This makes it a natural candidate for a DI quantum key distribution (QKD) protocol based on the CHSH game that does not require additional input. And maybe even in this case, we could observe an improved finite key rate over the standard CHSH game.

We may also consider how the local randomness of these NSB quantum correlations can be applied to cryptographic tasks with mistrustful parties. For example, in the DI protocol for certified deletion based on Mermin-Peres magic square [13], [26], if the deviation from the maximum score is δ then the secrecy guarantee scales like $9\sqrt{\delta}$. There might be some advantages with a protocol based on the CHSH game or any of its variants with zero constraints.

References

[1] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, "Quantum random number generation," *npj Quantum Information*, vol. 2, no. 1, p. 16 021, 2016.

[2] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," *Rev. Mod. Phys.*, vol. 89, p. 015 004, 1 2017.

[3] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, "Consequences and limits of nonlocal strategies," in *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.* IEEE, 2004, pp. 236–249.

[4] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, vol. 1, no. 3, pp. 195–200, 1964.

[5] D. Mayers and A. Yao, "Quantum cryptography with imperfect apparatus," in *Proceedings 39th Annual Symposium on Foundations of Computer Science (Cat. No.98CB36280)*, IEEE, 1998, pp. 503–509.

[6] I. Šupić and J. Bowles, "Self-testing of quantum systems: A review," *Quantum*, vol. 4, p. 337, 2020.

[7] 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.*, vol. 98, p. 230 501, 23 2007.

[8] D. P. Nadlinger, P. Drmota, B. C. Nichol, *et al.*, "Experimental quantum key distribution certified by bell's theorem," *Nature*, vol. 607, no. 7920, pp. 682–686, 2022.

[9] R. Colbeck, "Quantum And Relativistic Protocols For Secure Multi-Party Computation," PhD dissertation, Univ. Cambridge, 2006. eprint: 0911.3814.

[10] U. Vazirani and T. Vidick, "Certifiable quantum dice: or, true random number generation secure against quantum adversaries," in *STOC '12: Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, New York, NY, USA: Association for Computing Machinery, 2012, pp. 61–76.

[11] C. A. Miller and Y. Shi, "Robust Protocols for Securely Expanding Randomness and Distributing Keys Using Untrusted Quantum Devices," *J ACM*, vol. 63, no. 4, pp. 1–63, 2016.

[12] C. A. Miller and Y. Shi, "Randomness in nonlocal games between mistrustful players," *Quantum Inf Comput*, vol. 17, no. 7, p. 595, 2017.

[13] H. Fu and C. A. Miller, "Local randomness: Examples and application," *Phys. Rev. A*, vol. 97, p. 032 324, 3 2018.

[14] K.-S. Chen, G. N. M. Tabia, C. Jebarathinam, S. Mal, J.-Y. Wu, and Y.-C. Liang, *Quantum correlations on the no-signaling boundary: Self-testing and more*, 2022. arXiv: 2207.13850v2 [quant-ph].

[15] F. Dupuis, O. Fawzi, and R. Renner, "Entropy Accumulation," *Commun Math Phys*, vol. 379, no. 3, pp. 867–913, 2020.

[16] T. Metger, O. Fawzi, D. Sutter, and R. Renner, "Generalised entropy accumulation," in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, IEEE, 2022, pp. 844–850.

[17] R. Arnon-Friedman, F. Dupuis, O. Fawzi, R. Renner, and T. Vidick, "Practical device-independent quantum cryptography via entropy accumulation," *Nat Commun*, vol. 9, no. 459, pp. 1–11, 2018.

[18] P. J. Brown, S. Ragy, and R. Colbeck, "A Framework for Quantum-Secure Device-Independent Randomness Expansion," *IEEE Trans Inf Theory*, vol. 66, no. 5, pp. 2964–2987, 2019.

[19] T. Metger and R. Renner, "Security of quantum key distribution from generalised entropy accumulation," 2022. arXiv: 2203.04993 [quant-ph].

[20] P. Brown, H. Fawzi, and O. Fawzi, "Device-independent lower bounds on the conditional von Neumann entropy," 2021. arXiv: 2106.13692v2 [quant-ph].

[21] M. Navascués, S. Pironio, and A. Acín, "Bounding the Set of Quantum Correlations," *Phys Rev Lett*, vol. 98, no. 1, p. 010 401, 2007.

[22] M. Navascués, S. Pironio, and A. Acín, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations," *New J Phys*, vol. 10, no. 7, p. 073 013, 2008.

[23] W.-Z. Liu, M.-H. Li, S. Ragy, *et al.*, "Device-independent randomness expansion against quantum side information," *Nat Phys*, vol. 17, no. 4, pp. 448–451, 2021.

[24] E. Y.-Z. Tan, P. Sekatski, J.-D. Bancal, *et al.*, "Improved DIQKD protocols with finite-size analysis," *Quantum*, vol. 6, p. 880, 2022. eprint: 2012.08714v4.

[25] R. Ramanathan, M. Horodecki, H. Anwer, *et al.*, "Practical No-Signalling proof Randomness Amplification using Hardy paradoxes and its experimental implementation," 2018. arXiv: 1810.11648v3 [quant-ph].

[26] S. Kundu and E. Y.-Z. Tan, *Composably secure device-independent encryption with certified deletion*, 2022. arXiv: 2011.12704 [quant-ph].

Dynamics-based Witnesses of Nonclassicality and Entanglement

Lin Htoo Zaw[1] * Pooja Jayachandran[1] Valerio Scarani[1] [2]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[2] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*

Abstract. We recently introduced a series of protocols [1, 2, 3] that detect the nonclassicality or entanglement of suitable states of quantum systems, under the assumption that the measured dynamical observable undergoes a known time evolution. These works are based on an unexpected observation by Tsirelson about the quantum harmonic oscillator [4]: despite its time evolution being the same as its classical counterpart—a precession in phase space—its nonclassicality can be detected by probing its position at different times. Our protocols do not rely on sequential or simultaneous measurements, as only one randomly-chosen measurement is performed in each round. They are also more akin to Bell inequalities than to uncertainty relations: in particular, they do not admit false positives from classical theory. While the main focus will be for the case where the dynamics is that of a uniform precession, the extension of the protocol to systems with general bound dynamics will also be briefly covered.

Keywords: harmonic oscillator, quantum dynamics, negativity witness, entanglement witness

A common task in quantum mechanical experiments is to demonstrate that the controlled system is actually doing something quantum. This might be done by demonstrating some quantum feature of a single system, or by certifying entanglement over multiple systems.

Amongst the different types of controlled quantum systems, the harmonic oscillator is a mainstay: it appears in superconducting circuits, trapped ions, photonics, and optomechanics. Commonly reported quantum features of the harmonic oscillator tend to probe its discrete nature, like the nonzero ground state energy, or probe the incompatibility of observables, like with sequential measurements of position and momentum.

Meanwhile, the time evolution of the harmonic oscillator is the same in both classical and quantum theory. The classical observables $(x(t), p(t))$, and the corresponding quantum observables $(X(t), P(t))$ given in the Heisenberg representation, all satisfy

$$\begin{aligned} x(t) &= \cos(\omega t)x(0) + \sin(\omega t)p(0) \\ p(t) &= \cos(\omega t)p(0) - \sin(\omega t)x(0), \end{aligned} \tag{1}$$

with $(x,p) \to (X,P)$ in the quantum case. As such, one does not expect to find quantumness in the dynamics of the harmonic oscillator.

Surprisingly, an overlooked preprint by Tsirelson [4] showed that this was not the case. By simply asking "*How often is the coordinate of a harmonic oscillator positive?*", suitable quantum states can violate a classical bound set by the classical harmonic oscillator.

Our work. For single systems, we have expanded Tsirelson's original protocol to a family of protocols, and extended it to the case of spin angular momentum [1] and general bound dynamics [3]. For the multipartite scenario, we have shown that the protocol is an entanglement witness when applied to a normal mode of multiple harmonic oscillators [2].

A few key properties of our criteria are highlighted in the next page. Meanwhile, we summarise the protocols for harmonic oscillators here.

Protocol to Certify Nonclassicality of a Harmonic Oscillator. The protocol is performed for an odd $K > 1$, and consists of many independent rounds. In each round:

1. The system is prepared in some state. This state may not be identical in each round, and the interactions required to prepare the state do not need to be specified.
2. After the preparation is completed, the system is decoupled from everything else, and *undergoes the closed dynamics of a harmonic oscillator*, resulting in a uniform precession with period $T = 2\pi/\omega$.
3. A duration $t_k \in \{kT/K\}_{k=0}^{K-1}$ is randomly chosen. The system is then left to precess for a time t_k.
4. The position $x(t_k)$ is measured at the chosen time. Since the round ends here, the measurements can be destructive.

After many rounds, the average score

$$P_K := \frac{1}{K}\sum_{k=0}^{K-1}\left\{\Pr\left[x(t_k) > 0\right] + \frac{1}{2}\Pr\left[x(t_k) = 0\right]\right\}$$

is calculated. If $P_K > \mathbf{P}_K^c := \frac{1}{2}(1 + \frac{1}{K})$, then the system is quantum.

*zaw@l-lin.com

Protocol to Witness Entanglement of Two Harmonic Oscillators. The nonclassicality protocol is performed on the *centre-of-mass motion* of two coupled oscillators, which precesses uniformly as an effective oscillator. The centre-of-mass position is given by

$$x_\theta = \left(\frac{m_1}{m_2}\right)^{1/4} \cos(\theta) x_1 + \left(\frac{m_2}{m_1}\right)^{1/4} \sin(\theta) x_2.$$

Here m_j and x_j are respectively the mass and position of the jth oscillator. Meanwhile, θ is determined by the strength of the coupling between the two oscillators. There are two cases:

- If $\theta = \pi/4$, $P_K > \mathbf{P}_K^c$ implies that the two oscillators are entangled.
- If $\theta \neq \pi/4$, $P_K > \mathbf{P}_K^{\text{SEP}}(\theta)$ implies that the two oscillators are entangled, where $\mathbf{P}_K^{\text{SEP}}(\theta)$ can be found using semidefinite programming under some energy constraints [5].

Assumption of dynamics. The primary assumption of our criteria is that, while the protocol is being performed, the time evolution of the measured observables are known. As such, our criteria are especially useful when the dynamics themselves are not in question, but there are doubts about whether the system is actually behaving "quantumly". For example, this might be the case when one can be certain that a mesoscopic particle is trapped in a harmonic potential, but where full tomography is not an option due to the size of the system. Another example can be found in optical setups, where position measurements at different times, together with the assumption of harmonic dynamics, is equivalent to different settings of a quadrature measurement.

No simultaneous or sequential measurements. In each round of our protocol, only a single coordinate measurement is made, upon which the round ends. This is in contrast to usual tests of contextuality [6] and Leggett-Garg-type criteria [7], which require the simultaneous or sequential measurement of two or more observables in each round. In particular, our criteria do not utilise the noninvasive measurements required in similar Leggett-Garg tests of harmonic oscillators [8, 9], which avoids the "clumsiness" loophole entirely [10].

No false positives from classical theory. Many commonly-used witnesses of continuous-variable entanglement are based on the uncertainty relation, where the systems are said to be entangled if the measured standard deviations of some observables are below a certain threshold [11, 12, 13, 14]. This requires measuring the positions and momenta of both systems with a precision set by $\hbar$. At the precision of, say, human perception, two oscillators at equilibrium are described by $x_j = p_j = 0 \implies \Delta x_j = \Delta p_j = 0$. This would imply entanglement if the above witnesses are used naïvely. On the contrary, in our protocol, nonclassicality or entanglement is certified by showing that some observed value exceeds a classical bound, the maximum possible value achieved by the corresponding classical system. As such, our criteria do not admit false positives from classical theory by construction.

Detection of non-Gaussian states. Since the classical bound also holds for states that can be classically simulated—in particular, Gaussian states—they will not be certified as nonclassical by our criteria. As a consequence, our criteria only detect the entanglement of some non-Gaussian states. This is notable as non-Gaussian states are required for a quantum advantage in many protocols [15, 16, 17, 18], but commonly-used entanglement witnesses can be ineffective at detecting them [19].

Detection of states not detected by other entanglement witnesses. We have also explicitly shown the existence of families of entangled states that are detected by our criteria but not by any of the existing ones. A particular example is the entangled three-level cat state given by $|\Psi\rangle \propto \sum_{k=-1}^{1} |\alpha e^{i2\pi k/3}\rangle_1 \otimes |\alpha e^{i2\pi k/3}\rangle_2$. This state cannot be detected by the criteria of [11, 12] for any value of α, is detected by the criteria of [13, 14] for $1.23 \lesssim |\alpha| \lesssim 1.82$, and is detected by our criteria for $0.88 \lesssim |\alpha| \lesssim 1.23$.

References

[1] L. H. Zaw, C. C. Aw, Z. Lasmar, V. Scarani. Detecting quantumness in uniform precessions. *Phys. Rev. A*, 106(3):032222, 2022.

[2] P. Jayachandran, L. H. Zaw, V. Scarani. Dynamics-Based Entanglement Witnesses for Non-Gaussian States of Harmonic Oscillators *Phys. Rev. Lett.*, 130(16):160201, 2023.

[3] L. H. Zaw, V. Scarani. Dynamics-based quantumness certification of continuous variables with generic time-independent Hamiltonians arXiv:2212.06017, 2022.

[4] B. Tsirelson. How often is the coordinate of a harmonic oscillator positive? arXiv:quant-ph/0611147, 2006.

[5] L. H. Zaw. Script used in: Dynamic-based entanglement witnesses for non-Gaussian states of harmonic oscillators. `https://github.com/not-fred/arXiv-2210.10357`.

[6] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, J. Larsson. Kochen-Specker contextuality. *Rev. Mod. Phys.*, 94(4):045007, 2022.

[7] A. J. Leggett, A. Garg. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.*, 54(9):857–860, 1985.

[8] S. Bose, D. Home, S. Mal. Nonclassicality of the Harmonic-Oscillator Coherent State Persisting up to the Macroscopic Domain. *Phys. Rev. Lett.*, 120(21):210402, 2018.

[9] C. Mawby, J. J. Halliwell. Leggett-Garg tests for macrorealism in the quantum harmonic oscillator and more general bound systems. *Phys. Rev. A*, 105(2):022221, 2022.

[10] M. M. Wilde, A. Mizel. Addressing the Clumsiness Loophole in a Leggett-Garg Test of Macrorealism. *Found. Phys.*, 42(2):256–265, 2012.

[11] L.-M. Duan, G. Giedke, J. I. Cirac, P. Zoller. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.*, 84(12):2722, 2000.

[12] C. J. Zhang, H. Nha, Y. S. Zhang, G.-C. Guo. Detection of bound entanglement in continuous-variable systems. *Phys. Rev. A*, 82(3):032323, 2010.

[13] M. Hillery, M. S. Zubairy. Entanglement conditions for two-mode states. *Phys. Rev. Lett.*, 96(5):050503, 2006.

[14] H. Nha, J. Kim. Entanglement criteria via the uncertainty relations in su(2) and su(1,1) algebras: Detection of non-Gaussian entangled states. *Phys. Rev. A*, 74(1):012317, 2006.

[15] A. Mari, J. Eisert. Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient. *Phys. Rev. Lett.*, 94(23):230503, 2012.

[16] J. Niset, J. Fiurášek, N. J. Cerf. No-Go Theorem for Gaussian Quantum Error Correction. *Phys. Rev. Lett.*, 102(12):120501, 2009.

[17] J. Fiurášek. Gaussian Transformations and Distillation of Entangled Gaussian States. *Phys. Rev. Lett.*, 89(13):137904, 2002.

[18] G. Giedke, J. Ignacio Cirac. Characterization of Gaussian operations and distillation of Gaussian states. *Phys. Rev. A*, 66(3):032316, 2002.

[19] C. Weedbrook, S. Pirandola, R. García-Patrón, N. J. Cerf, T. C. Ralph, J. H. Shapiro, S. Lloyd Gaussian quantum information. *Rev. Mod. Phys.*, 84(2):621–669, 2012.

Minimal Port-based Teleportation

based on arXiv:2111.05499

Sergii Strelchuk[1] * **Michał Studziński**[2] †

[1]*DAMTP, Centre for Mathematical Sciences, University of Cambridge, Cambridge CB30WA, UK*
[2] *Institute of Theoretical Physics and Astrophysics, Faculty of Mathematics, Physics, and Informatics, University of Gdańsk, Wita Stwosza 57, 80-308 Gdańsk, Poland*

Abstract. We introduce the minimal set of requirements that define a feasible PBT protocol and construct a simple PBT protocol that teleports an unknown state of a qudit with success probability exponentially better than previously known schemes with the resource state consisting of N maximally entangled states. We define the corresponding efficient superdense coding protocols which transmit more classical bits with fewer maximally entangled states. Furthermore, we introduce rigorous methods for comparing and converting between different PBT protocols.

Keywords: quantum teleportation , port-based teleportation , universal programmable quantum processor, quantum channels simulations

1 Motivation

Quantum teleportation introduced first in [1] is the concept of sending an unknown state of the quantum system without physically transferring the medium found its use in an impressive range of applications [2]. Despite of many applications proposed the teleportation scheme has some limitations and one of the main is the necessity of unitary correction in the last step of the transmission procedure. However, there is also another teleportation scheme, introduced by Hiroshima and Ishizaka [3, 4], with properties that were previously believed to be unattainable. These protocols go under the name of port-based teleportation (PBT) and they possess a counter-intuitive property that appears to be at odds with non-signalling principle of quantum mechanics, namely, that the teleported state requires no unitary correction and is readily available for use after the sender performs a measurement and sends classical communication.

All PBT protocols found wide-ranging applications in cryptography and instantaneous non-local computation [5], they were instrumental in establishing a link between interaction complexity and entanglement in non-local computation and holog- raphy [6], they established a fundamental link between quantum communication complexity advantage and a violation of a Bell inequality [7], fundamental limitations for quantum channels discrimination by designing adaptive protocols called PBT stretching [8] and others [9, 10, 11, 12] including even continuous variable realisation [13].

2 Configuration for PBT

In the vanilla PBT setup parties share a large resource state consisting of N copies of the maximally entangled states $|\Psi^+\rangle^{\otimes N}$, where each singlet is a two-qudit state, called *port*. Alice performs a joint measurement on an unknown state ψ_C together with her half of the resource state and communicates the outcome to Bob. The outcome of the measurement indicates the subsystem where the state has been teleported to. To obtain the teleported state, Bob discards all ports except for the one indicated by Alice's outcome. There are two versions of the PBT protocol, depending on the exact set of measurements used by Alice. The first type, so-called *deterministic* teleportation (dPBT), is described by the set of N POVM elements $\mathcal{X} = \{\Pi_a\}_{a=1}^N$ (in a form of *square-root measurements (SRM)*). Upon measuring a-th element the teleported state ends up in the a-th port on Bob's side. He then traces out all but a-th subsystem which contains the teleported state. The second type, *probabilistic* PBT (pPBT), consists of a measurement with $N+1$ POVM elements $\{\Pi_a\}_{a=0}^N$ (different than for dPBT), where Π_0 indicates a failure of the teleportation. In this protocol, when Alice obtains the input $a \in \{1,\dots,N\}$, the parties proceed as above. When she obtains 0, they abort the protocol.

Due to the no-programming theorem [14] it is impossible to obtain perfect transmission with a finite number of shared ports. However, because of so many breakthrough applications, it was very important to learn what is the efficiency [1] of all variants of PBT schemes as a function of the number of shared ports N and local dimension d, and it was done in papers [15, 16, 17].

In both versions of the PBT scheme (pPBT and dPBT) we also distinguish so-called *optimized protocol*, where Alice before sending an unknown state optimizes simultaneously over the resource state and measurements. This procedure leads to a square improvement in N of the efficiency of dPBT and pPBT. Notice that after optimisation parties do not share a maximally entangled resource state but perform better.

*ss870@cam.ac.uk

†michal.studzinski@ug.edu.pl

[1] Entanglement fidelity for dPBT when parties send a half of the maximally entangled state. The probability of success for pPBT averaged over all input states. Both quantities depend on N and local dimension d.

3 Our goals

Despite of the huge progress in the area of PBT in recent years the topic of PBT-theory still offers many important questions and problems very important from fundamental point of view. In particular, we can ask about the following:

1. What other forms of transferring quantum information from one subsystem to another exist?
2. Are there many other distinct protocols which result in quantum state transfer akin to the ordinary teleportation and PBT-like protocols?
3. Can one treat the latter as a single class of protocols or, perhaps, there is a number of fundamentally different protocols within PBT?
4. Are there any practical implementations of the above addressed points?

Our work aims to address the above by providing novel conceptual insights. What is more in our work we intensively exploit nontrivial methods coming from representation theory of algebras, numerical methods supported by the Sage package, and tools coming from SDP theory. We strongly believe that this makes our results interesting not only from the fundamental point of view but also from the point of view of applied toolkits.

4 Main Results

Result 1 We introduce a *new protocol* that cannot be reduced to any known protocols and that satisfies this set of minimal requirements. This protocol is *exponentially more efficient than any known pPBT*, hence it cannot be obtained from any such protocol. consider the following sequence of steps outlined in the box below.

PBT protocol $\mathcal{P}$

INPUT
$n \geq 0$, a shared $2n - qubit$ state $\rho_{AB}^{(n)}$ between sender A and receiver B; a state ψ_C to be teleported; a set measurements $\{\mathcal{M}_{A,i}\}_{i=1}^{k}$, where without loss of generality we have $k = n+1$. The instantiated protocol is denoted as $\mathcal{P}_n(\{\mathcal{M}_{A,i}\}_{i=1}^{k}, \rho_{AB}^{(n)}, \psi_C)$.
ALGORITHM

- Alice performs a measurement $\mathcal{M}_{A,i}$ on $\rho_{AB}^{(n)} \otimes \psi_C$, obtaining outcome $i \in [1,\dots,k]$.
- Alice sends the index i to Bob by classical channel;

OUTPUT
If $i \in [1,\dots,k]$, then the teleported state is $\rho_{B_i}^{(n)}$. Otherwise, return "FAIL".

With each $\mathcal{P}_n(\{\mathcal{M}_{A,i}\}_{i=1}^{k}, \rho_{AB}^{(n)}, \psi_C)$ we associate a two-parameter estimate $Q(\mathcal{P}_n(\{\mathcal{M}_{A,i}\}_{i=1}^{k}, \rho_{AB}^{(n)})) = (F(\sigma_A, \rho_{B_i}^{(n)}), p)$ that describes the performance of a teleported state. The first parameter characterises the quality of teleported state and the second – the success of the teleportation. In the next step we constructed explicitly new PBT-like protocols, called the minimal PBT satisfying the above requirements. What is more we show that our protocol exponentially outperforms in probability of success all previously known PBT protocols, even when we use only resource state composed of maximally entangled states - no optimisation is needed for better performance. The corresponding entanglement fidelity is also evaluated and it is no worse than for previously known PBT schemes. This is contained in Table 1 and visualized on Figure 1 for higher dimensions.

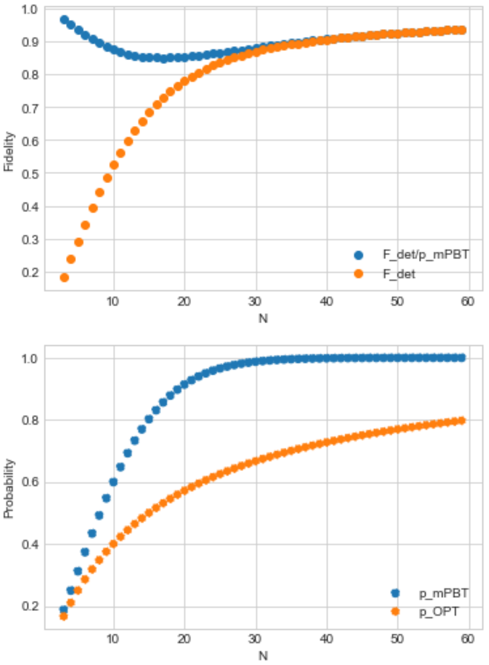


Figure 1: Fidelities (top) and probability of success (bottom) for $d = 4$. The mPBT scheme performed by SRM measurements (blue) outperforms the optimal qudit pPBT (orange), for which the probability of success is $p_{succ} = 1 - \frac{d^2-1}{N+d^2-1}$ [15]. When one considers resulting fielities, we see that the mPBT scheme (blue) is more efficient up to ~ 30 ports when one compares it with the non-optimal dPBT with SRM measurements (orange).

Result 2 We develop methods for *comparing port-based teleportation protocols* in terms of their resource states which enable one to estimate their entanglement per port and distinctness. Despite the fact that all PBT protocols share operational similarities, their underlying resource states are rather different. We discuss the properties of the resource states in all PBT protocols by evaluating their closeness. We show that by starting from maximally entangled states, the optimization operation applied by Alice has a large effect on their mutual distances. To quantify the distance, we will use the square-root fidelity $\sqrt{F}$. We show that in-

Teleportation protocol	Entanglement fidelity F	Average success probability p_{succ}
Non-optimised deterministic PBT	$F = 1 - \mathcal{O}(1/N)$	1
Optimised deterministic PBT	$F = 1 - \mathcal{O}(1/N^2)$	1
Non-optimised probabilistic PBT	1	$p_{succ} = 1 - \mathcal{O}(1/\sqrt{N})$
Optimised probabilistic PBT	1	$p_{succ} = 1 - \mathcal{O}(1/N)$
Minimal PBT (entangled resource)	$F = 1 - \mathcal{O}(1/N)$	$p_{succ} = 1 - \frac{N+2}{2^{N+1}}$

Table 1: Asymptotic behaviour of all known variants of PBT protocols compared with introduced mPBT – the qubit case. The mPBT offers exponentially better scaling in N compared to optimised pPBT for average probability success, even with the non-optimised resource state included in this table. For entanglement fidelity, mPBT offers the same scaling with the number of ports N.

creasing the number of shared ports $\sqrt{F}$ is decreasing, but exhibiting different asymptotic properties for different resource states. Mentioned behavior is depicted in Figure 2, however in the manuscript, we derive also analytical expressions for $\sqrt{F}$ in terms of group-theoretical quantities like multiplicities and dimensions in the Schur-Weyl duality.

Result 3 We further introduce methods for the *conversion between PBT protocols* and determine the conditions when it is possible to turn one type of PBT into another. In particular, we formulate an explicit procedure for converting dPBT into pPBT. We derive respective measurements and resulting efficiencies. We illustrate our findings by explicit construction of such conversion for known examples of PBT and we evaluate their efficiencies. We argue that conversion from pPBT to dPBT is not always possible - we formulate conditions for such conversion and we construct examples. Additionally, we also discuss how the newly introduced mPBT protocol lies in the landscape of all known PBT-like schemes, and how, and under what conditions it generates other schemes.

Result 4 We know that superdense coding protocols are dual to teleportation protocols. While ordinary superdense coding protocols are well-understood in the context of original teleportation, very little is known about their dual-PBT versions, with only one known example in [18]. We show how to take an arbitrary dPBT protocol with an established lower bound on fidelity and compute the corresponding performance of the superdense coding protocol. In particular, we find that there *exist superdense coding protocols* that are capable of transmitting the same amount of classical information as in [18], but using significantly less entanglement - we show that one can get square improvement in the number of shared ports N. The explicit construction of such schemes is also discussed. Similar results have been obtained recently and independently in [19].

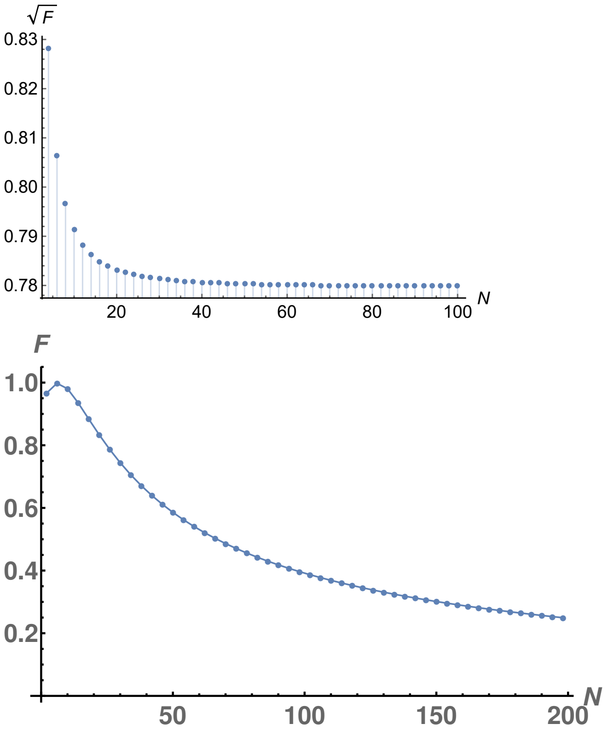


Figure 2: Top: Overlap between states for optimal dPBT and optimal pPBT calculated for qubits. We see that for these two states the overlap between them saturates on the value 0.778. Bottom: Overlap between states for non-optimal and optimal dPBT for qubits. The maximal value of the overlap which is $F = 0.9977$ is attained for $N = 6$. In the asymptotic limit, both states are orthogonal which was shown analytically. In both figures, we see completely different behavior of the overlaps between the resource states.

References

[1] Ch. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Physical Review Letters, 70(13):1895–1899 (1993).

[2] S. Pirandola, J. Eisert, Ch. Weedbrook, A. Furusawa, and S. L. Braunstein, Nature photonics, 9(10):641–652 (2015).

[3] S. Ishizaka and T. Hiroshima, Physical review letters, 101(24):240501 (2008).

[4] S. Ishizaka and T. Hiroshima, Physical Review A, 79(4):042306 (2009).

[5] S. Beigi and R. König, New Journal of Physics, 13(9):093036 (2011).

[6] A. May, Quantum, 6:864 (2022).

[7] H. Buhrman, L. Czekaj, A. Grudka, M. Horodecki, P. Horodecki, M. Markiewicz, F. Speel- man, and S. Strelchuk, Proceedings of the National Academy of Sciences, 113(12):3191–3196 (2016).

[8] S. Pirandola, R. Laurenza, C. Lupo, and J. L. Pereira, npj Quantum Information, 5(1):50 (2019).

[9] J. Pereira, L. Banchi, and S. Pirandola, Journal of Physics A: Mathematical and Theoretical, 54(20):205301 (2021).

[10] M. Sedlák, A. Bisio, and M. Ziman, Phys. Rev. Lett., 122:170502 (2019)

[11] M. T. Quintino, Quantum Views, 5:56 (2021).

[12] M. T. Quintino, D. Ebler, Quantum 6, 679 (2022)

[13] J. L. Pereira, L. Banchi, S. Pirandola, arxiv.org/abs/2302.08522

[14] M. A. Nielsen, and I. L. Chuang, Phys. Rev. Lett. 79(2), 321–324 (1997)

[15] M. Studziński, S. Strelchuk, M. Mozrzymas, and M. Horodecki, Scientific Reports, 7:10871 (2017).

[16] M. Mozrzymas, M. Studziński, S. Strelchuk, and M. Horodecki, New Journal of Physics, 20(5):053006 (2018)

[17] M. Christandl, F. Leditzky, Ch. Majenz, G. Smith, F. Speelman, and M. Walter, Communications in Mathematical Physics, 381:379–451 (2021)

[18] S. Ishizaka, arXiv:1506.01555

[19] E. Chitambar, and F. Leditzky, arxiv.org/abs/2302.14798

Quantum Compiling with Sparse Regularization

Taisei Nohara[1] * Itsuki Noda[2] † Satoshi Oyama[3] ‡

[1] *Graduate School of Information Science and Technology, Hokkaido University, Japan*
[2] *Faculty of Information Science and Technology, Hokkaido University, Japan*
[3] *School of Data Science, Nagoya City University, Japan*

Abstract. Quantum compiling is the task of translating an quantum operation into an executable quantum circuit. The circuit should be constructed using as few gates as possible to maximize arithmetic accuracy. In the use of current noisy intermediate-scale quantum devices, quantum compiling is one of the most important tasks because the effect of errors in each operation is severe. We investigated the sparse regularization methods suitable used in the field of mathmatical optimization as a tool to achieve quantum compiling, as we speculate that approximate circuits consisting of a small number of gates can be created by optimizing parameterized quantum circuits with loss functions based on classical sparse regularization. Implementation and testing of circuits that reduce the number of gates by using regularization demonstrated that methods using sparse regularization produce a sparse parameter set in parametrized quantum circuits.

Keywords: Quantum Compiling, Sparse Regularization, Quantum Machine Learning

1 Introduction

Quantum compiling is the task of translating quantum operations, where inputs and outputs are given as quantum states, into a quantum circuit containing feasible gates[1, 2]. In this task, it is necessary to output a quantum state that is as close as possible to the result of an appropriate unitary operation on the input quantum state. Quantum compiling is an important task in the execution of quantum circuits in noisy intermediate-scale quantum(NISQ) devices[3]. As quantum compiling requires outputs to be obtained by unitary operations, the optimization of gate placement is challenging.

As a recent approach to quantum compiling, the use of dynamic circuit models based on parameterized quantum circuits has attracted much research attention[4, 5, 6]. Parameterized quantum circuits are quantum circuit models that are optimized by iteratively updating the quantum circuit parameters[7]. The natural idea of constructing a quantum circuit based on the final parameters of each gate makes it easy to handle the compiling. In addition, since these parameters are classical numerical values and can thus be easily simulated on classical computers, applied research using simulators and actual equipment has been widely conducted[2].

Recently, Madden and Simonetto proposed using a classical sparse regularization method called group Lasso[8], which utilizes a sparse regularization method on a group basis consisting of multiple variables, to achieve quantum compiling[9]. Sparse regularization is used to optimize the value of a function $f(\boldsymbol{x}) \in \mathbb{R}$ so that many elements of multi-dimensional variable vector $\boldsymbol{x} \in \mathbb{R}^d$ become zero; i.e., $\boldsymbol{x}$ is sparse[10]. Application of sparse regularization produces a sparse solution in which only a few variables are non-zero, so it should enable extraction of information on the variables that truly contribute to the optimization of $f(\boldsymbol{x})$. General rotation gates (e.g., RX, RZ, CP) with a rotation angle of zero are equivalent to an identity gate, and such gates for which the parameters are zero can be 'removed' in the quantum compiling problem setting. In other words, sparse gate configurations should result from the application of sparse regularization. Therefore we focused on the difference in performance for each sparse regularization method as a tool to achieve quantum compiling and thereby obtain sparse circuits. In this study, we investigated the effect of applying classical regularization methods to sparse quantum compiling for more general circuits and gate sets.

2 Problem Setting

In a nutshell, the problem setting was designed for optimizing a general ansatz in a situation where a dataset consisting of input and output quantum state pairs that can be used multiple times is given, and the result of applying the ansatz to the input is as close to the output as possible. The details are described in the following sections. First, arbitrary one-qubit rotation gates and CNOT gates are set as the set of available gate sets. The main motivation for setting up the problem this way is that it utilizes a universal gate set and that the form is continuous-valued for any gate in order to obtain a sparse circuit by applying sparse regularization. The quantum circuit is set up as Figure 1 and alternates between a one-qubit rotation gate (local operation) and an entangle circuit (global operation). The configurations of the one-qubit rotation gate and entangle are shown in Figure 2, which shows that arbitrary unitary operations on one qubit can be expressed when the one-qubit rotation gate is configured as shown. The entangle circuit is constructed with continuous parameters, which play an important role in the application of sparse regularization.

Let $|\psi_i^{in}\rangle$ be the i-th input quantum state, and $|\psi_i^{out}\rangle$ be the i-th output quantum state. We define n as the number of data points in the dataset and $\boldsymbol{\theta} \in \mathbb{R}^d$ as parameters in the parametrized quantum circuit. For parametrized ansatz $U(\boldsymbol{\theta})$, we write $|\psi_i^{\boldsymbol{\theta}}\rangle$ as the output

*arahon243@eis.hokudai.ac.jp
†i.noda@ist.hokudai.ac.jp
‡oyama@ds.nagoya-cu.ac.jp

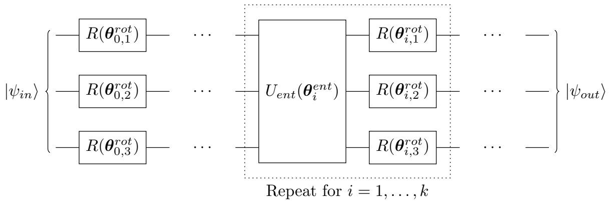
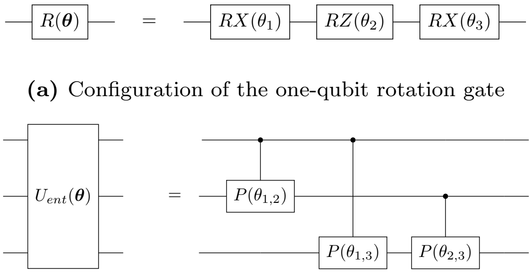

$|\psi_{in}\rangle$
$R(\boldsymbol{\theta}^{rot}_{0,1})$
$R(\boldsymbol{\theta}^{rot}_{0,2})$
$R(\boldsymbol{\theta}^{rot}_{0,3})$
$U_{ent}(\boldsymbol{\theta}^{ent}_{i})$
$R(\boldsymbol{\theta}^{rot}_{i,1})$
$R(\boldsymbol{\theta}^{rot}_{i,2})$
$R(\boldsymbol{\theta}^{rot}_{i,3})$
$|\psi_{out}\rangle$
Repeat for $i = 1, \dots, k$

Figure 1: Diagram of parameterized quantum circuit model of three-qubit system treated in this study



(a) Configuration of the one-qubit rotation gate

(b) Configuration of entangle circuit U_{ent}

Figure 2: Configuration of components used in parametrized quantum circuit

quantum state when we apply $U(\boldsymbol{\theta})$ to input quantum state $|\psi_i^{in}\rangle$; i.e., $|\psi_i^{\boldsymbol{\theta}}\rangle = U(\boldsymbol{\theta})|\psi_i^{in}\rangle$. The objective function for the accuracy of the outputs is defined by using the inner product of the quantum states:

$$\begin{aligned} f(\boldsymbol{\theta}) &= \frac{1}{n}\sum_{i=1}^{n}\left\{1 - \|\langle\psi_i^{out}|\psi_i^{\boldsymbol{\theta}}\rangle\|^2\right\} &(1)\\ &= \frac{1}{n}\sum_{i=1}^{n}\left\{1 - \|\langle\psi_i^{out}|U(\boldsymbol{\theta})|\psi_i^{in}\rangle\|^2\right\} &(2)\end{aligned}$$

When $f(\boldsymbol{\theta})$ takes values in the range $[0, 1]$ and $f(\boldsymbol{\theta}) = 0$, we can say that $U(\boldsymbol{\theta})$ is a suitable ansatz for performing a perfect unitary operation on the quantum states in the data set. Therefore, in the setting of the quantum compiling problem, the challenge is to find a value for parameter $\boldsymbol{\theta}$ that reduces the value of $f(\boldsymbol{\theta})$ as much as possible.

3 Controlled Phase Gate and Sparse Regularization

The Cphase gate is a two-qubit gate (Figure 3) that change the phase of the target bit with reference to the control bit. The amount of phase change is controlled by parameter θ, the determination of which is important in sparse regularization. This is because the number of CNOT gates $C(\theta)$ required to create a Cphase gate corresponding to parameter θ is

$$C(\theta) = \begin{cases} 0 & (\theta = 2k\pi, k \in \mathbb{Z}) \\ 1 & (\theta = (2k+1)\pi, k \in \mathbb{Z}) \\ 2 & (\text{otherwise}) \end{cases} \tag{3}$$

That is, by using sparse regularization, many θ values can be reduced to 0 or integer multiples of π, enabling the induced optimization to reduce the number of CNOT gates to be implemented.

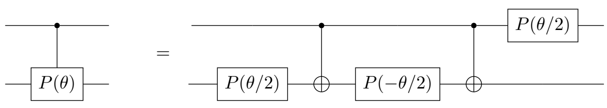


Figure 3: Cphase gate configuration

4 Regularization Methods

In general, in sparse regularization, the sum of the function to be optimized and the regularization term is the overall loss function, and the objective is to minimize the loss function. Let $g(\boldsymbol{\theta})$ be the regularization term for sparse regularization, then the overall loss function $L(\boldsymbol{\theta})$ is $L(\boldsymbol{\theta}) = f(\boldsymbol{\theta}) + g(\boldsymbol{\theta})$. There are various types of regularization terms, which are set by the user in accordance with the tradeoff between the desired function to be optimized and the degree of sparsity. This study focuses on two types of regularization, L1 regularization and minimax concave penalty (MCP) regularization. The form of each regularization term is shown in Figure 4.

L1 Regularization L1 regularization is one of the most fundamental methods in sparse regularization; the regularization term is defined as

$$g(\boldsymbol{\theta}) = \lambda\|\boldsymbol{\theta}\|_1 = \lambda\sum_i|\theta_i|. \tag{4}$$

where hyperparameter $\lambda \geq 0$ represents the strength of regularization. The higher the value of λ, the stronger the applied regularization. The closer the value of parameter θ_i to 0, the smaller the value of $g(\boldsymbol{\theta})$. Since $g(\boldsymbol{\theta})$ has an absolute value of slope $\lambda\text{sgn}(\theta_i)$ at the point $\theta_i \neq 0$ for each θ_i, L1 regularization induces the value of θ_i to approach 0.

MCP Regularization MCP regularization is a stronger regularization method[11]. Its regularization term is expressed as $g(\boldsymbol{\theta}) = \sum_i g_{MCP}(\theta_i)$, where $g_{MCP}(\theta)$ is defined as

$$g_{MCP}(\theta) = \begin{cases} \lambda|\theta| - \dfrac{\theta^2}{2\alpha} & (|\theta| \leq \alpha\lambda) \\ \dfrac{\alpha\lambda^2}{2} & (\text{otherwise}) \end{cases}. \tag{5}$$

where hyperparameter $\lambda \geq 0$ represents the strength of regularization, and hyperparameter α adjusts the range of parameter θ to enable regularization to be applied. This regularization term satisfies $g_{MCP}(\theta) \risingdotseq \lambda|\theta|$ at $\theta \risingdotseq 0$ and its slope is zero when $|\theta| > \alpha\lambda$. Because the absolute value of the slope increases as the value of θ_i approachs zero, MCP regularization adds stronger regularization compared with L1 regularization. In addition, in this problem setup, MCP regularization satisfies two properties.

Remark 1 *Let $\theta_1, \theta_2 \neq 0$ be any parameters in the parametrized quantum circuit. Then, the value of the MCP regularization term is reduced by consolidating θ_1 and θ_2. That is, the following inequality holds.*

$$g_{MCP}(\theta_1+\theta_2)+g_{MCP}(0) < g_{MCP}(\theta_1)+g_{MCP}(\theta_2) \quad (6)$$

As a result, MCP regularization reduces the number of non-sparse parameters. In addition, MCP regularization has another property.

Remark 2 *Let $\theta_1, \theta_2 \neq 0$ be parameters for successive parameterized gates of the same type in the quantum circuit. Then, the value of the overall loss function is reduced by consolidating θ_1 and θ_2. That is, the following inequality holds.*

$$L(\theta_1+\theta_2, 0) < L(\theta_1, \theta_2) \quad (7)$$

Here the parameters except θ_1 and θ_2 remain unchanged.

In other words, when MCP regularization is used, it induces the gates to merge and become sparse when there is a sequence of gates of the same type.

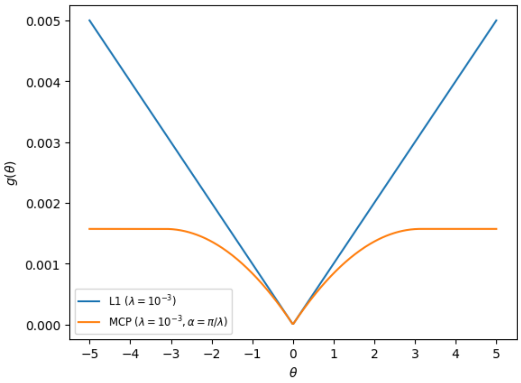


Figure 4: Regularization methods

It should be noted that the problem addressed in this study is the optimization of periodic functions. For example, $\theta = 0$ and $\theta = 2\pi$ are different in value but are equivalent as gate operations. It is therefore reasonable to design the regularization term so that it is also a periodic function. In this study, we introduce periodicity into regularization by re-designing the regularization term:

$$g_{periodic}(\theta) = \min_{k\in\mathbb{Z}} g(\theta + 2k\pi) \quad (8)$$

The value of $g_{periodic}(\theta)$ is defined as the minimum value of the infinite regularization term shifted by one cycle, so $g_{periodic}(\theta)$ trivially has periodicity.

5 Numerical Experiments

Numerical experiments were conducted to evaluate the effect of using each regularization term in the quantum compiling task. The target of the compiling task was a three-qubit quantum Fourier transform; 100 training and 100 test data points (2^3-dimensional complex vectors) were created for each term. The complex vectors that served as data input were generated from the 2^3-dimensional complex unit sphere, i.e., a Haar distribution. We set the depth of the circuit to 5 to guarantee the existence of a parameter set $\boldsymbol{\theta}$ such that $f(\boldsymbol{\theta}) = 0$. We initialized the parameters by sampling from a uniform distribution on $[-\pi, \pi]$ independently. When updating the parameters, the proximity gradient method was used instead of the usual gradient method because there is a partially impossible point in loss function $L(\boldsymbol{\theta})$. The learning rate was fixed at 0.5 for simplicity and was assumed to remain unchanged during the learning process. Hyperparameter α in MCP regularization was set as $\alpha = \pi/\lambda$. The changes in function values after 10,000 parameter updates are summerized in Table 1.

Table 1: Experimental results for sparse regularization methods

Method	Test loss	No. of CNOT gates
Baseline	$3.581 * 10^{-7}$	30
L1($\lambda = 10^{-3}$)	$7.693 * 10^{-5}$	20
MCP($\lambda = 10^{-3}$)	$1.870 * 10^{-5}$	16

As shown in Table 1, the method with L1 regularization and the method with MCP regularization reduced the number of CNOT gates and non-sparse gates while suppressing the increase in loss compared with the baseline. This indicates that methods using sparse regularization produce a sparse parameter set in parameterized quantum circuits. Furthermore, a comparison of the results for L1 regularization and MCP regularization shows that MCP regularization reduced the number of CNOT gates more. This is attributed to MCP regularization having a stronger effect for parameters to approach zero than L1 regularization when the λ parameters are aligned, as shown in Figure 4. Moreover, MCP regularization produced a somewhat sparser solution. That test loss for MCP regularization was lower than that for L1 regularization is a bit surprising because MCP has a disadvantage in the optimization of test loss due to the tradeoff between optimization and sparsity. This could have resulted from insufficient number of update steps for convergence. Nevertheless, the fact that a function with MCP regularization further reduced the loss function is non-trivial and interesting.

6 Conclusion

An optimization method using sparse regularization for parameterized quantum circuits was proposed as a quantum compiling method. Implementation and testing of parameterized quantum circuits that reduce the number of gates by using the regularization term, resulted in the attainment of a sparse parameter set, confirming that the number of gates was reduced.

Acknowledgments

This work was partially supported by JSPS KAKENHI Grant Number JP22K19820, and JST SPRING Grant Number JPMJSP2119.

References

[1] Marco Maronese, Lorenzo Moro, Lorenzo Rocutto, and Enrico Prati. Quantum compiling. *arXiv:2112.00187*, 2021.

[2] Sumeet Khatri, Ryan LaRose, Alexander Poremba, Lukasz Cincio, Andrew T. Sornborger, and Patrick J. Coles. Quantum-assisted quantum compiling. *Quantum*, 3:140, 2019.

[3] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[4] M. Bilkis, M. Cerezo, Guillaume Verdon, Patrick J. Coles, and Lukasz Cincio. A semi-agnostic ansatz with variable structure for quantum machine learning. *arXiv:2103.06712*, 2021.

[5] Matthias C. Caro, Hsin-Yuan Huang, M. Cerezo, Kunal Sharma, Andrew Sornborger, Lukasz Cincio, and Patrick J. Coles. Generalization in quantum machine learning from few training data. *Nature Communications*, 13(4919), 2022.

[6] Lukasz Cincio, Yiğit Subaşı, Andrew T Sornborger, and Patrick J Coles. Learning the quantum algorithm for state overlap. *New Journal of Physics*, 20(11):113022, 2018.

[7] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 4(4):043001, nov 2019.

[8] Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.

[9] Liam Madden and Andrea Simonetto. Best approximate quantum compiling problems. *ACM Transactions on Quantum Computing*, 3(2):1–29, 2022.

[10] Junya Otsuki, Masayuki Ohzeki, Hiroshi Shinaoka, and Kazuyoshi Yoshimi. Sparse modeling in quantum many-body problems. *Journal of the Physical Society of Japan*, 89(1):012001, 2020.

[11] Cun-Hui Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894 – 942, 2010.

Unified Quantum State Tomography and Hamiltonian Learning Using Transformer Models: A Language-Translation-Like Approach for Quantum Systems

Zheng An[1] * Jiahui Wu[1] † Muchun Yang[2 3] ‡ D. L. Zhou[2 3 4] § Bei Zeng[1] ¶

[1] *Department of Physics, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China*

[2] *Institute of Physics, Beijing National Laboratory for Condensed Matter Physics, Chinese Academy of Sciences, Beijing, China*

[3] *School of Physical Sciences, University of Chinese Academy of Sciences, Beijing, China*

[4] *Songshan Lake Materials Laboratory, Dongguan, Guangdong 523808, China*

Abstract. Schrödinger's equation is fundamental in characterizing quantum systems, with quantum state tomography and Hamiltonian learning being essential. This study introduces an attention-based transformer model that merges both tasks without changing architecture, effectively learning relationships between quantum states and Hamiltonian. We demonstrate effectiveness across various quantum systems, from simple 2-qubit cases to 2D antiferromagnetic Heisenberg structures. Our method streamlines data collection, improves scalability, and enables few-shot learning, potentially reducing resources required for characterizing and optimizing quantum systems. This research advances understanding of quantum systems and contributes to quantum computation development.

Keywords: Quantum information theory, Quantum metrology, Machine Learning, Natural Language Processing

[1] Quantum systems are governed by the Schrödinger equation, which plays a pivotal role in defining the relationship between the Hamiltonian structure and the states of the system. This relationship is central to understanding the behavior of quantum systems [1] and for applications such as quantum computing and communication [2, 3]. Moreover, the mapping between the Hamiltonian and the quantum states of a system is indispensable in quantum information science, as it enables us to predict the system's behavior [4, 5]. This knowledge is crucial in quantum computing applications, where Hamiltonian parameters are utilized to control and manipulate quantum systems for specific tasks [6, 7, 8]. Research in this domain can be bifurcated into two primary directions: Quantum State Tomography (QST) and Hamiltonian learning (see Fig. 1).

Quantum state tomography (QST) and Hamiltonian learning are essential techniques in quantum information science. QST comprehensively characterizes quantum states [9, 10, 11, 12, 13], while Hamiltonian learning estimates Hamiltonians for quantum computing and simulation [14, 15, 16, 17, 18, 19]. However, both methods face challenges such as computational complexity and data acquisition difficulties. Despite their individual progress, a unified approach that combines their advantages has not been developed.

Recent advancements in machine learning, particularly transformer architectures [20], have greatly impacted scientific research, including quantum information studies [21, 22, 23]. In this study, we introduce a novel approach that utilizes a language translation method to effectively address both quantum state tomography and Hamiltonian learning, uniting these two techniques in a unified model. The attention mechanism within the transformer model enables us to establish a language-translation-like strategy for mapping Hamiltonian param-

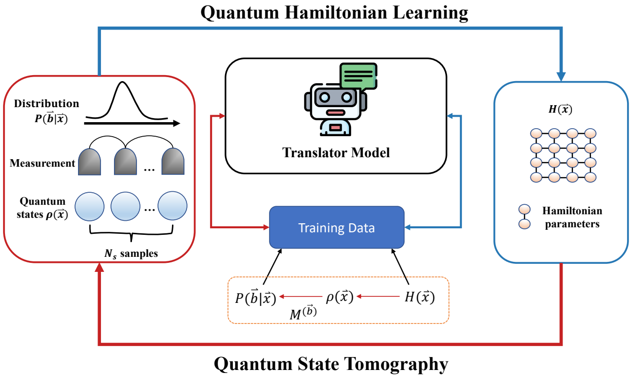


Figure 1: Bidirectional Translation Model for Quantum State Tomography and Hamiltonian Learning: Our model functions as an adaptable and efficient mediator between Quantum State Tomography and Hamiltonian Learning, effectively facilitating the interplay between the elucidation of quantum states derived from Hamiltonian parameters and the estimation of Hamiltonian parameters informed by observed ground states. The training data necessitates unidirectional generation from state tomography $H(\vec{x}) \to P(\vec{b}|\vec{x})$, proving to be advantageous for the concurrent training of both methodologies.

*anzheng@ust.hk

†jiahui.wu@connect.ust.hk

‡yang.muchun@iphy.ac.cn

§zhoudl72@iphy.ac.cn

¶zengb@ust.hk

[1] The full paper is available on arXiv: https://arxiv.org/abs/2304.12010

eters to quantum states. We apply our methodology to an extensive spectrum of quantum systems, ranging from 2-qubit cases to 2D antiferromagnetic Heisenberg models, and demonstrate the versatility of our approach by employing various QST methods.

In this study, we investigate problems within quantum state tomography and Hamiltonian learning, focusing on k-local Hamiltonians and the ground state of an n-qubit quantum system. The Hamiltonian is expressed as:

$$H(\vec{x}) = \sum_{i=1}^{m} x_i H_i, \tag{1}$$

where each H_i is an operator acting non-trivially on no more than k qubits, and x_i represents the local term parameter. We explore the $k = 2$ case.

We consider physical systems of n qubits and construct measurements from an $\mathbb{m}$ outcome single-qubit POVM $\mathcal{M} = \{M^{(b)}\}_b$. The n-qubit measurement is characterized by the tensor product of single-qubit POVM elements. The acquisition of training data involves a unidirectional generation process from state tomography, represented by $H(\vec{x}) \to P(\vec{b}|\vec{x})$. This approach simplifies data acquisition and enables simultaneous training of QST and Hamiltonian learning, improving efficiency and effectiveness in characterizing quantum states and estimating Hamiltonian.

We present a sophisticated Transformer-based model to establish a bidirectional relationship between Hamiltonian parameters and ground state measurement outcomes. The model comprises an encoder and decoder, converting continuous variable inputs into concise representations and generating target sequence outputs such as discrete Hamiltonian parameters or measurement expressions.

Our model employs an embedding neural network and self-attention, focusing on relevant input portions while generating output. We use a multilayer neural network as the embedding layer, transforming continuous Hamiltonian parameters or probability distributions of local measurements into learned vector representations. We discretize Hamiltonian parameters and employ the word embedding technique from natural language processing (NLP) to map each label to a learned vector representation.

We encode distinct local measurement outcomes into discrete tokens in a vocabulary list, allowing encoding of any measurement outcome or parameter for an n-qubit quantum system. The model produces probability distributions for ground states or Hamiltonian parameters.

No modifications to the architecture or parameters are needed, but careful selection and preparation of training data are crucial. We use teacher forcing and autoregression during training, with the objective to minimize average negative log-likelihood loss across the training data.

We present our algorithm to study quantum state tomography and Hamiltonian estimation problems. In a 2-qubit toy model,

$$H(\theta) = \cos(\theta) X_1 X_2 + \sin(\theta) Z_1 I_2, \tag{2}$$

we investigate the ability of POVM measurements to accurately represent the ground state, with results indicating high accuracy even without knowledge of Hamiltonian parameters. We also explore Hamiltonian learning, finding that the algorithm can generate appropriate Hamiltonian parameters given ground-state distribution.

In a 2D antiferromagnetic random Heisenberg model,

$$H(\mathbf{x}) = \sum_{\langle ij \rangle} \mathbf{x}_{ij} \left(X_i X_j + Y_i Y_j + Z_i Z_j \right), \tag{3}$$

we use classical shadow methodology to observe and retrieve physical observables. Our model shows high accuracy in predicting correlation functions, outperforming other methods. Furthermore, our approach effectively predicts Hamiltonians and their properties, offering valuable insights for quantum many-body systems.

Lastly, we implement a scalable few-shot learning strategy to predict properties of large-scale models using limited training data. Our method demonstrates the ability to generate accurate predictions for larger-scale models after being trained on a sparse dataset. This capability is achieved through extrapolation skills developed in the model, enabling it to generalize to configurations of greater scale and complexity.

In conclusion, we present a novel approach that employs language translation models to address quantum state tomography (QST) and Hamiltonian learning in a unified framework, an unexplored concept. Our method capitalizes on the inherent attention mechanism in transformer models, demonstrating adaptability, reduced computational demands, and scalability. By streamlining the data acquisition process and showcasing success across various quantum systems, our approach lays a solid theoretical foundation for practical quantum advantages through machine learning. Furthermore, its few-shot learning capabilities and the insights it provides into the relationship between Hamiltonian structure and quantum system behavior are critical for advancing quantum technologies. Overall, our work introduces a novel, unified, and scalable technique for QST and Hamiltonian learning, fostering a convergence between quantum and artificial intelligence technology development for near-term devices.

References

[1] Benedict Leimkuhler and Sebastian Reich. *Simulating hamiltonian dynamics*. Number 14. Cambridge university press, 2004.

[2] Michael A Nielsen and Isaac L Chuang. Quantum computation and quantum information. 2010.

[3] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018.

[4] C. L. Degen, F. Reinhard, and P. Cappellaro. Quantum sensing. *Rev. Mod. Phys.*, 89:035002, Jul 2017.

[5] David Gross, Yi-Kai Liu, Steven T. Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Phys. Rev. Lett.*, 105:150401, Oct 2010.

[6] Agnes Valenti, Evert van Nieuwenburg, Sebastian Huber, and Eliska Greplova. Hamiltonian learning for quantum error correction. *Phys. Rev. Res.*, 1:033092, Nov 2019.

[7] Ramon van Handel, John K Stockton, and Hideo Mabuchi. Modelling and feedback control design for quantum state preparation. *Journal of Optics B: Quantum and Semiclassical Optics*, 7(10):S179, sep 2005.

[8] Chenfeng Cao, Zheng An, Shi-Yao Hou, DL Zhou, and Bei Zeng. Quantum imaginary time evolution steered by reinforcement learning. *Communications Physics*, 5(1):57, 2022.

[9] G Mauro D'Ariano, Matteo GA Paris, and Massimiliano F Sacchi. Quantum tomography. *Advances in imaging and electron physics*, 128:206–309, 2003.

[10] David Gross, Yi-Kai Liu, Steven T Flammia, Stephen Becker, and Jens Eisert. Quantum state tomography via compressed sensing. *Physical review letters*, 105(15):150401, 2010.

[11] CF Roos, GPT Lancaster, M Riebe, H Häffner, W Hänsel, S Gulde, C Becher, J Eschner, F Schmidt-Kaler, and R Blatt. Bell states of atoms with ultralong lifetimes and their tomographic state analysis. *Physical review letters*, 92(22):220402, 2004.

[12] K Vogel and H Risken. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Physical Review A*, 40(5):2847, 1989.

[13] Tao Xin, Dawei Lu, Joel Klassen, Nengkun Yu, Zhengfeng Ji, Jianxin Chen, Xian Ma, Guilu Long, Bei Zeng, and Raymond Laflamme. Quantum state tomography via reduced density matrices. *Phys. Rev. Lett.*, 118:020401, Jan 2017.

[14] Xiao-Liang Qi and Daniel Ranard. Determining a local Hamiltonian from a single eigenstate. *Quantum*, 3:159, July 2019.

[15] Maxime Dupont, Nicolas Macé, and Nicolas Laflorencie. From eigenstate to hamiltonian: Prospects for ergodicity and localization. *Phys. Rev. B*, 100:134201, Oct 2019.

[16] Jianxin Chen, Zhengfeng Ji, Bei Zeng, and DL Zhou. From ground states to local hamiltonians. *Physical Review A*, 86(2):022339, 2012.

[17] Eyal Bairey, Itai Arad, and Netanel H. Lindner. Learning a local hamiltonian from local measurements. *Phys. Rev. Lett.*, 122:020504, Jan 2019.

[18] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17(8):931–935, 2021.

[19] X. Turkeshi, T. Mendes-Santos, G. Giudici, and M. Dalmonte. Entanglement-guided search for parent hamiltonians. *Phys. Rev. Lett.*, 122:150606, Apr 2019.

[20] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[21] Peter Cha, Paul Ginsparg, Felix Wu, Juan Carrasquilla, Peter L McMahon, and Eun-Ah Kim. Attention-based quantum tomography. *Mach. Learn.: Sci. Technol.*, 3(1):01LT01, 2021.

[22] Haoxiang Wang, Maurice Weber, Josh Izaac, and Cedric Yen-Yu Lin. Predicting properties of quantum systems with conditional generative models. *arXiv preprint arXiv:2211.16943*, 2022.

[23] Lu Zhong, Chu Guo, and Xiaoting Wang. Quantum state tomography inspired by language modeling. *arXiv preprint arXiv:2212.04940*, 2022.

Quantum Illumination with a Hetero-Homodyne Receiver and Sequential Detection

Maximilian Reichert[1 2] Quntao Zhuang[3 4] Jeffrey H. Shapiro[5] Roberto Di Candia[6 7 *]

[1] *Department of Physical Chemistry, University of the Basque Country UPV/EHU, Aptdo. 644, 48080 Bilbao, Spain*
[2] *EHU Quantum Center, University of the Basque Country UPV/EHU, Bilbao, Spain*
[3] *Ming Hsieh Department of Electrical and Computer Engineering, University of Southern California, Los Angeles, California 90089, USA*
[4] *Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA*
[5] *Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[6] *Department of Information and Communications Engineering, Aalto University, Espoo, 02150 Finland*
[7] *Dipartimento di Fisica, Università degli Studi di Pavia, Via Agostino Bassi 6, I-27100, Pavia, Italy*

Abstract. We propose a hetero-homodyne receiver for quantum illumination (QI) target detection. Unlike prior QI receivers, it uses a cascaded positive operator-valued measurement (POVM) that does not require a quantum interaction between QI's returned radiation and its stored idler. When used without sequential detection its performance matches the 3 dB quantum advantage over optimum classical illumination (CI) that Guha and Erkmen's [Phys. Rev. A **80**, 052310 (2009)] phase-conjugate and parametric amplifier receivers enjoy. When used in a sequential detection QI protocol, the hetero-homodyne receiver offers a 9 dB quantum advantage over a conventional CI radar, and a 3 dB advantage over a CI radar with sequential detection. Our work is a significant step forward toward a practical quantum radar for the microwave region, and emphasizes the potential offered by cascaded POVMs for quantum radar.

Keywords: Quantum illumination, quantum radar, microwave quantum sensing

Background

Quantum radars use resources unavailable to their classical counterparts, principally entanglement, to obtain improved remote-sensing performance at the same transmitted energy, see Refs. [1, 2, 3] for recent reviews. To date, the only quantum radar protocol whose target-detection performance is predicted to exceed that of its best classical competitor is Tan *et al.*'s quantum illumination (QI) [4]. QI with optimum reception offers a 6 dB quantum advantage in error-probability exponent for detecting a weakly-reflecting target embedded in high-brightness background noise. This advantage *only* occurs in a lossy, noisy setting that destroys the initial entanglement between QI's transmitted signal and its stored idler. In particular, Nair [5] has shown that in the absence of noise conventional coherent-state radar closely approximates the target-detection performance of the optimum quantum radar of the same transmitted energy. So, because daytime background light at near-visible wavelengths has extremely low brightness, Tan *et al.*'s QI attracted little interest from the radar community until Barzanjeh *et al.* [6] described how it might be used at microwave wavelengths, where high-brightness background noise is the norm and QI's quantum advantage could help in detecting stealth targets.

Tan *et al.*'s QI relies on the nonclassical phase-sensitive cross correlation between the brightness-N_S signal and idler beams produced by a spontaneous parametric downconverter (SPDC), viz., signal and idler consisting of $M \gg 1$ independent and identically-distributed mode pairs in two-mode squeezed-vacuum states. The TMSV's nonclassical cross correlation, $\sqrt{N_S(N_S+1)}$, greatly exceeds the classical limit, N_S, in low-brightness ($N_S \ll 1$) operation, and disappears as N_S grows without bound. Furthermore, because conventional interference techniques are incapable of detecting phase-sensitive correlation [1], the first proposed receivers [7] for obtaining *any* quantum advantage from QI used parametric amplifiers to convert phase-sensitive correlation into phase-insensitive correlation prior to detection by conventional techniques. These proposals—Guha and Erkmen's parametric amplifier (PA) and the phase-conjugate (PC) receivers—deliver *at most* a 3 dB quantum advantage in error-probability exponent, and to do so they require a quantum memory capable of losslessly storing the idler's high time-bandwidth product quantum state for the roundtrip radar-to-target-to-radar propagation delay. So far, however, only 20% quantum advantage has been demonstrated in optical wavelength (with high-brightness noise injection) [8] and microwave wavelength [9] table-top experiments.

The first explicit architecture for obtaining QI's full 6 dB quantum advantage was the feed-forward sum-frequency generation receiver [10], whose implementation requires an as yet unavailable single-photon nonlinearity as well as a quantum memory for idler storage. A more recent architecture, the correlation-to-displacement receiver [11], circumvents the need for a single-photon nonlinearity, but requires a lossless $M \times M$ programmable beam splitter with $M \gg 1$—which will be a daunting implementation task at microwave wavelengths—as well as the aforementioned quantum memory for idler storage. Were available technology capable of realizing such QI receivers, the ultimate performance for Tan *et al.*'s

[*] `roberto.dicandia@aalto.fi`

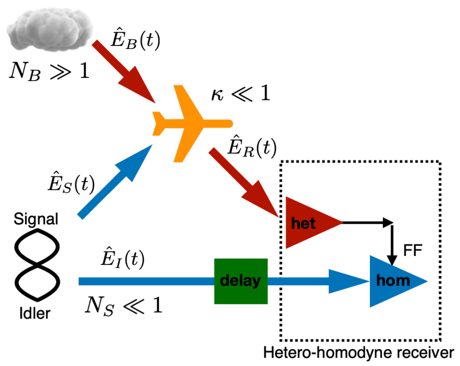


Figure 1: Sketch of the Tan *et al.*'s QI protocol with hetero-homodyne reception. The hetero-homodyne receiver measures the cross correlation between the returned radiation and the idler. Because the low-brightness TMSV mode pairs' phase-sensitive cross correlation $\sqrt{N_S(N_S+1)}$ greatly exceeds the classical limit N_S, QI outperforms classical illumination in this regime even though loss and noise have destroyed the TMSV states' initial entanglement. het: heterodyne. hom: homodyne. FF: feed-forward.

target-detection scenario would be obtained, because theory [12, 13, 14, 15, 16, 17, 18] has proven the optimality of the M mode-pair TMSV state for that setting.

Here, we report [19] a significant advance for microwave QI, and, more generally, emphasizes the potential offered by cascaded positive-operator valued (POVM) measurements for quantum radar. First, motivated by Shi *et al.* [11]'s coherence-to-displacement conversion and Shapiro's use of sequential detection [20] to break Nair's performance limit on noise-free target detection [5], we propose a hetero-homodyne receiver for QI, a cascaded POVM that, unlike prior QI receivers, does *not* need a quantum interaction between QI's returned radiation and its stored idler, see Fig. **??**. Our receiver achieves a 3 dB advantage over the optimum receiver for Tan *et al.*'s QI, i.e., 9 dB better than a conventional classical radar.

The Hetero-homodyne receiver

Figure **??** shows a schematic of quantum illumination with a hetero-homodyne receiver. A signal-idler system is initialized in an M mode-pair TMSV state, with each signal and idler mode containing $N_S \ll 1$ photons on average. The signal is sent to test for the presence of a weakly-reflecting (roundtrip transmissivity $\kappa \ll 1$) target embedded in high-brightness background noise ($N_B \gg 1$). The hetero-homodyne receiver measures the cross correlation between the returned radiation and the idler. It does so by first applying a heterodyne measurement to the returned signal, obtaining the complex value $\mathcal{M} \in \mathbb{C}$ as result. Finally, the idler is measure homodyne along the direction given by $\mathcal{M}^*$. The homodyne detector's output, conditioned on the heterodyne detector's output, is—with a convenient normalization—a measurement of the observable $\mathrm{Re}[E_R(t)\hat{E}_I(t)]$, which is the optimal non-collective observable for Gaussian quantum illumination [16].

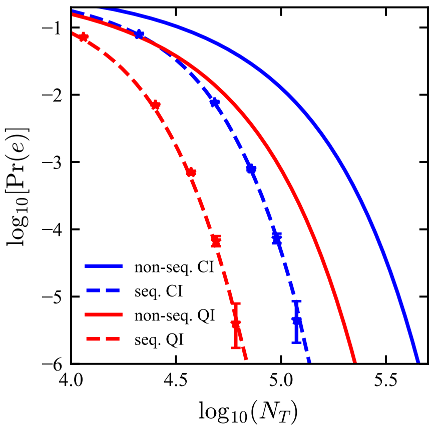


Figure 2: Classical and quantum illumination error probabilities versus average number of transmitted signal photons, N_T, assuming $N_S = 0.01$, $\kappa = 0.01$, and $N_B = 100$. The solid curves are theory results for non-sequential CI with homodyne detection and non-sequential QI with hetero-homodyne detection, respectively (see Ref. [19]). The points are simulated error probabilities for sequential CI with homodyne detection and sequential QI with hetero-homodyne detection with $M_S = 10^5$ for $P_{err,target} = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}$, and 10^{-5}. Each point was generated from 10^6 simulated experiments for each hypothesis.

Classical and quantum illumination with sequential detection

Wald [21] originated the sequential probability-ratio test (SPRT) as an alternative to the standard (LRT) used for fixed-length data. In SPRT states are sent sequentially to test the target region, and the transmission is halted when the desire error probability $P_{err,target}$ has been reached. To exhibit the benefits offered by sequential detection, we have assumed the low SNR regime, i.e., $N_S \ll 1$, $\kappa \ll 1$ and $N_B \gg 1$. We apply the SPRT protocol to CI using homodyne detection and to QI using hetero-homodyne detection. For each transmission, the CI transmitter will employ a coherent state $|\sqrt{M_S N_S}\rangle$. The QI transmitter, on the other hand, will employ M_S iid TMSV mode pairs with average photon number N_S in each signal and idler. For both systems we assume $M_S \gg 1$ and $\kappa M_S N_S / N_B \ll 1$ and evaluate the average number of transmissions—hence the average transmitted photon number—under the assumptions of equally-likely hypotheses and equal false-alarm and miss probabilities. We have found a 6 dB advantage in error-probability exponent over non-sequential detection with the same average transmitting power, see Fig. 2. Putting aside operational considerations regarding sequential detection's practicality for radar target detection—see Ref. [20] for some discussion of these considerations—this CI sequential-detection advantage matches the quantum advantage of the yet-to-be implemented optimum quantum receiver for QI without sequential detection. Thus, if sequential detection is suitable for radar target detection, there should be little interest in further pursuit of Tan *et al.* QI without sequential detection.

Discussion

Our paper has made two significant advances for microwave QI: (1) it proposed the hetero-homodyne receiver, whose non-sequential performance matches that of the Guha and Erkmen's receivers in QI's usual $N_S \ll 1$, $\kappa \ll 1$, $N_B \gg 1$, $M \gg 1$ operating regime (but exceeds their performance outside this regime) and does not require a quantum interaction between the idler and the returned signal; and (2) it showed that sequential detection, at low single-trial SNR in QI's usual operating regime, provides a 6 dB increase in error-probability exponent for both CI homodyne detection and QI hetero-homodyne detection as compared to their non-sequential counterparts. Advance (2) demonstrates that Tan *et al.*'s TMSV QI does *not* saturate what can be gained from using entanglement for target detection in a lossy, noisy environment despite the optimality proofs from Refs. [12, 13, 14, 15, 16, 17, 18]. At this point it behooves us to address some additional issues regarding the hetero-homodyne receiver and its use with sequential detection. On the negative side are its continuing need for knowledge of the target's phase delay and its requiring a quantum memory for idler storage. On the positive side for sequential hetero-homodyne QI is its less demanding bandwidth requirement.

In conclusion, we believe the hetero-homodyne receiver we have proposed both pushes microwave QI target detection closer to fruition and underscores the need for continued research into quantum radar.

References

[1] J. H. Shapiro, The quantum illumination story, IEEE Aerosp. Electron. Syst. Mag. **35**(4), 8-20 (2020).

[2] R. G. Torromé, N. B. Bekhti-Winkel, and P. Knott, Introduction to quantum radar, arXiv:2006.14238 [quant-ph].

[3] G. Sorelli, N. Treps, F. Grosshans, and F. Boust, Detecting a target with quantum entanglement, IEEE Aerosp. Electron. Syst. Mag. **37**(5), 68 (2022).

[4] 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. **101**, 253601 (2008).

[5] R. Nair, Discriminating quantum-optical beamsplitter channels with number-diagonal signal states: Applications to quantum reading and target detection, Phys. Rev. A **84**, 032312 (2011).

[6] S. Barzanjeh, S. Guha, C. Weedbrook. D. Vitali, J. H. Shapiro, and S. Pirandola, Microwave quantum illumination, Phys. Rev. Lett. **114**, 080503 (2015).

[7] S. Guha and B. I. Erkmen, Gaussian-state quantum-illumination receivers for target detection, Phys. Rev. A **80**, 052310 (2009).

[8] Z. Zhang, S. Mouradian, F. N. C. Wong, and J. H. Shapiro, Entanglement-enhanced sensing in a lossy and noisy environment, Phys. Rev. Lett. **114**, 110506 (2015).

[9] R. Assouly, R. Dassonneville, T. Peronnin, A. Bienfait, and B. Huard, Demonstration of quantum advantage in microwave quantum radar, arXiv:2211.05684 [quant-ph].

[10] Q. Zhuang, Z. Zhang, and J. H. Shapiro, Optimum mixed-state discrimination for noisy entanglement-enhanced sensing, Phys. Rev. Lett. **118**, 040801 (2017).

[11] H. Shi, B. Zhang, and Q. Zhuang, Fulfilling entanglement's benefit via converting correlation to coherence, arXiv:2207.06609 [quant-ph].

[12] G. De Palma and J. Borregaard, Minimum error probability of quantum illumination, Phys. Rev. A **98**, 012101 (2018).

[13] R. Nair and M. Gu, Fundamental limits of quantum illumination, Optica **7**, 771 (2020).

[14] R. Di Candia, H. Yiğitler, G. S. Paraoanu, R. Jäntti, Two-way covert communication in the microwave regime, PRX Quantum **2**, 020316 (2021).

[15] M. Bradshaw, L. O. Conlon, S. Tserkis, M, Gu, P. K. Lam. and Syed M. Assad, Optimal probes for continuous-variable quantum illumination, Phys. Rev. A **103**, 062413 (2021).

[16] M. Sanz, U. Las Heras, J. J. Garcia-Ripoll, E. Solano, R. Di Candia, Quantum estimation methods for quantum illumination, Phys. Rev. Lett. **118**, 070803 (2017).

[17] Z. Gong, N. Rodriguez, C. N. Gagatsos, S. Guha and B. A. Bash, Quantum-enhanced transmittance sensing, in *IEEE J. of Sel. Top. in Signal Process* (2022).

[18] R. Jonsson and R. Di Candia 2022, Gaussian quantum estimation of the loss parameter in a thermal environment, J. Phys. A: Math. Theor. **55**, 385301 (2022).

[19] M. Reichert, Q. Zhuang, J. H. Shapiro, and R. Di Candia, arXiv2303.18207 [quant-ph].

[20] J. H. Shapiro, First-photon target detection: Beating Nair's pure-loss performance limit, Phys. Rev. A **106**, 032415 (2022).

[21] A. Wald, Sequential tests of statistical hypotheses, Ann. Math. Stat. **16**, 117 (1945).

Hassle-free Extra Randomness from Quantum State's Identicalness with Untrusted Components

Hamid Tebyanian
Department of Mathematics University of York Heslington, York, YO10 5DD, United Kingdom
hamid.tebyanian@york.ac.uk

***Abstract*—Semi-device-independent (semi-DI) quantum random number generators (QRNGs) are gaining more awareness more and more, presenting a high level of security with an uncomplicated experimental requirement. In this paper, we study a semi-DI protocol based on a minimum error rate of energy-alike coherent states built on the prepare-and-measure scheme with a straightforward experimental requirement where the measurement device is untrusted. Furthermore, the security estimation is based on lower bounding the guessing probability, which is numerically optimized by utilizing semi-definite programming. Finally, a comparison of different encoding and decoding schemes is presented.**

***Index Terms*—Quantum random number generator, Assurance of quantum random number generators**

I. Introduction

Owning high-quality and secure randomness is a necessary step to initiate most of the cybersecurity protocols. The level of security is subject to several theoretical and experimental factors. Accordingly, it is essential to ensure that the random numbers are generated securely to prevent illegal access and obtain testable randomness. In general, random number generators can be classified into three main categories: pseudo-random number generator (PRNG), hardware random number generator (HRNG), and quantum random number generator (QRNG). The PRNG and HRNG are based on deterministic phenomena making them predictable. At the same time, randomness is a fundamental feature of quantum mechanics that originated from its probabilistic nature. Relying on the level of assumptions and experimental conditions, QRNGs can be divided into three sub-groups: device-dependent (DD), semi-device-independent (semi-DI), and device-independent (DI) QRNG. DD QRNGs are easy to implement and performant, while dishonest producers, imperfections, or any deviation from the ideal situation can compromise security. On the other hand, DI QRNGs offer the highest security, but the experimental realization of DI protocol is very challenging, causing them to be less practical. Semi-DI protocols, however, present an excellent trade-off between security and practicality, making them a perfect candidate for practical uses.

II. Protocol

This semi-DI protocol is based on the prepare-and-measure scheme where the preparation device is partially trusted, inspecting the single condition of the protocol, which is transmitting energy-alike states. On the other hand, no requirement is assumed on the measurement device, and it can be treated as a black box.

We consider the simplest encoding technique, binary state preparation, two coherent states represented by two circles in Fig. 1 (A), guaranteeing the states have similar energy; in this case, we can also bind the states' energy to be α-close to the vacuum state, discrimination of such states always comes with an error, see Fig. 1 (B). Basically, the state's indistinguishability detection imposes a minimal rate of unresolved events, namely error probability $P_e \geq 1 - \sqrt{1-\delta^2}$, where δ is the scalar product of the states ($\delta = \langle \psi_l | \psi_j \rangle$). As shown in Fig. 1 (B), the error probability is maximum when the states overlap is equal to one ($\mu = 0$), and it decreases with increasing the energy of the states. Otherwise stated, the ambiguity in states discrimination increase when the state's energy decreases; the closer to the vacuum state, the more ambiguity. Note that the states' energy cannot drop to zero as the system becomes single-choice, transmitting vacuum states all the time. The same reasoning applies to more inputs; as long as the state has the energy-alike constraint, the measurement comes with an error.

Suppose that the source takes an input i, chosen independently from the source and the detector, and prepares a physical system in one of the possible quantum states. Later, transmitted to the measurement part, where a detector, which could be either in the continuous or discrete variable domains, e.g., single-photon detector, homodyne, or heterodyne detector, returns an output string o. Randomness can be certified by analysing the input-output probabilities $p\left(i|o\right)$, given that the states obey the energy-alike constraint. As shown in [1,7,8,9], specific input-output correlations indicate genuine quantum randomness in the sense that the device's output cannot be perfectly predicted, whatever the underlying quantum representation cause it. In spirit, this is comparable to the violation of Bell inequalities which witness genuine randomness independently of the devices' implementation. Owning the measurement outcomes together with the inserted values to the preparation box, we can compute the input-output correlation $p\left(o|i\right)$:

$$p\left(o|i\right) = \sum_{\lambda} p\left(\lambda\right) \left\langle \psi_i^\lambda \left| \Pi_o^\lambda \right| \psi_i^\lambda \right\rangle \tag{1}$$

where ρ_i^λ are the propagated states, λ represents the possible strategies of an attacker, and Π_b^λ are the POVM determining the measurement method. The conditional min-entropy (CME)

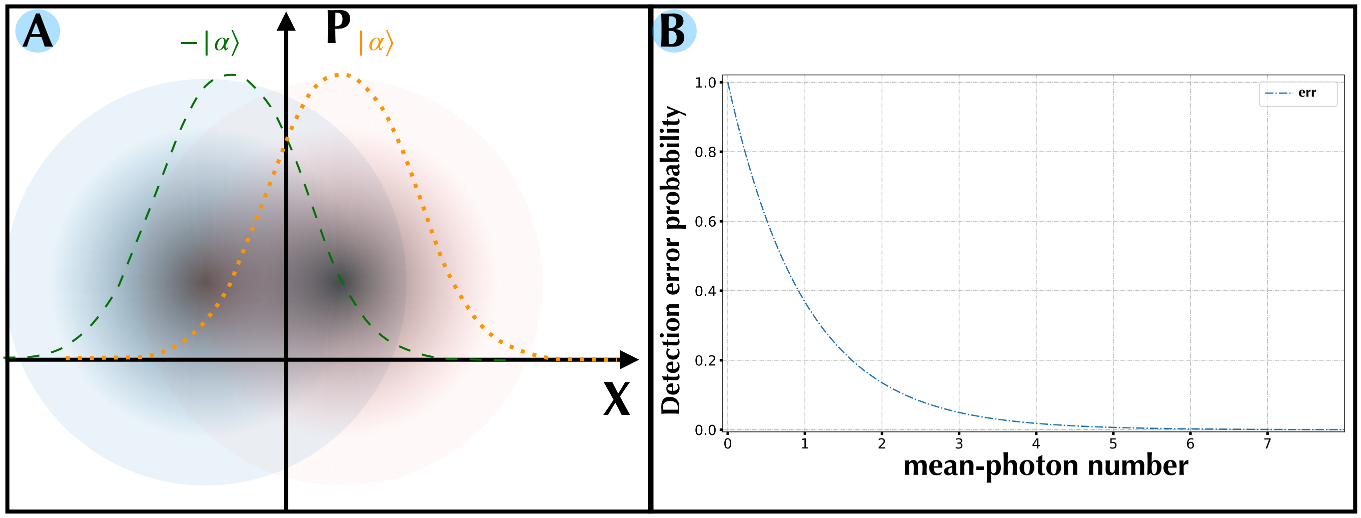


Fig. 1. (A): Schematic representation of coherent state with overlap, the dashed and dotted green and orange lines show the distribution of the states over X. (B): The detection error probability as a function of mean-photon number (μ) of states ($\mu = |\alpha|^2$).

is employed to estimate the system's entropy, put it differently, CME calculates the quantity of extractable genuine randomness which reads:

$$H_{\min} = -\log_2 (P_{guess}) = \sum_i p_i \sum_\lambda p_\lambda \max\left\{ \sum_o^{m-1} Tr\left[\rho_i^\lambda \Pi_o^\lambda\right]\right\} \tag{2}$$

p_i is the probability of transmitting i, while λ is arbitrary, P_{guess} is the guessing probability, which is the probability that an attacker can guess the outcome, given the input. P_{guess} should be optimized over all possible measurement and preparation strategies, making it complicated to be solved analytically. Following the approach presented in [2,3,4,5,6], we use a numerical tool (semi-definite programming) to solve the optimization problem and estimate the amount of extractable randomness.

III. Preparation and Measurement

The preparation device transmits quantum states with limited energy; this constraint can be seen in the context of energy-bound or overlap assumptions depending on the user's choice. The energy bound is a tighter bound which imposes an inevitable overlap of the prepared states. In any case, the source can be a weak coherent source or a time-bin single-photon source as long as the experiment energy-alike condition is satisfied.

Here we demonstrate the phase encoding scheme exploiting a weak coherent light, meaning that the states are encoded based on their phase. This provides the possibility to increase the number of inputs straightforwardly, as shown in Fig. 2 (A); infinite possible states can be encoded in this way, while the states are located within limits, dashed circle. On the measurement side, in general, any scheme can be employed; here, we study heterodyne detection as it gives information on both light field quadratures simultaneously, enabling tracking the states' phase. The heterodyne detection describes the probability density of getting detection proportional to $\frac{(X+iP)}{2}$ from an optimal simultaneous measurement of field quadratures X and P. This kind of measurement is undoubtedly non-ideal, considering that the field quadratures X and P do not commute. The heterodyne detection corresponding POVM reads;

$$\Pi\left(x_\phi\right) = \ket{\beta}\bra{\beta} \tag{3}$$

Where $\ket{\beta}$ is the coherent state with complex amplitude β. Having the POVM, we can compute the conditional probabilities:

$$p(\alpha) \;=\; 1/\pi \int |\,|\langle\alpha|\beta\rangle|^2\, d\beta^2 \tag{4}$$

Fig. 2(B) represents the conditional min-entropy as a function of the number of outcomes for the heterodyne detector for binary and ternary encoding schemes. As shown, the entropy improvement for the binary case, even for higher outcomes, is negligible, but for the ternary case, the gain is more evident.

Note that increasing the number of outcomes can be done in the post-processing stage without touching the actual experimental setup. Indicating that more randomness can be extracted from the same optical device only by changing the data processing stage.

IV. Conclusion

In conclusion, this paper studies a semi-DI QRNG with various encoding and decoding schemes, particularly phase encoding and heterodyne detection schemes were investigated in detail. It is shown that by increasing the number of inputs and outcomes, the extractable randomness increases accordingly, meaning more accessible randomness needles for changing any experimental component.

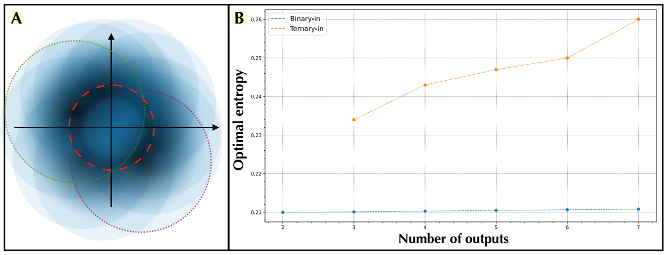


Fig. 2. (A): Encoding technique using weak coherent phase; in this way, one can generate infinite inputs using different phases to encode the states. The central dashed circle represents the experiment constraint, and the two dotted lines exemplify the binary phase keying scheme. (B): The conditional min-entropy as a function of the number of outcomes for heterodyne detection with binary and ternary phase shift keying encoding schemes.

V. Acknowledgement*

This work is supported by the Innovate UK Industrial Strategy Challenge Fund (ISCF), project 106374-49229 AQuRand (Assurance of Quantum Random Number Generators).

References

[1] Rusca, D., Tebyanian, H., Martin, A., Zbinden, Fast self-testing quantum random number generator based on homodyne detection, Applied Physics Letters 116 (2020).

[2] Avesani, M., Tebyanian, H., Villoresi, P., Vallone, G. Unbounded randomness from uncharacterized sources. *Communications Physics*. **5**, 273 (2022).

[3] Tebyanian, H., Zahidy, M., Avesani, M., Stanco, A., Villoresi, P. & Vallone, G. Semi-device independent randomness generation based on quantum state's indistinguishability. *Quantum Science And Technology*. **6**, 045026 (2021,9), https://doi.org/10.1088/2058-9565/ac2047

[4] Tebyanian, H., Avesani, M., Vallone, G., Villoresi, P. Semi-device-independent randomness from d-outcome continuous-variable detection. *Phys. Rev. A*. **104**, 062424 (2021).

[5] Brask, J., Martin, A., Esposito, W., Houlmann, R., Bowles, J., Zbinden, H. & Brunner, N. Megahertz-Rate Semi-Device-Independent Quantum Random Number Generators Based on Unambiguous State Discrimination. *Phys. Rev. Applied*. **7**, 054018 (2017,5), https://link.aps.org/doi/10.1103/PhysRevApplied.7.054018

[6] Tebyanian, et al., Time-bin Quantum Random Number Generator with Uncharacterized Devices.

[7] Tebyanian, H., Avesani, M., Vallone, G. & Villoresi, P. Semi-device-independent randomness from d-outcome continuous-variable detection. *Phys. Rev. A*. **104**, 062424 (2021,12), https://link.aps.org/doi/10.1103/PhysRevA.104.062424

[8] Avesani, M., Tebyanian, H., Villoresi, P. & Vallone, G. Semi-Device-Independent Heterodyne-Based Quantum Random-Number Generator. *Phys. Rev. Applied*. **15**, 034034 (2021,3), https://link.aps.org/doi/10.1103/PhysRevApplied.15.034034

[9] Zahidy, M., Tebyanian, H., Cozzolino, D., Liu, Y., Ding, Y., Morioka, T., Oxenløwe, L. & Bacco, D. Quantum randomness generation via orbital angular momentum modes crosstalk in a ring-core fiber. *AVS Quantum Science*. **4**, 011402 (2022),

Cross-Platform Comparison of Arbitrary Quantum Processes

Congcong Zheng[1] * Xutao Yu[1] † Kun Wang[2] ‡

[1] *State Key Lab of Millimeter Waves, Southeast University, Nanjing 211189, China*
[2] *Institute for Quantum Computing, Baidu Research, Beijing 100193, China*

Abstract. We propose a protocol to compare the performance of quantum processes implemented on spatially and temporally separated quantum platforms using Local Operations and Classical Communication (LOCC). Local unitaries are sampled and communicated to each platform to construct state preparation and measurement circuits. The generated probability distributions allow for direct estimation of the process fidelity. Remarkably, the experimental results reveal that the protocol can accurately compare the performance of the quantum processes implemented on different quantum computers, requiring significantly fewer measurements than those needed for full quantum process tomography. Our work serves as a novel application of the powerful randomized measurement toolbox and a catalyst for collaborative cross-platform comparison of quantum computers.

Keywords: random measurements, cross-platform comparison, quantum processes, LOCC

1 Summary

The current generation of noisy intermediate-scale quantum (NISQ) computers, despite their potential, are still hindered by quantum noise [1]. A great challenge is how to directly compare the performance of the quantum computers fabricated by different manufacturers, termed as *cross-platform comparison*. This task is especially relevant when we move towards regimes where comparing to classical simulations becomes computationally challenging.

In this work, by elaborating the core idea of [2], *we present a novel protocol for cross-platform comparing spatially and temporally separated quantum processes.* The protocol uses only single-qubit unitary gates and classical communication between quantum computers. This approach allows for accurate estimation of the performance of quantum devices manufactured in separate laboratories and companies using different technologies. We apply the protocol to compare the performance of five quantum devices from IBM and the "Qianshi" quantum computer from Baidu via the cloud. The experimental results reveal that our protocol accurately compares the performance of different quantum computers with significantly fewer measurements than quantum process tomography. Overall, our protocol serves as a novel application of the powerful randomized measurement toolbox [3]. A technical version can be found in arXiv:2303.13911.

2 Cross-Platform Comparison

The *max process fidelity* between two n-qubit quantum processes $\mathcal{E}_1$ and $\mathcal{E}_2$, implemented on different quantum platforms, is defined as [4]

$$F_{\max}(\mathcal{E}_1, \mathcal{E}_2) := \frac{\operatorname{Tr}[\eta_1\eta_2]}{\max\{\operatorname{Tr}[\eta_1^2], \operatorname{Tr}[\eta_2^2]\}}, \tag{1}$$

where η_i is the Choi state of $\mathcal{E}_i$ $(i = 1, 2)$. We first propose a protocol to estimate $F_{\max}$ that is conceptually straightforward yet experimentally challenging. Then, we make a novel modification to the protocol that employs randomized local input states and local measurements.

2.1 Ancilla-assisted protocol

In this section, we recover a conceptually simple approach for estimating the max process fidelity, which is illustrated in Figure 1(a)-(c). It employs the cross-platform state estimation protocol proposed in [2] as a subroutine. We refer to this protocol as the *ancilla-assisted cross-platform comparison* because it requires additional clean ancilla qubits to prepare the Choi state of the quantum process. To perform this protocol, a maximally entangled state is required as input, resulting in a two-fold overhead when comparing $2n$-qubit states instead of n-qubit states. Furthermore, the experimental challenge of preparing high-fidelity maximally entangled states can potentially affect the accuracy of the protocol.

2.2 Ancilla-free protocol

To overcome the limitations of the ancilla-assisted protocol, we propose an efficient and ancilla-free ap-

*zhengcongcong@seu.edu.cn
†yuxutao@seu.edu.cn
‡wangkun28@baidu.com

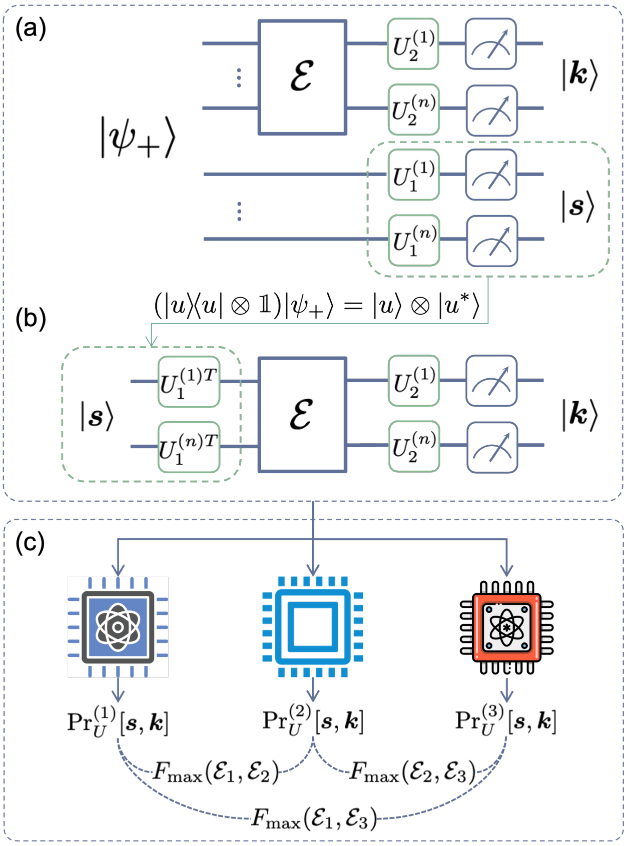

(a)
$|\psi_+\rangle$
$\mathcal{E}$
$U_2^{(1)}$
$U_2^{(n)}$
$U_1^{(1)}$
$U_1^{(n)}$
$|\boldsymbol{k}\rangle$
$|\boldsymbol{s}\rangle$
$(|u\rangle\langle u| \otimes \mathbb{1})|\psi_+\rangle = |u\rangle \otimes |u^*\rangle$
(b)
$|\boldsymbol{s}\rangle$
$U_1^{(1)T}$
$U_1^{(n)T}$
$\mathcal{E}$
$U_2^{(1)}$
$U_2^{(n)}$
$|\boldsymbol{k}\rangle$
(c)
$\mathrm{Pr}_U^{(1)}[\boldsymbol{s},\boldsymbol{k}]$
$\mathrm{Pr}_U^{(2)}[\boldsymbol{s},\boldsymbol{k}]$
$\mathrm{Pr}_U^{(3)}[\boldsymbol{s},\boldsymbol{k}]$
$F_{\max}(\mathcal{E}_1,\mathcal{E}_2)$
$F_{\max}(\mathcal{E}_2,\mathcal{E}_3)$
$F_{\max}(\mathcal{E}_1,\mathcal{E}_3)$

Figure 1: Two protocols to estimate the max process fidelity $F_{\max}$ between quantum processes on different platforms: (a) *Ancilla-assisted*: Prepare entangled state, execute target process, and perform randomized measurements using $U_1 \otimes U_2$. (b) *Ancilla-free*: Randomly sample basis $|\boldsymbol{s}\rangle$, execute U_1^T, execute target process, and perform randomized measurements using U_2. (c) Run circuits from (a) or (b) on platform $\mathcal{S}_i$ to obtain $\mathrm{Pr}_U^{(i)}[\boldsymbol{s},\boldsymbol{k}]$. Finally, $F_{\max}$ is inferred from probability distributions (see text).

proach for estimating the max process fidelity. We refer to the new protocol as the *ancilla-free cross-platform comparison* and it works as follows. Consider two n-qubit quantum processes $\mathcal{E}_1$ and $\mathcal{E}_2$ realized on different quantum platforms $\mathcal{S}_1$ and $\mathcal{S}_2$. The protocol, illustrated in Figure 1(b)-(c), consists of three main steps: sampling unitaries, running circuits, and post-processing.

Step 1. Sampling unitaries: Construct two n-qubit unitaries $U_i = \bigotimes_{k=1}^n U_i^{(k)}$, $i = 1, 2$, where each $U_i^{(k)}$ is identically and independently sampled from a single-qubit set satisfying unitary 2-design. The information of U_i is then communicated to both platforms via classical communication.

Step 2. Running circuits: On each platform $\mathcal{S}_i$ ($i = 1, 2$) initialize the quantum system to the computational states $|\boldsymbol{s}\rangle$ and apply the first unitary U_1 to $|\boldsymbol{s}\rangle$. Then, implement the quantum process $\mathcal{E}_i$ and apply the second unitary operation U_2. Afterward, a projective measurement is performed in the computational basis, yielding an outcome denoted as $\boldsymbol{k}$. Through repeated iterations of this procedure, two probability distributions are obtained: $\mathrm{Pr}^{(1)}_{K|\boldsymbol{s},U_1,U_2}$ and $\mathrm{Pr}^{(2)}_{K|\boldsymbol{s},U_1,U_2}$, representing the measurement outcomes for the fixed computational state and unitaries. By sampling the computational states and repeatedly sampling the unitaries, two probability distributions denoted as $\mathrm{Pr}^{(i)}_{K,S|U_1,U_2}$ are obtained, accounting for the sampled unitaries. For simplicity, we refer to $\mathrm{Pr}^{(i)}_{K,S|U_1,U_2}$ as $\mathrm{Pr}^{(i)}_U$.

Step 3. Post-processing: From the experimental data, we estimate the overlap between the Choi states η_i and η_j for $i, j = 1, 2$ as

$$\mathrm{Tr}[\eta_i\eta_j] = 4^n \sum_{\boldsymbol{s},\boldsymbol{s}',\boldsymbol{k},\boldsymbol{k}'\in\{0,1\}^n} (-2)^{-\mathcal{D}[\boldsymbol{s},\boldsymbol{s}']-\mathcal{D}[\boldsymbol{k},\boldsymbol{k}']} \times \overline{\mathrm{Pr}_U^{(i)}[\boldsymbol{s},\boldsymbol{k}]\mathrm{Pr}_U^{(j)}[\boldsymbol{s}',\boldsymbol{k}']}.$$

where $\overline{\cdots}$ denotes the ensemble average over the sampled unitaries U_1 and U_2. By setting different i and j, we can estimate Using the estimated quantities, we compute the max process fidelity $F_{\max}(\mathcal{E}_1, \mathcal{E}_2)$ in Eq. (1).

There are notable points regarding our protocol. Firstly, it enables comparison between experimentally implemented processes and theoretical simulations when classical simulation is available, facilitating experiment-theory comparisons. Secondly, our protocol estimates the process purity $\mathrm{Tr}[\eta_{\mathcal{E}}^2]$ of a quantum process $\mathcal{E}$. This efficient estimation is crucial for characterizing quantum processes. Lastly, the versatility of our protocol lies in its ability to extend to various metrics [4] based on the overlap $\mathrm{Tr}[\eta_{\mathcal{E}_1}\eta_{\mathcal{E}_2}]$ and purities $\mathrm{Tr}[\eta_{\mathcal{E}_i}^2]$ ($i = 1, 2$). This adaptability makes it applicable to diverse quantum computing scenarios.

3 Experiments

We utilize our ancilla-free protocol to assess the performance of `H` and `CNOT` gates implemented on **seven** distinct platforms freely accessible over the internet: *ibmq_quito* (IBM_1), *ibmq_oslo* (IBM_2), *ibmq_lima* (IBM_3), *ibm_nairobi* (IBM_4), *ibmq_manila* (IBM_5), *baidu_qianshi* (BD_1), and an *ideal simulator* (IDEAL) for theory comparison.

The performance of the single-qubit `H` gate is illustrated in Figure 2(a). Here, we create $2^1 \times N_U = 20$

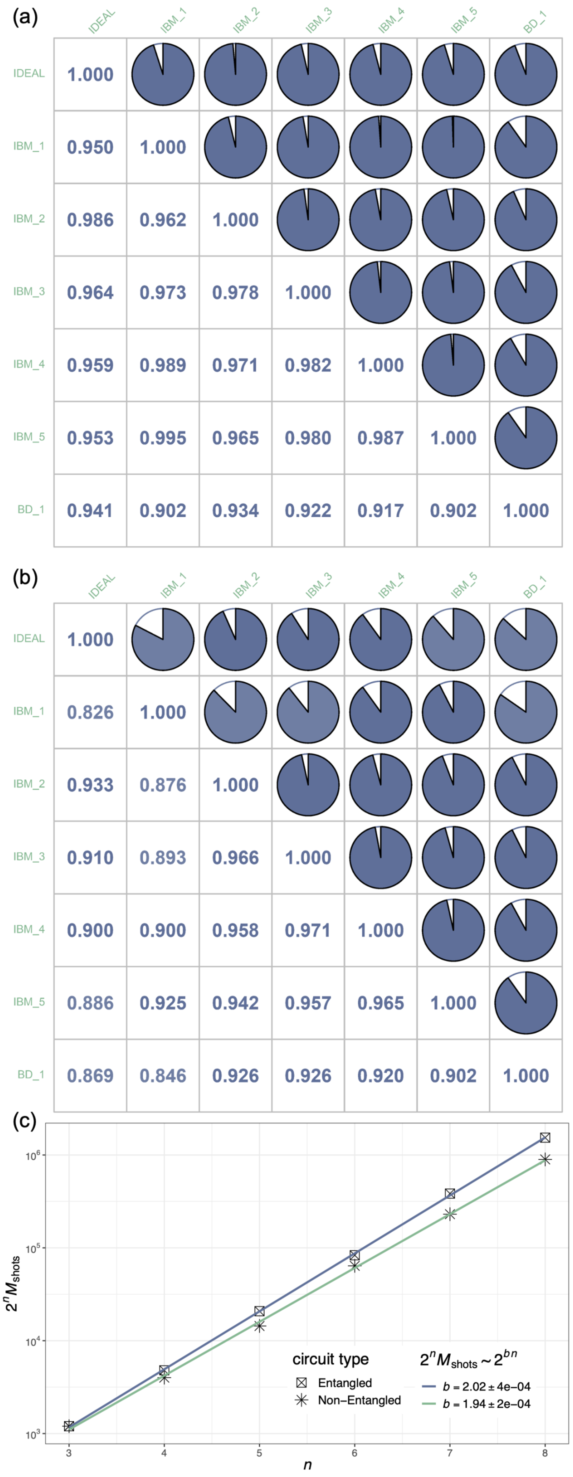


Figure 2: The performance matrices for the single-qubit `H` (a) and two-qubit `CNOT` (b) gates generated from seven different quantum platforms. The entry in the i-th row and j-th column of the matrix represents the max process fidelity between platform-i and platform-j. (c) Scaling of the minimal number of required experimental runs.

circuits and execute $M_{\text{shots}} = 500$ shots for each circuit. We also employ the same protocol to compare the performance of the `CNOT` gate, with $2^2 \times N_U = 400$ and $M_{\text{shots}} = 500$. The result is in Figure 2(b).

The experimental results make it clear that, while some devices may achieve fidelities that are comparable to the ideal simulator, there remains a significant discrepancy between them. *This emphasizes the importance of directly comparing the performance of quantum devices with each other.*

Scaling of sample complexity. We investigate the scaling of the required number of experimental runs, $2^n M_{\text{shots}}$, per unitary to estimate the max fidelity $\widetilde{F}_{\max}$ within an average statistical error of $\epsilon = 0.05$ while fixing N_U to 100. We employ our protocol to two types of quantum processes: (i) a highly entangled quantum process corresponding to an n-qubit GHZ state preparation circuit (Entangled) and (ii) a completely local quantum process composed of n single-qubit rotation gates (Non-Entangled). The result presented in Figure 2(c). The data shows that our protocol has a sample complexity that scales as $2^n N_U M_{\text{shots}} \sim 2^{bn}$ with $b \approx 2$. This scaling, despite exponential, is significantly less than full quantum process tomography (QPT), which has an exponent $b \geq 4$.

Acknowledgements. This work is partially supported by the Fundamental Research Funds for the Central Universities 2242022k60001.

References

[1] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[2] Andreas Elben, Benoît Vermersch, Rick van Bijnen, Christian Kokail, Tiff Brydges, Christine Maier, Manoj K Joshi, Rainer Blatt, Christian F Roos, and Peter Zoller. Cross-platform verification of intermediate scale quantum devices. *Physical Review Letters*, 124(1):010504, 2020.

[3] Andreas Elben, Steven T. Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller. The randomized measurement toolbox. *Nature Reviews Physics*, 5(1):9–24, December 2022.

[4] Alexei Gilchrist, Nathan K Langford, and Michael A Nielsen. Distance measures to compare real and ideal quantum processes. *Physical Review A*, 71(6):062310, 2005.

Faster quantum sampling of Markov chains in non-regular graphs with fewer qubits

Xinying Li[1 2] Yun Shang[1 3 *]

[1] *Institute of Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*

[2] *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[3] *NCMIS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190,China*

Abstract. Sampling from the stationary distribution is one of the fundamental tasks of Markov chain-based algorithms. Quantum sampling from Markov chains corresponds to preparing quantum states with amplitudes arbitrarily close to the square root of a stationary distribution instead of classical sampling from a stationary distribution. A new qsampling algorithm for all reversible Markov chains is constructed by discrete-time quantum walks. In non-regular graphs, the invocation of the quantum fast-forward algorithm accelerates previous state-of-the-art quantum sampling algorithms for both discrete-time and continuous-time cases, especially on sparse graphs. In regular graphs, our results match other quantum algorithms, and the reliance on the gap of Markov chains achieves quadratic speedup compared with classical cases. In some widely used graphs and a series of sparse graphs where stationary distributions are difficult to reach quickly, our algorithm is the first algorithm to achieve complete quadratic acceleration (without log factor) over the classical case without any limit.

Keywords: Quantum sampling, Markov chains, nonregular graphs, ancilla qubits, quantum fast forwarding

1 Introduction

Random walks or Markov chains which can simulate the dynamics of a particle moving randomly on some graphs are a powerful algorithmic model for classical computer science. It can be used to solve problems in machine learning, combinatorial optimization and network science. Similarly, quantum walks which simulate the quantum coherent dynamics of a particle moving on a graph present a powerful and universal framework for designing and creating new faster quantum algorithms.

Sampling from stationary distribution of Markov chain is one of the fundamental tasks of Markov chain-based algorithm, since many problems can be reduced to sampling [1], such as estimating the partition function in counting problems, calculating the importance of Internet pages (the goal of PageRank algorithm), etc.

Mixing time is the minimum time to achieve stationary distribution from any initial distribution. Currently, the quadratic speedup in quantum mixing time is only known for some special graphs [2] or stationary distributions [3]. No general results are obtained for the more general graphs corresponding to any reversible Markov chain.

So far the classical time complexity of reversible Markov chain is bounded by $\delta(P)^{-1}\log(\pi_{min}^{-1})$ [4, 5], where $\delta(P)$ denotes the eigenvalue gap of transition matrix P, π_g denotes the probability of vertex g in stationary distribution, $\pi_{min} := \min_{g\in V}\pi_g$. Correspondingly, the general quantum result can be described with an upper bound $\delta(P)^{-1/2}\pi_g^{-1/2}$ [6]. Although gap can reach quadratic speedup relative to classical case, the quantum mixing time is extremely dependent on π_{min}. Consider $G = (V, E)$ with large size n, then quantum result brings extra $\sqrt{n}$ cost compared with classical case and largely depends on the graphic size.

Several known results of preparing quantum counterpart of stationary distribution in special cases are listed as follows. A slowly evolving sequence of Markov chains are constructed in simulated annealing algorithm [7, 8, 9] and the total complexity is the product of $\sqrt{\delta(P)}$ and sequence length. By introducing a special Markov chain with the form of slowly evolving sequences, Orsucci et al. introduce a quadratic speedup, which improves the additional dependency of $\pi_{min}^{-1/2}$ from $\sqrt{n}$ to $\sqrt[4]{n}$ [10]. The time from single-vertex initial state $|g\rangle$ to stationary distribution with known π_g is $\Theta(\sqrt{t_{hit}})$ in continuous-time case [11], where t_{hit} is the classical hitting time bound of G for any vertex in V. The time to reach a long-time average probability distribution of quantum walk is also considered [12], which needs to prepare a limiting state, different from the quantum form stationary distribution.

These results are mainly focused on special Markov chains or special stationary distribution. However, in general case, whether the quantum state corresponding to the stationary distribution can be obtained with the quadratic acceleration is still an open question.

In this paper, a new qsampling algorithm for all reversible Markov chains is constructed by discrete-time quantum walks. Our work relies on two key technical methods. The first is the quantum interpolated walks method [13], which shows how to prepare a state having constant overlap with the stationary state. Specifically, an appropriate interpolation parameter s is chosen such that both the initial state and the stationary state have constant overlap with the 1-eigenvector (the eigenvector corresponds to the eigenvalue 1) of quantum interpolated

[*] shangyun@amss.ac.cn

walk operator. The second technique we use is quantum fast-forwarding [14, 15]. It can simulate the dynamics of classical random walks by using quantum walk operator, which keeps the stationary state evolving like classical process while weakens the effect of other states, namely distinguishing 1-eigenvectors from other eigenvectors. As a result, quadratic speedup relative to the classical case can be achieved.

2 Technical Overviews

We divide two cases to discuss. In general, only regular graphs have known π_g for a randomly chosen vertex g without requiring any additional information since all vertices have equal probability, but non-regular graphs with unknown π_g are often encountered. Then it is particularly important to prepare the stationary state for non-regular graphs. However, if using the reverse of search algorithm in [13] to determine unknown π_g, the total complexity will be $\Theta(\log(\pi_{min}^{-1})\sqrt{\mathrm{HT}}\log(\frac{1}{\varepsilon}))$, and increases at least additional $\log n$ compared with known π_g case. Here we construct a new algorithm to determine the value of π_g, which reduces the additional cost to $\Theta(\log(\pi_{min}^{-1}/n))$. The additional cost will become constant in most sparse graphs that satisfy $\pi_{min} = \Theta(1/n)$. Furthermore, by quantum fast-forwarding technique, a new reflection around the stationary state with fewer ancilla qubits is constructed to measure, which maintains quadratic speedup compared to classical scale δ^{-1} and reduces ancilla qubits from $\Theta(\frac{1}{\sqrt{\delta}}\log(\frac{1}{\varepsilon_2}))$ to $\Theta(\log(\frac{1}{\sqrt{\delta}})+\log\log(\frac{1}{\varepsilon_2}))$ compared with the method by phase estimation [7]. We believe that it may have independent applications.

In a word, according to π_g known or unknown, we divide two cases to construct a stationary state. For unknown π_g, we need to determine the value of π_g first. Then the rest can be reduced to preparing stationary state $|\pi\rangle$ with known π_g.

To construct a state ε-close to the target state $|\pi\rangle$ with constant success probability, we combine quantum fast-forwarding algorithm with the interpolated walk when constructing U_{main} operator (some kind of interpolated quantum walk unitary operator) according to initial state $|g\rangle$. An appropriate interpolated parameter s is chosen such that both the initial state and the stationary state have constant overlap with the 1-eigenvector of quantum interpolated walk operator $D(s)$. The method of quantum fast-forwarding is used to distinguish 1-eigenvector from other eigenvectors by using the quantum walk operator to simulate the dynamics of classical random walk, which reduces the dependence on π_{min} and ε. Besides, the number of ancilla qubits required is reduced as well. Then a state that has constant overlap with $|\pi\rangle$ will be prepared. At last, we apply a projection operator Π_π to generate a state that is ε-close to $|\pi\rangle$ with constant success probability.

3 Main results

Theorem 1 (Informal statement of main results) *For an ergodic reversible Markov chain P in $G(V,E)$, there exists an algorithm that generates a state ε-close to $|\pi\rangle$ with constant success probability from an initial state of the form $|g\rangle$, where g is any randomly selected vertex in graph G.*

1. For non-regular graphs (the case with unknown π_g), the complexity is $\Theta(\log(\pi_{min}^{-1}/n)\sqrt{\mathrm{HT}}\log(\varepsilon^{-1}))$ with $\Theta(\log(\pi_{min}^{-1}/n)(\log\sqrt{\mathrm{HT}}+\log\log(\frac{1}{\varepsilon})))$ ancilla qubits.

2. For regular graphs (the case with known π_g), the complexity is $\Theta(\sqrt{\mathrm{HT}}\log(\varepsilon^{-1}))$ with $\Theta(\log(\sqrt{\mathrm{HT}})+\log\log(\frac{1}{\varepsilon}))$ ancilla qubits.

Formal results can be found in Theorem 3 and Theorem 4 [16].

4 Results comparison and applications

Compared with the existing works, we build a qsampling algorithm that not only accelerates qsampling in non-regular graphs but also maintains the speed-up of existing quantum algorithms in regular graphs with fewer ancilla qubits. In non-regular graphs, the invocation of the quantum fast-forwarding algorithm accelerates previous state-of-the-art qsampling algorithms for both discrete-time and continuous-time cases, especially on sparse graphs. In regular graphs, our results match other quantum algorithms, and the reliance on the gap of Markov chains achieves quadratic speedup compared with classical cases. For both cases, we reduce the number of ancilla qubits required compared with the existing results. In some widely used graphs and a series of sparse graphs where stationary distributions are difficult to reach quickly, our algorithm is the first algorithm to achieve complete quadratic acceleration (without log factor) over the classical case without any limit. Our main results are summarized in Table 1 and applications on some sparse graphs can refer to Table 2.

References

[1] Aharonov D. & Ta-Shma A. Adiabatic quantum state generation and statistical zero knowledge. In *STOC*, 20–29(2003).

[2] Richter P. C. Quantum speedup of classical mixing processes. *Phys. Rev. A*, 76:042306(2007).

[3] Dunjko V.& Briegel H. Quantum mixing of markov chains for special distributions. *New J. Phys.*, 17(7):073004(2015)

[4] D. Aldous, Some inequalities for reversible markov chains, J. Lond. Math. Soc. **s2-25**(3), 564-576(1982).

[5] M. Jerrum, Counting, Sampling and Integrating: Algorithms and Complexity (2003).

Case		Our result	Best previous quantum results	
			Discrete-time[13][1]	**Continuous-time**[11]
non-regular (unknown π_g)	**Complexity**	$\Theta(\log(\pi_{min}^{-1}/n)\sqrt{\mathrm{HT}}\log(\frac{1}{\varepsilon}))$	$\Theta(\log(\pi_{min}^{-1})\sqrt{\mathrm{HT}}\frac{1}{\varepsilon})$	\
	Ancilla qubit	$\Theta\left(\log(\pi_{min}^{-1}/n)(\log\sqrt{\mathrm{HT}}+\log\log(\frac{1}{\varepsilon}))\right)$	$\Theta\left(\log(\pi_{min}^{-1})(\log\sqrt{\mathrm{HT}}+\log\frac{1}{\varepsilon})\right)$	\
regular (known π_g)	**Complexity**	$\Theta(\sqrt{\mathrm{HT}}\log(\frac{1}{\varepsilon}))$	$\Theta(\sqrt{\mathrm{HT}}\frac{1}{\varepsilon})$	$\Theta(\sqrt{\mathrm{HT}}\log(\frac{1}{\varepsilon}))$
	Ancilla qubit	$\Theta(\log\sqrt{\mathrm{HT}}+\log\log(\frac{1}{\varepsilon}))$	$\Theta(\log\sqrt{\mathrm{HT}}+\log\frac{1}{\varepsilon})$	$\Theta(\log\sqrt{\mathrm{HT}}\cdot\log(\frac{1}{\varepsilon}))$

Table 1: Summary of our main results. The corresponding classical complexity is $\Theta(\frac{1}{\delta}(\log(\pi_{min}^{-1})+\log(\frac{1}{\varepsilon})))$ for all graphs. Here δ denotes the eigenvalue gap of Markov chain P, $\pi_{min} := \min_{g\in V}\pi_g$ denotes the smallest probability component in stationary state, HT is the classical hitting time bound of G for any vertex in V, and ε denotes the error. Formal results can be found in Theorem 3 and Theorem 4 [16].

Graph	Classical mixing time	Our result	Best previous result
balanced r-tree	n	$\sqrt{n}\log n$	$\sqrt{n}\log^{3/2} n$
barbell	n^3	$n^{3/2}\log n$	$n^{3/2}\log^2 n$
circle	n^2	n	n
necklace	n^2	n	$n\log n$
two cliques glued at a single vertex	$n^2\log n$	n	$n\log n$
G(n, 1/n)	n	$\sqrt{n}$	$\sqrt{n}\log n$

Table 2: The tight bound quantitative results of many graphs. All above results are $\Theta(\cdot)$, where the detailed definition of the first four graphs and the corresponding hitting time and mixing time results can be found in [17], the fifth graphs can be found in [18] and the last graphs is in [19]. The corresponding best previous quantum results are from the result of [13, 11].

[6] Magniez F., Nayak A., Roland J. & Santha M. Search via quantum walk. *SIAM J. Comput.*, 40(1):142–164(2011).

[7] Wocjan P. & Abeyesinghe A. Speedup via quantum sampling. *Phys. Rev. A*, 78:042336(2008).

[8] Somma R. D., Boixo, S., Barnum H. & Knill E. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, 101 13:130504(2008).

[9] Harrow A. & Wei A. Y. Adaptive Quantum Simulated Annealing for Bayesian Inference and Estimating Partition Functions *SODA*, 193-212(2020).

[10] Orsucci D., Briegel H. & Dunjko V. Faster quantum mixing for slowly evolving sequences of Markov chains. *Quantum*, 2:105(2018).

[11] Shantanav C., Kyle L. & Jérémie R. Analog quantum algorithms for the mixing of markov chains. *Phys. Rev. A*, 102:022423(2020).

[12] Chakraborty S., Luh K. & Roland J. How fast do quantum walks mix? *Phys. Rev. Lett.*, 124 5:050501(2020).

[13] Krovi H., Magniez F., Ozols M. & Roland J. Quantum walks can find a marked element on any graph. *Algorithmica*, 74:851-907(2015).

[14] Apers S. & Sarlette A. Quantum fast-forwarding: Markov chains and graph property testing. *Quantum Inf. Comput.*, 19:181-213(2018).

[15] A. Ambainis, A. Gilyén, S. Jeffery, and M. Kokainis. Quadratic speedup for finding marked vertices by Quantum walks, 52nd Annual ACM SIGACT Symposium on Theory of Computing (2020).

[16] Xinying Li.& Yun Shang. Faster quantum sampling of Markov chains in non-regular graphs with fewer qubits *Physical Review A*, 107, 022432(2023).

[17] D. Aldous and J. A. Fill. Reversible Markov chains and random walks on graphs, Unfinished monograph (2002).

[18] S. Apers, A. Gily'en, and S. Jeffery. A Unified Framework of Quantum Walk Search, STACS **187**, **6**:1-13, (2021).

[19] Nachmias A. & Peres Y. Critical random graphs: Diameter and mixing time. *Annals of Probability*, 36:1267-1286(2008).

Hamiltonian formulation of unsupervised data clustering

David M. Seong[1] * Daniel K. Park[1 2] †

[1] *Department of Statistics and Data Science, Yonsei University, Seoul, 03722, Republic of Korea*
[2] *Department of Applied Statistics, Yonsei University, Seoul, 03722, Republic of Korea*

Abstract. Clustering is a fundamental task in data mining that aims to group data based on their similarities. Defining similarity is often ambiguous, making it challenging to determine the most appropriate objective function for describing the data. As a result, various clustering methods, such as the k-means algorithm and weighted max k-cut, have been developed, each with its own objectives and data-dependent performance. Additionally, these problems are known to be NP-complete. In this study, we propose a novel approach that formulates the clustering problem as a search for the ground state of a Hamiltonian. Within this framework, clustering objectives are tailored by designing the Hamiltonian. Moreover, our approach seamlessly incorporates constraints, unlike previous clustering algorithms that mainly focus on unconstrained problems. Our method provides an opportunity to leverage quantum simulation techniques for clustering problems, customized to the target data and underlying goals. We propose multiple Hamiltonians, each representing a distinct objective, and evaluate the clustering performance using well-known metrics such as the Silhouette score and Rand index through numerical simulations. The simulation results demonstrate the broad applicability of our method to various clustering problems.

Keywords: Quantum Machine Learning, Clustering, Quantum Optimization, Quantum Simulation

1 Introduction

Quantum machine learning (QML) offers new possibilities and approaches to address the various challenges in data mining, pushing the boundaries of existing methods. Among its potential applications, clustering stands out as a widely used technique in various domains of pattern recognition and data mining, such as image recognition [1], social network analysis [2], biological data analysis [3], customer segmentation in marketing [4], and anomaly detection [5]. However, clustering encounters several challenges [6,7]. Firstly, selecting an appropriate clustering objective without prior knowledge of the target dataset can be challenging. Moreover, most clustering algorithms are designed to handle datasets with a single data type, either categorical or numerical. Consequently, some data points may not fit well due to the inappropriate choice of similarity measures, resulting in cluster assignment errors. Thus, the selection of similarity measures and clustering objectives needs to be tailored to the specific characteristics of the dataset. Consequently, well-known clustering algorithms such as the k-means algorithm and the weighted max k-cut algorithm exhibit data-dependent performance. Another challenge is the scalability of clustering problems, as finding the global solution is known to be NP-complete.

This work introduces a novel approach to address the clustering problem by formulating it as a search for the ground state of a Hamiltonian. This formulation offers flexibility and a unified framework for clustering, allowing customization for different objectives and constraints. Specifically, we map the task of grouping N real data points $X \in \mathbb{R}^d$ into two clusters to a quadratic unconstrained binary optimization (QUBO) problem. The binary variable $z \in \{-1, +1\}^N$ is used to represent the assignment of data points to clusters, and a feature representation of data, $V = \phi(X) \in \mathbb{R}$, is utilized. The number of data points assigned to each cluster can then be computed as

$$N_{+1} = \sum_{i=1}^{N} \frac{1+z_i}{2} \qquad N_{-1} = \sum_{i=1}^{N} \frac{1-z_i}{2}. \tag{1}$$

These elements are the building blocks for constructing the desired objective function along with any necessary constraints. The corresponding Hamiltonian is obtained by replacing z_i with the Pauli Z operator and 1 with the identity operator on the i^{th} qubit. This Hamiltonian formulation of clustering offers the potential for more efficient solution finding on a quantum computer. It can be solved using existing quantum simulation techniques, such as quantum annealing [8], quantum approximate optimization algorithm [9], or variational quantum eigensolver [10–12], on a quantum device, leading to a potential quantum advantage.

To evaluate the performance of our approach, we compare it to the widely used k-means algorithm using the Iris dataset as a benchmark.

2 Methods

Intracluster Distance We begin by demonstrating that the well-known k-means algorithm can be reformulated as a Hamiltonian problem. The objective of the k-means clustering algorithm is to identify the optimal set of clusters that minimizes the sum of distances between each cluster's center (centroid) and the data points assigned to that cluster. The optimization problem can be

*europa0306@yonsei.ac.kr
†dkd.park@yonsei.ac.kr

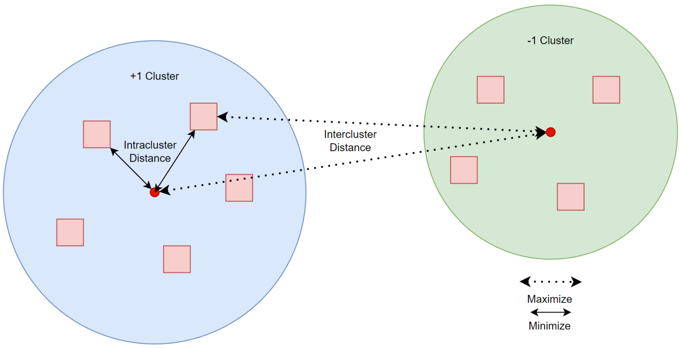


Figure 1: Illustration of the intracluster distance and the intercluster distance.

defined as follows:

$$\min_{z\in\{-1,+1\}^N} \sum_{i=1}^{N} \|V_i - \frac{1}{N_{+1}} \sum_{j}^{N} V_j \frac{1+z_j}{2} \|_2^2 \frac{1+z_i}{2} + \sum_{i=1}^{N} \|V_i - \frac{1}{N_{-1}} \sum_{j=1}^{N} V_j \frac{1-z_j}{2} \|_2^2 \frac{1-z_i}{2}. \quad (2)$$

Intracluster and Intercluster Distances The k-means algorithm focuses solely on intracluster distance, disregarding intercluster distance. Figure 1 illustrates the distinction between intracluster and intercluster distances. However, incorporating intercluster distance into the objective can enhance clustering results for certain datasets. This can be done easily in our framework. A proper formulation of this is achieved by adding the following term to the objective in Equation (2):

$$-\sum_{i=1}^{N} \|V_i - \frac{1}{N_{-1}} \sum_{j}^{N} V_j \frac{1-z_j}{2} \|_2^2 \frac{1+z_i}{2} - \sum_{i=1}^{N} \|V_i - \frac{1}{N_{+1}} \sum_{j=1}^{N} V_j \frac{1+z_j}{2} \|_2^2 \frac{1-z_i}{2}. \quad (3)$$

In this formulation, the intercluster distance is represented by the distance between the centroid of one cluster and a data point in another cluster.

Pairwise Distance In addition to constructing the objective based on the distance between centroids and data points, the pairwise distance between data points can also be taken into account. For example, one can add the following maximization problem to any objectives discussed thus far.

$$\max_{z\in\{-1,+1\}^N} \sum_{i=1}^{N} \sum_{j=1}^{N} \|V_i \frac{1+z_i}{2} - V_j \frac{1-z_j}{2} \|_2^2 \quad (4)$$

In this formulation, the distance between data points in different clusters is computed by considering the difference between their feature representations, denoted as V_i.

Alternatively, pairwise distance can be computed using pre-computed similarity values between data points based on standard measures like the Euclidean distance, instead of utilizing V_i representations. This leads to an optimization problem of the form:

$$\max_{z\in\{-1,+1\}^N} \sum_{i<j}^{N} w_{ij} z_i z_j. \quad (5)$$

Here, $w_{ij} \geq 0$ represents the similarity measure between the i^{th} and j^{th} data points. This is also known as the weighted max-cut problem on a graph.

Mixture Models A comprehensive objective can be formed by combining the discussed objectives to incorporate multiple considerations. For instance, a new objective can be defined as the search for the ground state energy of the Hamiltonian that combines the weighted max-cut objective and a customized Hamiltonian H_c. The hyperparameter $\lambda > 0$ controls the relative importance of each term in the objective:

$$-\sum_{i<j}^{N} w_{ij} Z_i Z_j + \lambda H_c. \quad (6)$$

Cardinality Constraint In certain applications, clustering with a designated number of data per cluster, known as a cardinality constraint, is desirable. In our framework, this constraint can be incorporated by augmenting the Hamiltonian with a penalty term:

$$H_c + \lambda(C - \sum_{i}^{N} z_i)^2. \quad (7)$$

Here, C represents the desired difference in the number of data points between two clusters.

3 Results

To evaluate the effectiveness of the customized Hamiltonians for clustering, we employed a brute force algorithm to determine the ground state of the Hamiltonian. The analysis involved 150 independent datasets, each consisting of 16 data points randomly sampled from the Iris dataset. We compared the performance of our method utilizing different Hamiltonians against the k-means algorithm. We used two standard metrics, namely the Silhouette score and the Rand index, to measure the quality of the clustering results. In terms of the Silhouette score, which serves as an internal validation metric, our method generally yielded lower scores compared to the k-means algorithm when utilizing the Euclidean distance as the similarity measure. However, in certain samples, the Hamiltonian approach surpassed the performance of the k-means algorithm, as indicated by the higher values of the external validation metric, the Rand index. This result suggests that our method with customized Hamiltonians have the potential to achieve more accurate clustering outcomes in specific scenarios.

Numerical analysis also showed that the Hamiltonian approach can successfully find clusters with the correct

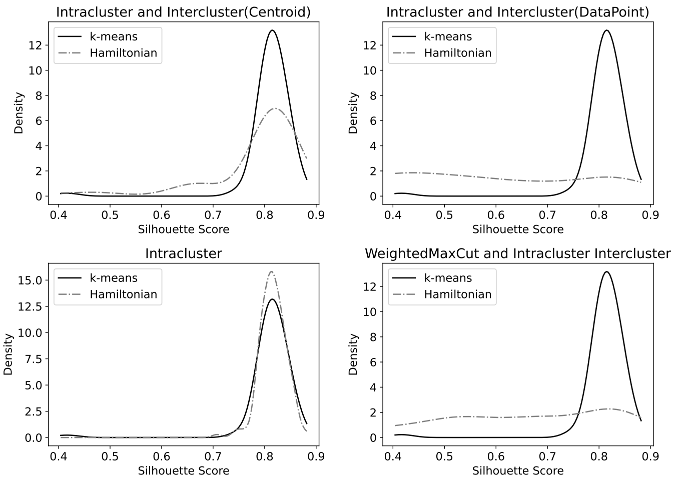


Figure 2: Comparison of k-means and customized Hamiltonian using Silhouette Score: KDE analysis

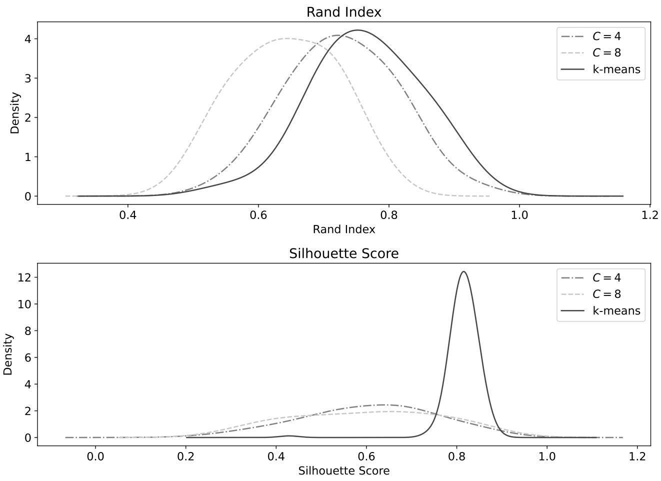


Figure 4: Rand Index and Silhouette Score for clustering with cardinality constraints. Note that the k-means algorithm does not support the incorporation of constraints. Despite yielding lower scores, the Hamiltonian-based clustering method effectively fulfills the specified constraints.

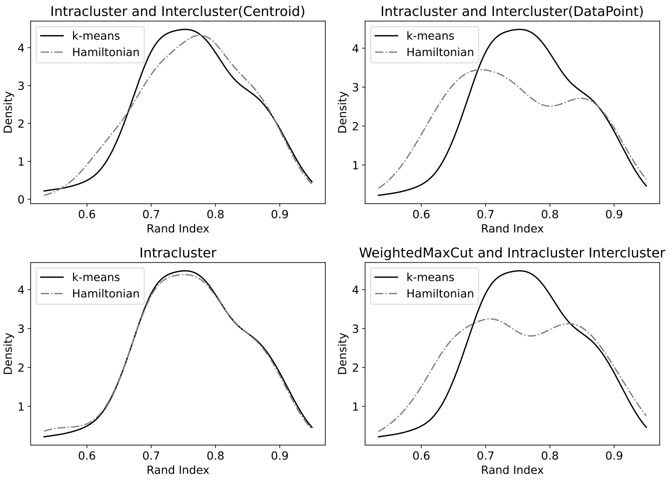


Figure 3: Comparison of k-means and customized Hamiltonian using Rand Index: KDE analysis

number of data as demanded by the cardinality constraint. We conducted tests on 24 and 12 data sets with $C = 4$ and $C = 8$, respectively, and found that all clusters identified by the Hamiltonian approach satisfied the constraint while achieving high Rand Index and Silhouette score. This feature is not supported by the k-means algorithm.

4 Conclusions and Discussions

The Hamiltonian formulation of diverse clustering objectives presented in this work enables customization for data- and context-specific objectives and constraints, offering a flexible and unified approach to address complex clustering problems. The hyperparameter λ can be fine-tuned using other machine learning techniques to further improve the performance of Hamiltonian-based clustering. The Hamiltonian approach in clustering benefits from the computational advantages offered by quantum simulation algorithms. Existing heuristic quantum simulation methods, such as quantum annealing, QAOA, and VQE, that are well-suited for the noisy intermediate scale quantum (NISQ) devices do not guarantee the finding of global minima in polynomial time. Nonetheless, the potential advantage of quantum simulators in finding local optima of a Hamiltonian in certain cases opens up opportunities for achieving quantum advantage in clustering. Additionally, in quantum annealing and QAOA, the annealing time or the quantum circuit depth can be controlled, allowing for flexibility in balancing runtime and clustering performance. The development of more efficient quantum simulation methods suitable for NISQ technology holds promise for further improvements in clustering. Another notable feature of the Hamiltonian approach is that the clustering performance is not constrained by initial values, whereas the performance of the k-means clustering algorithm is highly dependent on the initial guess of cluster centers.

Extending the Hamiltonian-based method to handle arbitrary k clusters is an important avenue for future research. One potential solution is to leverage a binary matrix representation, as proposed by Peng et al. [13], where each entry indicates whether a data point belongs to the k-th cluster or not. Determining the optimal number of clusters, denoted as k, without prior knowledge is challenging and greatly affects clustering results. Recent research has focused on auto-clustering [14] to address this issue. A promising future direction is to utilize the Hamiltonian approach to optimize the determination of k, enabling a data-driven and automated approach that reduces the need for manual determination of the number of clusters. Additionally, investigating the performance of this approach with larger datasets and more complex clustering problems would be valuable.

References

[1] Andrea Baraldi and Palma Blonda. A survey of fuzzy clustering algorithms for pattern recognition. i. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 29(6):778–785, 1999.

[2] Mark S Handcock, Adrian E Raftery, and Jeremy M Tantrum. Model-based clustering for social networks. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(2):301–354, 2007.

[3] Sara C Madeira and Arlindo L Oliveira. Biclustering algorithms for biological data analysis: a survey. *IEEE/ACM transactions on computational biology and bioinformatics*, 1(1):24–45, 2004.

[4] Jing Wu and Zheng Lin. Research on customer segmentation model by clustering. In *Proceedings of the 7th international conference on Electronic commerce*, pages 316–318, 2005.

[5] Shikha Agrawal and Jitendra Agrawal. Survey on anomaly detection using data mining techniques. *Procedia Computer Science*, 60:708–713, 2015.

[6] Anil K Jain. Data clustering: 50 years beyond k-means. *Pattern recognition letters*, 31(8):651–666, 2010.

[7] Dongkuan Xu and Yingjie Tian. A comprehensive survey of clustering algorithms. *Annals of Data Science*, 2:165–193, 2015.

[8] Mark W Johnson, Mohammad HS Amin, Suzanne Gildert, Trevor Lanting, Firas Hamze, Neil Dickson, Richard Harris, Andrew J Berkley, Jan Johansson, Paul Bunyk, et al. Quantum annealing with manufactured spins. *Nature*, 473(7346):194–198, 2011.

[9] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014.

[10] 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.

[11] Jarrod R McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, feb 2016.

[12] 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(9):625–644, 2021.

[13] Jiming Peng and Yu Wei. Approximating k-means-type clustering via semidefinite programming. *SIAM journal on optimization*, 18(1):186–205, 2007.

[14] Absalom E Ezugwu, Amit K Shukla, Moyinoluwa B Agbaje, Olaide N Oyelade, Adán José-García, and Jeffery O Agushaka. Automatic clustering algorithms: a systematic review and bibliometric analysis of relevant literature. *Neural Computing and Applications*, 33:6247–6306, 2021.

Single-copy nonlocality activation in a quantum photonic network

Luis Villegas-Aguilar[1 2 *] Kiarn Laverick[2] Farzad Ghafari[1] Emanuele Polino[1]
Marco T. Quintino[3] Sergei Slussarenko[1] Nora Tischler[1] Eric G. Cavalcanti[2]
Geoff J. Pryde[1]

[1] *Centre for Quantum Dynamics, Griffith University, Brisbane, Queensland 4111, Australia*
[2] *Centre for Quantum Dynamics, Griffith University, Gold Coast, Queensland 4222, Australia*
[3] *Sorbonne Université, CNRS, LIP6, F-75005 Paris, France*

Abstract. Despite entanglement being a prerequisite to observe Bell nonlocality, determining whether a single instance of an entangled quantum state can lead to nonlocal correlations remains an important, yet remarkably challenging problem. In this work, we experimentally demonstrate that a single copy of a photonic Bell-local state can have its nonlocality activated when it is embedded in a quantum network. Our four photon scheme, based on quantum broadcasting, provides a robust certification of the non-classical nature of mixed entangled states for noise proportions beyond what is possible for two parties.

Keywords: Bell nonlocality, quantum networks, quantum optics, single photons

1 Motivation

While all pure entangled states display nonlocal correlations [1, 2], some mixed entangled states [3, 4] are unable to violate a Bell inequality in the standard two-party scenario [5]. The motivation for understanding the relationship between entanglement and nonlocality is twofold: on one hand it provides foundational insights into quantum theory; on the other, nonclassical correlations are at the heart of many quantum technologies [6].

An area that has received significant attention in recent years is that of quantum networks, which seek to establish quantum communication channels between multiple distant nodes. Quantum networks [7] provide considerably different generalizations of nonlocality [8, 9, 10, 11] that can be exploited for multi-party secure communication [12] and cryptographic protocols [13, 14]. These developments highlight the importance of understanding and mitigating the impact of noise in quantum correlations in the context of networks [15].

An important discovery was that nonlocality is a resource that can be *activated*; that is, nonlocal behaviour can be recovered for local states when they are subjected to more exotic measurement procedures [16]. Quantum networks provide an avenue for activating Bell-local states [17, 18]. In a typical network scenario, one is allowed to distribute several copies of a local state independently to a number of distant parties, in a situation closely related to entanglement distillation [19]. By taking a sufficiently large number of copies, any distillable state can, in principle, be activated [20] in the asymptotic limit [21]. The associated resource overhead, however, quickly becomes impractical for experimental implementations.

In this work, we experimentally demonstrate a scheme for the activation of nonlocality in a photonic three-node network. Our demonstration requires only one instance of a noisy entangled state per measurement round, which is broadcast to two additional separated parties in a network structure, employing the recent theory results of Ref. [22]. Additionally, we certify our activation through a rigorous statistical analysis on our experimentally generated states. We numerically prove the existence of a local hidden variable model for our two-party states before the activating procedure.

Our demonstration does not rely on physical description of the measurement devices used and it can be regarded as a realistic, device-independent (DI) protocol with strengthened noise tolerance.

2 Results

We built a source of high-quality optical isotropic states $\rho_\alpha = \alpha|\Phi^+\rangle\langle\Phi^+| + (1-\alpha)\mathbb{I}_4/4$, where $|\Phi^+\rangle = (|HH\rangle + |VV\rangle)/\sqrt{2}$ is a maximally entangled state, $\mathbb{I}_4$ is the 4×4 the identity matrix, and the parameter $\alpha \in [0,1]$ represents the fractional purity of the state. Our experimental setup allowed for the quantum parameter α to be fully tuned by means of a controllable depolarizing channel. The average measured fidelity of our experimental states with the closest isotropic state was 0.993, with a maximum of 0.997 ± 0.003. These fidelities are comparable to the highest reported ones to date for similar states [23].

A conceptual representation of our activation demonstration is depicted in Fig. 1. Initially, a source prepared a two-qubit state. One of the subsystems was shared with one party, while the other underwent a local transformation applied by a quantum channel Ω, which broadcast it to two additional parties, forming a three-node network. This broadcasting operation was implemented via an optical controlled-NOT gate with an ancilla photon, plus local qubit operations. The gate relied on high-quality nonclassical interference [24] between single photons from independent sources, with an observed visibility of 0.97 ± 0.03.

The Bell inequality that establishes the classical limit for the probability distribution $p(a,b,c|x,y,z)$ in this sce-

[*] luis.villegasaguilar@griffithuni.edu.au

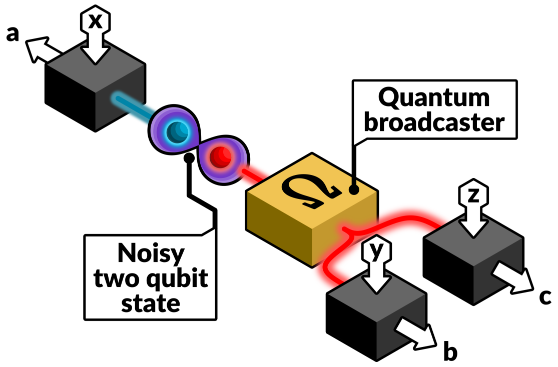


Figure 1: **Conceptual representation of the broadcast Bell scenario.** Half of of a noisy two-qubit state undergoes a quantum broadcasting operation, sharing its outcome with two separate parties, forming a three-node network. Each party receives input x, y and z, and perform local measurements on their qubits, yielding outcomes a, b and c, respectively.

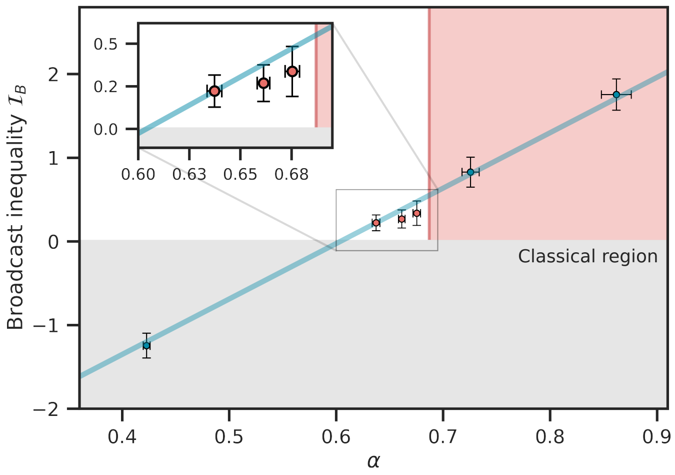


Figure 2: **Experimental broadcast nonlocality as a function of the quantum parameter** α**.** The parameter α represents pure-state fraction of our experimental isotropic states. The solid gray area represents the classical region above which a violation of inequality (1) occurs. The red area represents the upper bound ($\alpha \leq 0.687$) on Bell locality for isotropic states in the two-party scenario, points to the left of this region admit a local hidden variable (LHV) description. The diagonal blue line corresponds to theoretically predicted results. Red circles highlight data points for which activation occurs. The error bars represent two standard deviations total.

nario [22] is

$$\begin{aligned}\mathcal{I}_B =&\langle A_0B_0C_0\rangle + \langle A_0B_1C_1\rangle + \langle A_1B_1C_1\rangle - \langle A_1B_0C_0\rangle \\ &+ \langle A_0B_0C_1\rangle + \langle A_0B_1C_0\rangle + \langle A_1B_0C_1\rangle - \langle A_1B_1C_0\rangle \\ &- 2\langle A_2B_0\rangle + 2\langle A_2B_1\rangle - 4 \leq 0, \qquad (1)\end{aligned}$$

with $\langle A_xB_yC_z\rangle = \Sigma_{a,b,c=0,1}(-1)^{a+b+c}\ p(a,b,c|x,y,z)$ and analogously for the two-party correlator terms. A violation of this inequality certifies the presence of nonlocal correlations from the original source.

We tested this inequality for a number of experimentally produced states with varying degrees of noise as shown in Fig. 2. We show activation for a range of data points below $\alpha \leq 0.687$—the current upper bound for Bell locality of the isotropic state under projective measurements [25]—by violating the inequality $\mathcal{I}_B \leq 0$ with over two standard deviations. Noticeably, our strongest activating result was found at $\alpha = 0.637 \pm 0.004$, which is 12.5 standard deviations below the current theoretical upper bound for LHV models for the isotropic state under projective measurements. For this state, we recorded $\mathcal{I}_B = 0.268 \pm 0.11 > 0$, representing a clear violation of the local bound.

Additionally, we certify our experimental activation in a robust manner by ensuring that our original bipartite states admit a local hidden variable (LHV) description. The aim of our method is to express our experimental density matrices ρ_{exp} as a convex combination of a state with a known LHV and a separable state. This can be framed as a semidefinite programming (SDP) problem, a class of optimisation problems which can be solved computationally in an efficient and precise manner [26]. Using this method, we found a positive answer for our target experimental state, conclusively certifying the activation of its nonlocality by our protocol.

From a fundamental point of view, our results demonstrate that nonlocal behaviour can emerge from individual local states when studied in the context of quantum networks. From a practical perspective, they open up unexplored possibilities for quantum information processing tasks involving noisy states, an inevitable scenario for any realistic implementation of a future quantum internet [27, 28].

We note that, although stronger examples of nonlocality activation are known in multi-copy settings [29, 17], this can be prohibitively hard to achieve in practice when dealing with large ensembles of distributed, independent copies of a quantum state. For instance, to observe activation for a comparable noise proportion ($\alpha \sim 0.64$), one needs at least $N = 21$ copies of the isotropic state in a star network configuration [18].

References

[1] N. Gisin, “Bell's inequality holds for all non-product states,” *Physics Letters A*, vol. 154, pp. 201–202, Apr. 1991.

[2] N. Gisin and A. Peres, “Maximal violation of Bell's inequality for arbitrarily large spin,” *Physics Letters A*, vol. 162, pp. 15–17, Jan. 1992.

[3] R. F. Werner, “Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model,” *Physical Review A*, vol. 40, pp. 4277–4281, Oct. 1989.

[4] J. Barrett, “Nonsequential positive-operator-valued measurements on entangled mixed states do not al-

ways violate a Bell inequality," *Physical Review A*, vol. 65, no. 4, 2002.

[5] J. S. Bell, "On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, vol. 1, pp. 195–200, Nov. 1964.

[6] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, "Bell nonlocality," *Reviews of Modern Physics*, vol. 86, pp. 419–478, Apr. 2014.

[7] A. Tavakoli, A. Pozas-Kerstjens, M.-X. Luo, and M.-O. Renou, "Bell nonlocality in networks," *Reports on Progress in Physics*, vol. 85, p. 056001, Mar. 2022.

[8] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, "Definitions of multipartite nonlocality," *Physical Review A*, vol. 88, p. 014102, July 2013.

[9] A. Pozas-Kerstjens, N. Gisin, and A. Tavakoli, "Full Network Nonlocality," *Physical Review Letters*, vol. 128, p. 010403, Jan. 2022.

[10] I. Šupić, "Genuine network quantum nonlocality and self-testing," *Physical Review A*, vol. 105, no. 2, 2022.

[11] L.-Y. Hsu, "Genuine Bell locality and nonlocality in the networks," Sept. 2022.

[12] J. Ho, G. Moreno, S. Brito, F. Graffitti, C. L. Morrison, R. Nery, A. Pickston, M. Proietti, R. Rabelo, A. Fedrizzi, and R. Chaves, "Entanglement-based quantum communication complexity beyond Bell nonlocality," *npj Quantum Information*, vol. 8, pp. 1–7, Feb. 2022.

[13] A. Pickston, J. Ho, A. Ulibarrena, F. Grasselli, M. Proietti, C. L. Morrison, P. Barrow, F. Graffitti, and A. Fedrizzi, "Experimental network advantage for quantum conference key agreement," July 2022.

[14] F. Grasselli, G. Murta, H. Kampermann, and D. Bruß, "Boosting device-independent cryptography with tripartite nonlocality," Sept. 2022.

[15] B. C. Coutinho, W. J. Munro, K. Nemoto, and Y. Omar, "Robustness of noisy quantum networks," *Communications Physics*, vol. 5, pp. 1–9, Apr. 2022.

[16] S. Popescu, "Bell's Inequalities and Density Matrices: Revealing "Hidden" Nonlocality," *Physical Review Letters*, vol. 74, pp. 2619–2622, Apr. 1995.

[17] A. Sen(De), U. Sen, Č. Brukner, V. Bužek, and M. Żukowski, "Entanglement swapping of noisy states: A kind of superadditivity in nonclassicality," *Physical Review A*, vol. 72, p. 042310, Oct. 2005.

[18] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acín, "Quantum networks reveal quantum nonlocality," *Nature Communications*, vol. 2, p. 184, Feb. 2011.

[19] C. H. Bennett, "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels," *Physical Review Letters*, vol. 76, no. 5, pp. 722–725, 1996.

[20] D. Cavalcanti, A. Acín, N. Brunner, and T. Vértesi, "All quantum states useful for teleportation are nonlocal resources," *Physical Review A*, vol. 87, p. 042104, Apr. 2013.

[21] L. Masanes, "Asymptotic Violation of Bell Inequalities and Distillability," *Physical Review Letters*, vol. 97, no. 5, 2006.

[22] J. Bowles, F. Hirsch, and D. Cavalcanti, "Single-copy activation of Bell nonlocality via broadcasting of quantum states," *Quantum*, vol. 5, p. 499, July 2021.

[23] N. Tischler, F. Ghafari, T. J. Baker, S. Slussarenko, R. B. Patel, M. M. Weston, S. Wollmann, L. K. Shalm, V. B. Verma, S. W. Nam, H. C. Nguyen, H. M. Wiseman, and G. J. Pryde, "Conclusive Experimental Demonstration of One-Way Einstein-Podolsky-Rosen Steering," *Physical review letters*, vol. 121, no. 10, pp. 100401–100401, 2018.

[24] C. K. Hong, Z. Y. Ou, and L. Mandel, "Measurement of subpicosecond time intervals between two photons by interference," *Physical Review Letters*, vol. 59, pp. 2044–2046, Nov. 1987.

[25] S. Designolle, G. Iommazzo, M. Besançon, S. Knebel, P. Gelß, and S. Pokutta, "Improved local models and new Bell inequalities via Frank-Wolfe algorithms," Mar. 2023.

[26] S. Boyd and L. Vandenberghe, *Convex Optimization.* Cambridge University Press, 2004.

[27] H. J. Kimble, "The quantum internet," *Nature*, vol. 453, pp. 1023–1030, June 2008.

[28] S. Wehner, D. Elkouss, and R. Hanson, "Quantum internet: A vision for the road ahead," *Science*, vol. 362, p. eaam9288, Oct. 2018.

[29] C. Palazuelos, "Superactivation of Quantum Nonlocality," *Physical Review Letters*, vol. 109, p. 190401, Nov. 2012.

Resource reduction in an error correction code using quantum multiplexing

Nicolò Lo Piparo[1] * Shin Nishio[1 2 3] William J. Munro[4] Kae Nemoto[1 2 3]

[1] *Quantum Information Science and Technology Unit, Okinawa Institute of Science and Technology Graduate University, Onna-son, Kunigami-gun, Okinawa904-0495, Japan.*

[2] *Department of Informatics, School of Multidisciplinary Sciences, SOKENDAI (The Graduate University for Advanced Studies), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo,101-8430, Japan.*

[3] *National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430, Japan.*

[4] *NTT Basic Research Laboratories and NTT Research Center for Theoretical Quantum Physics, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi, Kanagawa, 243-0198, Japan.*

Abstract. Quantum error correction (QEC) is a fundamental tool for the implementation of quantum computation and quantum communication systems. Due to losses, high-dimensional QEC codes are necessary to guarantee the successful transmission of information between users separated by large distances. However, such codes require an enormous number of two-qubit gates (CNOT) for their initial encoding, making their realization quite unpractical. Here, we apply the technique of quantum multiplexing to the encoding of high dimension Quantum Reed-Solomon (QRS) codes. We show that we can drastically reduce the number of CNOT gates required for creating the initial codewords of a QRS code.

Keywords: quantum error correction codes, quantum multiplexing, quantum gates.

1 Introduction

Quantum communication systems will drastically change the way users share information in the near future [1]. Such systems are expected to enhance the performance of the current classical technologies providing the basis for distributing quantum information systems, such as quantum key distribution [2], distributed quantum computation [3], quantum remote sensing [4] and the quantum internet [5]. However, the performance of such systems is greatly affected by channel loss, limiting their potential on their classical counterpart. Quantum error correcting codes (QECCs) address this issue as they can correct both gate and loss errors. Among the numerous variety of QECCs the quantum Reed-Solomon (QRS) code [6] is a remarkable code as it allows to correct qudit loss errors, which are a fundamental issue in long-distance quantum communication. However, the number of physical resources required for an optimal performance of this code over large distances is quite demanding in terms of two-qubit gates (CNOT) needed for the construction of the codewords. Recently it has been shown that applying the technique of quantum multiplexing to these codes allow to reduce drastically the number of physical resources (both in terms of photons and qubits for communication) [7]. Quantum multiplexing exploits multiple degrees of freedoms of a photon to carry the information. Here we show that using quantum multiplexing in the encoding of the codewords of a QRS code can lead to a drastic reduction of the number of CNOT gates when compared to other encoding methods. This can be achieved by using only linear optical elements, such as beam splitters and optical switches. Our method is quite versatile and can be extend to other quantum tasks.

2 Encoding the Quantum Reed-Solomon code

The QRS code has remarkable properties of correcting qudit loss errors [6] and, when the quantum multiplexing technique is applied to it, leads to a drastic reduction of physical resources [7]. In the $[[d,\,1,\,d-k+1]]_d$ QRS code one logic qudit of dimension d is encoded into d physical qudits, in such a way that can tolerate the loss of $d-k$ or less qudits, where d is a prime number. We describe how to estimate the number of gates required to encode a d dimensional QRS code with the example of Fig. 1(a), where $d=5$. Then we encode a 5 dimensional logic qudit using 5 physical qudits each of dimension 5 to create the $[[5,1,3]]_5$ QRS code. This can be done by initializing the first qudit $|\psi\rangle$ to a superposition of states, i.e., $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle + \delta|3\rangle + \epsilon|4\rangle$ and all the other qudits to $|0\rangle$, where $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 + |\epsilon|^2 = 1$. Then a series of SUM gates between two qudits are performed as shown in Fig. 1(a) and a discrete Fourier transform (DFT) gate is applied to a single qudit, which creates a superposition of states given by: $\mathrm{DFT}(|j\rangle) = \frac{1}{\sqrt{d}}\sum_{k=0}^{d-1} e^{2\pi i(jk/d)}\,|k\rangle$, where $|k\rangle$ is the kth Fock state and $|j\rangle$ is the jth phase state. The SUM gate is the generalization of the CNOT gate for $d-$dimensional qudits and is given by $\mathrm{SUM}(|\mathrm{A}\rangle, |\mathrm{B}\rangle) = |\mathrm{A}\rangle\,|(\mathrm{A}+\mathrm{B})[\mathrm{mod\,d}]\rangle$, where A and B are integers less than or equal to $d-1$. It is important for efficiency comparisons to determine the number of SUM gates needed to create a $[[d,1,(d+1)/2]]_d$ code using a $d-$dimensional qudit. It can be shown that this number increases quadratically with the dimension of the QRS code as shown in Fig. 1(b). Intuitively, this can be explained by considering that we require $(d^2-d-2)/2$ SUM gates between the first qudit as control and the $d-1$ qudits as targets and then $d-1$ SUM gates between the

*nicolo.lopiparo@oist.jp

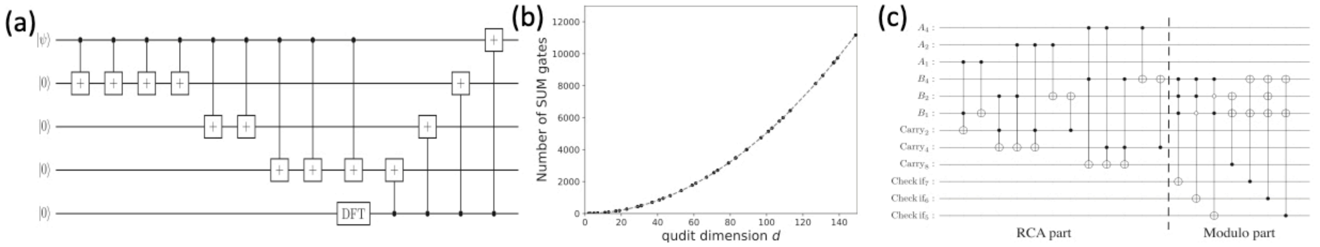


Figure 1: (a) The encoding circuit for the $[[5, 1, 3]]_5$ QRS code. Each qudit has a dimension of 5 while the two qudit SUM gate is represented by a dot corresponding to the control qudit connecting a "+" symbol inside a box representing the target qudit. The DFT box represents the DFT gate applied to a single qudit initialized to 0. (b) The number of two qudit SUM gates versus the dimension of the QRS code. (c) The circuit of the encoding of a 5-dimensional QRS code in which each qudit has been substituted by 3 qubits.

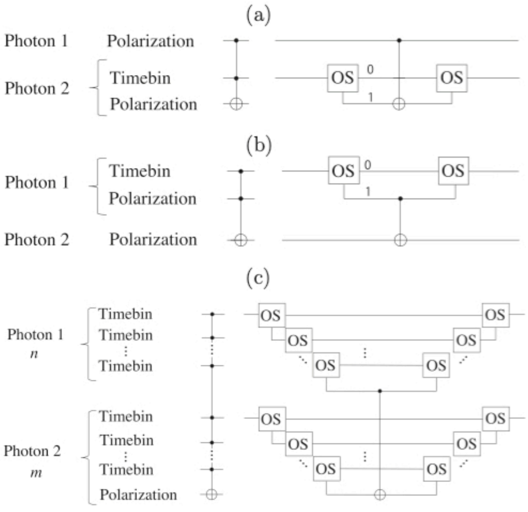


Figure 2: (a) and (b) Quantum circuit showing the C_2X gate between two photons and its optical representation (right) in which the time-bin DOF are split into and then recombined through optical switches (OSs). (c) The generalization of the C_2X gate to $C_{n+m-1}X$ gate.

qudit in which the DFT gate has been applied as control and the rest of the qudits as targets. The sum of these two terms gives $(d^2 + d - 4)/2$. Moreover, within each SUM gate there are several CNOT gates making the implementation of higher dimensional QRS codes more challenging. A more convenient way of encoding each logical $d-$ dimensional qudit is to use k qubits, where $2^{k-1} < d < 2^d$. This will allow us to show the advantages of applying the quantum multiplexing technique on the encoding of QRS codes.

There are several ways to construct SUM gates, but considering that essentially it is a modulo adder, it is possible to decompose our SUM gates into two parts. The first is an adder circuit given by a Ripple carry adder (RCA), which performs the following transformation: $\text{RCA}(|A\rangle, |B\rangle) = |A\rangle \, |(A + B)[\text{mod}\, 2^k]\rangle$. The second part is a modulo conversion that performs the following transformation: $\text{Mod}(|A\rangle \, |(A + B)[\text{mod}\, 2^k]\rangle) = |A\rangle \, |(A + B)[\text{mod}\, d]\rangle$. Figure 1(c) shows the SUM gate of a 5-dimensional QRS code between two qudits both encoded by 3 qubits ($2^3 > 5$). This corresponds to the addition of two binary number, A and B, each represented by 3 qubits whose digits are denoted as A_i and B_i (where i = 4,2,1 in Fig. 1(c)), respectively. Now the RCA module adds sequentially two digits, A_i and B_i, with an ancilla qubit (that we refer as carry qubit) that stores the possible overflow of the sum of A_{i-1} and B_{i-1}. The RCA unit is then made by a series of small adders subunits and the entire RCA unit will require then 3 ancillae, one for each digit of the numbers A and B. Next, the modulo circuit converts the state stored in B to the $|(a + b)[\text{mod}\, d]\rangle$ state in the case the sum of the inputs is between 5 and 8, no conversion is needed when the sum is between 0 and 4. To this aim ancillae qubits that we refer as check-if qubits are flagged (flipped from $|0\rangle$ to $|1\rangle$) when the sum of the inputs in each corresponding value is between 5 and 8. To perform such an operation a C_3X gate is applied between the qubits of B as control qubits and the corresponding check-if qubit as target qubit. The check-if qubit will then be the control of a multi-target CNOT gate that converts the B qubit from 5, 6, 7 and 8 to 0, 1, 2, and 3, respectively. This example can be easily extended to an arbitrary qubit dimension k required to encode a d dimensional QRS code. In the modulo conversion part the C_kX gate can be decomposed into $4(k-2)$ C_2X gates as shown in [8], which can further decomposed into 6 CNOT gates plus single qubit gates. We refer to this decomposition as "general decomposition" and we will compare it with the more efficient multiplexing decomposition in the next section.

3 Quantum Multiplexing and its performance

The key aspect of the Quantum Multiplexing (QM) technique is the use of multiple degrees of freedoms (DOFs) carried by a photon [10] and the possibility of performing CNOT gates between two different DOFs corresponding to two photons [11]. Figure 2 shows the key idea used in this work to perform multi-control gates between two photons using only one CNOT gate. Figure 2(a), shows how, using only optical switches (OSs), to perform a Toffoli (C_2X) gate, having as controls the polarization DOF of a photon and the time-bin DOF of

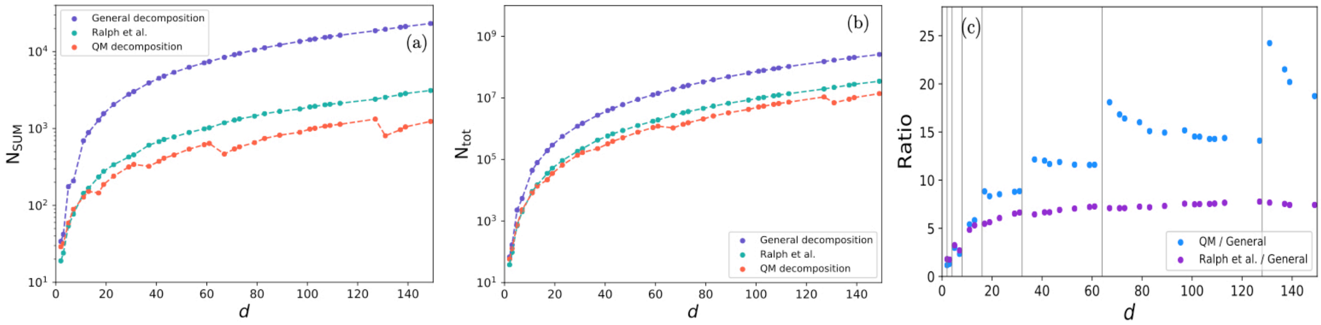


Figure 3: (a) Number of CNOT gates required to construct a single SUM gate versus the dimension d of the code. The blue curve is based on the general decomposition, the green curve is based on Ralph *et al*. [9] and the red curve is based on the quantum multiplexing decomposition of Fig. 2 (b). The total number of CNOT gates required to construct the entire encoder of the $[[d,\, 1,\, d-k+1]]_d$ QRS code versus d. (c) The ratio between the blue curve and the red(magenta) curve plotted in (a), respectively. The grey vertical line correspond to 2^m integer values of m.

a second photon and, as target qubit, the polarization DOF of the second photon. This can be accomplished by splitting the time-bin DOF of the second photon into two different spatial modes and then applying one CNOT gate between the relevant modes (in this example the CNOT gate is applied between the polarization DOF of the first photon and the mode labeled as “1” of the second photon). Then the time-bin modes are recombined into a single spatial mode using another OS. Figure 2(b) shows that this can be done also when both controls are in the same photon and the target is in another photon. This procedure can be generalized then to the case of a gate having n controls in the first photon and m controls in the second photon (see Fig. 2 (c)). The general decomposition of this gate would require a large number of direct CNOT gates whereas, with the QM method, only one CNOT gate is needed (provided that a large number of OSs are available). In our example so far we have considered polarization and time-bin DOFs, however, other modes are also possible, such as frequency-modes and orbital angular momentum.

We determine the number of CNOT gates in a SUM gate, N_{SUM}, required to encode a $d-$dimensional QRS code using the general decomposition and the quantum multiplexing method. Figure 3(a) shows N_{SUM} versus d for the general decomposition (blue curve) and when QM is applied (red curve). Both curves are almost indistinguishable for small values of d, however, they separate largely for higher values of d. In particular at $d = 139$ $N_{SUM} = 21182$ for the general decomposition and only $N_{SUM} = 1049$ when QM is in use. Figure 3(b) shows the total number of CNOT gates, N_{tot}, to construct the entire decoder. We also compare our results with the ones obtained in [9] in which the authors realize C_kX gates using one $k-$dimensional qudit, $2k-1$ two-qubit gates, and single-qudit gates. Our results still show an advantage, in terms of two-qubit gates reduction, than the scheme in Ref. [9]. To quantify the reduction of the number of gates we calculate the ratio between the QM approach with the general decomposition and the ratio of the approach of Ref. [9] with the general decomposition (see Fig. 3(c)). The advantage when the QM is in use is small for small dimensional codes (~ 7 times less CNOT gates) but it reaches 24 times less CNOT gates at $d = 131$. Hence, applying QM leads to a drastic reduction in the number of CNOT gates.

4 Conclusions

In this work we have applied the quantum multiplexing method to the encoding of QRS codes. We showed that especially for high-dimensional codes the improvement, in terms of CNOT gates reduction, is very high. Moreover, our approach can be also applied to other error correction codes and algorithms that require a big number of gates, such as the Grover's algorithm and the quantum walk algorithm. We believe that this work shows an important approach to reduce the costs of the implementation of quantum technologies in the near future.

Acknowledgments

This work was supported by Moonshot Grant No. JPMJMS226C, JSPS KAKENHI Grants No. JP21H04880 and No. JP22J20882, the MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant No. JPMXS0118069605, and the JST Moonshot R&D Grant No. JPMJMS2061.

References

[1] W. J. Munro, K. Azuma, K. Tamaki, and K. Nemoto. Inside quantum repeaters. *IEEE J. Sel. Top. Quantum Electron*, 21:78, 2015.

[2] C. H. Bennett and G. Brassard. Quantum Cryptography: Public key distribution and coin tossing. *Theor. Comput. Sci.* , 560:7, 2014.

[3] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello. Distributed quantum computation over noisy channels. *Phys. Rev. A*, 59:4249, 1999.

[4] E. Shanda. Physical Fundamentals of Remote Sensing. *Springer, New York*, 2012.

[5] H. J. Kimble. The quantum internet. *Nature (London)*, 453:1023, 2008.

[6] M. Grassl, W. Geiselmann, and T. Beth. Quantum Reed-Solomon codes. *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (Springer, New York)*, 231:244, 1999.

[7] N. Lo Piparo, M. Hanks, C. Gravel, K. Nemoto and William J. Munro. Resource reduction for Distributed Quantum Information Processing Using Quantum Multiplexed Photons. *Phys. Rev. Lett.*, 124:210503, 2020.

[8] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, 75:062313, 2006.

[9] T. C. Ralph, K. Resch, and A. Gilchrist. Efficient Toffoli gates using qudits. *Phys. Rev. A*, 73:062313, 2006.

[10] N. Lo Piparo, W. J. Munro, and K. Nemoto. Quantum Multiplexing. *Phys. Rev. A*, 99:022337, 2019.

[11] N. Lo Piparo, M. Hanks, K. Nemoto, and W. J. Munro. Aggregating quantum networks. *Phys. Rev. A*, 102:052613, 2020.

Depth analysis of variational quantum algorithms for the heat equation

Nikita Guseynov[1 2] * Andrey Zhukov [1] † Walter Pogosov [1 2 3] ‡

[1] *Dukhov Research Institute of Automatics (VNIIA), Moscow, 127030, Russia*
[2] *Moscow Institute of Physics and Technology (MIPT), Dolgoprudny, 141700, Russia*
[3] *Institute for Theoretical and Applied Electrodynamics, Russian Academy of Sciences, Moscow, 125412, Russia*

Abstract. We consider three approaches to solve the heat equation on a quantum computer. Using the direct variational method we minimize the expectation value of a Hamiltonian with its ground state being the solution of the problem under study. Typically, an exponential number of Pauli products in the Hamiltonian decomposition does not allow for the quantum speed up to be achieved. The Hadamard test based approach solves this problem, however, the performed simulations do not evidently prove that the Ansatz circuit has a polynomial depth with respect to the number of qubits. The Ansatz tree approach exploits an explicit form of the matrix what makes it possible to achieve an advantage over classical algorithms. In our numerical simulations with up to $n = 11$ qubits, this method reveals the exponential speed up.

Keywords: Ansatz tree approach, variational algorithms, heat equation, quantum computing, Hadamard test

1 Introduction

Quantum computing is a promising technology based on the principles of quantum mechanics Feynman (1982). The main motivation is outperform the state-of-the-art classical algorithms and achieve the so called quantum supremacy Zhong et al. (2020). Well-known examples are the quantum search algorithm Grover (1996) and Shor's factorization algorithm Shor (1999), which are both superior to the classical ones. Quantum computers can also be useful in various linear algebra problems. A remarkable example is the Harrow, Hassidim, and Lloyd (HHL) algorithm for solving systems of linear equations Harrow et al. (2009); Montanaro and Pallister (2016). Because the classical algorithms generally have the polynomial complexity in the matrix size N, HHL algorithm provides the exponential speedup in the case of sparse matrices. This algorithm however requires a large-depth quantum circuit composed of highly accurate quantum gates. Both these requirements are problematic in the present era of noisy intermediate-scale quantum (NISQ) computers Preskill (2018). Error correction codes Devitt et al. (2013); Noh and Chamberland (2020); Egan et al. (2021) and error mitigation techniques Temme et al. (2017); Endo et al. (2018); Zhukov and Pogosov (2022) could potentially overcome these problems, however, the state-of-the-art quantum devices lack enough number of qubits to work in the fault-tolerant regime.

2 Results

The paper studies the implementation of three variational quantum algorithms for solving the heat equation presented in the finite difference form. This problem is reduced to the solution of the system of linear equations arising at each discrete step of the time evolution.

In the first approach (direct variational method) the expectation value of the Hamiltonian (2) is minimized on some class of probe functions. The Hamiltonian is constructed in a way that its ground state corresponds to the solution of the system of linear equations.

$$|x\rangle = A^{-1} |b\rangle \,. \tag{1}$$

It can be readily shown Xu et al. (2021) that this solution $|x\rangle$ is the ground state of the Hamiltonian

$$H = A^{+}(I - |b\rangle \langle b|)A, \tag{2}$$

We performed proof-of-principles quantum computation with the matrix of size 4×4 using the real quantum processor of IBM Q project. The direct variational algorithm demonstrates a fundamental possibility of solving the system of linear equations (3) on a quantum computer. However, the exponential number of Pauli products in the matrix decomposition does not allow one to achieve the

*guseynov.nm@gmail.com
†zugazoid@gmail.com
‡walter.pogosov@gmail.com

quantum speedup (superiority over classical algorithms). In some cases it is possible to effectively sample over these products if we know the distribution of the decomposition coefficients, but this requires a separate study.

$$Ax = b, \tag{3}$$

$$A(c) = \begin{pmatrix} -2-c & 1 & 0 & \cdots & 0 & 1 \\ 1 & -2-c & 1 & \cdots & 0 & 0 \\ 0 & 1 & -2-c & \cdots & 0 & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ 0 & 0 & \cdots & -2-c & 1 & 0 \\ 0 & 0 & \cdots & 1 & -2-c & 1 \\ 1 & 0 & \cdots & 0 & 1 & -2-c \end{pmatrix}, \tag{4}$$

$$x = \begin{pmatrix} U_0^{\tau+1} \\ U_1^{\tau+1} \\ \vdots \\ U_{N-1}^{\tau+1} \end{pmatrix}, \qquad b = \begin{pmatrix} b_0^{\tau} \\ b_1^{\tau} \\ \vdots \\ b_{N-1}^{\tau} \end{pmatrix}. \tag{5}$$

The second approach (Hadamard test approach) is based on the minimization of the expectation value of the same Hamiltonian, but the problem of the exponential number of Pauli products is eliminated by using the Hadamard test Huang et al. (2021). A numerical simulation of the algorithm was performed with up to $n = 8$ qubits using three different entanglers or Ansatzs. The results show that it can be possible to achieve the quantum superiority, but the simulations with more qubits are required to definitively confirm this issue. It is also important to identify an effective entangler for the investigated problem. With this approach, three types of Ansatzes were tested: the Hardware Efficient, Checkerboard and the Digital-Analog Ansatz. The best results were obtained for the Checkerboard Ansatz, as it gives a more uniform entanglement. In addition, by increasing the grid parameter c, see, one decreases the number of required layers in the Ansatz. An exponential acceleration of up to eight qubits was demonstrated for this entangler. However, we argue that the considered number of qubits is not enough for an unambiguous conclusion about the advantage of the algorithm over the classical one.

The third type of approach (Ansatz tree approach) minimizes the l_2 norm (6), rather than the expectation value of the Hamiltonian.

$$L_R(x) = \|Ax - |b\rangle\|_2^2 = x^\dagger A^\dagger Ax - 2Re\{x^\dagger A\,|b\rangle\} + 1; \tag{6}$$

The algorithm is based on the unitary decomposition of the matrix (4). For the heat equation it turns out to be advantageous to switch to the Fourier representation by using the quantum Fourier transform. In the Fourier representation, the matrix becomes diagonal with a sinusoidal spectrum. Then we used a technique that allows us to replace the spectrum of this matrix by a piecewise-quadratic function, which, at the level of the original discretized problem, corresponds to the elimination of high-frequency oscillations of the solution, justified from the physical point of view. This makes it possible to radically reduce the number of Pauli products in the matrix decomposition. The simulation of the algorithm with up to eight qubits was performed and the complexity of the algorithm was estimated. The complexity is determined by the depth of the algorithm. The results show that the depth depends polynomially on the number of qubits for certain values of the grid parameter c. This reveals the fundamental ability of the Ansatz Tree Approach to demonstrate the quantum superiority for the heat equation.

Thus, the third approach can be considered as the most promising. The reason is that Ansatz Tree Approach makes use of the explicit form of the matrix (4), unlike the other algorithms discussed, which use the universal entanglers.

References

R. P. Feynman, Int. J. Theor. Phys **21** (1982).

H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, Y. Hu, et al., Science **370**, 1460 (2020).

L. K. Grover, in *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing* (1996), pp. 212–219.

P. W. Shor, SIAM Review **41**, 303 (1999).

A. W. Harrow, A. Hassidim, and S. Lloyd, Phys. Rev. Lett. **103**, 150502 (2009).

A. Montanaro and S. Pallister, Phys. Rev. A **93**, 032324 (2016).

J. Preskill, Quantum **2**, 79 (2018).

S. J. Devitt, W. J. Munro, and K. Nemoto, Rep. Prog. Phys. **76**, 076001 (2013).

K. Noh and C. Chamberland, Phys. Rev. A **101**, 012316 (2020).

L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, et al., Nature **598**, 281 (2021).

K. Temme, S. Bravyi, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180509 (2017).

S. Endo, S. C. Benjamin, and Y. Li, Phys. Rev. X **8**, 031027 (2018).

A. Zhukov and W. Pogosov, Quant. Inf. Proc. **21**, 1 (2022).

X. Xu, J. Sun, S. Endo, Y. Li, S. C. Benjamin, and X. Yuan, Science Bulletin **66**, 2181 (2021).

H.-Y. Huang, K. Bharti, and P. Rebentrost, New J. of Phys. **23**, 113021 (2021).

Low-rank quantum state preparation

Israel F. Araujo[1] * Carsten Blank[2] Ismael C. S. Araújo[3] Adenilton J. da Silva[3]

[1] *Department of Statistics and Data Science, Yonsei University, Seoul 03722, Republic of Korea*
[2] *Data Cybernetics, Landsberg 86899, Germany*
[3] *Centro de Informática, Universidade Federal de Pernambuco, Recife, 50740-560, Pernambuco, Brazil*

Abstract. Quantum state preparation, the process of encoding data into a quantum state, is a fundamental step in quantum computing. However, for non-structured data, this process exhibits exponential complexity with respect to the number of qubits involved. To address this challenge, several approaches have been proposed, including variational methods that train fixed-depth circuits with manageable complexity. Despite their potential, these methods have limitations, such as the absence of a back-propagation technique and the presence of barren plateaus. In this work, we present an innovative algorithm designed to reduce the depth of the state preparation circuit by transferring computational complexity to a classical computer. This approach allows for the initialization of quantum states that can be either exact or approximations. Remarkably, we demonstrate that the approximate initialization outperforms the initialization of the original state on current quantum processors. Through experimental evaluation, we provide evidence that our proposed method enables more efficient initialization of probability distributions in a quantum state.

Keywords: Quantum computing, entanglement, Schmidt decomposition, state preparation, approximate state preparation

1 Introduction

Quantum devices can execute information processing tasks that classical computers cannot perform efficiently [1]. In some instances, this leads to exponential advantages in solving systems of linear equations [2] and principal component analysis [3]. Additionally, there are known advantages in Monte-Carlo sampling [4, 5] in which a squared increase of convergence can be attained. Furthermore, quantum machine learning applications [6–8] may exhibit heuristic advantages. For all these applications, the initialization of a n-qubit quantum state, commonly called quantum state preparation, is an important step in quantum information processing.

Encoding a N-dimensional (complex) vector requires n-qubits with $N = 2^n$ and quantum circuits with $O(2^n)$ controlled-NOT (CNOT) gates [9, 10] these gates are the building blocks of today's quantum computers. Therefore, several works focus on the development of algorithms that supposes data-efficient initialization, as all above-mentioned quantum advantages could be undone when the conversion of classical data to quantum data becomes a bottleneck.

There are several quantum state preparation algorithms [11–14] with a lower bound of $O(2^n)$ CNOT gates to prepare an arbitrary quantum state with n qubits. Attempts to prepare quantum states more efficiently include a divide-and-conquer strategy that exchanges circuit depth by circuit width [14, 15], probabilistic approaches [16, 17], and strategies to initialize approximated quantum states [18–20]. Most recently, there has been an increasing focus on developing methods that are tailored to specific classes of quantum states. This approach recognizes that not all quantum states are equal, and that different types of states may require different preparation techniques to achieve optimal results. By tailoring the state preparation algorithm to the specific characteristics of the state of interest, these methods can often achieve better fidelity and gate complexity than more general approaches. Examples of this include uniform [21], sparse [13] and probability distribution [18] states. However, there is no clear understanding of which classes of quantum states can be created efficiently.

Method	Script	Entangled		Separable	
		CNOTs	Depth	CNOTs	Depth
Low Rank	[22]	30998	53644	0	2
PB [12]	[22]	38814	71580	38811	71575
Isometry [23]	[24]	32752	65505	32752	65505
Multiplexor [25]	[24]	65504	131025	65504	131025

Table 1: Depth and number of CNOTs comparison between LRSP and other state preparation algorithms. The column "Entangled" indicates that the state is non-separable, and the column "Separable" indicates that the state is completely separable (a product state).

The main goal of this article is to define a deterministic state preparation algorithm that creates circuits with depth as a function of entanglement, and show that for mapping the classical data into a Noisy Intermediate-Scale Quantum (NISQ) device, one indeed wants to be in the regime of low-entanglement. After finalizing the state preparation, one can increase the amount of entanglement (exponential in the number of qubits) by applying unitary evolution. This path could lead to the solution of the so-called loading problem (LP) of classical-quantum information processing.

The proposed approach's advantage is that it can accelerate quantum applications that require the initialization of quantum states, in particular on noisy devices. Using an approximation of the quantum state, the result is that the error introduced by the approximation is smaller than the error to encode the original state. The fundamental

*ifa@yonsei.ac.kr

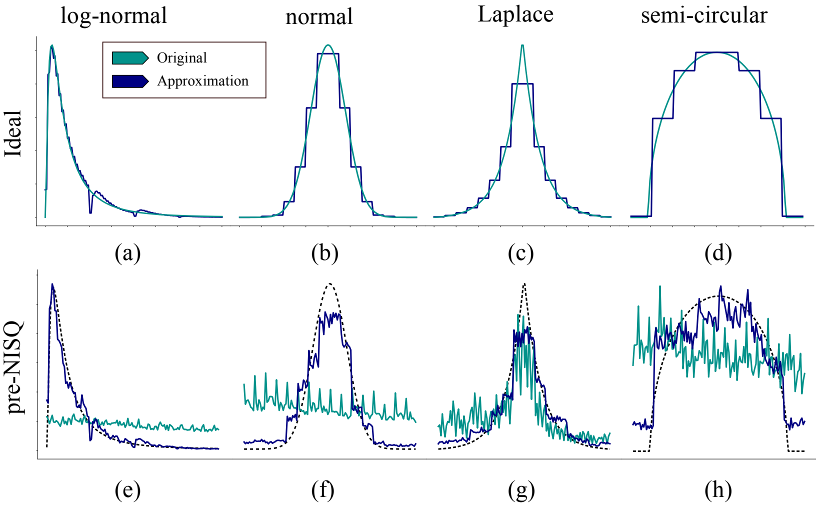


Figure 1: Visualization of probability distributions encoded on the amplitudes of a 7-qubit quantum state. (a)-(d) Ideal values of the exact (green line) and approximate (blue line) distributions. (e)-(h) Values of the exact and approximate distributions encoded via BAA and estimated by measurements on ibm_perth device. In the actual device, the encoding of the approximations performs closer to the ideal (black dotted line) than the exact distribution encoding. Each result is an average of 10 runs with 8192 shots.

cause of this behavior is the difference in the number of noisy operations necessary between the circuits to encode the original versus the approximate state. But also apart from this obvious advantages in the NISQ-era, the proposed approach reduces the complexity to initialize a low-entangled quantum state into fault-tolerant quantum devices.

2 Results

This work introduces a novel algorithm called the Low-Rank State Preparation (LRSP) algorithm, which is based on the Plesch & Brukner (PB) state preparation algorithm [12]. The PB algorithm is connected to the Schmidt decomposition, which allows representing a given quantum state $|\psi\rangle$ as a superposition of Schmidt coefficients and corresponding orthonormal basis states in two quantum subsystems, $\mathcal{H}_A$ and $\mathcal{H}_B$. The Schmidt decomposition of $|\psi\rangle$ is expressed as:

$$|\psi\rangle = \sum_{i=1}^{k} \sigma_i |i_A\rangle |i_B\rangle \tag{1}$$

Here, σ_i represents the Schmidt coefficients, $|i_A\rangle \in \mathcal{H}_A$ and $|i_B\rangle \in \mathcal{H}_B$ are orthonormal bases, and $1 \leq i \leq \min(\dim(\mathcal{H}_A), \dim(\mathcal{H}_B))$.

The PB algorithm consists of four main steps. The first step involves performing the Schmidt decomposition on a classical computer. In the second step, the algorithm initializes a quantum state in the first register as $\sum_i \sigma_i |i\rangle |0\rangle$, aiming to incorporate the Schmidt coefficients into the state. Following that, in the third step, $\lfloor n/2 \rfloor$ Controlled-NOT (CNOT) gates are applied to create the state $\sum_i \sigma_i |i\rangle |i\rangle$. Finally, in the last step, the algorithm applies the unitary operation U to the first register and the transpose of V (denoted as V^T) to the second register, where $U|i\rangle = |i_A\rangle$ and $V^T|i\rangle = |i_B\rangle$.

The proposed LRSP differs from PB algorithm when the Schmidt measure $m = \lceil \log_2(k) \rceil < \lfloor n/2 \rfloor$ and also by the use of isometries instead of full unitaries. Theorem 1 establishes the CNOT gates count needed when a low-rank representation of a state can be found using the Schmidt decomposition, as well as approximating the state by truncating the Schmidt coefficients.

Theorem 1 (Low-Rank State Preparation)**.** *Given Eqn. (1) with the Schmidt measure $m = \lceil \log_2(k) \rceil$, the low-rank state preparation has a complexity of*

condition	CNOT count
$0 \leq m < n_A$	$O(2^{m+n_B})$
$m = n_A$	$O(2^n)$

Proof. When acting on s qubits, a quantum state preparation typically requires $2^s - s - 1$ CNOTs [11], a unitary decomposition ${}^{23}/_{48}(2^{2s}) - {}^{3}/_{2}(2^s) + {}^{4}/_{3}$ [25], and an isometry decomposition $2^{m+s} - {}^{1}/_{24}(2^s) + O(s^2)2^m$ [23]. Let $|\psi\rangle$ be a n-qubit quantum state with a Schmidt decomposition where subsystem $\mathcal{H}_A$ has n_A qubits ($1 \leq n_A \leq \lfloor n/2 \rfloor$), subsystem $\mathcal{H}_B$ has $n_B = n - n_A$ qubits and k is the Schmidt rank. Considering the complete LRSP circuit, the overall number of CNOT gates is represented by:

- $0 \leq m < n_A$

$$\underbrace{2^m - m - 1 +}_{\text{phase 1}} \underbrace{m}_{\text{phase 2}} + \underbrace{2^{m+n_A} - \frac{1}{24}2^{n_A}}_{\text{phase 3 (isometry)}} + \underbrace{2^{m+n_B} - \frac{1}{24}2^{n_B}}_{\text{phase 4 (isometry)}}$$

- $m = n_A$ and $n_A < n_B$

$$\underbrace{2^{n_A} - n_A - 1 +}_{\text{phase 1}} \underbrace{n_A}_{\text{phase 2}} + \underbrace{\frac{23}{48}2^{2n_A} - \frac{3}{2}2^{n_A} + \frac{4}{3}}_{\text{phase 3 (unitary)}} + \underbrace{2^n - \frac{1}{24}2^{n_B}}_{\text{phase 4 (isometry)}}$$

- $m = n_A$ and $n_A = n_B$

$$\underbrace{2^{n_A} - n_A - 1 +}_{\text{phase 1}} \underbrace{n_A}_{\text{phase 2}} + \underbrace{2\left(\frac{23}{48}2^n - \frac{3}{2}2^{n_A} + \frac{4}{3}\right)}_{\text{phases 3 and 4 (unitaries)}}$$

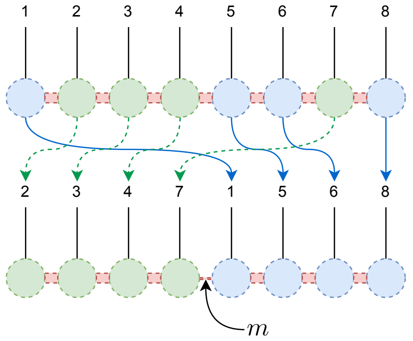


Figure 2: Schematic of the logical swap of qubits. By rearranging the qubits it is possible to find a bipartition such that the bond dimension between the partition blocks is small, which means that these blocks are not strongly entangled. The indices indicate the position of the qubit in the quantum circuit, and the color (blue or green) indicate the bipartition block.

The phases brackets indicate the contribution from each phase of the LRSP procedure to the number of CNOTs. Phase 1 is a state preparation, phase 2 a sequence of CNOT gates, phases 3 and 4 are isometry or unitary decompositions. These equations are bounded by the results of Theorem 1.

2.1 Low-Rank Approximation

The LRSP algorithm also allows a low-rank approximation limiting the Schmidt rank in exchange for an error. The fidelity loss can be used to quantify the loss by the approximation.

Definition 1. *Given the low-rank parameter r, the approximated state is denoted as $\left|\psi^{(r)}\right\rangle = \sum_{i=1}^{r} \sigma_i \left|i_A\right\rangle \left|i_B\right\rangle$ with coefficients for $1 \leq r \leq k$, i.e. $\sigma_j = 0$, $r < j \leq k$.*

It is possible to partially ($r > 1$) or completely ($r = 1$) disentangle subsets of qubits while the introduced fidelity loss $l(r, |\psi\rangle) := (1 - |\langle\psi, \psi^{(r)}\rangle|^2) = \sum_{i=r+1}^{k} |\sigma_i|^2$ scales with the Schmidt coefficients that are dropped. The remaining coefficients must be normalized.

The task of finding the optimal configuration for disentangling arbitrary bipartition blocks is not immediately straightforward. However, by rearranging the qubits between these blocks, it becomes possible to identify a configuration that exhibits a low bond-dimension. This low bond-dimension allows for a more accurate approximation and potential disentanglement of the blocks (see Figure 2).

This approach can be applied recursively on the resulting two blocks separately. By using a search algorithm with a specified maximal approximation error, the optimal approximation of any desired quantum state can be found. This includes scenarios such as vector encoding for matrix inversion using the HHL algorithm [3], loading data into a quantum machine learning model [6,7], or utilizing quantum simulation for stochastic processes [26]. The resulting search algorithm, known as the Bounded Approximation error Algorithm (BAA), is described in Section 2.2.

Table 1 compares the LRSP circuit depth and number of CNOTs with previous state preparation algorithms for both separable and non-separable 15-qubit quantum states.

2.2 Bounded Approximation Algorithm

By design, the low-rank approximation only applies to bipartite systems, yet it can be used hierarchically to enable the analysis of multipartite quantum systems [27] by a recursive algorithm. One then can recursively apply this approach on the two resulting partitions separately, i.e., find in the best rearranging of sites to minimize a bond dimension. This method leads to a tree search algorithm. With a given maximal approximation error, one can then find the optimal approximation of a quantum state. It is a bounded approximation error state preparation algorithm (BAA) that has a classical exponential run-time with respect to the number of qubits of the state as an upper bound. As it is a branch-and-bound algorithm using breadth-first search, the complexity usually converges faster [28,29], especially when the maximum allowed error is low (< 0.1) and the algorithm terminates at the first levels of the search tree. The full set of pseudocode which describes the algorithm is printed in the main text's supplementary information.

The BAA proves to be efficient when working with specific classes of quantum states found, e.g., in the quantum finance [13,18,21] area. Indeed, encoding probability distributions on the amplitudes of a 7-qubit quantum state show that even a modest fidelity loss (≤ 0.02) of the approximation using the BAA results in a significant reduction in the number of CNOTs. The result of the BAA discretization is shown in Figure 1. It shows, that the current devices have a hard time to create the state when not approximated, but the end-to-end result of an approximation is closer on the contrary.

3 Conclusion

With the help of Low-Rank State Preparation algorithm, significant improvements in the complexity of two-qubit entangling gates can be attained. The proposed algorithm exploits this fact by connecting state preparation complexity with the entanglement structure of the quantum state. Indeed, we highlight that classical data can be rearranged such that the entanglement structure attains an easier to approximate structure, as can be seen in Figure 2. This directly bestows the data with a topological structure of qubit systems. In particular, by logically swapping qubits between partitions, one can find more local behavior. This makes it easier to simulate classically, but also, by virtue of the above-mentioned circuit designs, less complex to create on a quantum computer. To showcase the immediate advantages of this approach, we present example applications in loading probability density function and application of quantum machine learning in the main text. We conclude that this method can help in the near term to make end-to-end calculations more precise. The technical version of this work is available on arXiv:2111.03132.

References

[1] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[2] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009.

[3] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, 2014.

[4] Ashley Montanaro. Quantum speedup of Monte Carlo methods. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 471(2181):20150301, 2015.

[5] Patrick Rebentrost, Brajesh Gupt, and Thomas R. Bromley. Quantum computational finance: Monte Carlo pricing of financial derivatives. *Physical Review A*, 98(2):022321, 2018.

[6] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195, 2017.

[7] Maria Schuld and Francesco Petruccione. *Supervised Learning with Quantum Computers*. Springer Publishing Company, Incorporated, 1st edition, 2018.

[8] Carsten Blank, Daniel K. Park, June-Koo Kevin Rhee, and Francesco Petruccione. Quantum classifier with tailored quantum kernel. *npj Quantum Information*, 6(1):41, 2020.

[9] Scott Aaronson. Read the fine print. *Nature Physics*, 11:291–293, 2015.

[10] Frank Leymann and Johanna Barzen. The bitter truth about gate-based quantum algorithms in the NISQ era. *Quantum Science and Technology*, 5(4):044007, 2020.

[11] Ville Bergholm, Juha J Vartiainen, Mikko Möttönen, and Martti M Salomaa. Quantum circuits with uniformly controlled one-qubit gates. *Physical Review A*, 71(5):052330, 2005.

[12] Martin Plesch and Časlav Brukner. Quantum-state preparation with universal gate decompositions. *Physical Review A*, 83(3):032302, 2011.

[13] Emanuel Malvetti, Raban Iten, and Roger Colbeck. Quantum circuits for sparse isometries. *Quantum*, 5:412, 2021.

[14] Israel F. Araujo, Daniel K. Park, Francesco Petruccione, and Adenilton J. da Silva. A divide-and-conquer algorithm for quantum state preparation. *Scientific Reports*, 11:6329, 2021.

[15] Israel F. Araujo, Daniel K. Park, Teresa B. Ludermir, Wilson R. Oliveira, Francesco Petruccione, and Adenilton J. da Silva. Configurable sublinear circuits for quantum state preparation. *arXiv preprint arXiv:2108.10182*, 2021.

[16] Xiao-Ming Zhang, Man-Hong Yung, and Xiao Yuan. Low-depth quantum state preparation. *Phys. Rev. Research*, 3:043200, 2021.

[17] Daniel K. Park, Francesco Petruccione, and June-Koo Kevin Rhee. Circuit-based quantum random access memory for classical data. *Scientific Reports*, 9:3949, 2019.

[18] Christa Zoufal, Aurélien Lucchi, and Stefan Woerner. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Information*, 5:103, 2019.

[19] Kouhei Nakaji, Shumpei Uno, Yohichi Suzuki, Rudy Raymond, Tamiya Onodera, Tomoki Tanaka, Hiroyuki Tezuka, Naoki Mitsuda, and Naoki Yamamoto. Approximate amplitude encoding in shallow parameterized quantum circuits and its application to financial market indicator, 2021.

[20] Gabriel Marin-Sanchez, Javier Gonzalez-Conde, and Mikel Sanz. Quantum algorithms for approximate function loading, 2021.

[21] Fereshte Mozafari, Heinz Riener, Mathias Soeken, and Giovanni De Micheli. Efficient boolean methods for preparing uniform quantum states. *IEEE Transactions on Quantum Engineering*, 2:1–12, 2021.

[22] Israel F. Araujo, Carsten Blank, Adenilton da Silva, Ismael Cesar, and Leon Silva. Quantum computing library (qclib). `https://github.com/qclib/qclib`, 2022.

[23] Raban Iten, Roger Colbeck, Ivan Kukuljan, Jonathan Home, and Matthias Christandl. Quantum circuits for isometries. *Physical Review A*, 93(3):032318, mar 2016.

[24] Gadi Aleksandrowicz, Thomas Alexander, Panagiotis Barkoutsos, Luciano Bello, et al. Qiskit: An Open-source Framework for Quantum Computing, January 2019.

[25] Vivek V Shende, Stephen S Bullock, and Igor L Markov. Synthesis of quantum-logic circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 25(6):1000–1010, 2006.

[26] Carsten Blank, Daniel K Park, and Francesco Petruccione. Quantum-enhanced analysis of discrete stochastic processes. *npj Quantum Information*, 7(1):1–9, 2021.

[27] Peiyuan Teng. Accurate calculation of the geometric measure of entanglement for multipartite quantum states. *Quantum Information Processing*, 16(7):181, 2017.

[28] Kurt Mehlhorn and Peter Sanders. *Algorithms and data structures: the basic toolbox*. Springer, Berlin, 2008.

[29] Steven S. Skiena. *The algorithm design manual*. Springer, London, 2nd ed edition, 2008. OCLC: ocn228582051.

Practical advantage of quantum machine learning in ghost imaging

Tailong Xiao[1] * Xinliang Zhai[1] † Xiaoyan Wu[1] ‡ Jianping Fan[2] § Guihua Zeng[1] ¶

[1] *State Key Laboratory of Advanced Optical Communication Systems and Networks, and Center for Quantum Sensing and Information Processing, Shanghai Jiao Tong University, Shanghai 200240, China*
[2] *Department of Computer Science, University of North Carolina-Charlotte, Charlotte, North Carolina 28223, USA*

Abstract. We investigate the practical advantage of quantum machine learning in ghost imaging by overcoming the limitations of classical methods in blind object identification and imaging. We propose two hybrid quantum-classical machine learning algorithms and a physical-inspired patch strategy to allow distributed quantum learning with parallel variational circuits. In light of the algorithm, we conduct experiments for imaging-free object identification and blind ghost imaging under different physical sampling rates. The numerical results showcase that quantum learning can restore high-quality images but classical learning fails. Our work explores a physics-related application capable of practical quantum advantage, which highlights the prospect of quantum computation in the machine learning field.

Keywords: Quantum machine learning, ghost imaging, variational quantum algorithm

1 Introduction

One of the challenges in quantum computation is to identify practical applications or problems where quantum algorithms can outperform classical ones [1, 2]. Although various experiments have demonstrated quantum advantage, such as random circuit sampling [3], Boson sampling [4, 5], and quantum walks [6], they have limited practical applications in the near term. In the noisy intermediate-scale quantum (NISQ) era [7], quantum noise cannot be fully eliminated, and the number of qubits is limited. Therefore, the focus is on demonstrating the advantage of quantum hardware-based computation in terms of sample or time complexity compared to classical counterparts [8]. Several research studies aim to clarify the advantage of quantum machine learning (QML), such as the rigorous speedup for discrete logarithm problems using the quantum kernel method [9] and the theoretical advantage for identifying quantum states/processes using QML [10]. These works demonstrate the possibility of achieving theoretical quantum advantage in NISQ devices using QML to solve specific problems.

GI retrieves an image by using two correlated beams, the reference beam, and the object beam [11, 12, 13, 14]. The reference beam is typically captured by a spatial resolution detector and does not interact with the object. In contrast, the object beam records the object information by a bucket detector that lacks spatial resolution. Figure 1a shows the experimental setup. Compared to a traditional array camera, GI is a time-for-space imaging method. Due to its inevitable time-consuming sampling, obtaining high-quality images at a low sampling rate is significant, under which the dimension of the bucket signals can be greatly reduced. The signals contain information about the physical process, leading to potential advantages when using QML. The bucket signals can be directly used for downstream tasks, such as imaging-free recognition [16], tracking [17], and segmentation [18].

In this work, we propose a hybrid QML algorithm for GI systems to demonstrate their practical advantages in physically-inspired imaging systems [19]. We investigate two challenging applications: object identification and imaging, which can be regarded as classification and regression problems in the machine learning field. We collect experimentally detected signals from the GI system to train QML models using a physical-inspired patch strategy to divide high-dimensional measured signals into low-dimensional pieces for accessible data encoding by current NISQ devices. We also build classical neural networks with an approximate number of trainable parameters to benchmark the performance fairly. In the identification and imaging applications, our hybrid QML methods are shown to be superior to their corresponding classical machine learning methods. We investigate the generalization capability of QML when reducing the training samples and quantify the quantum advantage using a capacity measure of QML from the perspective of information geometry. Furthermore, we study the impact of quantum noise in the QML method on the imaging application. Our results demonstrate the substantial advantage of QML algorithms in the GI system through rigorous quantitative analysis, highlighting their potential advantages in physically-related systems.

2 Results

As shown in Figure 1a, the single-beam ghost imaging technique retrieves an image by utilizing the correlation between the modulated patterns displayed on the digital DMD and the bucket signals collected by the detector. In this technique, a laser beam initially illuminates the DMD, which is loaded with various modulated patterns. The modulated light field then propagates towards the object, and after interacting with the scene, the transmitted or reflected light is collected by a bucket detector (i.e., a single-pixel detector). This physical process essentially encodes the scene information optically via the illu-

*tailong_shaw@sjtu.edu.cn
†zhaixl@sjtu.edu.cn
‡xiaoyanwu@sjtu.edu.cn
§jfan@uncc.edu
¶ghzeng@sjtu.edu.cn

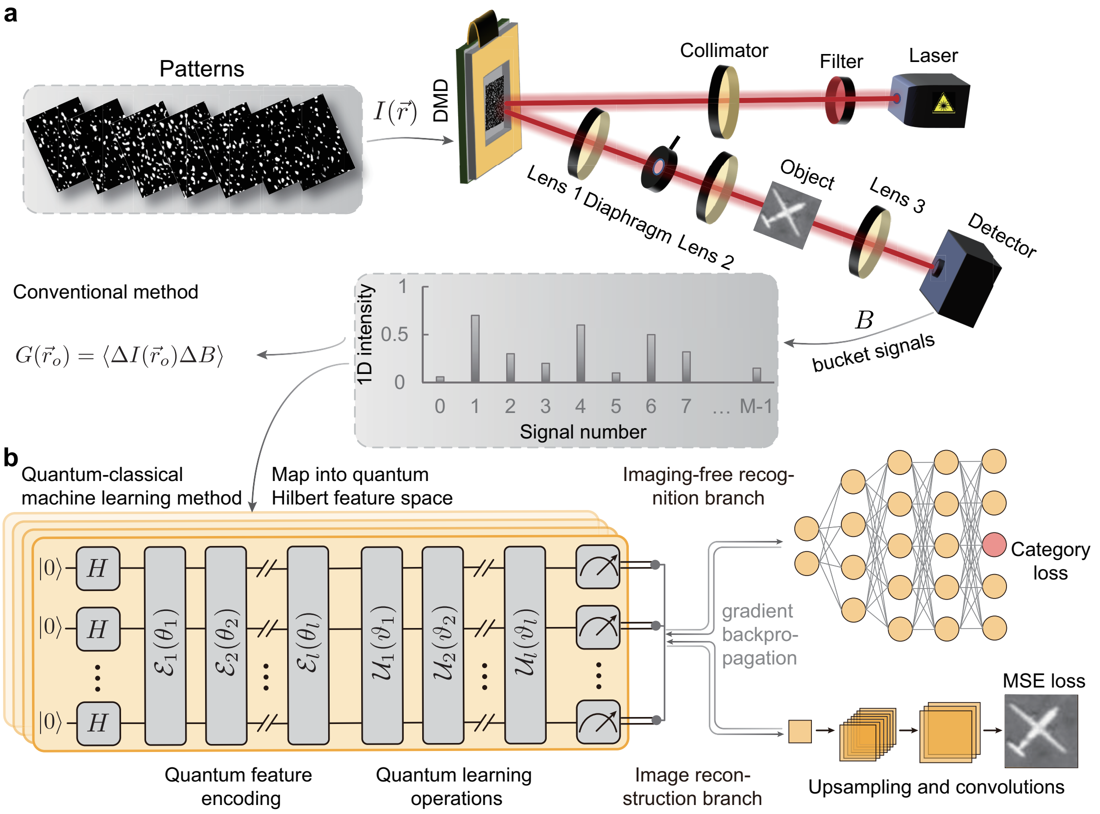

a
Patterns
$I(\vec{r})$
DMD
Collimator
Filter
Laser
Lens 1
Diaphragm
Lens 2
Object
Lens 3
Detector
Conventional method
$G(\vec{r}_o) = \langle \Delta I(\vec{r}_o)\Delta B \rangle$
1D intensity
1
0.5
0
0 1 2 3 4 5 6 7 ... M-1
Signal number
B
bucket signals
b
Quantum-classical machine learning method
Map into quantum Hilbert feature space
Imaging-free recognition branch
$|0\rangle$
H
$\mathcal{E}_1(\theta_1)$
$\mathcal{E}_2(\theta_2)$
$\mathcal{E}_l(\theta_l)$
$\mathcal{U}_1(\vartheta_1)$
$\mathcal{U}_2(\vartheta_2)$
$\mathcal{U}_l(\vartheta_l)$
gradient backpropagation
Category loss
MSE loss
Quantum feature encoding
Quantum learning operations
Image reconstruction branch
Upsampling and convolutions

Figure 1: Hardware-efficient quantum machine learning enhanced ghost imaging. (a) The experimental setup for ghost imaging in which the patterns can be randomly sampled or optimized according to the poster processing and a single-pixel detector is used to measure the object-interacted light field modulated by the prescribed illumination patterns. (b) Hybrid quantum machine learning algorithm by combining the artificial neural network or convolution neural network enhances object identification and the quality of object imaging.

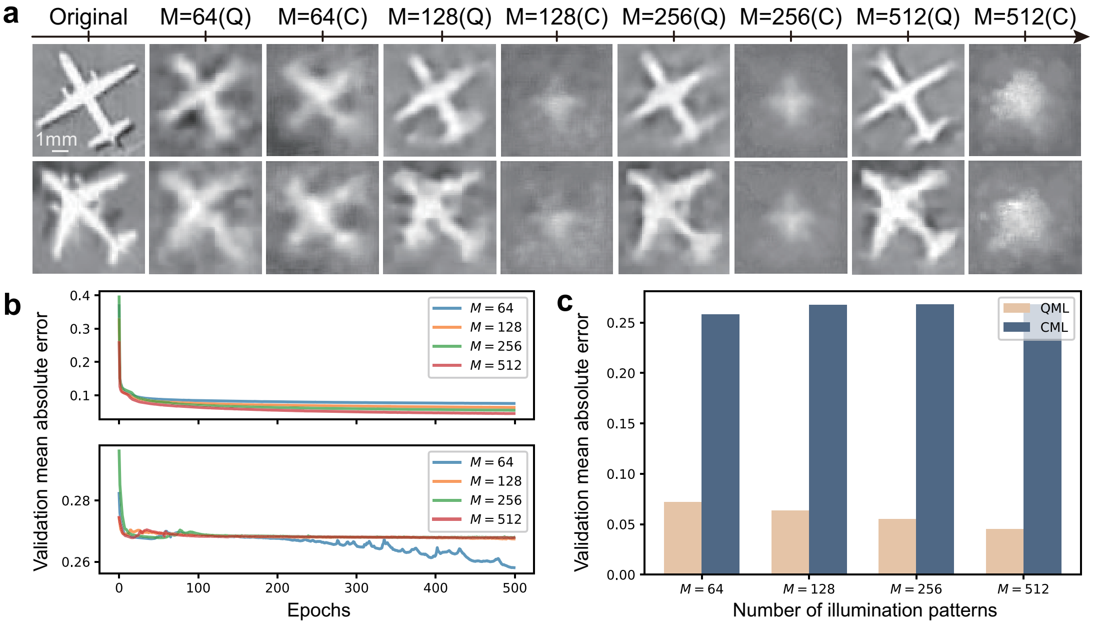


Figure 2: Blind ghost imaging (GI) based on quantum machine learning (QML) and classical machine learning (CML). (a) QML-based GI and CML-based GI under different illumination patterns $M = 64, 128, 256, 512$. (b) The validation means absolute error (MAE) of QML (top panel) and CML (bottom panel) with a different number of patterns. (c) The final validation MAE of QML and CML under different M.

mination light field. Consequently, the collected bucket signal B can be mathematically represented as follows

$$B = \int I(\vec{r}_0)S(\vec{r}_0)\mathrm{d}\vec{r}_0, \tag{1}$$

where $I(\vec{r}_0)$ is the intensity distribution imprinted on the object plane, and $S(\vec{r}_0)$ is the intensity transmission or reflection function of the object. In practical computational GI scenes, to reduce imaging time, the sampling is always limited and the illumination $I\left(\vec{r}_o\right)$ is approximated by the loaded patterns $I\left(\vec{r}\right)$. Thus, the recovered image is given by

$$G(\vec{r}_0) \approx \frac{1}{M}\sum_{i=1}^{M}\Delta I_i\left(\vec{r}\right)\Delta B_i, \tag{2}$$

where $\Delta I\left(\vec{r}_o\right)$ and ΔB is the intensity fluctuations of illuminations and the bucket signals, respectively. The notation $\langle\cdot\rangle$ presents the ensemble average. In Eq. (2), M needs to be at least equal to the imaging pixel number N_p. To accelerate the progress, GI is performed under $M \ll N_p$. In this case, the image reconstruction of GI is an underdetermined optimization problem

$$\hat{G} = \arg\min_{G}\left\|\mathbf{I}G - B\right\|_2^2, \tag{3}$$

where $\mathbf{I}$ is the measurement matrix consisting of the modulated illumination patterns $I\left(\vec{r}\right)$. The conventional method to directly solve the optimization problem is still hard, especially in cases where the sample rate is insufficient. We propose using QML to solve the inverse problem. QML, specifically its hybrid version (see Methods), shows promise in enhancing learning capabilities and reducing neural network size. We use patch strategy to encode the high-dimensional data with parallel quantum circuit. The basis for patch processing of bucket signals is rooted in the independence of each bucket signal from the others.

The backbone of the hybrid quantum-classical machine learning algorithm consists of a classical artificial neural network and a parameterized quantum circuit (PQC) as Figure 1b shows. The classical state $\vec{x}$ can be encoded by

$$|\psi_x\rangle = \left(\prod_{i=1}^{L}\mathcal{E}_i(\theta_i, \vec{x})\right)|0\rangle^{\otimes M}, \tag{4}$$

where L denotes the number of encoding layers, the parameters $\{\theta_i\}$ are the variational quantum parameters. Formally, the final quantum state evolved by the learning operations is given by

$$|\psi_f\rangle = \prod_{i=1}^{L}\mathcal{U}_i(\vartheta_i)|\psi_x\rangle, \tag{5}$$

where $\mathcal{U}_i$ denotes the variational learning operation, ϑ_i is the trainable parameter in ith learning layer. To obtain the classical information of the final quantum state, we require measuring the observable i.e. $\langle O\rangle = \langle\psi_f|O|\psi_f\rangle$. We regard the expectation values of the observable as the feature representation to make predictions.

To study the potential advantage of QML ghosting imaging, we conduct the optical experiments in Figure 1a to collect the bucket signals. We adopt the remote sensing images as our dataset. The images have a low signal-to-noise ratio and the contrast ratio of the images is also low. Reconstructing remote-sensing images is a challenging task in the field of imaging processing. By using a GI system, we collect the bucket signals as the input of the QML/CML and then output the reconstructed images. By using the MSE loss function, we can calculate the gradients of loss over each training parameter. As we can see, the intuitive comparison from reconstructed images between QML and CML in Figure 2a demonstrates that QML outperforms CML in terms of the resolution and outline of the images. Here, we do not concentrate on the resolution of reconstructed images. From Figure 2b, we present the validation mean absolute error (MAE) of QML and CML. We can find that the MAE of the QML is much smaller than the CML. The ultimate MAE in the last epoch (500) of both CML and QML is presented in Figure 2c. The final MAE of QML is $3.6, 4.2, 4.9, 5.9$ fold smaller than the MAE of CML for $M = 64, 128, 256, 512$, respectively. Increasing the number of illumination patterns from $M = 64$ to 512, the MAE of the QML is decreased linearly. On the contrary, the MAE of the CML does not decrease, which indicates that the network did not converge well.. In principle, increasing the number of patterns can increase the GI quality. Even though we find that QML results are still better than DGI under the same sample rate.

3 Discussion

In summary, we apply QML to practical GI systems to demonstrate its advantages experimentally and theoretically. We propose a hardware-efficient hybrid QML framework based on shallow variational quantum circuits and quantitatively demonstrate its practical advantages in classic GI task. We exploit a highly flexible physical-inspired patch strategy that is applicable for current NISQ devices when handling large-scale classical dataset. The strategy also makes the large-scale classical simulation of QML in the GI system possible.

Through collecting the experimental dataset with different sampling rates in object imaging tasks, QML-enhanced blind GI can fully make use of the information of a large sampling rate and reconstruct the object images with high PSNR. In contrast, CML cannot simultaneously learn the illumination patterns and the feature information of the object such that it cannot reconstruct a high-PSNR image. We attribute the superior performance of QML in part to the exponentially larger quantum-featured Hilbert space, which provides a more powerful learning capability in high-dimensional spaces.

Although other researches use QML in classical machine learning fields but achieve no obvious practical advantage, the application of real physical-related GI system amplifies the advantage of the QML algorithm. Our study presents a practical and crucial application for the QML field and also highlights the point that QML is likely to be suitable for processing physically system-generated datasets. In future work, we will study the connections of QML and sparse encoding and other applications of QML in the GI system.

References

[1] Leonard Wossnig, Zhikuan Zhao, and Anupam Prakash. Quantum linear system algorithm for dense matrices. *Physical Review Letters*, 120(5):050502, 2018.

[2] Austin P Lund, Michael J Bremner, and Timothy C Ralph. Quantum sampling problems, bosonsampling and quantum supremacy. *npj Quantum Information*, 3(1):1–8, 2017.

[3] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[4] Juan M Arrazola, Ville Bergholm, Kamil Brádler, Thomas R Bromley, Matt J Collins, Ish Dhand, Alberto Fumagalli, Thomas Gerrits, Andrey Goussev, Lukas G Helt, et al. Quantum circuits with many photons on a programmable nanophotonic chip. *Nature*, 591(7848):54–60, 2021.

[5] Justin B Spring, Benjamin J Metcalf, Peter C Humphreys, W Steven Kolthammer, Xian-Min Jin, Marco Barbieri, Animesh Datta, Nicholas Thomas-Peter, Nathan K Langford, Dmytro Kundys, et al. Boson sampling on a photonic chip. *Science*, 339(6121):798–801, 2013.

[6] Qingling Zhu, Sirui Cao, Fusheng Chen, Ming-Cheng Chen, Xiawei Chen, Tung-Hsun Chung, Hui Deng, Yajie Du, Daojin Fan, Ming Gong, et al. Quantum computational advantage via 60-qubit 24-cycle random circuit sampling. *Science bulletin*, 67(3):240–245, 2022.

[7] Seth Lloyd and Christian Weedbrook. Quantum generative adversarial learning. *Physical Review Letters*, 121(4):040502, 2018.

[8] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[9] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 17(9):1013–1017, 2021.

[10] Hsin-Yuan Huang, Michael Broughton, Jordan Cotler, Sitan Chen, Jerry Li, Masoud Mohseni, Hartmut Neven, Ryan Babbush, Richard Kueng, John Preskill, et al. Quantum advantage in learning from experiments. *Science*, 376(6598):1182–1186, 2022.

[11] Todd B Pittman, YH Shih, DV Strekalov, and Alexander V Sergienko. Optical imaging by means of two-photon quantum entanglement. *Physical Review A*, 52(5):R3429, 1995.

[12] Ryan S Bennink, Sean J Bentley, and Robert W Boyd. "two-photon" coincidence imaging with a classical source. *Physical Review Letters*, 89(11):113601, 2002.

[13] Alessandra Gatti, Enrico Brambilla, Morten Bache, and Luigi A Lugiato. Ghost imaging with thermal light: comparing entanglement and classicalcorrelation. *Physical Review Letters*, 93(9):093602, 2004.

[14] Yanhua Zhai, Xihao Chen, Da Zhang, and Lingan Wu. Two-photon interference with true thermal light. *Physical Review A*, 72(4):043805, 2005.

[15] Alejandra Valencia, Giuliano Scarcelli, Milena D'Angelo, and Yanhua Shih. Two-photon imaging with thermal light. *Physical Review Letters*, 94(6):063601, 2005.

[16] Sadao Ota, Ryoichi Horisaki, Yoko Kawamura, Masashi Ugawa, Issei Sato, Kazuki Hashimoto, Ryosuke Kamesawa, Kotaro Setoyama, Satoko Yamaguchi, Katsuhito Fujiu, et al. Ghost cytometry. *Science*, 360(6394):1246–1251, 2018.

[17] Qiwen Deng, Zibang Zhang, and Jingang Zhong. Image-free real-time 3-d tracking of a fast-moving object using dual-pixel detection. *Optics Letters*, 45(17):4734–4737, 2020.

[18] Haiyan Liu, Liheng Bian, and Jun Zhang. Image-free single-pixel segmentation. *Optics & Laser Technology*, 157:108600, 2023.

[19] Tailong Xiao, Xinliang Zhai, Xiaoyan Wu, Jianping Fan and Guihua Zeng. Practical advantage of quantum machine learning in ghost imaging. *Communications physics, to be published*, 2023.

Comparison of several quantum receivers for coherent-state signals by using classical reliability function

Ken Masaki[1] * Tiancheng Wang[2] † Souichi Takahira[1] ‡ Shogo Usami[1] §
Tsuyoshi Sasaki Usuda[3] ¶

[1] *Graduate School of Science and Engineering, Meijo University, Aichi 468-8502, Japan.*
[2] *Faculty of Informatics, Kanagawa University, Kanagawa 221-8686, Japan.*
[3] *Faculty of Information Science and Technology, Aichi Prefectural University, Aichi 480-1198, Japan.*

Abstract. There is a gap between properties of quantum and classical reliability functions. That is, the upper bound of quantum reliability function diverges, whereas that of classical one does not. We are interested in the origin of the gap. We have focused on the fact that classical channel matrices of classical-quantum communications depend on what kind of quantum receiver is used. In this paper, we investigate the properties of classical reliability functions for classical-quantum communications on quantum receivers. As a result, we show the optimal quantum receiver is not superior in low rates.

Keywords: Reliability function, Quantum communication, Quantum measurement

1 Introduction

Quantum information technology is the application of quantum nature to technology [1, 2]. A coherent state [3] is known as a promising medium for practical applications of quantum communication and cryptography. Since the state is approximately realized by laser light, it is often referred to as classical light, but it is also regarded as a typical quantum medium. In fact, even using the same coherent states, the performance of an optimal classical receiver such as heterodyne and homodyne receivers and that of an optimal quantum receiver are quite different [1, 2]. Recently, an upper bound on the quantum reliability function has been clarified, and it has been shown that the classical and quantum reliability functions [4, 5, 6] also show qualitatively different results. That is, for coherent-state signals, the classical reliability function based on the optimal classical receiver always takes finite values, whereas the quantum reliability function based on the optimal quantum collective decoding diverges [7, 8].

The purpose of this research is to consider the above-mentioned differences between the quantum and classical from various viewpoints. First, in quantum communication, the channel matrix changes depending on what kind of quantum receiver is used [1, 2]. That is, we note that the channel matrix can be controlled. Divergence of the upper bound of the classical reliability function never occurs when any input transitions to any output. It only occurs when there is a “lack of transition.” This “lack of transition,” that is, the existence of a zero component in the channel matrix, occurs even for very basic signals, such as a BPSK coherent-state signal, when a certain quantum receiver is used. As simple examples, there are the Kennedy receiver [9] and an unambiguous quantum receiver [10, 11, 12].

In this paper, in addition to these receivers, we consider the Helstrom receiver which is called the optimal quantum receiver and investigate the properties of the upper bound of the classical reliability function for a given channel matrix corresponding to each receiver.

2 Preparation

2.1 Basis of quantum communication

Let $A = \{0, 1, \ldots, M-1\}$ be an alphabet which is a set of the classical information, ξ_i be the *a priori* probability of i, and $\xi = \{\xi_i \mid i \in A\}$. For M-ary classical-quantum communication, a sender transmits a quantum state $|\psi_i\rangle \in \mathcal{H}$ corresponding to i. The set of quantum states $\{|\psi_i\rangle \in \mathcal{H} \mid i \in A\}$ is referred to as a quantum signal.

Here, consider the case where the receiver input is $|\psi_i\rangle$. At the receiver, a quantum measurement is performed and a classical information j is obtained as a measurement outcome. Then the conditional probability of the measurement outcome j when the

* 233426018@ccmailg.meijo-u.ac.jp
† wang@kanagawa-u.ac.jp
‡ takahira@meijo-u.ac.jp
§ susami@meijo-u.ac.jp
¶ usuda@ist.aichi-pu.ac.jp

input state is $|\psi_i\rangle$ is

$$P(j|i) = \text{Tr } |\psi_i\rangle\langle\psi_i|\hat{\Pi}_j, \tag{1}$$

where POVM: $\Pi = \{\hat{\Pi}_j \mid j = 0, 1, \ldots, N-1\}$ is a quantum measurement. The matrix $[P(j|i)]$ whose (i, j)-component is obtained by the above conditional probability corresponds to a channel matrix in the classical information theory. In this case, the channel is determined by the quantum measurement and is therefore called the quantum measurement channel [2].

2.2 Classical reliability function

Let $P_{\mathrm{e}}^{\mathrm{opt}}(n, R)$ be the optimal error probability for a sufficiently long codeword length n and a given coding rate R. Then the reliability function $E(R)$ is conceptually defined as the error exponent [4, 5]:

$$P_{\mathrm{e}}^{\mathrm{opt}}(n, R) = e^{-nE(R)}. \tag{2}$$

Although it is difficult to compute the reliability function directly, its upper and lower bounds are known. Let $[W(j|i)]$ be the channel matrix of a classical channel. Then the so-called sphere-packing upper bound is given by

$$E_{\mathrm{CU}}(R) = \max_{0\le s}\max_{\xi}\left[\nu(s, \xi) - sR\right], \tag{3}$$

where

$$\nu(s, \xi) = -\ln\sum_{j\in B}\left(\sum_{i\in A}\xi_i W(j|i)^{\frac{1}{1+s}}\right)^{1+s}, \tag{4}$$

and A and B are the input and output alphabets. In this paper, we compute $E_{\mathrm{CU}}(R)$ by substituting $P(j|i)$ in Eq.(1) into $W(j|i)$. And hereafter, if there is no confusion, the upper bound will simply be referred to as the (classical) reliability function.

3 Quantum receiver for BPSK signals

Here, we explain the three quantum receivers treated in this paper. We also confirm that the "lack of transition" mentioned in Sec.1. In this paper, we consider a BPSK coherent-state signal (BPSK signal) $\{|\psi_0\rangle, |\psi_1\rangle\} = \{|\alpha\rangle, |-\alpha\rangle\}$ as a quantum signal. Here, α is a complex amplitude.

3.1 Helstrom receiver

The Helstrom receiver is the optimal quantum receiver that minimizes the average probability of error. The conditional probabilities that constitute the channel matrix are as follows:

$$P(0|0) = P(1|1) = 1 - \varepsilon, \tag{5}$$
$$P(1|0) = P(0|1) = \varepsilon, \tag{6}$$

where

$$\varepsilon = \frac{1}{2}\left(1 - \sqrt{1 - e^{-4|\alpha|^2}}\right). \tag{7}$$

3.2 Kennedy receiver

The Kennedy receiver [9] was proposed by Kennedy in 1973 as a suboptimal quantum receiver for a BPSK signal. The conditional probabilities for the Kennedy receiver are

$$P(0|0) = 1, \tag{8}$$
$$P(1|0) = 1 - P(0|0) = 0, \tag{9}$$
$$P(0|1) = e^{-4|\alpha|^2}, \tag{10}$$
$$P(1|1) = 1 - P(0|1) = 1 - e^{-4|\alpha|^2}. \tag{11}$$

From Eq.(9), there is a "lack of transition."

3.3 Unambiguous receiver

An unambiguous receiver is a receiver that introduces a 'failure' in measurement and guarantees 'no error' if the measurement succeeds [10, 11, 12]. When a BPSK signal is employed, the conditional probabilities for the receiver are

$$P(0|0) = P(1|1) = 1 - p_?, \tag{12}$$
$$P(1|0) = P(0|1) = 0, \tag{13}$$
$$P(?|0) = P(?|1) = p_? = e^{-2|\alpha|^2}, \tag{14}$$

where the outcome '?' corresponds to a 'failure.' From Eq.(13), there is a "lack of transition."

4 Result

4.1 Property of classical reliability function

Fig.1 shows the classical reliability functions $E_{\mathrm{CU}}(R)$ corresponding to the three kinds of quantum receivers. Here, maximization with respect to s and ξ on the right-hand side of Eq.(3) was performed numerically. Denote the capacity by C, then $E_{\mathrm{CU}}(R) = 0$ for $R > C$. The capacities by three receivers are as follows:

Helstrom Receiver:	$C^{(\mathrm{He})} = 0.66380,$
Kennedy Receiver:	$C^{(\mathrm{Ke})} = 0.64728,$
Unambiguous Receiver:	$C^{(\mathrm{Un})} = 0.59934.$

That is, $C^{(\mathrm{Un})} < C^{(\mathrm{Ke})} < C^{(\mathrm{He})}$. Thus, for example, when $R = 0.65$, $E_{\mathrm{CU}}^{(\mathrm{He})}(R) > 0$ and $E_{\mathrm{CU}}^{(\mathrm{Ke})}(R) = E_{\mathrm{CU}}^{(\mathrm{Un})}(R) = 0$. And the classical reliability function using the Helstrom receiver is largest when $R \approx C$. However, the superiority cannot be maintained when R becomes low. Actually, in an intermediate range of R, we have,

$$E_{\mathrm{CU}}^{(\mathrm{He})}(R) < E_{\mathrm{CU}}^{(\mathrm{Ke})}(R) < E_{\mathrm{CU}}^{(\mathrm{Un})}(R),$$

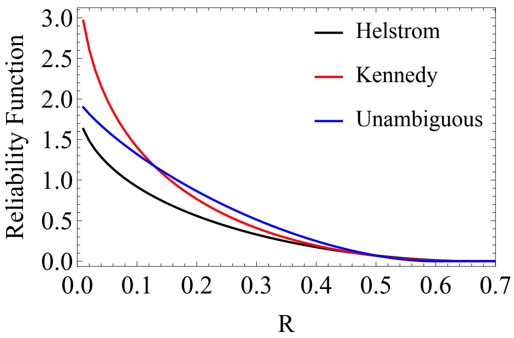


Figure 1: Classical reliability function $E_{\text{CU}}(R)$

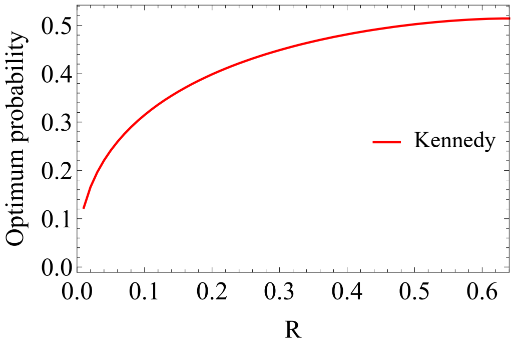


Figure 2: Optimal *a priori* probability of classical reliability function in the Kennedy receiver

and we also have,

$$E_{\text{CU}}^{(\text{He})}(R) < E_{\text{CU}}^{(\text{Un})}(R) \ll E_{\text{CU}}^{(\text{Ke})}(R)$$

when $R \approx 0$. Although no divergence of the classical reliability function was observed for either receiver, that of the Kennedy receiver has a steep change at low rates. Therefore, at lower rates, the decoding error probability of the Kennedy receiver is expected to lower sharply even at short codeword lengths.

4.2 Optimal *a priori* probability

Since the quantum measurement channels for the Helstrom and unambiguous receivers are symmetric, the optimal *a priori* probabilities are uniform [7]. As for the Kennedy receiver, the optimal *a priori* probability ξ_0 of the signal $|\psi_0\rangle = |\alpha\rangle$ is shown in Fig.2. Note $\xi_1 = 1 - \xi_0$.

From Fig.2, when the rate R is large and close to the capacity, the optimal *a priori* probabilities are almost uniform. However, when the rate is low, the optimal *a priori* probability of $|\psi_0\rangle = |\alpha\rangle$ becomes small, and it is suggested to send more $|\psi_1\rangle = |-\alpha\rangle$. This result can be explained as follows: Since $P(1|0) = 0$ from the channel matrix, it is obvious that when 1 is received, the receiver knows the transmitted signal is 1 without error, so sending more 1 will reduce the average probability of error.

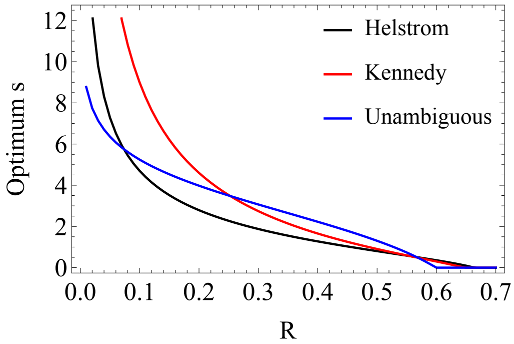


Figure 3: Optimal parameter of s in $\nu(s,\xi) - sR$

4.3 Optimal parameter of s

The maximization of $\nu(s,\xi) - sR$ with respect to s was performed numerically. Fig.3 shows the optimal s for each receiver. For any receiver, the optimal s is $s \leq 1$ at high rates. In this case, the upper and lower bounds of the classical reliability function coincide, indicating that the true value of the classical reliability function itself is obtained. In contrast, the divergence of the classical reliability function occurs if $\nu(s,\xi) - sR$ increases with increase in s and it has no peak due to s. From Fig.3, the optimal s for the Kennedy receiver has a very large value at low rates. Thus, although there is no divergence for the Kennedy receiver, the optimal s is much larger. Consequently, the classical reliability function was considerably larger than those for the other two receivers.

5 Conclusion

In this paper, we investigate the properties of the classical reliability function from the corresponding channel matrices for three kinds of quantum receivers; the Helstrom receiver, the Kennedy receiver, and the unambiguous receiver. Although the Helstrom receiver provides the largest capacity, it is not always the best receiver in terms of the upper bound of the classical reliability function. That shows, the Kennedy and the unambiguous receivers may have lower error probability than the optimal quantum receiver in low rates.

Acknowledgments: This work has been supported in part by JSPS KAKENHI Grant Number 20H00581, 20K20397, 21K04064, 22K20437, and The Nitto Foundation.

References

[1] C. W. Helstrom, *Quantum detection and estimation theory*, Academic Press, New York, (1976).

[2] O. Hirota, *Optical Communication Theory – Basis of Quantum Theory –*, Morikita Publishing, Tokyo, (1985). (in Japanese)

[3] R. J. Glauber, "Coherent and incoherent states of the radiation field," Phys. Rev. **131**, no.6, pp.2766-2788, (1963).

[4] R. G. Gallager, *Information Theory and Reliable Communication*, John Wiley & Sons, Inc., New York, (1968).

[5] M. V. Burnashev and A. S. Holevo, "On reliability function of quantum communication channel," Probl. Peredachi Inform. **34**, no.2, pp.1-13, (1998).

[6] M. Dalai, "Lower bounds on the probability of error for classical and classical-quantum channels," IEEE Trans. on Inform. Theory **59**, 12, pp.8027-8056, (2013).

[7] K. Kato, "A note on the reliability function for M-ary PSK coherent state signal," Tamagawa Univ. Quant. ICT Res. Inst. Bulletin **8**, 1, pp.21-25, (2018).

[8] E. Morimoto, R. Nakagawa, T. Wang, and T. S. Usuda, "Characteristics of upper bound of quantum reliability function for M-ary PSK signals and application to security evaluation of KCQ quantum cryptographic protocol," Proc. of The 44th Symposium on Information theory and its Applications (SITA), pp.92-97, (2021). (in Japanese)

[9] R. S. Kennedy, "A near-optimum receiver for the binary coherent state quantum channel," Quarterly Progress Report No.108, Research Lab. of Electrons, M.I.T., pp.219-225, (1973).

[10] I. D. Ivanovic, "How to differentiate between non-orthogonal states," Phys. Lett. **A123**, issu.6, pp.257-259, (1987).

[11] D. Dieks, "Overlap and distinguishability of quantum states," Phys. Lett. **A126**, issu.5-6, pp.303-306, (1988).

[12] A. Peres, "How to differentiate between non-orthogonal states," Phys. Lett. **A128**, issu.1-2, p.19 (1988).

[13] K. Masaki, T. Wang, S. Usami, and T. S. Usuda, "Comparison of quantum and classical properties of communication using coherent-states," Proc. of 2022 Tokai-Section Joint Conference on Electrical, Electronics, Information, and Related Engineering, E3-1, (2022). (in Japanese)

Efficient simulation of Helstrom measurement for supervised learning

Wooseop Hwang[1] * Daniel K. Park[2] [3] Israel F. Araujo[3] Carsten Blank[4] †

[1] *University, Oxford OX1 3PH, United Kingdom*
[2] *Department of Applied Statistics, Yonsei University, Seoul 03722, Republic of Korea*
[3] *Department of Statistics and Data Science, Yonsei University, Seoul 03722, Republic of Korea*
[4] *Data Cybernetics, Landsberg 86899, Germany*

Abstract.

The Helstrom measurement, an optimal method for discriminating between two quantum states with minimum error, presents new opportunities for binary classification in machine learning. Previous empirical studies have indicated that the prediction performance of Helstrom-based classifiers is affected by the number of identical copies of the quantum state used to encode the training data. However, due to the exponential growth the runtime with the number of copies in existing Helstrom classifiers, these studies have been limited to only a small number of copies. This paper presents an efficient method to simulate the Helstrom classifier for any number of copies by utilizing the relationship between the Helstrom measurement and fidelity. Numerical experiments on six standard datasets, involving up to 200 copies, clearly demonstrate the significant impact of the number of quantum copies on prediction performance. The best instances of the Helstrom classifier consistently outperform or match those of the fidelity classifier in all cases. This finding suggests that the number of copies should be optimized as a hyperparameter to achieve optimal classifier performance. Our method provides an efficient means of optimizing this parameter, thereby enhancing the Helstrom-based classifier beyond previous capabilities.

Keywords: Quantum state discrimination, Quantum-inspired machine learning, Optimal quantum measurement

1 Introduction

Quantum Machine Learning (QML) has shown promising applications and the potential to provide exponential speed-up over classical algorithms [1, 2, 3]. One of the advantages of QML is its ability to exploit quantum properties to perform classification tasks, which can be framed as quantum state discrimination problems. In previous work by Sergioli et al., the authors introduced the use of Helstrom measurement as a binary classifier [4]. They explored the fact that multiple copies of a quantum state provide additional information compared to that encoded into a single copy, improving classification performance [5]. In this paper, we propose an efficient method to classically simulate the Helstrom classifier on the number of quantum copies.

The Helstrom measurement [6], is known as the most ideal one-shot measurement that can be implemented in two states discrimination tasks [7]. Consider a labelled training data set:

$$D = \{(x_1, y_1), \dots, (x_M, y_M)\} \subseteq \mathbb{R}^d \times \{0, 1\}. \tag{1}$$

A quantum feature map $\phi(x)$ is responsible for mapping this data set to Hilbert space as $\phi(x) = |\phi(x)\rangle = |a\rangle$. There exists numerous schemes that can be used as feature maps such as amplitude encoding, Hamiltonian encoding etc. [8]. In this paper, amplitude encoding is used to efficiently simulate the encoding process. Using amplitude encoding, the data set can be mapped to

$$D' = \{(a_1, y_1), \dots, (a_M, y_M)\} \subseteq \mathbb{C}^d \times \{0, 1\}. \tag{2}$$

By defining two subsets $D_0 = \{a : (a, 0) \in D'\}$ and $D_1 = \{b : (b, 1) \in D'\}$, quantum centroids can be constructed as,

$$\rho = \frac{1}{M_a} \sum_a |a\rangle \langle a| \tag{3}$$

$$\sigma = \frac{1}{M_b} \sum_b |b\rangle \langle b| . \tag{4}$$

where M_a is the number of elements in D_0, and M_b is the number of elements in D_1. Under such conditions, the Helstrom operator is defined as

$$m_{\text{fid}} = \rho - \sigma. \tag{5}$$

Equation (5) can be rewritten in terms of eigenbasis as

$$m_{\text{fid}} = \sum_j \lambda_j |d_j\rangle\langle d_j|, \tag{6}$$

where λ_j and $|d_j\rangle$ are the eigenvalues and eigenstates of m_{fid} respectively. Then, the Helstrom projection is given by [7]

$$m_{\text{hel}} = \sum_j \text{sgn}\left[\lambda_j\right] |d_j\rangle \langle d_j| = \Pi_+ - \Pi_- \tag{7}$$

with the projectors

$$\Pi_+ = \sum_{\lambda_j \geq 0} |d_j\rangle \langle d_j| \tag{8}$$

$$\Pi_- = \sum_{\lambda_j < 0} |d_j\rangle \langle d_j| . \tag{9}$$

The classification score of a classifier is strongly related to the measurement outcome of the m_{fid} and m_{hel}

*wooseop.hwang@exeter.ox.ac.uk
†blank@data-cybernetics.com

operators. For a given test data c, fidelity and Helstrom classifiers can be constructed as

$$f_{\text{fid}}(c) = \langle c|m_{\text{fid}}|c\rangle$$
$$f_{\text{hel}}(c) = \langle c|m_{\text{hel}}|c\rangle.$$

By using Equations (5) and (7), the classifiers can be written as

$$\begin{aligned} f_{\text{fid}}(c) &= \sum_j \lambda_j |\langle c|d_j\rangle|^2 && (10)\\ &= \frac{1}{M_a}\sum_a |\langle c|a\rangle|^2 - \frac{1}{M_b}\sum_b |\langle c|b\rangle|^2 && (11)\\ f_{\text{hel}}(c) &= \sum_j \operatorname{sgn}[\lambda_j]\ |\langle c|d_j\rangle|^2. && (12) \end{aligned}$$

To reduce the complexity of simulating the Helstrom classifier, we propose that the eigenvalues of a system with k copies can be computed by Equation (13).

$$\lambda_{a,b} = \pm\sqrt{1 - |\langle a|b\rangle|^{2k}}. \qquad (13)$$

Now, for a system with k copies, define $f_{\text{fid}}^{a,b,k}$ as

$$f_{\text{fid}}^{a,b,k}(c) = |\langle c|a\rangle|^{2k} - |\langle c|b\rangle|^{2k},$$

it can be shown from (13) that there exists only two eigenvalues $\pm\lambda_{a,b}$ with $\lambda_{a,b} > 0$. Then consider,

$$\begin{aligned} \frac{1}{\lambda_{a,b}} f_{\text{fid}}^{a,b,k}(c) &= \frac{1}{\lambda_{a,b}}(|\langle c|a\rangle|^{2k} - |\langle c|b\rangle|^{2k})\\ &= \lambda_{a,b}^{-1}\lambda_{a,b}(|\langle c|d_+\rangle|^{2k} - |\langle c|d_-\rangle|^{2k})\\ &= |\langle c|d_+\rangle|^{2k} - |\langle c|d_-\rangle|^{2k}\\ &= \operatorname{sgn}[\lambda_+]\,|\langle c|d_+\rangle|^{2k} + \operatorname{sgn}[\lambda_-]\,|\langle c|d_-\rangle|^{2k}\\ &= f_{\text{hel}}^{a,b,k}(c), \end{aligned}$$

where $\lambda_\pm$ corresponds to $\pm\lambda_{a,b}$. Hence, the general relationship between the Helstrom and fidelity classifier is

$$f_{\text{hel}}(c) = \frac{1}{M_a}\frac{1}{M_b}\sum_a\sum_b \frac{1}{|\lambda_{a,b}|} f_{\text{fid}}^{a,b,k}(c). \qquad (14)$$

By substituting (13) in (14), we obtain

$$f_{\text{hel}}^{k}(c) = \frac{1}{M_a}\frac{1}{M_b}\sum_a\sum_b \frac{1}{\sqrt{1 - |\langle a|b\rangle|^{2k}}} f_{\text{fid}}^{a,b,k}(c). \qquad (15)$$

At large k, $\lambda_{a,b}$ becomes ±1, and the Helstrom and fidelity classifier become equivalent to each other.

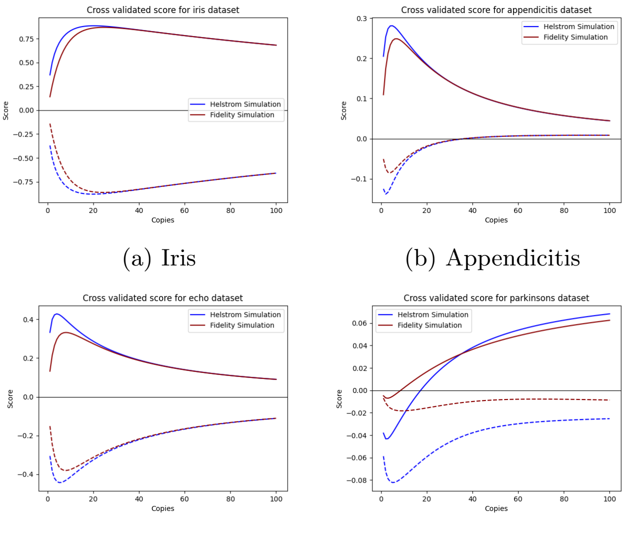
2 Numerical simulations

In this section, the simulation results for Helstrom and fidelity classifiers, and classical classifiers are presented. The classical classifiers used are: Nearest Neighbors, Linear Support Vector Classification (SVC), Radial Basis Function (RBF) SVC, Decision Tree Classifier, Random Forest Classifier, Multi-layer Perceptron Classifier, Ada Boost Classifier, Naive Bayes Classifier, Quadratic Discriminant Analysis, and Logistic Regression.

A five-fold cross-validation was performed on the entire data, which was split into five segments. The simulation was conducted on six different datasets: appendicitis [9], echocardiogram [10], hepatitis [11], iris [12], Parkinson [13], and wine [14]. Amplitude encoding was used to embed the data, and Equation (15) allowed for efficient simulation of the Helstrom and fidelity classifiers for large numbers of copies. As the datasets are unbalanced, the f1 score was used as the metric for comparing prediction performance. The number of copies of the system was the main parameter considered in the study. The classification score, which is illustrated in the figure 1, is the output of the classifiers defined in (12) and (15). Based on the sign of the classification score, we assigned a class to the given test data by using the following scheme:

$$y_{\text{pred}} = \begin{cases} 0, & \text{if } f(c) > 0\\ 1, & \text{otherwise} \end{cases}$$

where the $f(x)$ corresponds to the classifiers, and c refers to the input test data. The f1 score was computed by using y_{pred} and the actual value for each test data y_{test}.



(a) Iris

(b) Appendicitis

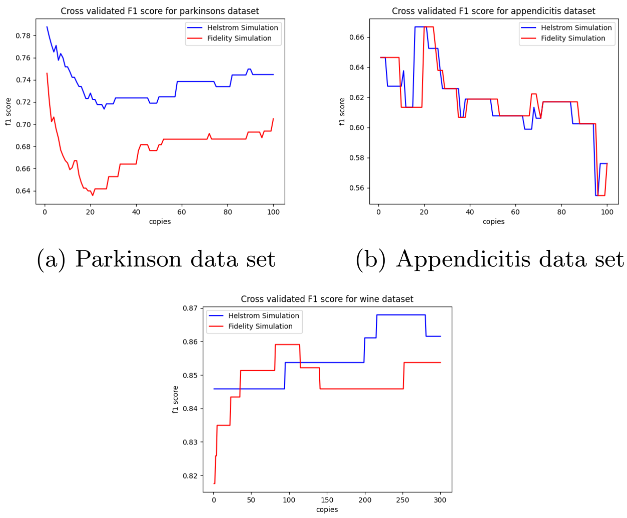
(c) Echo-cardiogram

(d) Parkinson

Figure 1: Classification score for different data sets

2.1 Overview

The simulation results for cross validated classification score are presented in the figures (1), with the classification score representing the output of equations (11) and (15). The fidelity and Helstrom scores are shown as brown and blue lines, respectively. The solid lines are used for class 0, and the dotted lines are used for class 1. The results indicate that the Helstrom classifiers generate better scores than fidelity, with the average score for class 0 closer to 1 and class 1 closer to -1. Additionally, there is a clear sweet spot in the number of copies where the classification score is maximized, and it declines beyond that point. It is worth noting that the Helstrom



(a) Parkinson data set

(b) Appendicitis data set

(c) Wine data set

Figure 2: Cross validated f1 score

score converges to the fidelity as the number of copies increases, as expected.

The simulation results for cross validated f1 score are displayed in Figure 2. The Parkinson dataset demonstrates a clear out-performance of Helstrom as shown in Figure 2a. For the remaining datasets, Helstrom performed at least as well as fidelity. The findings also indicate that the prediction performance is influenced by the number of copies.

Figure 3 presents the prediction performance comparison of classical classifiers and the Helstrom and fidelity classifiers over all six datasets. The results show that for datasets such as iris and echocardiogram, the Helstrom and fidelity classifiers performed as well as the classical classifiers. On the other hand, there was a clear out-performance of Helstrom for the hepatitis and appendicitis datasets. However, for the Parkinson dataset, the performance was lacking. Table 1 reports the number of copies required to achieve the maximum f1 score. These results indicate that the performance of the Helstrom and fidelity classifiers is highly dependent on the dataset being analyzed.

Data set	appendicitis	echo	hepatitis	iris	parkinsons	wine
Helstrom	16	78	34	1	1	216
Fidelity	20	81	33	1	1	82

Table 1: Copies at which the maximum f1 scores were obtained. The second row illustrates the optimal copies for the Helstrom classifier. The last row presnets the optiaml copies for the fidelity classifier.

3 Conclusion

The performance of the Helstrom and fidelity classifiers can be affected by the number of copies. For the data sets we analyzed, we observe a peak in the classification score before declining. Therefore, choosing the appropriate number of copies is crucial to obtain the best prediction performance.

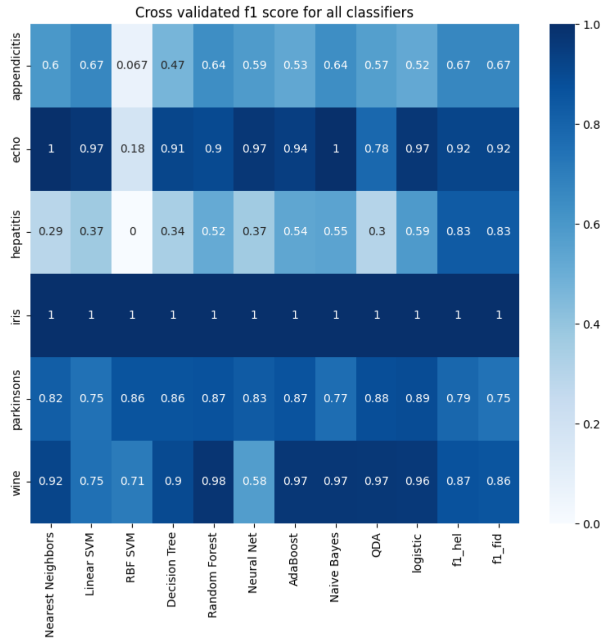


Figure 3: Comparison of f1 scores with other classifiers

The proposed method allows for an efficient simulation of the Helstrom classifier on any number of copies. This is achieved by computing the eigenvalues of the Helstrom operator directly using the overlap between pairwise combination of train data of different classes, instead of using eigen-decomposition. Experimental results demonstrate that the Helstrom classifier performs equally well as the fidelity classifier. Moreover, the performance of both classifiers in terms of prediction is influenced by the number of copies, which is not a monotonic effect and varies depending on the dataset. Therefore, it is recommended to treat the number of copies as a hyper-parameter. The efficient simulation of quantum classifiers at large numbers of copies allows for a more thorough exploration of the hyper-parameter space.

One potential avenue for future research is the implementation of the kernel trick to theoretically compare the prediction performance of fidelity and Helstrom classifiers. This approach would involve comparing the geometric differences between two different kernels, as shown in previous work on quantum classifiers [15]. Additionally, further improvements to our results may be possible by exploring different embeddings that consist solely of Clifford gates, which can be simulated classically. This is important since the prediction performance is dependent on the data embedding [16].

References

[1] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. Prediction by linear regression on a quantum computer. *Phys. Rev. A*, 94:022342, Aug 2016.

[2] Peter Wittek. *Quantum Machine Learning: What Quantum Computing Means to Data Mining.* 08 2014.

[3] Nana Liu and Patrick Rebentrost. Quantum machine learning for quantum anomaly detection. *Phys. Rev. A*, 97:042315, Apr 2018.

[4] Giuseppe Sergioli, Roberto Giuntini, and Hector Freytes. A new quantum approach to binary classification. *PLOS ONE*, 14(5):1–14, 05 2019.

[5] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete. Discriminating states: The quantum chernoff bound. *Phys. Rev. Lett.*, 98:160501, Apr 2007.

[6] Carl W. Helstrom. Quantum detection and estimation theory. *Journal of Statistical Physics*, 06 1969.

[7] Joonwoo Bae and Leong-Chuan Kwek. Quantum state discrimination and its applications. *Journal of Physics A: Mathematical and Theoretical*, 48(8):083001, jan 2015.

[8] M. Schuld and F. Petruccione. *Machine Learning with Quantum Computers.* Quantum Science and Technology. Springer International Publishing, 2021.

[9] Joseph D Romano, Trang T Le, William La Cava, John T Gregg, Daniel J Goldberg, Praneel Chakraborty, Natasha L Ray, Daniel Himmelstein, Weixuan Fu, and Jason H Moore. Pmlb v1.0: an open source dataset collection for benchmarking machine learning methods. *arXiv preprint arXiv:2012.00058v2*, 2021.

[10] UCI Machine Learning Repository. Echocardiogram, 1989.

[11] UCI Machine Learning Repository. Hepatitis, 1988.

[12] R.A. Fisher. Iris, 1988.

[13] Max Little. Parkinsons. UCI Machine Learning Repository, 2008. DOI: 10.24432/C59C74.

[14] UCI Machine Learning Repository. Wine, 1991.

[15] Hsin-Yuan Huang, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, and Jarrod R. McClean. Power of data in quantum machine learning. *Nature Communications*, 12(1), may 2021.

[16] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. Quantum embeddings for machine learning, 2020.

Magic state injection on Heavy-hexagon structure

Younhun Kim[1] * Hansol Kim[1] † Wonjae Choi[1] ‡ Younhun Kwon[1] §

[1] *Department of Applied Physics, Hanyang University(ERICA), Ansan 15588, Republic of Korea*

Abstract. Magic state is necessary to build a Non-Clifford logic operator. Up to now, a magic state injection has been proposed only in the lattice structure. However, IBM quantum computer has the heavy-hexagon structure. In this work, we propose a magic state injection in the heavy-hexagon structure. We show that in the heavy-hexagon structure magic state of distance of 5 can be built in the logical error rate of 0.382% with success probability of 29.27% when the single qubit error rate is $p = 0.05\%$.

Keywords: Surface code, Lattice structure, Heavy-hexagon structure, Magic state

1 Introduction

Present quantum computers show the property that as the depth of the quantum circuits is deepened, the errors increase, and the quantum computers become useless. To avoid errors in quantum computer, the quantum error correction codes(QEC codes) have been suggested[1, 2]. The well-known QEC code is the stabilizer code which recently revealed the reducing errors in Google's quantum computer[3, 4]. The typical QEC code among stabilizer codes is the surface code where the physical qubits are located in a lattice structure, and each qubit has a maximum of four neighboring qubits. In surface code, the measurement round to stabilizer should be performed to detect errors locally. The syndromes are obtained from the result of measurement, and the correction of errors can be done through a decoding algorithm[5, 6, 7].

A set of universal logic operators should be constructed to build a quantum circuit. The Clifford logic operators can be obtained through transversal gates[2, 8], but Non-Clifford logic operators cannot be obtained through transversal gate[9]. To build Non-Clifford logic operators, one should prepare magic state. For example, to perform $\hat{T}_L$, one needs to have $|M\rangle_L = |0\rangle_L + e^{i\pi/4}|1\rangle_L$. Through the combination of magic state and Clifford logic operator $\hat{T}_L$ can be performed[10].

To build a magic state $|M\rangle_L$, we make a projection of magic state $|M\rangle$ in the level of physical qubits, onto the logical quantum state. This process is called magic state injection. The method for the magic state injection has been proposed in the lattice structure[11, 12, 13]. Since the magic state injection is performed not at the level of logical quantum state but at the level of the physical qubit, the process for quantum error correction should be included[14].

However, one of the most promising quantum computers, such as IBM quantum computers, uses the heavy-hexagon structure. Therefore, in the heavy-hexagon structure, a new method for the magic state injection should be needed because the maximum number of nearest neighbors in the lattice structure is four, but the maximum number of nearest neighbors in the heavy-hexagon structure is three[17, 18, 19, 20]. Therefore, in this work, we propose a method for the magic state injection in the heavy-hexagon structure.

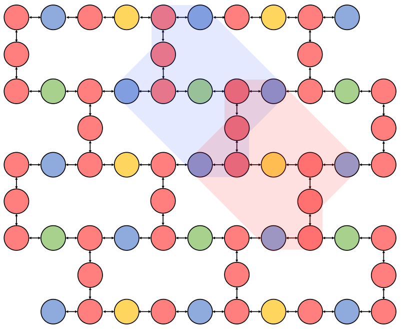


Figure 1: The arragement of qubits for quantum error correction in the heavy-hexagon structure. The blue, green(yellow), and red nodes denote the data, syndrome, and flag qubits, respectively. The blue plane displays Z stabilizer, and the red plane displays X stabilizer.

2 QEC code on heavy-hexagon structure and Magic state injection

In this section, we explain the quantum error correcting code in the heavy-hexagon structure. Fig.1 shows the arrangement of qubits in the heavy-hexagon structure, where the blue, green(yellow), and red nodes denote the data, syndrome, and flag qubits, respectively. The stabilizer consists of four data qubits around the syndrome qubit in the form of $Z^{\otimes 4}$ or $X^{\otimes 4}$. The blue plane displays Z stabilizer, and the red plane displays X stabilizer. Since in the heavy-hexagon structure, the maximum number of interaction with nearest qubits is three, we use six additional flag qubits.

To detect an error in data qubits, we should build the measurement circuit for stabilizer. Fig.2 shows the measurement circuit of X stabilizer and Z stabilizer in the heavy-hexagon structure.

The protocol for Magic state injection consists of two stages. In the first stage we setup the initial state of physical qubits and perform the projection on the logical quantum state, to build $|M\rangle_L$ with distance of d_1. The $|M\rangle_L$, projected on the logical quantum state, can be protected from errors by the quantum error correc-

*hpoqh@hanyang.ac.kr

†khshk18@hanyang.ac.kr

‡marchenw@hanyang.ac.kr

§yyhkwon@hanyang.ac.kr

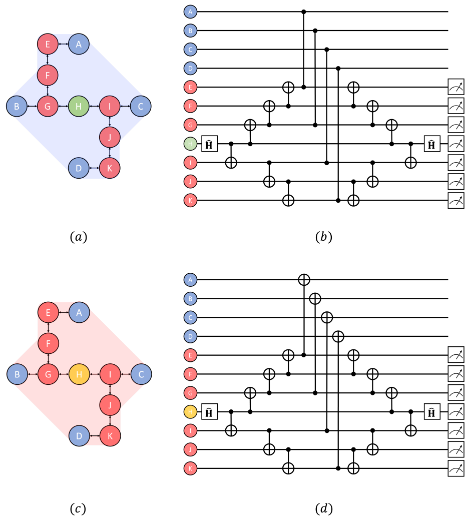


Figure 2: (a)The qubit arrangement of Z stabilizer in in the heavy-hexagon structure (b)The measurement circuit of Z stabilizer (c)The qubit arrangement of X stabilizer in in the heavy-hexagon structure (d)The measurement circuit of X stabilizer

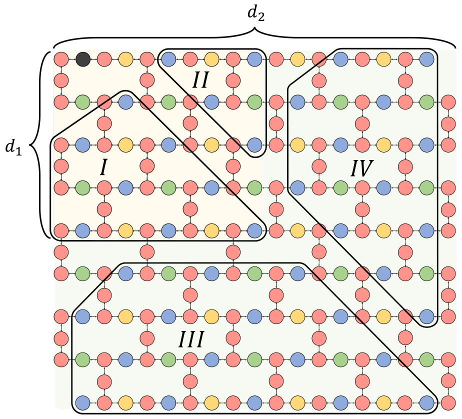


Figure 3: The region of I,II,III, and IV for performing the magic state injection.

tion. In the second stage d_1 Magic state is extended to d_2. The black node in the top left of Fig.3 denotes the magic state in the level of physical qubit. For the magic state injection the remaining nodes are divided into the regions of I,II,III, and IV[11]. The each node is used as the data qubit of d_1 or d_2 in the magic state.

2.1 Stage I

- As the first step of Stage I, we determine the initial quantum states of physical qubits located in the region I and II, indicated in Fig.3. The data qubits in the region I are prepared as $|+\rangle$, and the data qubits in the region II are prepared as $|0\rangle$. Meanwhile, for the round of measurement of stabilizer, syndrome and flag qubit are prepared as $|+\rangle$ and $|0\rangle$, respectively. And a single qubit state $|M\rangle$ is prepared in upper-left side.

- As the second step of Stage I, we perform twice the round of measurement of stabilizer. In each round of measurement X stabilizer and Z stabilizer are measured independently. After the measurement, syndrome qubit and flag qubit are initialized for the next round.

- As the final step of Stage I, we check whether $|M\rangle$ is projected to the logical quantum state as the result of measurement of stabilizer. If there is no error from the result of measurement obtained by two rounds of measurement of stabilizer, $|M\rangle$, which is the quantum state of physical qubit, should be projected on the logical quantum state. If even a single error is detected, we should return to the first step.

2.2 Stage II

- As the first step of Stage II, to extend the distance of magic state to d_2, we add qubits in region III and IV. The physical qubits in the region III are prepared as $|+\rangle$ and the physical qubits in the region IV are prepared as $|0\rangle$.

- As the second step of Stage II, we perform the measurement of stabilizer to the whole region of d_2. By performing d_2 times of round of measurement of stabilizer, we obtain the syndrome. If errors are detected, we correct the errors by a decoding algorithm. Through this process, we can prepare $|M\rangle_L$ of d_2-size. And we can perform non-Clifford logical operators even in the heavy-hexagon structure.

In Stage I the quantum state $|M\rangle_L$ consists of $|M\rangle = \alpha|0\rangle + \beta|1\rangle$. The quantum state in the first step of Stage I, when the data qubits consisting of Z_L are considered, becomes $|0\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data}(\ |1\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data})$. Let us denote $|\phi\rangle_{data}$ as the quantum state of data qubits except d_1 qubits consisting of Z_L. Therefore, the total quantum state before the round of measurement of stabilizer can be written as follows:

$$\begin{aligned}|\psi\rangle = {} & \alpha|0\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data} \otimes |\phi\rangle_{data} \\ & +\beta|1\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data} \otimes |\phi\rangle_{data}\end{aligned} \tag{1}$$

If there is no error when two times of round of measurement of stabilizer are performed, $|\psi\rangle$ is projected on the logical quantum state with $+1$ eigenvalue to every stabilizer. We denote the i-th element of total set of stabilizer as S_i. Then the following expression provides $|\psi\rangle_L$ in terms of $|\psi\rangle$ and S_i.

$$\begin{aligned}|\psi\rangle_L &= \prod_i (I+S_i)|\psi\rangle \\ &= \alpha\prod_i (I+S_i)|0\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data} \otimes |\phi\rangle_{data} \\ &+ \beta\prod_i (I+S_i)|1\rangle_{magic} \otimes |0\rangle^{d_1-1}_{data} \otimes |\phi\rangle_{data} \\ &= \alpha|0\rangle_L + \beta|1\rangle_L\end{aligned} \tag{2}$$

The error caused in preparing the magic state can be regarded mainly for three reasons.The first one is that the

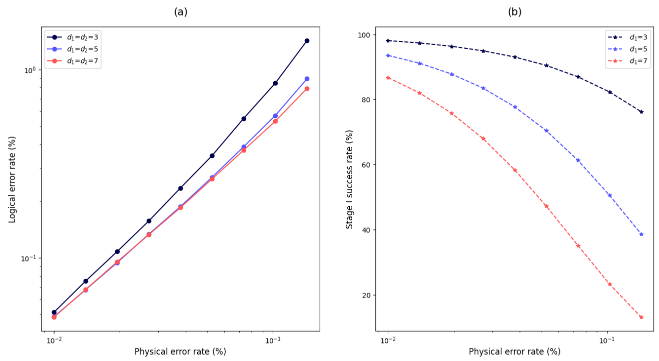


Figure 4: (a)The logical error rate of surface code to the magic state in the lattice structure. (b)The success probability of Stage I of surface code to the magic state in the lattice structure.

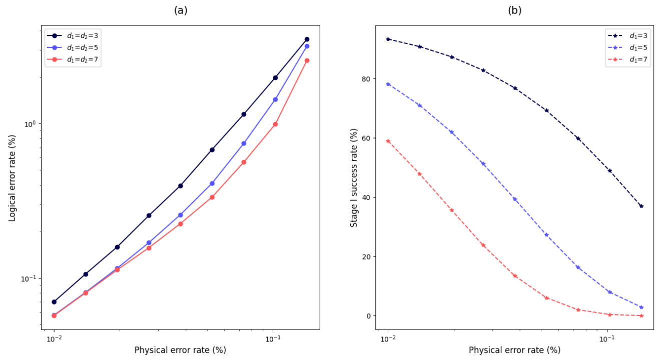


Figure 5: (a)The logical error rate of surface code to the magic state in the heavy-hexagon structure. (b)The success probability of Stage I of surface code to the magic state in the heavy-hexagon structure.

single qubit error in the preparation of physical qubits of the magic state causes an error in the magic state. The second one is that because one should use CNOT(CZ) operation between flag qubit and magic state of level of the physical qubit in the measurement circuit of stabilizer, if there is an error in the two-qubit operations, there can be an error in the magic state. The last one is that even though every data qubit consisting of Z_L should be the zero state, there can be an error in the magic state if there is a flip.

3 Numerical results

We consider a depolarization error model for a numerical analysis. In the model when the error rate of a single physical qubit is p the error in $\{X, Y, Z\}$ becomes identically $p/3$. The two qubit operator in this work is only CNOT operator and CZ operator consists of the combination of Hadamard gate and CNOT operator. And the error rate of 15-two qubit operator $\{I, X, Y, Z\}^{\otimes 2}/\{I \otimes I\}$ becomes $p/15$. The error rate in reset of $|0\rangle$ is $2p/3$. When performing the round of measurement of stabilizer, we consider an idling error in every data qubit. The error in every gate is an independent and identical distribution.

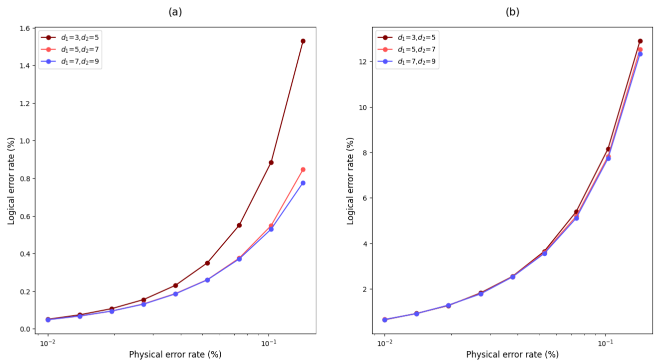


Figure 6: The performance of magic state injection protocol in the case of $d_2 = d_1 + 2$. (a) The performance of magic state injection protocol in the lattice structure. (b)The performance of magic state injection protocol in the heavy-hexagon structure.

As p varies, we obtain the logical error rate to quantum error correction of magic state injection in lattice and heavy-hexagon structure and get the success probability to Stage I. We evaluate the logical error rate to quantum error correction of magic state, by comparing the quantum state of $|M\rangle$ with the quantum state of $|M\rangle_L$ obtained through quantum error correction in Stage II. We get the success probability to Stage I, from the ratio between the total sample and the result of +1 from the syndrome in the round of measurement of stabilizer. The simulation code is made by the Stim code and we obtain the syndrome samples by applying the depolarization error model[21]. The Pymatching algorithm is applied to the syndrome samples for a decoding algorithm[22].

Fig.4 and Fig.5 show the results when the magic state injections are performed in the lattice and heavy-hexagon structure, respectively. Here the value of d_1 is $\{3, 5, 7\}$. And the value of d_2 is equal to that of d_1. Fig.6 shows the result of magic state injection in the case of $d_2 = d_1 + 2$. The number of samples is 3×10^6.

In the case of $d_1 = d_2$, as the distance becomes larger, the logical error rate to the magic state becomes less not only in the lattice structure but also in the heavy-hexagon structure. In the case of $d_1 = d_2 = 3$, at the value of $p = 0.05\%$ the logical error rate to the magic state in the lattice and heavy-hexagon structure become 0.328% and 0.627% respectively. Meanwhile, in the case of $d_1 = d_2 = 7$, at the value of $p = 0.05\%$ the logical error rate to the magic state in the lattice and heavy-hexagon structure become 0.247% and 0.314% respectively. In the case of $d_1 = d_2 = 5$, the success probability of Stage I in the heavy-hexagon structure becomes 29.27% and the logical error rate to the magic state in the heavy-hexagon structure becomes 0.382%. And we can see that as p increases, the success probability of Stage I decreases rapidly both in the lattice and heavy-hexagon structure. In Fig.6 we note that when d_2 is not d_1, as the value of d_1 becomes larger, the logical error rate to the magic state in the lattice and heavy-hexagon structure becomes less.

4 Conclusions

In this work, we proposed a method to perform a protocol for magic state injection in a heavy-hexagon structure. Further, we compared the logical error rate of magic state in the heavy-hexagon structure with that in the lattice structure. And we showed that the magic state can be built even in the heavy-hexagon structure.

Acknowledgment

This work is supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF2022R1F1A1064459) and Creation of the Quantum Information Science R&D Ecosystem (Grant No. 2022M3H3A106307411) through the National Research Foundation of Korea (NRF) funded by the Korean government (Ministry of Science and ICT).

References

[1] G. Daniel. Stabilizer codes and quantum error correction. California Institute of Technology, 1997.

[2] A. G. Fowler, A. C. Whiteside, and L. CL. Hollenberg. Towards practical classical processing for the surface code. Physical review letters 108.18:180501, 2012.

[3] Google Quantum AI Exponential suppression of bit or phase errors with cyclic error correction. Nature 595.7867:383-387, 2021.

[4] Google Quantum AI Suppressing quantum errors by scaling a surface code logical qubit. Nature 614.7949:676-681, 2023.

[5] J. Edmonds Paths, trees, and Flowers. Canadian Journal of Mathematics. 17, 449–467, 1965.

[6] V. Kolmogorov. Blossom V: A new implementation of a minimum cost perfect matching algorithm. Mathematical Programming Computation 1, 43–67, 2009.

[7] A. G. Fowler Minimum weight perfect matching of fault-tolerant topological quantum error correction in average O(1) parallel time. Quantum Information and Computation. 15, 145–158, 2015.

[8] C. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter. Surface code quantum computing by lattice surgery. New Journal of Physics, 14(12), 123011, 2012.

[9] B. Eastin, and E. Knill. Restrictions on transversal encoded quantum gate sets. Physical review letters 102.11:110502, 2009.

[10] S. Bravyi, and A. Kitaev. Universal quantum computation with ideal Clifford gates and noisy ancillas. Physical Review A 71.2:022316, 2005.

[11] Y. Li. A magic state's fidelity can be superior to the operations that created it. New Journal of Physics 17.2:023037, 2015.

[12] S. Singh, et al. High-fidelity magic-state preparation with a biased-noise architecture. Physical Review A 105.5:052410, 2022.

[13] L. Lao, and B. Criger. Magic state injection on the rotated surface code. Proceedings of the 19th ACM International Conference on Computing Frontiers. 2022.

[14] S. Bravyi, J. Haah. Magic-state distillation with low overhead. Physical Review A 86.5:052329, 2012.

[15] C. Rigetti, and M. Devoret. Fully microwave-tunable universal gates in superconducting qubits with linear couplings and fixed transition frequencies. Physical Review B 81.13:134507, 2010.

[16] J. M. Chow, et al. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. Physical review letters 107.8:080502, 2011.

[17] E. J .Zhang, et al. High-performance superconducting quantum processors via laser annealing of transmon qubits. Science Advances 8.19:eabi6690, 2022.

[18] C. Chamberland, et al. Topological and subsystem codes on low-degree graphs with flag qubits. Physical Review X 10.1:011022, 2020.

[19] A. Wu, et al. Mapping Surface Code to Superconducting Quantum Processors. arXiv preprint arXiv:2111.13729, 2021.

[20] Y. Kim, J. Kang, and Y. Kwon. Design of Quantum error correcting code for biased error on heavy-hexagon structure. arXiv preprint arXiv:2211.14038, 2022.

[21] C. Gidney. Stim: A fast stabilizer circuit simulator. Quantum. 5, 497, 2021.

[22] O. Higgott, and C. Gidney. Sparse Blossom: correcting a million errors per core second with minimum-weight matching. arXiv preprint arXiv:2303.15933, 2023.

Geometric Quantum Machine Learning

Martín Larocca[1] * Louis Schatzki[1 2] Quynh T. Nguyen[1 3] Paolo Braccia[1]
Michael Ragone[1] Patrick J. Coles[4] Frédéric Sauvage[1] M. Cerezo[1 †]

[1] *Los Alamos National Laboratory, Los Alamos, NM 87545, USA*
[2] *University of Illinois Urbana-Champaign, Illinois 61820, USA*
[3] *Harvard University, Cambridge, Massachusetts 02138, USA*
[4] *Normal Computing Corporation, New York, New York, USA*

Abstract. Quantum Machine Learning (QML) presents exciting opportunities for accelerating data analysis. Yet, current architectures, while broadly applicable, often encounter substantial hurdles related to trainability, generalizability, and scalability. This limitation has prompted the exploration of problem-specific architectures, which could potentially mitigate these challenges. Among the promising avenues is the adoption of Geometric Deep Learning principles in the creation of Geometric QML (GQML) models. In essence, the idea in GQML is to exploit symmetry to focus the search resources into the relevant region of model space. In this talk, we will introduce various methodologies for constructing these models and highlight their benefits.

Keywords: Quantum Machine Learning, Geometric Deep Learning, Representation Theory

1 Introduction

Quantum Machine Learning (QML) models aim at learning from data encoded in quantum states. Recently, it has been shown that models with little to no inductive biases are likely to have trainability and generalization issues, especially as we aim at scaling them to real world problems. As such, it is fundamental to develop models that encode as much information, or assumptions, as available about the problem at hand. This often untapped source of information comes predominantly as knowledge of the underlying symmetries of the problem.

In the last years, the study of symmetries has revolutionized ML. In particular, the field of Geometric Deep Learning (GDL) [1] has thrived in conceptualizing and extending the ideas behind the success of convolutional neural networks (CNNs). The main insight was to recognize that the translation invariance of CNNs was key in their ability to scale to high-dimensional datasets,and that the same principle could be extended to more general groups of symmetries, resulting in the development of so-called group CNNs [2, 3, 1, 4] that have already found applications in many domains of science.

The aim of this work is to show that (and how) one can exploit symmetries in a problem to construct novel QML architectures with sharp geometric priors [5, 6, 7]. As a first step, we characterize the space of group-invariant QML models [5]. For instance, we apply such characterization to identify solutions to problems of purity and entanglement learning. In more general cases, however, such a direct approach is not practical and one needs to parameterize the relevant search space to turn the learning problem into an optimization one. To address this problem, we recur to equivariant quantum neural networks (EQNNs) which provide a constructive apporach to group invariance. We lay down a general theory for EQNNs [6] and demonstrate their merits in practical tasks [7].

2 Results

Group-invariant QML models. Consider a generic task of supervised QML aiming at learning an unknown function f provided a dataset $\mathcal{S} = \{(\rho_i, y_i)\}_{i=1}^N$ consisting of quantum states ρ_i and labels $y_i = f(\rho_i)$. One starts by positing a family of parameterized models $h_{\boldsymbol{\theta}}$ that could recover f for an adequate values of $\boldsymbol{\theta}$. A prototypical choice in the realm of QML are models consisting of a quantum neural network (QNN) $U(\boldsymbol{\theta})$ (e.g., a quantum circuit with parameterized gates) operating on k copies of an input state ρ, followed by the measurement of an observable O. That is, models of the form $h_{\boldsymbol{\theta}}^{(k)}(\rho) = \mathrm{Tr}[U(\boldsymbol{\theta})(\rho^{\otimes k})U^\dagger(\boldsymbol{\theta})O]$. Once the model structure is defined, one proceeds by optimizing the parameters $\boldsymbol{\theta}$ to minimize discrepancies between model predictions and labels.

In many situations of interest, information about the task is not limited to the dataset, but also en-

* larocca@lanl.gov
† cerezo@lanl.gov

Task	$\mathfrak{G}$	$\mathcal{C}^{(1)}(\mathfrak{G})$	$\mathcal{C}^{(2)}(\mathfrak{G})$
Purity	$\mathbb{U}(d)$	$\mathbb{1}_d$	$\mathbb{1}_d \otimes \mathbb{1}_d, \mathrm{SWAP}$
Ent.	$\bigotimes_{j=1}^n \mathbb{U}(2)$	$\mathbb{1}_d$	$\bigotimes_{j=1}^n \{\mathbb{1}_4^j, \mathrm{SWAP}^{\mathrm{j}}\}$

Table 1: For the purity and entanglement tasks, we provide the symmetry group $\mathfrak{G}$ and elements spanning the 1st and 2nd order commutant. Here, SWAP the operator swapping the two d-dimensional copies of ρ, and $\mathbb{1}_4^j$ ($\mathrm{SWAP}^{\mathrm{j}}$) is the identity (SWAP) acting on the jth qubit of each of the $k = 2$ copies.

compasses some knowledge about f. In particular, one often knows that the labels produced by f remain invariant under some set of operations $\mathfrak{G}$ on the input, $f(V_g \rho V_g^\dagger) = f(\rho)$ for $g \in \mathfrak{G}$, in which case we say f is $\mathfrak{G}$-invariant. The set of such transformations adopts the structure of a group, acting on quantum states through a unitary representation V_g for all $g \in G$. Given the a-priori knowledge of such invariance, it is natural to require that the model $h_{\boldsymbol{\theta}}$ should be, by-design, $\mathfrak{G}$-invariant for any $\boldsymbol{\theta}$.

For our purposes, the symmetries of $\mathfrak{G}$ are captured by its *kth order commutant* $\mathcal{C}^{(k)}(\mathfrak{G}) = \{W \in \mathbb{C}^{dk \times dk} \mid [W, V_g^{\otimes k}] = 0\,,\ \forall V_g \in \mathfrak{G}\}$, which is the vector space of all complex matrices commuting with k-fold tensor products of elements of $\mathfrak{G}$. The case $k = 1$ copies corresponds to the standard commutant of $\mathfrak{G}$. At a mathematical level, invariance of $h_{\boldsymbol{\theta}}^{(k)}$ can be achieved by restricting $\widetilde{O}(\boldsymbol{\theta}) = U^\dagger(\boldsymbol{\theta}) O U(\boldsymbol{\theta})$ to belong to $\mathcal{C}^{(k)}(\mathfrak{G})$ (see Proposition 1 in Ref. [8]). Notably, the structure of $\mathcal{C}^{(k)}(\mathfrak{G})$ can be directly deduced from how $V_g^{\otimes k}$ block-decomposes into *irreducible representations (irreps)*. Given that many groups occurring in the realm of quantum physics have been thoroughly studied, we can often leverage representation-theory results to readily characterize the space of possible $\widetilde{O}(\boldsymbol{\theta})$, as is now exemplified.

Toy problems. As a warm-up excercise, we propose to analize two examplary QML tasks: learning the purity and entanglement of states. For the purity task the labels of the dataset are given as $f(\rho_i) = 1(0)$ if ρ_i is pure (mixed), while for the entanglement one we assume that f is some multipartite entanglement measure. In both cases, the group $\mathfrak{G}$ of corresponding symmetries is reported in Table. 1. Then, recalling the irreducible nature of $\mathbb{U}(d)$ (i.e., it does not admit any non-trivial block structure), and by virtue of the commutation theorem for tensor products [9, 10], we find that $\mathcal{C}^{(1)}(\mathfrak{G})$ is trivial in both cases (i.e., only uninformative $\mathfrak{G}$-invariant models can be realized). From an ML perspective, $h_{\boldsymbol{\theta}}^{(1)}$ is ultimately a linear model (in the entries of ρ) and thus is ill-fitted to approximate non-linear functions such as the purity or entanglement. However, non-linearities can be introduced when allowing the models to act on several copies of ρ, with $k = 2$ copies sufficing for our purposes. Invoking the Schur-Weyl duality, we obtain the 2nd order commutants. For the purity task, inspecting the non trivial part of $\mathcal{C}^{(2)}(\mathfrak{G})$ one finds that choosing $\widetilde{O}(\boldsymbol{\theta}) = \mathrm{SWAP}$ yields a model $h_{\boldsymbol{\theta}}^{(2)}(\rho) = \mathrm{Tr}[\rho^2]$ solving the task. For the entanglement case, $\mathcal{C}^{(2)}(\mathfrak{G})$ is spanned by 2^n elements and, while not all choices are useful, it is remarkable that we can recover in an elegant way all the multipartite entanglement measures proposed in [11, 12, 13, 14, 15, 16, 17] as special choices of $\mathcal{C}^{(2)}(\mathfrak{G})$.

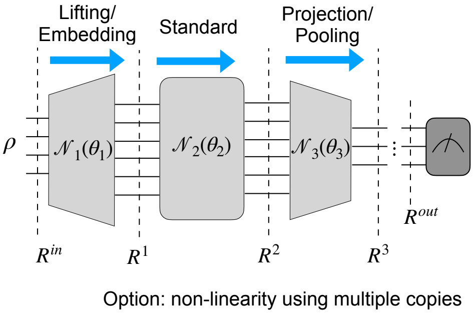

Lifting/
Embedding
Standard
Projection/
Pooling
ρ
$\mathcal{N}_1(\theta_1)$
$\mathcal{N}_2(\theta_2)$
$\mathcal{N}_3(\theta_3)$
R^{in}
R^1
R^2
R^3
R^{out}
Option: non-linearity using multiple copies

Figure 1: General equivariant QNNs are composed of equivariant layers satisfying Eq. (1) acting on potentially different spaces and representations.

Designing parameterized geometric quantum machine learning models. So far, we have characterized the *ultimate form* that a $\mathfrak{G}$-invariant model $h_{\boldsymbol{\theta}}^{(k)}$ can adopt: a $\tilde{O}(\boldsymbol{\theta}) \in \mathcal{C}^{(k)}(\mathfrak{G})$. However, this neither prescribes how to parameterize $U(\boldsymbol{\theta})$, nor tells us how to choose the observable O that realizes $\widetilde{O}(\boldsymbol{\theta})$. Furthermore, while for sufficiently large symmetry group (or equivalently sufficiently small commutants), one may be able to directly identify a solution of the problem, in general the space of invariant models is too large to directly reason over. We now explore a more constructive approach towards the design of $\mathfrak{G}$-invariant models.

Composable invariant models. Consider a model as composed of M maps, $h_{\boldsymbol{\theta}} = \mathcal{N}_M(\boldsymbol{\theta}_M) \circ \cdots \circ \mathcal{N}_1(\boldsymbol{\theta}_1)$, where each map can represent the action of a layer U_l of U, a measurement, post-processing, or encoding classical data into quantum states in the first place. Although imposing *invariance* at the level of these maps effectively enforces invariance of the model, this is too restrictive. A more relaxed approach towards invariant models involves the concept of *equivariance* [2, 3, 1]. A map $\mathcal{N} : \mathcal{A} \mapsto \mathcal{B}$ is

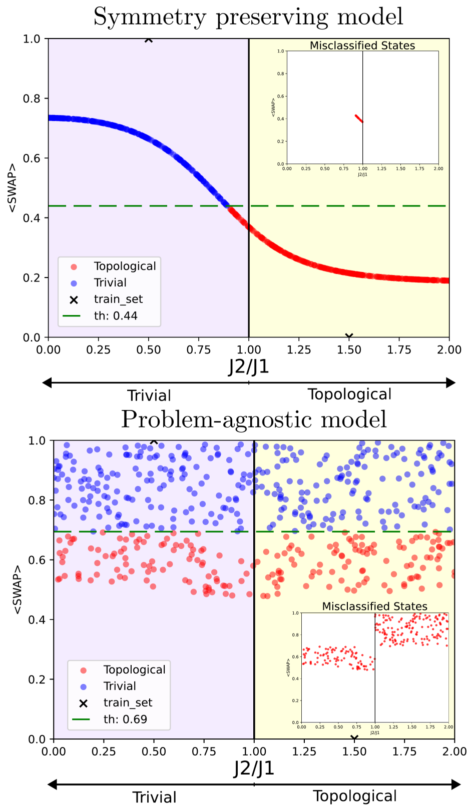


Figure 2: Comparison of an equivariant QCNN (left) with a non-equivariant one (right) for classification of ground states of the XXX model.

called $\mathfrak{G}$-equivariant if it it satisfies:

$$\mathcal{N} \circ R^{\mathcal{A}}(g) = R^{\mathcal{B}}(g) \circ \mathcal{N},\ \forall g \in \mathfrak{G} \tag{1}$$

with representations $R^{\mathcal{A}}$ and $R^{\mathcal{B}}$ of the symmetries $g \in \mathfrak{G}$ on the spaces $\mathcal{A}$ and $\mathcal{B}$. In turn, it can be verified that $\mathfrak{G}$-equivariance of all $\mathcal{N}^{(m<M)}$ along with $\mathfrak{G}$-invariance of $\mathcal{N}^{(M)}$ ensures $\mathfrak{G}$-invariance of the composed overall model.

Equivariant QNNs. Given that our goal is to develop parameterized quantum models, we focus on the case when $\mathcal{N}$ is a completely-positive trace-preserving channel (CPTP). Eq. (1) recovers the concept of the commutant developed thus far in the case of unitaries, but generalizes to quantum channels over arbitrary input and output spaces. As seen in Fig. 1 these can be used to compose general EQNNs, and to realize in quantum models the main components of GDL architectures including pooling, lifting and standard convolutions. In Ref. [6] we provide a set of complementary techniques to discover $\mathfrak{G}$-equivariant channels, one of which is presented now. First, in the case of finite groups, we show that Eq. (1) only needs to be imposed on a subset $S \subseteq \mathfrak{G}$ of generators. Left with these $|S|$ constraints, the task of solving Eq. (1) can be mapped to a nullspace problem that can be solved symbolically or numerically [18]. Explicitly, for all $g \in S$ we need to solve $(\mathbb{1} \otimes R^{\mathcal{A}}(g)^t - R^{\mathcal{B}}(g) \otimes \mathbb{1})\text{vec}(\mathcal{N}) = 0\,,$ where vec$(\cdot)$ denotes vectorization. The space of all equivariant channels is obtained by intersecting the nullspaces corresponding to each $g \in S$ and further restricting it to CPTP maps.

Application: $\mathbb{SU}$-equivariant QCNN for ground-state classification. To demonstrate the importance of EQNNs, we consider the task of classifying ground states ρ_i of the a *bond-alternating* XXX Heisenberg model $H = \sum_{i=1}^{n} J_i(X_iX_{i+1} + Y_iY_{i+1} + Z_iZ_{i+1})$, where $J_i = J_1(J_2)$ for i even(odd). The ground states of this $\mathfrak{G} = \mathbb{U}(2)$-invariant Hamiltonian -under the representation $R(U \in \mathbb{U}(2)) = U^{\otimes n}$- can be found in two phases: a trivial (topological) one whenever $J_2/J_1 < 1 (> 1)$. This allows us to build a dataset $\{\rho_i, y_i\}$ where and we assign labels $y_i = 0$ and 1 to each phase. For the classification we employ a quantum convolutional neural network (QCNN) [19, 20].

While QCNNs, in analogy to CNNs, were envisioned to respect translation invariance we show that symmetries unique to quantum systems can, and should, be accounted for. Using the nullsapce approach described earlier we devise a corresponding symmetry-respecting architecture. The convolutional layers of the $\mathfrak{G}$-invariant QCNN are composed of 2-qubit gates generated by parametrized SWAPs on pairs of neighbouring qubits, while in each pooling layer we trace out half of the qubits.

As seen in Fig. 2(a), the EQCNN achieves accurate classification with as few as $N = 2$ training data points, while the non-equivariant version is found to perform almost as poorly as random guessing.

3 Discussion.

Geometric Quantum Machine Learning is a nascent field that intends to exploit symmetry present in learning tasks to enhance QML models [5, 6, 7]. We envision that soon enough the field of *Geometric Quantum Machine Learning* will be a thriving and exciting field. Merits of the constructions proposed were explored in toy examples and also practical tasks demonstrating a substantial separation in performances between invariant and non-invariant architectures. As a very nascent field, we expect many more developments including more systematic studies of EQNNs performances, stronger theoretical connections with the already well-developed field of GDL, and understanding limits imposed by practical implementations.

Link to the Manuscripts:
https://arxiv.org/abs/2205.02261
https://arxiv.org/abs/2210.08566
https://arxiv.org/abs/2210.07980

References

[1] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, arXiv preprint arXiv:2104.13478 (2021).

[2] T. Cohen and M. Welling, in *International conference on machine learning* (PMLR, 2016) pp. 2990–2999.

[3] R. Kondor and S. Trivedi, in *International Conference on Machine Learning* (PMLR, 2018) pp. 2747–2755.

[4] A. Bogatskiy, S. Ganguly, T. Kipf, R. Kondor, D. W. Miller, D. Murnane, J. T. Offermann, M. Pettee, P. Shanahan, C. Shimmin, *et al.*, arXiv preprint arXiv:2203.06153 (2022).

[5] M. Larocca, P. Czarnik, K. Sharma, G. Muraleedharan, P. J. Coles, and M. Cerezo, Quantum **6**, 824 (2022).

[6] 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).

[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] M. Larocca, F. Sauvage, F. M. Sbahi, G. Verdon, P. J. Coles, and M. Cerezo, PRX Quantum **3**, 030341 (2022).

[9] M. A. Rieffel and A. Van Daele, Bulletin of the London Mathematical Society **7**, 257 (1975).

[10] C. B. Mendl and M. M. Wolf, Communications in Mathematical Physics **289**, 1057 (2009).

[11] G. K. Brennen, arXiv preprint quant-ph/0305094 (2003).

[12] D. A. Meyer and N. R. Wallach, Journal of Mathematical Physics **43**, 4273 (2002).

[13] P. Rungta, V. Bužek, C. M. Caves, M. Hillery, and G. J. Milburn, Physical Review A **64**, 042315 (2001).

[14] V. S. Bhaskara and P. K. Panigrahi, Quantum Information Processing **16**, 1 (2017).

[15] J. L. Beckey, N. Gigena, P. J. Coles, and M. Cerezo, Phys. Rev. Lett. **127**, 140501 (2021).

[16] A. R. R. Carvalho, F. Mintert, and A. Buchleitner, Phys. Rev. Lett. **93**, 230501 (2004).

[17] A. Wong and N. Christensen, Physical Review A **63**, 044301 (2001).

[18] W. Bosma, J. Cannon, and C. Playoust, J. Symbolic Comput. **24**, 235 (1997), computational algebra and number theory (London, 1993).

[19] I. Cong, S. Choi, and M. D. Lukin, Nature Physics **15**, 1273 (2019).

[20] A. Pesah, M. Cerezo, S. Wang, T. Volkoff, A. T. Sornborger, and P. J. Coles, Physical Review X **11**, 041011 (2021).

Exponential separations between classical and quantum learners

Casper Gyurik[1] * Vedran Dunjko[1]

[1] *applied Quantum algorithms (aQa), Leiden University, The Netherlands*

Abstract. We explore provable quantum speedups in learning problems, addressing the challenge of finding scenarios where quantum algorithms exponentially outperform classical counterparts. Existing quantum learning advantages are limited to artificial cryptography-inspired datasets. We highlight the importance of subtle differences in definitions, and present two new learning separations where the classical hardness lies in *identifying* rather than evaluating the data-generating function. We discuss computational hardness assumptions required to prove learning separations for quantum-generated data, and show quantum advantage in condensed matter and high-energy physics. Additionally, we examine the classical shadow paradigm and its relation to learning separations for quantum-generated data.

Keywords: Quantum machine learning, Computational learning theory, Complexity theory

Technical manuscript: `arxiv.org/abs/2208.06339` The results described below are based on version 2 of the technical manuscript, which will appear on arXiv soon.

1 Introduction

Quantum machine learning (QML) is a bustling field with the potential to deliver quantum enhancements for practically relevant problems. An important goal of the community is to find practically relevant learning problems for which one can prove that quantum learners have an exponential advantage over classical learners. In this paper, we study how to achieve such exponential separations between classical and quantum learners for problems with classical data in the efficient probably approximately correct (PAC) learning framework. The first thing we address is that there is no single definition of what precisely constitutes a *learning* separation. In particular, when trying to come up with a definition there are many choices to be made, and various choices make sense depending on the use-cases. For instance, as we explain, a significant difference arise if the emphasize is on the task of *identifying* or *evaluating* the functions that are generating the data. This ambiguity can lead to conflating the task of learning in an intuitive sense with a purely computational task. To address this issue, we provide multiple definitions of a learning separation, and we discuss in which cases the tasks involve learning in an intuitive sense. Moreover, we study existing learning separations [8, 12] and carefully delineate where the classical hardness of learning lies and the types of learning separations they achieve. Furthermore, we provide new examples of learning separations where the classical hardness lies more in learning in an intuitive sense rather than evaluating the functions to be learned.

Next, we turn our attention to the folklore in the community that states that quantum machine learning most likely to have advantages when the data quantum-generated. For instance, it is believed that quantum learners are more likely to offer an advantage in predicting the phases of physical systems rather than distinguishing between images of dogs and cats. This is because genuinely quantum-generated data typically has some BQP-hard function underlying it. However, it is not immediately clear how these BQP-hard functions can give rise to a learning separation. In other words, if we assume a complexity-theoretic separation like $\mathsf{BPP} \neq \mathsf{BQP}$, how can we construct a learning separation from the fact that the labeling function is BQP-hard? In this paper, we address this question by exploring the additional complexity-theoretic assumptions required to build such a learning separation. Moreover, we provide several examples of how learning separations can be constructed from physical systems, such as the Bose-Hubbard model [2], the antiferromagnetic Heisenberg and antiferromagnetic XY model [11], the Fermi-Hubbard model [10], supersymmetric systems [1], interacting bosons [13], interacting fermions [7].

2 Results

The main contributions of our paper are as follows:

Formalizing quantum advantage in learning theory:

- We clarify the finer points regarding the possible definitions of a learning separation by highlighting that there are various ways of defining them. Above all, we explore the distinction between the task of identifying (i.e., giving a specification of) the correct labeling function versus the task of evaluating it (i.e., computing its output), and we explain that these differences have a significant impact on whether a problem exhibits a learning separation.
- We outline computational hardness assumptions that one can leverage to establish learning separations in the efficient PAC learning framework. In particular, we define the complexity class $\mathsf{HeurBPP/samp}$ that aims to capture all classically learnable functions. Moreover, using ideas from [6] we relate our new complexity class to a familiar though unexplored complexity class by showing that $\mathsf{HeurBPP/samp} \subseteq \mathsf{HeurP/poly}$.

*`c.f.s.gyurik@liacs.leidenuniv.nl`

Learning separations with efficient data generation:

- We discuss known learning separations [8, 12], and we and provide a fine-grained analysis of where the classical hardness of learning stems from. While discussing these learning separations, we find that the ones available in literature largely rely on the classical hardness of *evaluating* the function generating the data on unseen points, as opposed to the hardness of *identifying* it. We elaborate how the identification problem can be what is needed in practice, and we address this gap by proving two new learning separations where the classical hardness primarily lies in identifying the function generating the data.

Learning separations without efficient data generation:

- We show how leveraging stronger complexity-theoretic assumptions can lead to learning separations where the data is generated by a genuinely quantum process. Our main contribution is Theorem 1, which outlines a method of establishing learning separations from BQP-complete functions. We also provide two lemmas, Lemmas 2 and 3, which introduce natural assumptions under which the criteria in Theorem 1 are satisfied, which gives rise to Corollary 4. Finally, we show how Theorem 1 can be used to build learning separations from problems in quantum many-body physics.

Theorem 1 *Consider a family of concept classes* $\{\mathcal{C}_n\}_{n\in\mathbb{N}}$ *and distributions* $\{\mathcal{D}_n\}_{n\in\mathbb{N}}$ *such that*

Quantum learnability:

(a) Every $c_n \in \mathcal{C}_n$ *can be evaluated on a quantum computer in time* $\mathcal{O}(\mathrm{poly}(n))$.

(b) There exists a polynomial p *such that for every* $n \in \mathbb{N}$ *we have* $|\mathcal{C}_n| \leq p(n)$.

Classical non-learnability:

(c) There exists a family $\{c_n\}_{n\in\mathbb{N}}$*, where* $c_n \in \mathcal{C}_n$*, such that* $(\{c_n\}_{n\in\mathbb{N}}, \{\mathcal{D}_n\}_{n\in\mathbb{N}}) \notin \mathsf{HeurP/poly}$.

Then, $L = (\{\mathcal{C}_n\}_{n\in\mathbb{N}}, \{\mathcal{D}_n\}_{n\in\mathbb{N}})$ *exhibits a* CC/QQ *learning separation.*

Lemma 2 *If there exists a* $(L, \mathcal{D}) \notin \mathsf{HeurP/poly}$ *with* $L \in \mathsf{BQP}$*, then for every* $L' \in \mathsf{BQP}$-complete[5] *there exists a family of distributions* $\mathcal{D}' = \{\mathcal{D}'_n\}_{n\in\mathbb{N}}$ *such that* $(L', \mathcal{D}') \notin \mathsf{HeurP/poly}$.

Lemma 3 *lemmatwo If* $L \notin \mathsf{P/poly}$ *and* L *is polynomially random self-reducible with respect to some distribution* $\mathcal{D}$*, then* $(L, \mathcal{D}) \notin \mathsf{HeurP/poly}$.

Corollary 4 *If there exists an* $L \in \mathsf{BQP}$ *such that* $L \notin \mathsf{P/poly}$ *and it is random self-reducible, then every* BQP-*complete problem gives rise to a* CC/QQ *separation.*

Discussion:

- To connect our work to some of the related results in the field [5, 6, 9, 3], we discuss selected topics related to learning separations with classical data:
  - We discuss the milestone work of Huang et al. [5] and how their classical machine learning methods based on the classical shadow framework relate to learning separations with quantum-generated data (i.e., those from Theorem 1). In particular, we highlight their limitations by showing the existence of a family of Hamiltonians whose groundstates properties it cannot predict based on cryptographic assumptions.
  - We discuss a specific example (i.e., evaluating parameterized quantum circuits) that exemplifies how access to data radically enhances what can be efficiently evaluated.
  - We discuss how two physically-motivated problems (i.e., Hamiltonian learning, and identifying order parameters and phases of matter) naturally fit in a PAC learning setting where the learner is constrained to output hypotheses from a fixed hypothesis class, and how potential separations there are different from the non-fixed hypothesis class setting

3 Impact and importance

Our paper holds significant importance in the field of quantum information and computation as it addresses the fundamental question of identifying learning tasks where quantum algorithms can provide a provable exponential speedup over classical algorithms. By examining the subtle differences in definitions and the specific tasks that require the use of a quantum computer, we contribute to the understanding of the computational power and limitations of quantum learning. Furthermore, our exploration of learning problems with provable quantum speedups, particularly in scenarios involving quantum-generated data, has implications for various domains, including condensed matter physics and high-energy physics. Our findings provide valuable insights into the potential quantum advantages in natural settings and pave the way for future advancements in quantum machine learning.

References

[1] Chris Cade and P Marcos Crichigno. Complexity of supersymmetric systems and the cohomology problem. *arXiv 2107.00011*, 2021.

[2] Andrew M Childs, David Gosset, and Zak Webb. The bose-hubbard model is QMA-complete. In *International Colloquium on Automata, Languages, and Programming*. Springer, 2014.

[5] With respect to many-to-one reductions (as is the case for, e.g., quantum linear system solving [4]).

[3] Jeongwan Haah, Robin Kothari, and Ewin Tang. Optimal learning of quantum hamiltonians from high-temperature gibbs states. *arXiv 2108.04842*, 2021.

[4] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103, 2009.

[5] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V Albert, and John Preskill. Provably efficient machine learning for quantum many-body problems. *Science*, 377, 2022.

[6] HY Huang, M Broughton, M Mohseni, R Babbush, S Boixo, H Neven, and JR McClean. Power of data in quantum machine learning (2020). *Nature Communications*, 2021.

[7] Yi-Kai Liu, Matthias Christandl, and Frank Verstraete. Quantum computational complexity of the n-representability problem: Qma complete. *Physical review letters*, 98(11):110503, 2007.

[8] Yunchao Liu, Srinivasan Arunachalam, and Kristan Temme. A rigorous and robust quantum speed-up in supervised machine learning. *Nature Physics*, 2021.

[9] Thomas O'Brien, LevC Ioffe, Yuan Su, David Fushman, Hartmut Neven, Ryan Babbush, and Vadim Smelyanskiy. Quantum computation of molecular structure using data from challenging-to-classically-simulate nuclear magnetic resonance experiments. *arXiv 2109.02163*, 2021.

[10] Bryan O'Gorman, Sandy Irani, James Whitfield, and Bill Fefferman. Electronic structure in a fixed basis is qma-complete. *arXiv 2103.08215*, 2021.

[11] Stephen Piddock and Ashley Montanaro. The complexity of antiferromagnetic interactions and 2d lattices. *Quantum Information & Computation*, 17(7-8):636–672, 2017.

[12] Rocco Servedio and Steven J Gortler. Equivalences and separations between quantum and classical learnability. *SIAM Journal on Computing*, 2004.

[13] Tzu-Chieh Wei, Michele Mosca, and Ashwin Nayak. Interacting boson problems can be qma hard. *Physical review letters*, 104, 2010.

Quantum Merlin-Arthur proof systems for synthesizing quantum states

Hugo Delavenne[2] [1] François Le Gall[1] Yupan Liu[1] Masayuki Miyamoto[1]

[1]*Graduate School of Mathematics, Nagoya University*
[2]*ENS Paris-Saclay, Université Paris-Saclay*

Abstract. Complexity theory typically focuses on the difficulty of solving computational problems using classical inputs and outputs, even with a quantum computer. In the quantum world, it is natural to apply a different notion of complexity, namely the complexity of synthesizing quantum states. We investigate a state-synthesizing counterpart of the class NP, referred to as stateQMA, which is concerned with preparing certain quantum states through a polynomial-time quantum verifier with the aid of a single quantum message from an all-powerful but untrusted prover. This is a subclass of the class stateQIP recently introduced by Rosenthal and Yuen (ITCS 2022), which permits polynomially many interactions between the prover and the verifier. Our main result consists of error reduction of this class and its variants with an exponentially small gap or a bounded space, as well as how this class relates to other fundamental state synthesizing classes, i.e., states generated by uniform polynomial-time quantum circuits (stateBQP) and space-uniform polynomial-space quantum circuits (statePSPACE). Additionally, we demonstrate that stateQCMA achieves perfect completeness. Our proof techniques are based on the quantum singular value transformation introduced by Gilyén, Su, Low, and Wiebe (STOC 2019), and its adaption to achieve exponential precision with a bounded space. The full version of this paper is available from arXiv:2303.01877.

Keywords: quantum state synthesis, state complexity theory, quantum Merlin-Arthur proof systems

Classical and quantum complexity theory typically concentrates on the computational difficulty of solving problems with *classical* inputs and outputs. However, quantum computers have the ability to handle not only classical problems, but also quantum tasks, such as synthesizing quantum states. The most famous example is preparing ground states of a physical system [17, 24], which even dates back to Feynman's original ideas [7]. Analogous tasks are also commonplace in quantum cryptography and generalized notions of the pseudorandomness, such as quantum money [1] and pseudorandom quantum states [11]. This motivates the study of complexity of synthesizing quantum states.

In [2], Aaronson investigated the concept on quantum state complexity, leading to the *state synthesis problem.* This problem involves generating a quantum state ρ from the all-zero state based on a quantum circuit with a succinct description acting on n qubits with the depth up to exponential. The resulting state ρ is supposed to be close to the designated *target state* $|\psi\rangle$[1]. This problem is solvable in (quantum) polynomial space (PSPACE), i.e., a quantum computer running in exponential time but using polynomially many gates can generate a state that well approximates the target state.

Quantum computers are seemingly not capable of solving any PSPACE problem in polynomial time, while *polynomially many* messages interactive protocols with the help of an all-powerful and *untrusted* prover (known as *interactive proofs*, IP) captures the full computational power of polynomial-space computation, referred to as the celebrated IP = PSPACE theorem [20, 27]. A recent line of works [9, 22, 26] initializes the study on the *interactive* state synthesis problem. Rosenthal and Yuen [26] denote the polynomial-space-preparable state families as statePSPACE[2] and show that such state families are preparable by interactive synthesis protocols, which belongs to the class stateQIP. Afterwards in [9], the authors explore the state synthesis problem by taking advantage of fairly powerful and *trusted* oracles. Recently, Metger and Yuen [22] manage to prove the equivalence between state families that are preparable using polynomial space and those generated by interactive synthesis protocols, that is, stateQIP = statePSPACE, which is the state-synthesizing counterpart of the IP = PSPACE theorem (and its quantum analogue [10]).

However, there is currently a lack of fine-grained characterizations of computationally easier state families, viz., state families that are efficiently preparable (e.g., efficiently simulating view of quantum statistical zero-knowledge [28]), or state families that are synthesizable via simply one-message interactive protocols (e.g., efficient verification of pure quantum states in the adversarial scenario [29]). This opens up opportunities for our main results.

1 Main results

In this work, we are particularly interested in state families that are preparable by *one-message* protocol, denoted as stateQMA, which is obviously a subclass of stateQIP. Let us first devote to defining stateQMA informally. For a state family in stateQMA is a family $\{|\psi_n\rangle\}_{n\in\mathbb{N}}$ indexed by *natural numbers* such that there is a verifier that has the following properties, verifying whether the target state $|\psi_n\rangle$ corresponds to a given input 1^n is well-approximated. The verifier's computation, which is a polynomial-size *unitary* quantum cir-

[1] We measure the closeness between ρ and $|\psi\rangle$ by the trace distance $\mathrm{td}(\rho, |\psi\rangle\langle\psi|) := \frac{1}{2}\|\rho - |\psi\rangle\langle\psi|\|_1$.

[2] The definition of statePSPACE is a bit subtle: although all quantum states can be well-approximated by an exponentially long gate sequence owing to the Solovay-Kitaev theorem [15], this exponential gate sequence is not necessarily *space-uniform.*

cuit[3], takes a quantum-proof state $|w\rangle$ (with no limitations on the preparation method) and ancillary qubits in the state $|0\rangle$ as input. After performing the verification circuit, a designated output qubit will be measured on the computational basis, and the verifier accepts if the measurement outcome is 1. If the verifier accepts, the verification circuit has prepared the resulting state ρ_w on the remaining qubits that is a good approximation of the target state $|\psi_n\rangle$ (if the verifier rejects, the resulting state could be anything). The acceptance probability is viewed *the success probability* for approximately preparing $|\psi_n\rangle$.

More precisely, the state family is in the class $\mathsf{stateQMA}_\delta[c,s]$ for some $0 \leq s < c \leq 1$ and $\delta \geq 0$, if the resulting state ρ_w is δ-close to the target state $|\psi_n\rangle$ provided that the verifier accepts with probability at least s (soundness condition); and additionally there exists a quantum witness that makes the verifier accepts with probability at least c (completeness condition).

It is evident that $\mathsf{stateQMA}_\delta[c,s] \subseteq \mathsf{stateQMA}_{\delta'}[c',s']$ if $c' \leq c$, $s' \geq s$ and $\delta' \geq \delta$. However, how crucially does $\mathsf{stateQMA}_\delta[c,s]$ depend on its parameters? For commonplace complexity classes, viz. BQP, QMA, QIP, etc., the dependence on such parameters is very weak: the class remains the same so long as the completeness c and soundness s differ by at least some inverse polynomial. This is known as *error reduction*, which typically involves performing the verification circuit in parallel and taking the majority vote.

However, error reduction for $\mathsf{stateQMA}$ requires a more nuanced approach. A simple parallel repetition of the verification circuit ends with a tensor product of the resulting state that evidently differs from the original state family. Therefore, error reduction for $\mathsf{stateQMA}$ does need to preserve not only the quantum witness state $|w\rangle$, but also the resulting state ρ_w, referred to as the *doubly-preserving error reduction* in Theorem 1.

Theorem 1 (Doubly-preserving error reduction for stateQMA – informal) *For any $c(n) - s(n) \geq 1/\mathrm{poly}(n)$ and $0 \leq c(n), s(n) \leq 1$, we have*

$$\mathsf{stateQMA}_\delta[c(n), s(n)] \subseteq \mathsf{stateQMA}_\delta[1-2^{-l(n)}, 2^{-l(n)}].$$

Nevertheless, applying Theorem 1 to a *polynomial-space-bounded* variant of $\mathsf{stateQMA}[c,s]$, which we denote by $\mathsf{stateQMA_U PSPACE^{off}}$[4], will result in *exponential* space. To address this, we generalize Theorem 1 in a manner that preserves the polynomial space complexity. Here in the class $\mathsf{stateQMA_U PSPACE^{off}}$, the verifier's computation stays polynomially space-bounded but may take *exponential time* and the gap between the completeness c and the soundness s is at least some inverse-exponential.

Theorem 2 (Doubly-preserving error reduction for stateQMA$_\mathsf{U}$PSPACE$^\mathsf{off}$ – informal) *For any $c(n) - s(n) \geq \exp(-\mathrm{poly}(n))$ and $0 \leq c(n), s(n) \leq 1$, we have*

$$\mathsf{stateQMA_U PSPACE^{off}}_\delta[c(n), s(n)] \subseteq$$
$$\mathsf{stateQMA_U PSPACE^{off}}_\delta\left[1-2^{-l(n)}, 2^{-l(n)}\right].$$

We note that Theorem 1 is a state-synthesis analogue of the witness-preserving error reduction for QMA [21, 23]. Likewise, Theorem 2 shares similarities with error reduction for unitary quantum computations [4] in the context of synthesizing states. Along the line of Marriott and Watrous [21], we demonstrate that logarithmic-size quantum witness states are useless for $\mathsf{stateQMA}$, and this variant is referred to as $\mathsf{stateQMA}[\log]$. Here $\mathsf{stateBQP}$ is defined as a subclass of $\mathsf{statePSPACE}$ with only polynomially many quantum gates.

Corollary 3 $\mathsf{stateQMA}_\delta[\log] = \mathsf{stateBQP}_\delta$.

Resembling the approach of Fefferman and Lin [5], we demonstrate that a variant of $\mathsf{stateQMA}$ that admits an exponentially small gap between completeness and soundness, known as $\mathsf{statePreciseQMA}$, is contained in $\mathsf{statePSPACE}$. Surprisingly, Corollary 4 shows that the distance parameter δ remains *unchanged*, while a similar $\mathsf{statePSPACE}$ containment following from [22] will worsen the distance parameter δ, namely $\mathsf{stateQMA}_\delta \subseteq \mathsf{statePSPACE}_{\delta+1/\mathrm{poly}(n)}$.

Corollary 4 $\mathsf{statePreciseQMA}_\delta \subseteq \mathsf{statePSPACE}_\delta$.

Furthermore, we prove that $\mathsf{stateQCMA}$, which is a variant of $\mathsf{stateQMA}$ in which the optimal quantum witness state is classical (i.e., a binary string) for the completeness condition, can archive the perfect completeness. This result is analogous to the $\mathsf{QCMA} = \mathsf{QCMA}_1$ theorem [13, 14] for synthesizing quantum states.

Theorem 5 (stateQCMA achieves perfect completeness – informal) *For any $c(n) - s(n) \geq 1/\mathrm{poly}(n)$ and $0 \leq c(n), s(n) \leq 1$, we have $\mathsf{stateQCMA}_\delta[c,s] \subseteq \mathsf{stateQCMA}_\delta[1,s']$ for some s' such that $1 - s'(n) \geq 1/\mathrm{poly}(n)$.*

In addition, it is worth noting that Theorem 5 also straightforwardly extends to $\mathsf{statePreciseQCMA}$.

2 Proof techniques

The proof of Theorem 1 and Theorem 2 employs the quantum linear algebra techniques developed by Gilyén, Su, Low, and Wiebe [8], specifically the quantum singular value discrimination.

Error reduction for stateQMA by manipulating singular values. To elaborate on the intuition, we begin by briefly reviewing the witness-preserving error reduction for QMA [21, 23]. Consider a QMA verification circuit V_x that takes a quantum witness state $|w\rangle$ (on the register W) and ancillary qubits in the state $|0\rangle$ as input. The corresponding acceptance probability is

[3] In particular, extending to general quantum circuits does not change the class $\mathsf{stateQMA}$ owing to the principle of deferred measurement. However, such extensions do not immediately work for space-bounded $\mathsf{stateQMA}$.

[4] We emphasize that $\mathsf{stateQMA_U PSPACE^{off}}$ is not a state-synthesizing counterpart of the class $\mathsf{NPSPACE}$.

$\||1\rangle\langle 1|_{\text{out}}V_x|w\rangle|\bar{0}\rangle\|_2^2$, which is equal to a quadratic form $\langle w|M_x|w\rangle$ where the matrix $M_x := \langle\bar{0}|V_x^\dagger|1\rangle\langle 1|_{\text{out}}V_x|\bar{0}\rangle$. It is not hard to see the maximum acceptance probability of V_x is the largest eigenvalue of M_x. We then view $M_x = \Pi_{\text{in}}\Pi\Pi_{\text{in}}$ as a product of Hermitian projectors Π_{in} and Π where $\Pi_{\text{in}} = I_{\mathsf{W}} \otimes |\bar{0}\rangle\langle\bar{0}|$ and $\Pi = V_x^\dagger|1\rangle\langle 1|_{\text{out}}V_x$. Remarkably, there exists an orthogonal decomposition of the Hilbert space, which the projectors Π_{in} and Π act on, into *one-dimensional* and *two-dimensional* common invariant subspaces. This elegant decomposition property is referred as to the Jordan lemma[5] [12]. Marriott and Watrous [21] then take advantage of the Jordan lemma and present error reduction for QMA that preserves the quantum witness state.

However, this error reduction technique does not automatically preserve the resulting state, as required in stateQMA, we thus need a more sophisticated technique, namely the quantum singular value transformation [8]. This technique generalizes the qubitization technique introduced by Low and Chuang [19] that inspired by the aforementioned decomposition property. Moving on to the maximum acceptance probability of a stateQMA verifier V_n, it corresponds to the square root of the largest singular value of the matrix $A_n = \Pi_{\text{out}}V_n\Pi_{\text{in}}$ where $\Pi_{\text{out}} := |1\rangle\langle 1|_{\text{out}}$ is the final measurement. In Section 3.2 of [8], the authors extend the Jordan lemma to the singular value scenarios. In particular, $\text{Img}(\Pi_{\text{in}})$ and $\text{Img}(\Pi_{\text{out}})$ can be decomposed into one-dimensional or two-dimensional common invariant subspaces. Now let us focus on the specific case of stateQMA, we notice that the right singular vectors of A_n correspond to the quantum witness state $|w\rangle$, as well as the left singular vectors correspond to the resulting state ρ_w. Therefore, we result in *doubly-preserving error reduction* for stateQMA (Theorem 1) by manipulating the singular values accordingly[6].

It is noteworthy that Theorem 1 differs from Theorem 38 in [8] since our construction is based on the *projected unitary encoding* (e.g., the presented matrix A_n) instead of the block-encoding. Furthermore, for establishing Theorem 2, we make use of an *exponential-degree* approximation polynomial of the sign function that all coefficients within the *exponential precision* are computable in PSPACE [22]. We additionally observe that the proof techniques in [22] can be straightforwardly adapted to *projected unitary encodings* instead of the block-encodings originally utilized in their work.

Applications of error reduction for stateQMA. Along the line of Theorem 3.13 in [21], with Theorem 1, it seems to straightforwardly make for Theorem 3. Nevertheless, the resulting state raises some concern upon initial inspection. We fortunately circumvent this caveat by a careful analysis. Specifically, utilizing the error reduction for stateQMA, we begin with a verifier with completeness $1 - 2^{-p(n)}$ and soundness $2^{-p(n)}$ where p is a polynomial of n. Then we replace the short quantum witness state $|w\rangle$ with a completely mixed state I_{W}, which gives us a computation meeting the soundness condition such that the soundness s is preserved and the gap between the completeness and the soundness shrinks to some inverse-polynomial of n. Although the new resulting state $\rho_{I_{\mathsf{W}}}$ may greatly differ from ρ_w, the definition of stateQMA guarantees that $\rho_{I_{\mathsf{W}}}$ is also close to the target state because the acceptance probability of the verifier with I_{W} is greater than the soundness s. This proof also easily extends to Corollary 4 employing Theorem 2. In addition, it is noteworthy that stateBQP achieves perfect completeness with a worsening distance parameter δ'. By incorporating error techniques for both stateBQP and state$_{\mathsf{U}}$PSPACE, the difference between the new distance parameter δ' and the original one can be made *exponentially small*. Furthermore, we remark that stateBQP is not trivially contained in stateQMA, still this is effortless. We therefore complete the other direction in Corollary 3.

stateQCMA achieves perfect completeness. Our proof for Theorem 5 takes inspiration from [14, 13], but it requires several modifications. Note that our concern in stateQCMA is not only the maximum acceptance probability but also the resulting state after performing the verification circuit and the final measurement. To meet these requirements, we must choose a specific universal gateset $\mathcal{S}$ such that $\mathcal{S}$ can generate a dense subgroup of $\text{SU}(2^n)$, and all quantum states generated by these gates in $\mathcal{S}$ have rational entries. For this reason, we opt for the "Pythagorean gateset" [14, 3]. To ensure that the resulting state is indeed close to the target state, we slightly adjust the construction outlined in [13].

3 Discussion and open problems

Reduction and completeness in state-synthesizing complexity theory. In the context of state-synthesizing complexity theory, including prior works [9, 22, 26] and our own, the concepts of *reduction* and *completeness* have not been defined. However, these concepts hold significant importance in (quantum) complexity theory. The immediate challenge lies in appropriately defining these concepts, such as reduction, in a manner that ensures the resulting states exhibit reasonable behavior before and after the application of the reduction.

The computational power of statePreciseQMA. Although Corollary 4 establishes that a statePSPACE containment of statePreciseQMA, the reverse inclusion, namely statePSPACE $\subseteq$ statePreciseQMA, remains an open problem. The main challenge lies in adapting existing proof techniques that demonstrate PSPACE $\subseteq$ PreciseQMA [5, 6, 16], as these techniques heavily rely on notions of *completeness* or *reduction* for the class PSPACE.

[5] See [25] for the detailed statement of the Jordan lemma, as well as a simple proof.

[6] Concretely speaking, the analysis of error reduction based on majority votes essentially corresponds to obtaining tail bounds for the Binomial distribution. By leveraging the central limit theorem, it becomes sufficient to estimate tail bounds for the normal distribution, referred to as the error function $\text{erf}(x)$. The approximation polynomials of the sign function in [18] then achieve this task.

References

[1] Scott Aaronson. Quantum copy-protection and quantum money. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 229–242. IEEE, 2009. 1

[2] Scott Aaronson. The complexity of quantum states and transformations: from quantum money to black holes. *arXiv preprint arXiv:1607.05256*, 2016. 1

[3] Marcos Crichigno and Tamara Kohler. Clique Homology is QMA_1-hard. *arXiv preprint arXiv:2209.11793*, 2022. 3

[4] Bill Fefferman, Hirotada Kobayashi, Cedric Yen-Yu Lin, Tomoyuki Morimae, and Harumichi Nishimura. Space-efficient error reduction for unitary quantum computations. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, 2016. 2

[5] Bill Fefferman and Cedric Lin. Quantum Merlin Arthur with exponentially small gap. *arXiv preprint arXiv:1601.01975*, 2016. 2, 3

[6] Bill Fefferman and Cedric Yen-Yu Lin. A complete characterization of unitary quantum space. In *9th Innovations in Theoretical Computer Science Conference (ITCS 2018)*, 2018. 3

[7] Richard P Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982. 1

[8] András Gilyén, Yuan Su, Guang Hao Low, and Nathan 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)*, pages 193–204, 2019. 2, 3

[9] Sandy Irani, Anand Natarajan, Chinmay Nirkhe, Sujit Rao, and Henry Yuen. Quantum search-to-decision reductions and the state synthesis problem. In *37th Computational Complexity Conference*, 2022. 1, 3

[10] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. Qip= pspace. *Journal of the ACM (JACM)*, 58(6):1–27, 2011. 1

[11] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In *Advances in Cryptology–CRYPTO 2018: 38th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 19–23, 2018, Proceedings, Part III 38*, pages 126–152. Springer, 2018. 1

[12] Camille Jordan. Essai sur la géométrie à n dimensions. *Bulletin de la Société mathématique de France*, 3:103–174, 1875. 3

[13] Stephen P Jordan, Hirotada Kobayashi, Daniel Nagaj, and Harumichi Nishimura. Achieving perfect completeness in classical-witness quantum merlin-arthur proof systems. *Quantum Information & Computation*, 12(5-6):461–471, 2012. 2, 3

[14] Stephen P Jordan and Daniel Nagaj. QCMA with one-sided error equals QCMA with two-sided error. *arXiv preprint arXiv:1111.5306*, 2011. 2, 3

[15] Alexei Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997. 1

[16] Yulong Li. A simple proof of preciseqma= pspace. *arXiv preprint arXiv:2206.09230*, 2022. 3

[17] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996. 1

[18] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by uniform spectral amplification. *arXiv preprint arXiv:1707.05391*, 2017. 3

[19] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019. 3

[20] Carsten Lund, Lance Fortnow, Howard Karloff, and Noam Nisan. Algebraic methods for interactive proof systems. *Journal of the ACM (JACM)*, 39(4):859–868, 1992. 1

[21] Chris Marriott and John Watrous. Quantum Arthur–Merlin games. *computational complexity*, 14(2):122–152, 2005. 2, 3

[22] Tony Metger and Henry Yuen. $\mathsf{stateQIP} = \mathsf{statePSPACE}$, 2023. 1, 2, 3

[23] Daniel Nagaj, Pawel Wocjan, and Yong Zhang. Fast amplification of QMA. *Quantum Information & Computation*, 9(11):1053–1068, 2009. 2

[24] David Poulin and Pawel Wocjan. Preparing ground states of quantum many-body systems on a quantum computer. *Physical review letters*, 102(13):130503, 2009. 1

[25] Oded Regev. Witness-preserving amplification of QMA, 2006. 3

[26] Gregory Rosenthal and Henry Yuen. Interactive proofs for synthesizing quantum states and unitaries. In *13th Innovations in Theoretical Computer Science Conference (ITCS 2022)*, 2022. 1, 3

[27] Adi Shamir. $\mathsf{IP} = \mathsf{PSPACE}$. *Journal of the ACM (JACM)*, 39(4):869–877, 1992. 1

[28] 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

[29] Huangjun Zhu and Masahito Hayashi. Efficient verification of pure quantum states in the adversarial scenario. *Physical review letters*, 123(26):260504, 2019. 1

Tighter And Stronger Quantum Speed Limits

Shrobona Bagchi[1] * Abhay Srivastav[2] † Arun Kumar Pati[2] ‡

[1] *Center for Quantum Information, Korea Institute of Science and Technology, Seoul, 02792, Korea*
[2] *Quantum Information and Computation Group, Harish-Chandra Research Institute, Chhatnag Road, Jhunsi, Allahabad 211019, India and Homi Bhabha National Institute, Anushaktinagar, Training School Complex, Mumbai 400085, India*

Abstract. We derive various quantum speed limits for unitary evolution for the case of general quantum states using the stronger uncertainty relation for mixed quantum states and tighter uncertainty relation for general quantum states. These bounds are proved to be stronger and tighter than many earlier bounds in the literature, which renders them useful in the arena of quantum metrology and potential applications in quantum information processing tasks. In the process we also generalize the tiger uncertainty relation for pure quantum states to that of mixed quantum states and prove its better performance theoretically. It is then shown that these bounds can be optimized over different choices of operators for obtaining even better bounds. We illustrate these with many examples and show their better performance with respect to at least three existing bounds for general quantum states and many different choices of Hamiltonians that are useful in different quantum information processing tasks. **A part of the technical version of this work is available in arXiv:2211.14561 and under review in peer revived journal and another part is under review in peer reviewed journal.**

Keywords: Tighter and Stronger Quantum Speed Limits, Time Energy uncertainty relation

1 Introduction

The uncertainty relations have helped us to reveal the behavior of the microscopic world in many different ways. At first the uncertainty principle was discovered by Werner Heisenberg who heuristically provided a lower bound to the product of the error and disturbance for two canonically conjugate quantum mechanical observables [1]. On the other hand, the uncertainty relations are also capable of capturing the intrinsic restrictions in preparation of quantum systems, which are termed as the preparation uncertainty relations [2]. This interpretation was quite fruitful for the uncertainty relations like position-momentum, angular position-angular momentum uncertainty relations etc. However, the energy-time uncertainty relation [3, 4] is different from the above stated uncertainty relations because time is not treated as an operator in quantum mechanics [5]. In fact in textbook quantum mechanics time is not treated as a quantum mechanical observable, but as a classical parameter with no inherent quantum uncertainty in it [3]. In terms of various forms of uncertainty relations, Robertson formulated an uncertainty relation for two arbitrary quantum-mechanical observables. This relation is a type of the preparation uncertainty relation and expresses the impossibility of jointly sharp preparation of any two incompatible observables. However, the Robertson uncertainty relation do not completely express the incompatibility nature of two non-commuting observables in terms of uncertainty quantification. The stronger variations of the uncertainty relations have been proved which capture the notion of incompatibility more efficiently [6]. However, time, not being a quantum observable, energy-time uncertainty relation lacked a good interpretation as such like for those of other quantum mechanical observables like the position and the momentum.

Mandelstam and Tamm (MT) derived an uncertainty relation which is called an energy-time uncertainty relation which follows from the Robertson uncertainty relation when we take into account the initial quantum state and the Hamiltonian as the corresponding quantum mechanical operators [7]. An interpretation of this uncertainty relation was given in terms of the so called quantum speed limit [6]. In the existing literature, there are several other approaches to obtain quantum speed limits [6]. As an example, the Margolus-Levitin bound gives the average of energy in the quantum speed limit [7]. The quantum speed limit bounds have also been extended to the case of mixed quantum states undergoing unitary evolution [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24], entangled quantum states [25, 26, 27] and also open quantum system dynamics [28, 29, 30].

Not only of fundamental importance, but quantum speed limit bounds have practical applications also. The quantum speed limit bounds have proven to be very useful in quantifying the maximal rate of quantum entropy production [32, 33], the maximal rate of quantum communication [34], the maximal rate of quantum information processing [35], and in many other avenues. As a result, these motivate us to find better quantum speed limit bounds that can be useful.

In the first part of this work, we find tighter form of quantum speed limit. We show that the new bound provides a tighter expression of quantum speed limit compared to the existing bounds such as the MT bound [7]. This bound can also be optimized over many orthonormal basis vector sets, as in the case of tighter uncertainty relations [31]. We find various analytical examples that shows the performance of our bound over some of the other bounds. As a byproduct we generalize tighter uncertainty relation to mixed quantum states analytically.

*shrobona@kist.re.kr
†abhaysrivastav@hri.res.in
‡arunpati2008@gmail.com

In the second part of this work, we use the stronger uncertainty relation, then generalised to the case of mixed quantum states to derive a stronger form of quantum speed limit for mixed quantum states undergoing unitary evolution. We show that the new bound provides a stronger expression of quantum speed limit compared to the MT like bound for mixed quantum states. This bound can also be optimized over many operators. We then find various analytical examples for mixed states and some example Hamiltonians that shows the better performance of our bound over the MT like bound for mixed quantum states and the bounds for mixed states.

A part of the technical version of this work is available in arXiv:2211.14561 and under review in peer revived journal and another part is under review in peer reviewed journal.

2 Main Results

As per our main results, we prove two main different types of quantum speed limits for unitary evolution of general quantum states and prove analytically that they outperform the MT bound for pure quantum states and three other different important quantum speed limit bounds for mixed quantum states with numerical data. As a byproduct we also generalize the tighter uncertainty relation for pure quantum states to mixed quantum states and prove analytically that it outperforms the Robertson uncertainty relation. After this we work out various examples that include different cases of random Hamiltonians, Heisenberg spin chain, perfect state transfer Hamiltonian, two qubit CNOT Hamiltonian that are usally very useful in various quantum information and processing tasks. In the next paragraphs, we state our main results in the form of theorems and their proofs.

3 Main Results: Tighter and stronger quantum speed limit for general quantum states

Theorem 1 *For a quantum state $\rho(t)$, the speed of unitary evolution generated by the Hamiltonian H is bounded by the following inequality*

$$\tau \geq \left[\frac{\hbar}{\Delta H}\left(\cos^{-1}(\sqrt{\mathrm{Tr}(\rho_0\rho_\tau)}) - \cos^{-1}(\sqrt{\mathrm{Tr}(\rho_0^2)})\right) + \frac{1}{\sqrt{\mathrm{Tr}(\rho_0^2)}\Delta H}\int_0^\tau \frac{K(t)dt}{\cos\frac{s_0(t)}{2}\sqrt{1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2}}}\right].$$

All the above terms will be clear in the proof as follows. Consider two non-commuting operators A and B, the tighter uncertainty relation for a mixed state ρ is given by

$$\Delta A \Delta B \geq \sum_n \sqrt{|\mathrm{Tr}(\bar{A}\rho\bar{A}\bar{B}_n^{\psi}\rho\bar{B}_n^{\psi})|}. \tag{1}$$

where

$\bar{A} = A - \mathrm{Tr}(\rho A)\mathbb{I}$, $\bar{B} = B - \mathrm{Tr}(\rho B)\mathbb{I}$, $\bar{B}_n^{\psi} = |\psi_n\rangle\langle\psi_n|\bar{B}$, and $\{|\psi_n\rangle\}$ form a complete orthonormal basis. Now with some algebra we get

$$\Delta A \Delta B \geq \Big[\sum_n \sqrt{|\mathrm{Tr}(\bar{A}\rho\bar{A}\bar{B}_n^{\psi}\rho\bar{B}_n^{\psi})|} - |\mathrm{Tr}(\bar{A}\rho\bar{B})|\Big] + |\mathrm{Tr}(\bar{A}\rho\bar{B})|. \tag{2}$$

We will now analyze the term $|\mathrm{Tr}(\bar{A}\rho\bar{B})|$. For this we use a more convenient notation as $|\mathrm{Tr}(\bar{A}\rho\bar{B})| = |\langle\bar{A}\bar{B}\rangle|$. We note that the following equation holds for all mixed quantum states and where the expectation values denoted by the angled brackets are with respect to the mixed quantum state ρ, i.e.,

$$|\langle\bar{A}\bar{B}\rangle|^2 = \frac{1}{4}|\langle[A,B]\rangle|^2 + |\frac{1}{2}\langle\{A,B\}\rangle - 2\langle A\rangle\langle B\rangle|^2.$$

Since both the terms on the R.H.S are positive, we have

$$|\langle\bar{A}\bar{B}\rangle| \geq \frac{1}{2}|\langle[A,B]\rangle|. \tag{3}$$

Using Eq.(2) and the inequality from the above equation we get

$$\Delta A \Delta B \geq \frac{1}{2}|\langle[A,B]\rangle| + K(t), \tag{4}$$

where $K(t) = \Big[\sum_n \sqrt{|\mathrm{Tr}(\bar{A}\rho\bar{A}\bar{B}_n^{\Psi}\rho\bar{B}_n^{\Psi})|} - |\mathrm{Tr}(\bar{A}\rho\bar{B})|\Big]$ is positive semidefinite. Let us now take the operators A and B as follows $A = \rho(0)$ and $B = H$, and $\rho \equiv \rho(t) = e^{-iHt}\rho(0)e^{iHt}$. The variance of the operator A is then given by

$$\begin{aligned}\Delta A^2 &= \mathrm{Tr}(\rho(0)^2\rho(t)) - (\mathrm{Tr}(\rho(0)\rho(t)))^2 \\ &= \mathrm{Tr}(\rho_0^2\rho_t) - (\mathrm{Tr}(\rho_0\rho_t))^2,\end{aligned} \tag{5}$$

where we have used the notation $\rho(0) \equiv \rho_0$ and $\rho(t) \equiv \rho_t$. We can now take the following parametrization

$$\langle A\rangle = \mathrm{Tr}(\rho_0\rho_t) = \mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2}. \tag{6}$$

Now, using the equation of motion for the average of A, we get

$$\left|\hbar\frac{d}{dt}\langle A\rangle\right| = |\langle[A,H]\rangle|,$$

where the averages are all with respect to the mixed quantum state ρ and A has no explicit time dependence. Using Eq.(6) then, we get

$$\left|\frac{d\langle A\rangle}{dt}\right| = \mathrm{Tr}(\rho_0^2)\frac{\sin s_0(t)}{2}\frac{ds_0}{dt}. \tag{7}$$

Therefore, putting the values of A and B explicitly in the above derived equations we get

$$\Delta A \Delta H \geq \mathrm{Tr}(\rho_0^2)\frac{\hbar\sin s_0(t)}{4}\frac{ds_0}{dt} + K(t). \tag{8}$$

Now let us analyse the structure of ΔA^2 as follows

$$\Delta A^2 = \mathrm{Tr}(\rho_0^2\rho_t) - (\mathrm{Tr}(\rho_0\rho_t))^2. \tag{9}$$

Doing some algebra we obtain

$$\sqrt{\mathrm{Tr}(\rho_0^2)}\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}\Delta H \geq \Delta A\Delta H. \tag{10}$$

The above inequality using Eq.(8) becomes

$$\sqrt{\mathrm{Tr}(\rho_0^2)}\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}\Delta H \\ \geq \mathrm{Tr}(\rho_0^2)\frac{\hbar\sin s_0(t)}{4}\frac{ds_0}{dt}+K(t). \tag{11}$$

Now, integrating the above equation with respect to t and $s_0(t)$ we obtain the new quantum speed limit bound for mixed quantum states as follows

$$\tau \geq \frac{\hbar\sqrt{\mathrm{Tr}(\rho_0^2)}}{4\Delta H}\int_{s_0(0)}^{s_0(\tau)}\frac{\sin s_0(t)}{\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}}ds_0 \\ +\frac{1}{\sqrt{\mathrm{Tr}(\rho_0^2)}\Delta H}\int_0^{\tau}\frac{K(t)}{\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}}dt.$$

The first term on the right hand side can be integrated in the analytical form, so that we get the following relation

$$\tau \geq \frac{\hbar\sqrt{\mathrm{Tr}(\rho_0^2)}}{4\Delta H}\Big[-4\frac{\sin^{-1}(\sqrt{\mathrm{Tr}(\rho_0^2)}\cos\frac{s_0(t)}{2})}{\sqrt{\mathrm{Tr}(\rho_0^2)}}\Big]_{s_0(0)}^{s_0(\tau)} \\ +\frac{1}{\sqrt{\mathrm{Tr}(\rho_0^2)}\Delta H}\int_0^{\tau}\frac{K(t)}{\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}}dt.$$

Now, we know that $\cos\frac{s_0(0)}{2}=1$ and $\sqrt{\mathrm{Tr}(\rho_0^2)}\cos\frac{s_0(\tau)}{2}=\sqrt{\mathrm{Tr}(\rho_0\rho_\tau)}$. Thus, putting these value in the above equation and simplifying, we get

$$\tau \geq \frac{\hbar}{\Delta H}\Big[\sin^{-1}(\sqrt{\mathrm{Tr}(\rho_0^2)})-\sin^{-1}(\sqrt{\mathrm{Tr}(\rho_0\rho_\tau)})\Big] \\ +\frac{1}{\sqrt{\mathrm{Tr}(\rho_0^2)}\Delta H}\int_0^{\tau}\frac{K(t)}{\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}}dt.$$

As before, $K(t)$ is always greater than or equal to zero in all cases. The optimized version can be expressed as optimization over the bases.

Theorem 2 *The time evolution of a general mixed quantum state governed by a unitary operation generated by a Hamiltonian is given by the following equation*

$$\tau \geq \tau_{SQSLM}=\frac{\sqrt{\mathrm{Tr}(\rho_0^2)}}{2\Delta H}\times \\ \int_{s_0(0)}^{s_0(\tau)}\frac{\sin s_0(t)}{(1-R(t))\cos\frac{s_0(t)}{2}\sqrt{(1-\mathrm{Tr}(\rho_0^2)\cos^2\frac{s_0(t)}{2})}}ds_0,$$

where τ_{SQSLM} stands as a short form for the stronger quantum speed limit for mixed quantum states and we have the following definitions of the quantities expressed in the above equation

$$s_0(t)=2\cos^{-1}|\sqrt{\frac{\mathrm{Tr}(\rho(0)\rho(t))}{\mathrm{Tr}(\rho_0^2)}}|, \\ \Delta H=\mathrm{Tr}(H^2\rho)-(\mathrm{Tr}(H\rho))^2 \\ R(t)=\frac{1}{2}|\mathrm{Tr}(\rho^{\frac{1}{2}}(\frac{A}{\Delta A}\pm i\frac{B}{\Delta B})\sigma)|^2, \\ \text{where } \mathrm{Tr}(\rho^{\frac{1}{2}}\sigma)=0 \text{ and } ||\sigma||_2=1,$$

where we have $||\sigma||_2=(\sum_{n\in I}\langle e_n|\sigma\sigma^{\dagger}|e_n\rangle)^{\frac{1}{2}}$, $\{|e_n\rangle\}$ forming a complete orthonormal basis in Hilbert space H, $\sigma\in L^2(H)$, i.e., σ belongs to the set of all Hilbert Schmidt linear operators.

The proof proceeds in similar way as theorem 1, however we have some different starting point.

4 Discussion of main results

In the first part of this work, we have derived a tighter quantum speed limit that outperforms the MT bound and the stronger speed limit bound in many cases. We have used the tighter uncertainty relations to derive our bound. We have obtained the necessary and sufficient conditions for this bound to reduce to that of the MT bound for pure two qubit states as an example and argued that we can derive similar conditions for higher dimensions and mixed steps. Hereafter we have shown numerically using random Hamiltonians obtained from Gaussian Unitary ensemble and some other analytical examples involving interacting quantum systems, that our bound performs better than the MT and some other existing bounds in many cases. Since we have shown that our bound is always better than the MT bound in all cases, therefore all the cases where the MT bound performs better than the Margolus Levitin bound, our new bound also performs better than the Margolus Levitin bound in those cases.

In the second part of this work, we have derived a stronger quantum speed limit for mixed quantum states using the mixed state generalization of stronger preparation uncertainty relations. We have shown that this bound reduces to that of the pure states under appropriate conditions. Thereafter, we have discussed methods to derive the suitable operators that allows us to calculate our bound. Hereafter we have shown numerically using random Hamiltonians obtained from Gaussian Unitary ensemble that our bound performs better than the mixed state version of the MT bound. Also, we have then shown using many suitable analytical examples of Hamiltonians useful in important quantum information and computation tasks that the stronger quantum speed limit bound derived here for mixed quantum states also perform better than the MT like bound and also two more existing quantum speed limit bounds for mixed quantum states existing in the current literature.

References

[1] W. Heisenberg.Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik.

Zeitschrift fur Physik **43**, 172,1927.

[2] H. P. Robertson, Robertson, H.P. (1929), Phys. Rev. **34**, 163,1929.

[3] Y. Aharonov, D. Bohm.Time in the Quantum Theory and the Uncertainty Relation for Time and Energy. Phys. Rev.**122**, 1649,1961.

[4] Y. Aharonov, S. Massar, S. Popescu.Measuring Energy, Estimating Hamiltonians, and the Time-Energy Uncertainty Relation. Phys. Rev. A **66**, 052107,2002.

[5] P. Busch. The Time-Energy Uncertainty Relation, in "Time in Quantum Mechanics", eds. J. G. Muga, R. Sala Mayato, I.L. Egusquiza, (Springer-Verlag, Berlin, 2002) pp. 69-98. 2nd rev. ed. 2008, pp. 73-105, also at arXiv:quant-ph/0105049.

[6] L. Maccone and A. K. Pati.Stronger Uncertainty Relations for All Incompatible Observables. Phys. Rev. Lett. **113**, 260401,2014.

[7] L. Mandelstam and I. Tamm. The Uncertainty Relation Between Energy and Time in Non-relativistic Quantum Mechanics. J. Phys. (USSR) **9**, 249,1945.

[8] S. Deffner, S. Campbell. Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control. J. Phys. A: Math. Theor. **50**, 453001,2017.

[9] N. Margolus and L. B. Levitin.The maximum speed of dynamical evolution. Physica D, **120**, 188,1998.

[10] A. Uhlmann. An energy dispersion estimate. Phys. Lett. A, **161**, 329,1992.

[11] J. B. Uffink. The rate of evolution of a quantum state. Am. J. Phys. 61, 935,1993.

[12] A. K. Pati.Uncertainty relation of Anandan-Aharonov and Intelligent states. Phys. Lett. A 262, 296,1999.

[13] V. Giovannetti, S. Lloyd, and L. Maccone.The role of entanglement in dynamical evolution. Europhys. Lett. (EPL) 62, 615,2003.

[14] V. Giovannetti, S. Lloyd, and L. Maccone.The speed limit of quantum unitary evolution. J. Opt. B: Quantum and Semiclassical Optics 6, S807,2004.

[15] S. Luo and Z. Zhang. On Decaying Rate of Quantum States. Lett. Math. Phys. 71, 1, 2005.

[16] J. Batle, M. Casas, A. Plastino, and A. R. Plastino.Connection between entanglement and the speed of quantum evolution. Phys. Rev. A 72, 032337,2005.

[17] M. Andrews.Bounds to unitary evolution. Phys. Rev. A 75, 062112,2007.

[18] C. Zander, A. R. Plastino, A. Plastino, and M. Casas.Entanglement and the speed of evolution of multi-partite quantum systems. J. Phys. A: Mathematical and Theoretical 40, 2861, 2007.

[19] B. Zielinski and M. Zych.Generalization of the Margolus-Levitin bound. Phys. Rev. A 74, 034301, 2006.

[20] H. F. Chau.Tight upper bound of the maximum speed of evolution of a quantum state. Phys. Rev. A 81, 062133, 2010.

[21] F. Frowis.Kind of entanglement that speeds up quantum evolution. Phys. Rev. A 85, 052127, 2012.

[22] S. Ashhab, P. C. de Groot, and F. Nori.Speed limits for quantum gates in multiqubit systems. Phys. Rev. A 85, 052327, 2012.

[23] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho.Quantum Speed Limit for Physical Processes. Phys. Rev. Lett. 110, 050402, 2013.

[24] D. Mondal and A. K. Pati.Quantum Speed Limit For Mixed States Using Experimentally Realizable Metric. Phys. Lett. A 380, 1395, 2016.

[25] V. Giovannetti, S. Lloyd, and L. Maccone. The role of entanglement in dynamical evolution. EPL **62**, 615, 2003.

[26] K. Svozil, L. B. Levitin, T. Toffoli, and Z. Walton.Maximum Speed of Quantum Gate Operation . Int. J. Theo. Phys., **44**, 965, 2005.

[27] C. Zander, A. R. Plastino, A. Plastino, and M. Casas.Entanglement and the speed of evolution of multi-partite quantum systems. J. Phys. A: Math. Theor., **40**, 2861, 2007.

[28] M. M. Taddei, B. M. Escher, L. Davidovich, and R. L. de Matos Filho, Phys. Rev. Lett., **110**,050402, 2013.

[29] A. del Campo, I. L. Egusquiza, M. B. Plenio, and S. F. Huelga.Quantum Speed Limits in Open System Dynamics. Phys. Rev. Lett., **110**, 050403, 2013.

[30] S. Deffner and E. Lutz.Quantum Speed Limit for Non-Markovian Dynamics. Phys. Rev. Lett., **111**, 010402, 2013.

[31] D. Mondal, S. Bagchi, A. K. Pati.Tighter uncertainty and reverse uncertainty relations. Phys. Rev. A **95**, 052117 ,2017.

[32] S. Deffner and E. Lutz.Generalized Clausius Inequality for Nonequilibrium Quantum Processes. Phys. Rev. Lett. 105, 170402,2010.

[33] S. Das, S. Khatri, G. Siopsis, and M. M. Wilde.Fundamental limits on quantum dynamics based on entropy change. J. Math. Phys. 59, 012205, 2018.

[34] J. D. Bekenstein.Energy Cost of Information Transfer. Phys. Rev. Lett. 46, 623, 1981.

[35] S. Lloyd.Ultimate physical limits to computation. Nature 406, 1047, 1981.

Quantum support vector data description for anomaly detection

Leo H. Oh[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. Anomaly detection is a critical problem in data analysis and pattern recognition, finding applications in various domains. This paper presents quantum support vector data description (QSVDD), an unsupervised learning algorithm designed for anomaly detection. QSVDD utilizes a shallow-depth quantum circuit to learn a hypersphere that tightly encloses normal data, tailored for the constraints of noisy intermediate-scale quantum (NISQ) computing. Simulation results on the MNIST image dataset demonstrate that QSVDD outperforms both quantum autoencoder and deep learning-based approaches. Notably, QSVDD offers the advantage of training an extremely small number of model parameters, which grow logarithmically with the number of input qubits. This enables efficient learning with a simple training landscape, presenting a compact and efficient quantum machine learning model for anomaly detection.

Keywords: Quantum computing, Quantum machine learning, Anomaly detection

1 Introduction

Quantum machine learning (QML) leverages quantum information theory to overcome the fundamental limitations of classical counterparts in addressing various data analysis tasks [1–3]. QML algorithms have achieved notable progress in the field of binary classification, which is a fundamental problem in pattern recognition. These algorithms demonstrate the potential to overcome the limits of classical approaches in terms of runtime, trainability, model capacity, and prediction accuracy [4–6]. Anomaly detection (AD) is yet another important branch of pattern recognition that finds applications in various domains, including finance [7–9], bioinformatics [10, 11], manufacturing [12], computer vision [13,14], and high energy physics [15]. However, AD poses greater challenges compared to binary classification, primarily due to the lack of anomalies by definition. Consequently, training AD models necessitates the use of unlabeled data and unsupervised learning techniques.

Several QML methods have been proposed to address AD [16–21]. However, these algorithms face limitations. Some require expensive subroutines like the quantum linear solver [22] and matrix exponentiation [23], making them unsuitable for Noisy Intermediate-Scale Quantum (NISQ) computing [24] computing. Others focus on data compression through quantum autoencoders (QAE) rather than AD itself.

To overcome these limitations, we introduce an end-to-end QML algorithm, specifically designed for AD and suitable for NISQ computing. Inspired by deep support vector data description (SVDD) [25–27], which has served as a foundation for several state-of-the-art classical AD approaches, our quantum AD algorithm is named quantum support vector data description (QSVDD). QSVDD trains a parameterized quantum circuit, using the SVDD objective function constructed from a set of quantum measurements, on normal data. While the parameterized quantum circuit can take various forms, we leverage the quantum convolutional neural network (QCNN) [28] due to its desirable properties such as efficient learning with a small number of parameters [29], absence of barren plateaus [30, 31], and excellent prediction performance [32]. Through numerical experiments on the MNIST image dataset, we demonstrate that QSVDD with the QCNN circuit outperforms QAE-based AD and classical deep learning-based AD approaches. Notably, the number of parameters required for training the model grows logarithmically with the number of input qubits. Hence, our algorithm offers an extremely compact and effective machine learning model for AD.

2 Quantum support vector data description

2.1 Loss Function

SVDD is based on learning a feature map $\Phi : \mathbb{R}^d \to \mathbb{R}^{d'}$, where $d' < d$, using a training set $\{\boldsymbol{x}_i\}_{i=1,\ldots,m}$. The goal is to minimize the radius of the hypersphere that contains $\{\Phi(\boldsymbol{x}_i)\}_{i=1,\ldots,m}$. The loss function is defined as:

$$L_c(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m} ||\Phi(\boldsymbol{x}_i, \boldsymbol{\theta}) - \boldsymbol{c}||^2 + \lambda R(\boldsymbol{\theta}), \tag{1}$$

where $\boldsymbol{c}$ represents the center of the hypersphere, which is defined before the optimization process. The second term in the equation is the regularization term that prevents overfitting. To identify anomalies in test data $\tilde{\boldsymbol{x}}$, the following condition is used:

$$||\Phi(\tilde{\boldsymbol{x}}, \arg\min_{\boldsymbol{\theta}} L(\boldsymbol{\theta})) - \boldsymbol{c}||^2 > b, \tag{2}$$

where b is a predetermined threshold.

QSVDD is the quantum counterpart of the previously described procedure, utilizing a quantum computer with n qubits to optimize the feature map Φ. In QSVDD, $\Phi(\boldsymbol{x}, \boldsymbol{\theta}) = g(U(\boldsymbol{\theta})|\psi(\boldsymbol{x})\rangle)$, where $|\psi(\boldsymbol{x})\rangle$ represents the quantum data encoding, and $U(\boldsymbol{\theta})$ is a parameterized quantum circuit. These elements, along with the function $g : \mathbb{C}^{2^n} \to \mathbb{R}^{d'}$, are discussed in detail in subsequent

*leo9123@yonsei.ac.kr
†dkd.park@yonsei.ac.kr

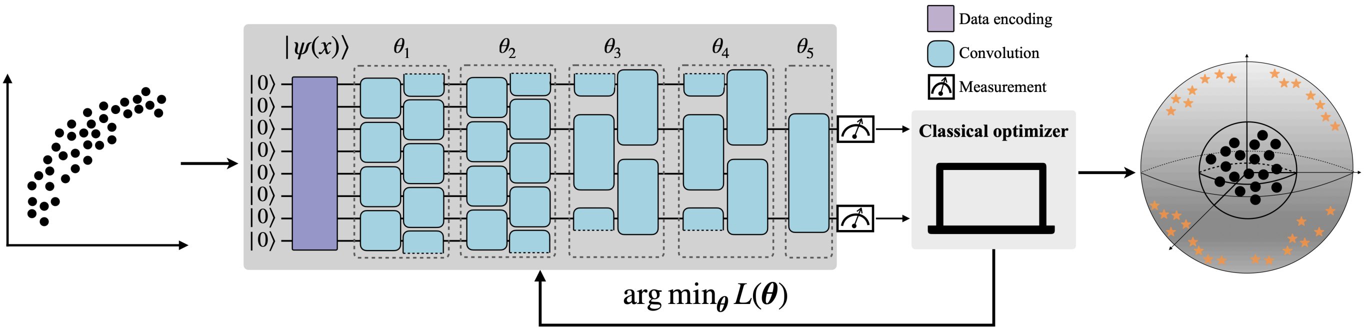


Figure 1: QSVDD consists of four main components: data encoding, variational quantum circuit, measurement, and optimization. The classical input data is encoded into quantum data with n qubits, where n depends on the number of features. This encoding produces the state $|\psi(x)\rangle$. Subsequently, the quantum data undergoes a QCNN circuit. In the circuit diagram, the dashed line on the convolutional gate indicates its connection through the top and bottom wires. During the measurement process, the expectation values of Pauli observables are computed, and these values are utilized in the optimization step. The trainable parameters within the variational quantum circuit are optimized using a classical optimizer to minimize the loss function.

sections. In the quantum version, since the trainable parameters of a quantum circuit typically range from 0 to 2π, the regularization term is unnecessary. Therefore, the QSVDD loss function is defined as:

$$L_q(\boldsymbol{\theta}) = \frac{1}{m}\sum_{i=1}^{m} ||\Phi(\boldsymbol{x}_i, \boldsymbol{\theta}) - \mathbf{c}||^2. \qquad (3)$$

Empirical observations consistently show that setting $\boldsymbol{c} = \mathbf{0}$ yields good results.

2.2 Data encoding

To perform a QML algorithm on classical data, classical data must be first mapped to a quantum state using a quantum feature map represented by the function $\Psi : \mathcal{X} \to \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space for quantum states [29, 33, 34]. The process of encoding classical data into a quantum state using a quantum feature map is described by the function $x \in \mathcal{X} \to |\psi(x)\rangle \in \mathcal{H}$. This encoding can be achieved through a unitary transformation of the initial state $|0\rangle^{\otimes n}$. Consequently, we obtain quantum feature mapped data $U_\psi(x)|0\rangle^{\otimes n} = |\psi(x)\rangle$, where n represents the number of qubits, and $U_\psi(x)$ denotes the unitary transformation.

QSVDD is capable of working with any quantum feature map. For this study, we employ amplitude encoding as an illustrative example. Amplitude encoding allows classical images to be encoded into an n-qubit system, where $n = O(\log N)$. Specifically, $|\psi(x)\rangle = \frac{1}{||x||}\sum_{i=1}^{N} x_i|i\rangle$, where $N = 2^n$, and $|i\rangle$ corresponds to the i-th computational basis state. This choice of encoding, along with QCNN, allows for a doubly-exponential reduction in the number of trainable parameters, making it the most compact QML model. Thus the use of amplitude encoding tests the performance of the QSVDD algorithm under the most extreme condition.

2.3 Variational quantum circuit

The encoded data undergoes a variational quantum circuit (VQC), $U(\boldsymbol{\theta})$, that contains trainable parameters $\boldsymbol{\theta}$. To make training more efficient, we utilize a QCNN circuit and divide $\boldsymbol{\theta}$ into five groups, denoted as $\boldsymbol{\theta} = \{\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, \boldsymbol{\theta_3}, \boldsymbol{\theta_4}, \boldsymbol{\theta_5}\}$. Each group corresponds to a set of parameters in the VQC structure in Figure 1. Within each group, we implement a convolutional circuit represented by a small blue box in Figure 1, consisting of $R_i(\theta)$ rotation gates where i denotes the axis of Bloch sphere, CNOT gates, and $U_3(\theta, \phi, \delta) = R_z(\phi)R_x(-\pi/2)R_z(\theta)R_x(\pi/2)R_z(\delta)$ gates. Since this convolutional circuit can generate the $SU(4)$ group, the pooling layer in certain QCNN models does not use parameterized gates [29].

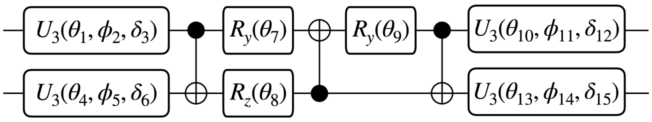
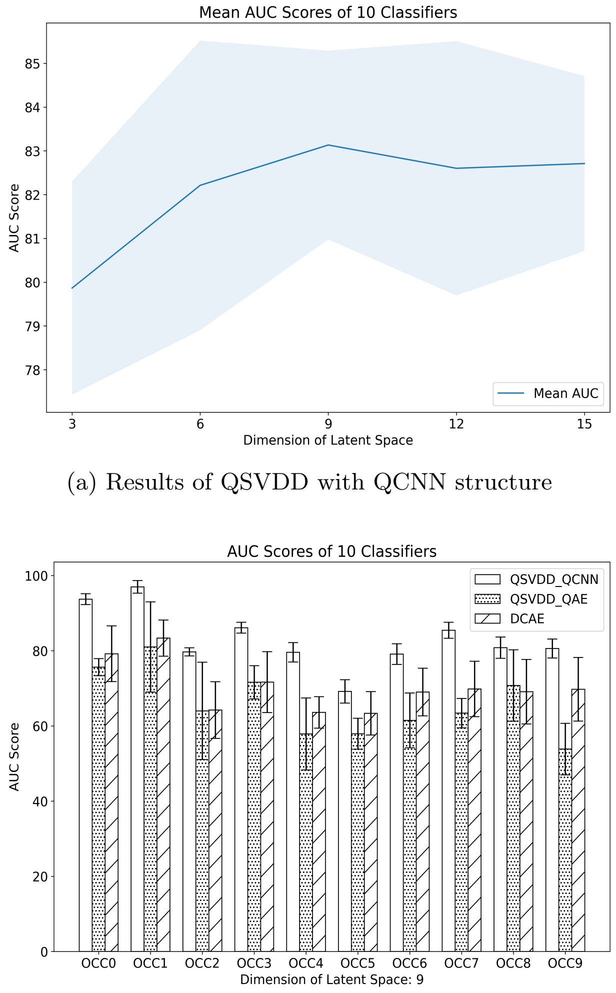


Figure 2: The convolutional circuit used in our study.

QCNN structure has the number of parameters for the optimization that grows as $O(\log n)$, where n is the number of qubits [29]. This implies that with QCNN structure and amplitude encoding stated in Section 2.2, we can reduce the number of parameters double-exponentially with the dimension of the classical input data. This is a great advantage in QML since it induces a simplified structure to optimize. This can also help prevent overfitting and enhance the VQC's generalization ability.

2.4 Measurement

The function $g : \mathbb{C}^{2^n} \to \mathbb{R}^{d'}$ is responsible for mapping the quantum data to a lower-dimensional latent space, where the hypersphere is determined. In this latent space, the maximum of d' is given by $\max(d') = 4^{n_o} - 1$, where n_o is the number of output qubits in the QCNN circuit. This represents the number of real parameters for an n_o-qubit density matrix. To implement the mapping function g, we utilize the expectation values of n_o-qubit Pauli observables. Each expectation value serves as a



(a) Results of QSVDD with QCNN structure

(b) Comparison with the other methods

Figure 3: AUC scores

coordinate within the feature space. Consequently, the dimensionality of the latent space is determined by the number of Pauli observables selected by the user, ranging from 1 to $4^{n_o} - 1$.

2.5 Optimization

The QSVDD loss function in Equation 3 is optimized by using a classical optimizer. The optimization process is initialized at $\boldsymbol{\theta_0}$, representing the initialized parameters sampled from the standard normal distribution. To evaluate the training performance of our model, we set the center as the origin of the given latent space, aiming to identify $\Phi(\cdot)$ that is effective in anomaly detection.

Before presenting numerical experiments, it is worth noting that QSVDD provides considerable flexibility in selecting its components, such as the data encoding, the ansatz, the measurement operators, and the optimizer. This freedom allows for customization to suit specific tasks. In this paper, we utilize amplitude encoding as the data encoding, and the QCNN and QAE ansatz as the variational quantum circuits for our numerical studies. Additionally, we employ the set of Pauli measurements as the measurement operators and Adam optimizer.

3 Numerical experiments

Our experimental findings highlight the exceptional performance of QSVDD when compared to the QAE method [21] and the deep convolutional autoencoder (DCAE), a common approach in deep learning for anomaly detection [27,35]. To ensure a fair comparison, we adjusted the number of parameters in each model accordingly. The QCNN structure utilized 75 parameters, the QAE structure utilized 78, and the DCAE structure utilized 92.

In our experiment, we evaluated the anomaly detection performance using the AUC score and compared them across different dimensionalities of the latent space. The results demonstrated that the optimal performance of QSVDD, employing the QCNN structure, was achieved when the latent space dimension was set to 9. Therefore, we compared the methods using a latent space dimension of 9, as depicted in Figure 3. Notably, QSVDD exhibited outstanding AUC performance with consistently small standard deviations across five repeated experiments. This low variability suggests a high level of robustness and reliability in the obtained results.

4 Conclusion

In this paper, we have presented the QSVDD algorithm, which utilizes a VQC framework for anomaly detection. Our approach involves training a shallow-depth quantum circuit to learn a hypersphere that tightly encloses normal data, with the flexibility to adapt to feature space dimensions ranging from 3 to 15. This flexibility is achieved by directly incorporating Pauli measurements into the loss function component and optimizing them using a classical optimizer. The flexibility of QSVDD enables its application to specific tasks. Additionally, QSVDD's simplified structure enabled by the QCNN circuit results in logarithmic parameter growth with qubit count, making it more favorable than other QML models. This scalability is particularly valuable for NISQ devices, as it enables more efficient anomaly detection with limited resources.

In summary, our work introduces a novel quantum anomaly detection algorithm by optimizing variational parameters through the QCNN circuit and the QSVDD loss function. By leveraging the power of quantum computing, QSVDD allows us to tackle challenging problems in the feature space that are difficult to address efficiently in the input space.

References

[1] M. Cerezo, Guillaume Verdon, Hsin-Yuan Huang, Lukasz Cincio, and Patrick J. Coles. Challenges and opportunities in quantum machine learning. *Nature Computational Science*, pages 1–10, 2022.

[2] Maria Schuld, Ilya Sinayskiy, and Francesco Petruccione. An introduction to quantum machine learning. *Contemporary Physics*, 56(2):172–185, 2015.

[3] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[4] Marcello Benedetti, Erika Lloyd, Stefan Sack, and Mattia Fiorentini. Parameterized quantum circuits as machine learning models. *Quantum Science and Technology*, 5(1):019601, 2020.

[5] 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(9):625–644, 2021.

[6] 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.

[7] Clifton Phua, Vincent Lee, Kate Smith, and Ross Gayler. A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv:1009.6119*, 2010.

[8] Shing-Han Li, David C Yen, Wen-Hui Lu, and Chiang Wang. Identifying the signs of fraudulent accounts using data mining techniques. *Computers in Human Behavior*, 28(3):1002–1013, 2012.

[9] Mohamad Jeragh and Mousa AlSulaimi. Combining auto encoders and one class support vectors machine for fraudulant credit card transactions detection. In *2018 Second World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4)*, pages 178–184. IEEE, 2018.

[10] Kristen Feher, Jenny Kirsch, Andreas Radbruch, Hyun-Dong Chang, and Toralf Kaiser. Cell population identification using fluorescence-minus-one controls with a one-class classifying algorithm. *Bioinformatics*, 30(23):3372–3378, 2014.

[11] Seonwoo Min, Byunghan Lee, and Sungroh Yoon. Deep learning in bioinformatics. *Briefings in bioinformatics*, 18(5):851–869, 2017.

[12] Luis Martí, Nayat Sanchez-Pi, José Manuel Molina, and Ana Cristina Bicharra Garcia. Anomaly detection based on sensor data in petroleum industry applications. *Sensors*, 15(2):2774–2797, 2015.

[13] Babak Saleh, Ali Farhadi, and Ahmed Elgammal. Object-centric anomaly detection by attribute-based reasoning. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 787–794, 2013.

[14] Yuequan Bao, Zhiyi Tang, Hui Li, and Yufeng Zhang. Computer vision and deep learning–based data anomaly detection method for structural health monitoring. *Structural Health Monitoring*, 18(2):401–421, 2019.

[15] Katherine Fraser, Samuel Homiller, Rashmish K. Mishra, Bryan Ostdiek, and Matthew D. Schwartz. Challenges for unsupervised anomaly detection in particle physics. *Journal of High Energy Physics*, 2022(3):66, Mar 2022.

[16] Nana Liu and Patrick Rebentrost. Quantum machine learning for quantum anomaly detection. *Phys. Rev. A*, 97:042315, Apr 2018.

[17] Zihua Chai, Ying Liu, Mengqi Wang, Yuhang Guo, Fazhan Shi, Zhaokai Li, Ya Wang, and Jiangfeng Du. Quantum anomaly detection of audio samples with a spin processor in diamond. *arXiv preprint arXiv:2201.10263*, 2022.

[18] Korbinian Kottmann, Friederike Metz, Joana Fraxanet, and Niccolò Baldelli. Variational quantum anomaly detection: Unsupervised mapping of phase diagrams on a physical quantum computer. *Phys. Rev. Research*, 3:043184, Dec 2021.

[19] Alona Sakhnenko, Corey O'Meara, Kumar JB Ghosh, Christian B Mendl, Giorgio Cortiana, and Juan Bernabé-Moreno. Hybrid classical-quantum autoencoder for anomaly detection. *Quantum Machine Intelligence*, 4(2):27, 2022.

[20] Vishal S. Ngairangbam, Michael Spannowsky, and Michihisa Takeuchi. Anomaly detection in high-energy physics using a quantum autoencoder. *Phys. Rev. D*, 105:095004, May 2022.

[21] Gunhee Park, Joonsuk Huh, and Daniel Kyungdeock Park. Variational quantum one-class classifier. *Machine Learning: Science and Technology*, 2022.

[22] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103:150502, Oct 2009.

[23] Seth Lloyd, Masoud Mohseni, and Patrick Rebentrost. Quantum principal component analysis. *Nature Physics*, 10(9):631–633, 2014.

[24] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[25] David MJ Tax and Robert PW Duin. Support vector data description. *Machine learning*, 54:45–66, 2004.

[26] Bo Liu, Yanshan Xiao, Longbing Cao, Zhifeng Hao, and Feiqi Deng. Svdd-based outlier detection on uncertain data. *Knowledge and information systems*, 34:597–618, 2013.

[27] Lukas Ruff, Robert Vandermeulen, Nico Goernitz, Lucas Deecke, Shoaib Ahmed Siddiqui, Alexander Binder, Emmanuel Müller, and Marius Kloft. Deep one-class classification. In *International conference on machine learning*, pages 4393–4402. PMLR, 2018.

[28] Iris Cong, Soonwon Choi, and Mikhail D. Lukin. Quantum convolutional neural networks. *Nature Physics*, 15(12):1273–1278, December 2019.

[29] Tak Hur, Leeseok Kim, and Daniel K Park. Quantum convolutional neural network for classical data classification. *Quantum Machine Intelligence*, 4(1):3, 2022.

[30] 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):4812, Nov 2018.

[31] Arthur Pesah, M. Cerezo, Samson Wang, Tyler Volkoff, Andrew T. Sornborger, and Patrick J. Coles. Absence of barren plateaus in quantum convolutional neural networks. *Phys. Rev. X*, 11:041011, Oct 2021.

[32] Leonardo Banchi, Jason Pereira, and Stefano Pirandola. Generalization in quantum machine learning: A quantum information standpoint. *PRX Quantum*, 2:040321, Nov 2021.

[33] Israel F Araujo, Daniel K Park, Francesco Petruccione, and Adenilton J da Silva. A divide-and-conquer algorithm for quantum state preparation. *Scientific reports*, 11(1):1–12, 2021.

[34] Seth Lloyd, Maria Schuld, Aroosa Ijaz, Josh Izaac, and Nathan Killoran. Quantum embeddings for machine learning. *arXiv preprint arXiv:2001.03622*, 2020.

[35] Raghavendra Chalapathy, Aditya Krishna Menon, and Sanjay Chawla. Anomaly detection using one-class neural networks. *arXiv preprint arXiv:1802.06360*, 2018.

Mitigation and Amplification under Generic Symmetry-Protected Thermalisation

Harshank Shrotriya[1] * Midhun Krishna[2] † Leong-Chuan Kwek[1] ‡
Varun Narasimhachar[3] § Sai Vinjanampathy[1 2 4] ¶

[1] *Centre for Quantum Technologies, National University of Singapore, Block S15, 3 Science Drive 2, 117543*
[2] *Department of Physics, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India*
[3] *Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, Singapore 138632*
[4] *Centre of Excellence in Quantum Information, Computation, Science and Technology, Indian Institute of Technology Bombay, Powai, Mumbai 400076, India.*

Abstract. When quantum systems are thermalised with constraints, the non-equilibrium steady state (NESS) may differ in its internal energy from the corresponding equilibrium state. Such a difference is called amplification (resp. mitigation) when the NESS has more (resp. less) energy than the corresponding equilibrium state, and is useful in engine design. In this talk, we present the general theory of amplification threshold with a view to non-equilibrium baths with strong symmetry, where both the Hamiltonian and Lindblad operators commute with a symmetry operator. We relate the amplification threshold to Landauer's erasure, towards a generic theory for strong symmetry-protected thermal machines.

Keywords: Open System Thermalisation, Strong Symmetry, Amplification and Mitigation

Symmetries of open quantum systems are powerful tools that provide control over the systems' dynamics. They come in two varieties depending on whether the symmetry is at the level of the Hilbert space or the Liouville space—namely, strong and weak symmetry. Open quantum systems with strong symmetries have multiple steady states associated with the dynamics and are hence capable of executing non-thermalising dynamics by "remembering" some information about the initial state of the system. The symmetries of open quantum systems have been employed in various practical applications, including quantum error correction [1], symmetry reduction [2] and controlling quantum transport [3]. But their implications in quantum thermodynamics still remain mostly unexplored. We investigate the role of strong symmetry in the amplification and mitigation of the action of a thermalising bath [4, 5, 6]. Multiple steady states associated with strong symmetric dynamics lead to partial thermalisation wherein thermalisation occurs within invariant subspaces of the symmetry (symmetry blocks), but thermal equilibriation of the initial state's overall populations in these subspaces is prevented by the symmetry. We call such a process symmetry-protected thermalisation (SPT).

We compare the quantum thermodynamics of the steady states attained by generic thermalisation and SPT using standard information theoretic definitions involving the non-equilibrium free energy. We interpret the initial state information protected by the SPT to be held in a classical register in the possession of a Maxwell's demon. This interpretation is elucidated in the two figures presented here. The demon perspective assists in providing a Landauer erasure interpretation for the amplification or mitigation of the thermal bath's effects in the presence of strong symmetry.

1 Review of non-Equilibrium Free Energy

The free energy of a generic quantum state with respect to an ambient thermal bath at temperature T is defined to be $F = E - TS$. Here E is the system's average internal energy and S is the von Neumann entropy of its state. For the equilibrium state at temperature T, $\rho^\beta = e^{-\beta H}/Z_\beta$, the free energy reduces to $F(\rho^\beta) = -KTln(Z_\beta)$, where $\beta := (k_\text{B}T)^{-1}$ is the inverse temperature of the bath,and $Z_\beta = Tr(e^{-\beta H})$ is the partition function of the equilibrium state. The free energy of an arbitrary non-equilibrium state ρ_ness can be expressed in terms of the equilibrium free energy as $F(\rho_\text{ness}) = F(\rho^\beta) + TS[\rho_\text{ness}||\rho^\beta]$ with $S[\rho||\sigma] = -S(\rho) - Tr(\rho \ln \sigma)$ being the relative entropy between the two states.

2 Amplification and Mitigation

The thermalisation process takes an arbitrary initial state of the system to the thermal state in equilibrium with the thermal bath driving the process. We seek to compare steady states attained by the generic thermalisation process and SPT. Consider two scenarios, one wherein the initial state ($\rho^{\beta_0} = \sum_n p_n^{\beta_0} \rho_n^{\beta_0}$) in equilibrium with a thermal bath at inverse temperature β_0 undergoes generic thermalisation to thermal state at inverse temperature β ($\rho^{\beta_B} = \sum_n p_n^{\beta_B} \rho_n^{\beta_B}$) and the other where SPT takes it to a non-equilibrium steady state $\rho_\text{ness} = \sum_n p_n^{\beta_0} \rho_n^{\beta_B}$ where the initial probabilities $p_n^{\beta_0}$ to be in symmetric subspace n are preserved.

*harshank.s@u.nus.edu
†midhunkrishna@iitb.ac.in
‡cqtklc@nus.edu.sg
§varun.achar@gmail.com
¶sai@phy.iitb.ac.in

Amplification (mitigation) refers to the scenario where the change in thermodynamic quantities is more (less) in SPT relative to generic thermalization. The difference in internal energies of steady states attained after both the procedure can be expressed as,

$$E_{\text{ness}} - E^{\beta_B} = T_B(S(\rho_{\text{ness}}) - S(\rho^{\beta_B}) + S[\rho_{\text{ness}}||\rho^{\beta_B}]). \quad (1)$$

We understand the amplification and mitigation effects of SPT using Landauer's erasure principle below.

3 Erasure Principle

The erasure principle was originally proposed by Landauer [7] to resolve the Maxwell demon paradox by associating a heat cost $k_B T \ln 2$ with erasure of every bit of classical information. We study a scenario where classical information is encoded in an ensemble of quantum states as described in [8]. This ensemble of states interacts sequentially with a bath which thermalises each of the states to the thermal state thus erasing the encoded classical information. It follows from the Erasure principle that this erasure has to be associated with an entropy cost $\Delta S_{\text{erasure}}$ such that

$$\Delta S_{\text{erasure}} = \Delta S_{\text{system}} + \Delta S_{\text{bath}} \quad (2)$$

where ΔS_{system} and ΔS_{bath} are calculated for the underlying thermalisation process. We employ the erasure principle to gain insights into the conditions leading to the amplification and mitigation effects of the bath. To that end, we interpret the thermalisation pathways as compositions of intermediate steps as shown in Fig. 1.

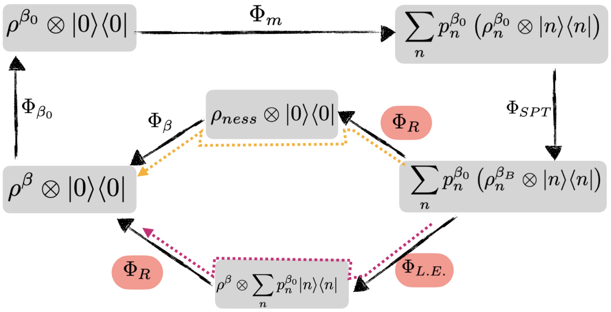

$\rho^{\beta_0} \otimes |0\rangle\langle 0|$
Φ_m
$\sum_n p_n^{\beta_0} (\rho_n^{\beta_0} \otimes |n\rangle\langle n|)$
Φ_{β_0}
Φ_β
$\rho_{ness} \otimes |0\rangle\langle 0|$
Φ_R
Φ_{SPT}
$\rho^\beta \otimes |0\rangle\langle 0|$
$\sum_n p_n^{\beta_0} (\rho_n^{\beta_B} \otimes |n\rangle\langle n|)$
Φ_R
$\rho^\beta \otimes \sum_n p_n^{\beta_0} |n\rangle\langle n|$
$\Phi_{L.E.}$

Figure 1: A flowchart of two complete therodynamic cycles involving different procedures to erasure and reset the information gained from a measurement which acquires information about the eigenvalues of the strong symmetry operator as discussed in the text.

4 Central Result

The first step is the measurement step Φ_m which can be thought of as an external observer acquiring the classical information associated with probabilities $\{p_i\}$. Also, note that this step is accompanied by a decrease in the entropy of the system since $S(\rho^{\beta_0}) \leq \sum_i p_i S(\rho_i^{\beta_0})$ and the difference is exactly the information theoretic entropy $H(p)$ associated with the information acquired by the external observer. Second step is the Symmetry Protected Thermalisation Φ_{SPT} where each of the subspace restricted states $\rho_i^{\beta_0}$ is thermalised to $\rho_i^{\beta_B}$ while conserving the probabilities $\{p_i\}$. After this point, the two pathways diverge, geenralised thermalisation consists of the erasure step Φ_{LE} where each state in the ensemble is thermalised to the full thermal state ρ^{β_B} effectively erasing the information associated with the probabilities $\{p_i\}$.

One can note from above that the entropy of erasure associated with Φ_{LE} is only part of the generalized thermalisation pathway leading to the following remark.

Remark 1 *Total entropy change asociated with the generalized thermalisation pathway is always higher than symmetry protected thermalisation pathway and the difference can be interpreted as an entropy of erasure of classical information encoded in the probabilities* $\{p_i\}$.

In the full draft [9], we describe how entropy of erasure of the final step can be manipulated to obtain the relation of amplification/mitigation boundary as shown in (1). In order to physically interpret this result, we start with independently analyzing the different kinds of entropies associated with the erasure step

$$\{p_i, \rho_i^{\beta_B}\} \longrightarrow \rho^{\beta_B} \quad (3)$$

which are $\Delta S_{\text{erasure}}$, ΔS_{system} and ΔS_{bath} with $\Delta S_{\text{erasure}} = \Delta S_{\text{system}} + \Delta S_{\text{bath}}$.

Now, for the transformation of the system associated with the erasure step above ΔS_{system} is fixed and positive. Further, one can interpret ΔS_{bath} in relation to the internal energy change of the system with the opposite sign since

$$\Delta S_{\text{bath}} = \frac{\Delta Q_{\text{bath}}}{T_B} = -\frac{\Delta Q_{\text{sys}}}{T_B} = -\frac{\Delta U_{\text{sys}}}{T_B}. \quad (4)$$

where $\Delta U_{\text{sys}} = \text{Tr}[(\rho_{\text{ness}} - \rho^{\beta_B})H]$. Depending on $\Delta S_{\text{erasure}}$, three scenarios could occur:

- $\Delta S_{\text{erasure}} = \Delta S_{\text{system}}$: We call this scenario as that of optimal erasure since the entropy increase accompanied by erasure of information is exactly equal to the entropy increase of the system (note that erasure entropy is always positive).

- $\Delta S_{\text{erasure}} > \Delta S_{\text{system}}$: Since the erasure entropy is higher than entropy increase of the system, the extra entropy is supplied to the bath thus causing an amplification of the internal energy in the presence of symmetry protected thermalisation.

- $\Delta S_{\text{erasure}} < \Delta S_{\text{system}}$: It follows that in this case entropy of bath decreases since it supplies heat to the system as a result causing mitigation of the internal energy in the presence of symmetry protected thermalisation.

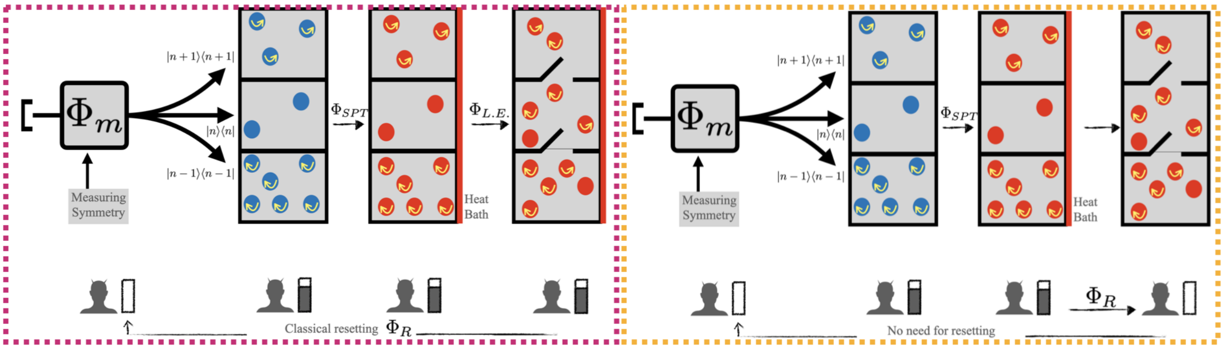


Figure 2: A schematic representation of the flowchart in fig. 1. On the left is an initial measurement that separates the quantum system into different symmetry sectors, indicated by the circles with differently oriented arrows in the first set of boxes. In the middle set of boxes, these states are allowed to thermalise within their symmetry sectors. In the last column of boxes to the left, the thermalisation continues while the symmetries are relaxed. This produces a standard Gibbs equilibrium at temperature β. On the right hand side, three panels describe the same process, except that the thermalisation is restricted only to the symmetry sectors. Here the final state is a NESS as described in the text.

5 Discussion

The amplification vs mitigation behaviour of thermodynamic quantities during thermalisation is of much interest due to the potential for enhancing the performance of quantum thermal machines and quantum batteries. This behaviour has been studied in the context of a pair of indistinguishable two-level systems (TLS) [5], which has been later extended to a collection of an arbitrary number of such TLSs [4]. Furthermore, a further generalisation to a collection of indistinguishable multilevel systems has also been reported [10]. However, all these studies have been restricted to a special case of strong symmetry, namely permutational symmetry, where the systems undergo collective dynamics due to indistinguishability. In this work, we show a general theory of the amplification and mitigation action of thermal baths on quantum systems in the presence of arbitrary strong symmetries encompassing all previous scenarios. We employ Maxwell's demon perspective to understand symmetry-protected thermalisation allowing an interpretation of the amplification and mitigation in terms of Landauer's erasure principle. The new thermalisation protocol, namely symmetry-protected thermalisation, by reservoir engineering, provides versatile control the steady state attained by the dynamics.

References

[1] Philippe Faist, Sepehr Nezami, Victor V Albert, Grant Salton, Fernando Pastawski, Patrick Hayden, and John Preskill. Continuous symmetries and approximate quantum error correction. *Physical Review X*, 10(4):041018, 2020.

[2] Berislav Buča and Tomaž Prosen. A note on symmetry reductions of the lindblad equation: transport in constrained open spin chains. *New Journal of Physics*, 14(7):073007, 2012.

[3] D Manzano and PI Hurtado. Harnessing symmetry to control quantum transport. *Advances in Physics*, 67(1):1–67, 2018.

[4] Camille L Latune, Ilya Sinayskiy, and Francesco Petruccione. Thermodynamics from indistinguishability: Mitigating and amplifying the effects of the bath. *Physical Review Research*, 1(3):033192, 2019.

[5] CL Latune, I Sinayskiy, and F Petruccione. Energetic and entropic effects of bath-induced coherences. *Physical Review A*, 99(5):052105, 2019.

[6] Camille L Latune, Ilya Sinayskiy, and Francesco Petruccione. Collective heat capacity for quantum thermometry and quantum engine enhancements. *New Journal of Physics*, 22(8):083049, 2020.

[7] Rolf Landauer. Irreversibility and heat generation in the computing process. *IBM journal of research and development*, 5(3):183–191, 1961.

[8] Koji Maruyama, Franco Nori, and Vlatko Vedral. Colloquium: The physics of maxwell's demon and information. *Reviews of Modern Physics*, 81(1):1, 2009.

[9] Harshank Shrotriya, Midhun Krishna, Leong-Chuan Kwek, Varun Narasimhachar, and Sai Vinjanampathy. Mitigation and amplification under generic symmetry-protected thermalisation. *[In Preparation]*.

[10] Benjamin Yadin, Benjamin Morris, and Kay Brandner. Thermodynamics of permutation-invariant quantum many-body systems: A group-theoretical framework. *arXiv preprint arXiv:2206.12639*, 2022.

Design and Implementation of a GRU Model Based on Quantum Circuits

Xiao Shi[1 2] Yun Shang[1 3 *] Tianen Chen[1 2]

[1] *Institute of Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*

[2] *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*

[3] *NCMIS, MDIS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190,China*

Abstract. GRU(Gated Recurrent Units), a popular variant of recurrent neural networks, has been widely employed in natural language processing and time series forecasting tasks due to its ability to capture long-term dependencies. In this paper, we propose QGRU, a novel classical-quantum hybrid algorithm that incorporates GRU's memory and forgetting mechanisms into variational quantum circuits. QGRU achieves accurate forecasts with fewer qubits in the hidden layer and consistently outperforms other quantum algorithms on diverse real datasets in classification tasks. This study showcases the potential of QGRU in temporal data processing and suggests exciting prospects for further research in quantum computing applications

Keywords: GRU, variational quantum circuits, Quantum GRU, temporal data processing, classification

1 INTRODUCTION

Quantum machine learning (QML) harnesses quantum computing and machine learning for diverse domains [1, 2]. Exploiting quantum features like entanglement and superposition, QML offers exponential speedup in optimization and simulation. Some QML algorithms show proven speedups, but many use parameterized quantum circuits and large Hilbert spaces. Noisy intermediate-scale quantum (NISQ) devices enable QML implementation, spurring advances in this emerging field.

Recurrent neural networks (RNN) [3] process sequential data, capturing temporal correlations and dependencies. With a looped structure, RNN uses previous outputs as inputs for later time steps, handling sequences with varying lengths. RNN applies to language modeling, speech recognition, machine translation, and more. Their success motivates using RNN to learn quantum evolutionary dynamics from sequential experimental data.

RNN processes sequential data, but suffers from vanishing gradients, hindering long-term dependencies. Researchers have proposed solutions, such as LSTM and GRU. Compared to LSTM, GRU has fewer parameters and is easier to train, while matching LSTM results.

GRU [4] defines the reset gate r_t and update gate z_t based on the previous hidden state h_{t-1} and current input x_t as follows:

$$r_t = \sigma(W_{xr}X_t + W_{hr}H_{t-1} + b_r) \tag{1}$$

$$z_t = \sigma(W_{xz}X_t + W_{hz}H_{t-1} + b_z) \tag{2}$$

Once the reset gate r_t and update gate z_t are obtained, the candidate hidden state is defined using the reset gate as follows:

$$H'_t = tanh(X_tW_{xh} + (R_t \odot H_{t-1})W_{hh} + b_h) \tag{3}$$

In this equation, the symbol $\odot$ denotes element-wise multiplication. The output of the activation function *tanh* scales the data to the range of -1 to 1. It is evident that H'_t mainly contains the current input x_t and selectively incorporates the data of H_{t-1} to it. Next, GRU employs the update gate z_t to compute the final result of the current unit H_t and pass it to the next unit. The expression for H_t is given by

$$H_t = (1 - z_t) \odot H_{t-1} + z_t \odot H'_t \tag{4}$$

The range of z_t is between 0 and 1. A value close to 1 indicates that more information is being "remembered", while a value closer to 0 indicates more "forgetting". In other words, GRU selectively forgets information from the input based on the weight z_t. Therefore, it uses a single gate to achieve both forgetting and selective memory.

2 METHOD

Our quantum QRNN framework is described as follows. The QRNN consists of three parts: the data encoding layer, the variational layer, and the measurement layer. In the proposed QRNN framework, the data encoding layer utilizes parameterless quantum gates to transform classical data into quantum states. At time step t, the input data $\vec{x}_t = (x_t^1, x_t^2, ..., x_t^n)$ is encoded into the quantum state $|x_t\rangle$. Then, $U(\theta)$ is applied to the quantum state $|H_{t-1}\rangle \otimes |x_t\rangle$, where $|H_{t-1}\rangle$ incorporates all the information from previous inputs x_i. $U(\theta)$ introduces parameters through single-qubit rotation gates and entanglement through CNOT gates, resulting in the output quantum state $|\psi_h\rangle = U(\theta)(|\psi_{h-1}\rangle \otimes |x_t\rangle)$. The gate U is composed of CNOT gates and single-qubit rotation gates. Subsequently, the last x_t qubits are measured to introduce nonlinearity and reduce the degrees of freedom, producing the quantum state $|H_t\rangle$, which is then passed to the next time step, as shown in Figure 1(a). In the

[*] shangyun@amss.ac.cn

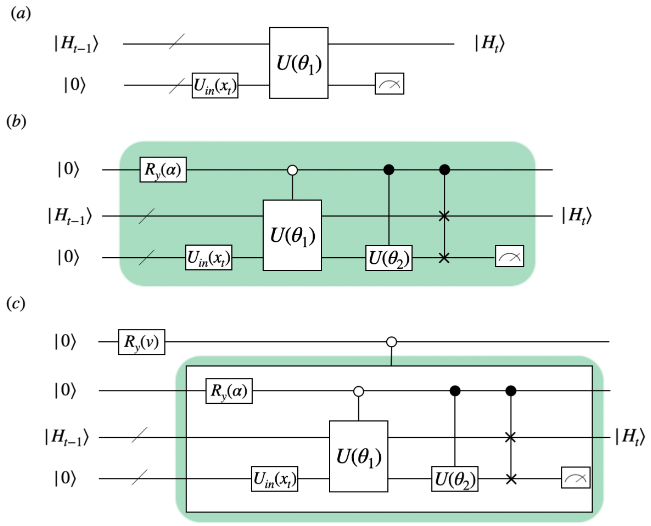


Figure 1: At time t, (a) represents the quantum circuit of QRNN, (b) represents a part of the circuit in QGRU that implements the reset gate, and (c) shows the complete circuit of QGRU.

final time step, the amplitudes of the hidden layer qubits are measured, and the final output is obtained through classical fully connected layers.

We have borrowed the ideas of update and forget mechanisms from the classical GRU algorithm and introduced them into our quantum circuit. Specifically, we assume that the quantum state $|\psi_i\rangle_{AB} = |\psi_i\rangle_A \otimes |\psi_i\rangle_B$ represents the quantum state on $n_h + n_x$ qubits, where subsystem A consists of n_h qubits and subsystem B consists of n_x qubits. Here we consider the case where the update and reset gates only contain biases. We set $|\psi_i\rangle_A = |H_{t-1}\rangle$ to represent the previous hidden state, and $|\psi_i\rangle_B = |x_t\rangle$ to represent the current input. In the circuit shown in Figure 1(b), we introduce an auxiliary qubit to achieve the reset mechanism of the quantum state, that is, the final quantum state contains information of $|\psi_i\rangle_B$ and selectively adds $|\psi_i\rangle_A$ to the current hidden state. After passing through the quantum circuit, the initial state $|0\rangle \otimes |\psi_i\rangle_{AB}$ undergoes an evolution and transforms into the final state $\alpha|0\rangle \otimes (U_{AB}(\theta_1)|\psi_i\rangle_{AB}) + \sqrt{1-\alpha^2}|1\rangle \otimes ((U_B(\theta_2) \otimes I_A)|\psi_i\rangle_{BA})$. Where α refers to the parameters of the $R_y(\alpha)$ gate in the first qubit of Figure 1(b), $U_{AB}(\theta_1)$ represents the quantum gate operation on the combined system AB with the parameter θ_1, $U_B(\theta_2)$ represents the quantum gate operation on system B with the parameter θ_2, and I_A represents the identity operation on system A. This expression describes the entangled state resulting from the evolution of the initial state through the quantum circuit.

After resetting the ancilla qubit to $|0\rangle$ and measuring the quantum bit corresponding to subsystem B, we obtain a candidate hidden quantum state

$$|\widetilde{H}\rangle = \alpha tr_B(U_{AB}(\theta_1)|\psi_i\rangle_{AB}) + \sqrt{1-\alpha^2}U_B(\theta_2)|\psi_i\rangle_B. \tag{5}$$

When $\alpha = 0$, $|\widetilde{H}\rangle$ simplifies to $U_B(\theta_2)|\psi_i\rangle_B$. When $\alpha = 1$, $|\widetilde{H}\rangle$ simplifies to the quantum state $tr_B(U_{AB}(\theta_1)|\psi_i\rangle_{AB})$. Otherwise, $|\widetilde{H}\rangle$ is a superposition of the two. As we can see, the candidate quantum state $|\widetilde{H}\rangle$ here implements the reset function in QGRU.

After obtaining the candidate hidden state $|\widetilde{H}\rangle$, we introduce an additional auxiliary qubit to realize the update mechanism. As shown in Figure 1(c), the auxiliary qubit is first acted upon by the $R_y(v)$ gate and then used to control the quantum circuit in Figure 1(b). After resetting the auxiliary qubit to 0, we obtain the final quantum state.

$$\begin{aligned}|H_t\rangle &= v|\widetilde{H}\rangle + \sqrt{1-v^2}|\psi_i\rangle_A \\ &= v\alpha\ tr_B(U_{AB}(\theta_1)|\psi_i\rangle_{AB}) + v\sqrt{1-\alpha^2}U_B(\theta_2)|\psi_i\rangle_B \\ &+ \sqrt{1-v^2}|\psi_i\rangle_A\end{aligned} \tag{6}$$

Noting that the value of v ranges from 0 to 1, a gate signal closer to 1 represents a greater amount of "memorized" data, while a signal closer to 0 represents more "forgotten" data. By employing a single gate, both forgetting and memory selection are achieved simultaneously. Subsequently, $|H_t\rangle$ is taken as the hidden quantum state input for the next block.

3 RESULTS AND DISCUSSION

We initiate function simulations to validate the performance of QGRU in time series forecasting. Cosine and second-order Bessel functions are utilized as illustrative examples. Predicting the output of the (N+1)-th term based on the values of the previous N terms, where N=4, we examine QRNN and QGRU with 1 hidden layer neurons, requiring 2 and 4 qubits, respectively. For classical data, amplitude encoding is employed to convert it into quantum states. In Table 1, showcasing their mean square errors (MSE) on the test dataset, alongside the classical GRU. Comparing with QRNN, it is evident that QGRU achieves a lower MSE. Furthermore, even when compared to classical GRU with five hidden layers, QGRU's performance only exhibits a slight difference. These results indicate that QGRU outperforms QRNN in terms of predictive accuracy and demonstrates competitive performance with classical GRU models.

Table 1: The mean squared error of QRNN and QGRU on the test set of sine function and Bessel function.

Function	QRNN loss	QGRU loss	GRU loss
cos	3.67×10^{-4}	2.61×10^{-4}	6.0510^{-5}
Bessel	$2.95 * 10^{-4}$	$1.17 * 10^{-4}$	$1.06 * 10^{-4}$

Next, we apply our model to classification tasks. We used the method mentioned in Ref. [5] to generate a simple circular decision boundary binary dataset (x^k, y^k), which is a nonlinear dataset where $x^k \in R^2$ represents the data point and $y^k \in \{0, 1\}$ represents the label. The visualization of the dataset is shown in Figure 2. In a classification task, unlike function approximation, we typically add a softmax function F to the output layer.

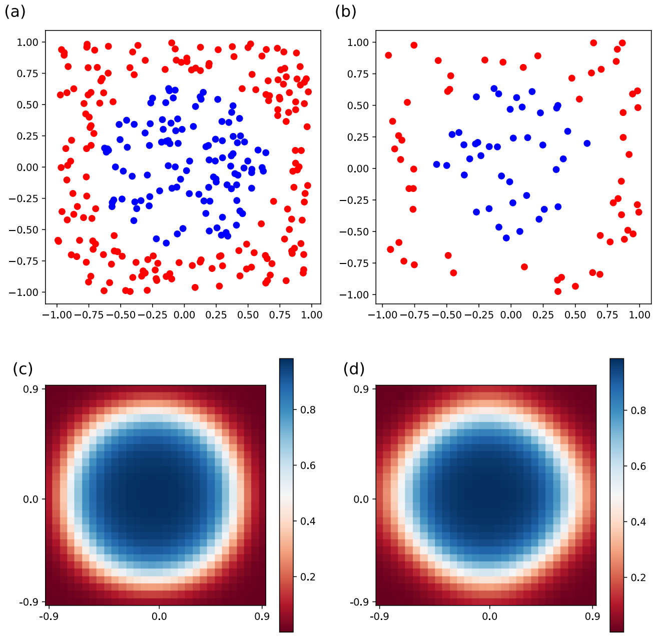


Figure 2: The figure shows a training set with 300 data points in (a) and a test set with 100 data points in (b). Blue points belong to class 0 and red points belong to class 1. (c) shows the decision boundary learned by QRNN, while (d) shows the decision boundary learned by QGRU.

The classification results are shown in Figure 2, and we can see that both QRNN and QGRU achieve 100% classification accuracy in the nonlinear classification task.

In our experiments, QRNN and QGRU were evaluated for classification tasks on real-world datasets. We divided the data into a training set (70%) and a test set (30%). To compare their performance, we also considered other quantum classification algorithms, such as circuit-centric quantum classifiers and ADQC. For QRNN and QGRU, we set the number of hidden layer qubits to 1, requiring two qubits for QRNN and four qubits for QGRU. To ensure fairness, both circuit-centric quantum classifiers and ADQC employed rotation encoding, thus also requiring four qubits. Similarly, the classical GRU model had 5 qubits in its hidden layer. Notably, QGRU achieved 100% accuracy on the iris dataset, surpassing other algorithms by 2.22% to 7.78%. Additionally, QGRU demonstrated superior performance on the Bupa and Wine datasets compared to existing algorithms. We further validated our results using the USPS image dataset. Each image was treated as a temporal sequence with 16 features. To save quantum resources, we used amplitude encoding instead of rotation encoding. For QRNN and QGRU, with 16 features requiring 4 qubits for encoding, the number of hidden qubits was set to 4. We selected two classes ("0" and "1") from the USPS dataset and randomly chose 1,000 images from the training set for training, while evaluating the classification results using all 621 images from the test set. Notably, QGRU achieved a significantly higher classification accuracy of 96.78% compared to QRNN's 88.41%. These results highlight the improved classification performance of QGRU, attributed to the introduction of update and forget mechanisms.

Table 2: Performance of 5 algorithms on 3 real-world datasets.

Datasets Methods	Iris	Bupa	Wine
Circuit-centric [6]	97.78%	65.38%	53.70%
ADQC [7]	95.56%	70.19%	64.81%
GRU[4]	95.56%	66.35%	62.96%
QRNN	93.33%	64.42%	64.81%
QGRU	100%	73.08%	68.52%

4 CONCLUSION

We compare QRNN and QGRU for time series prediction. Our experiments show that our QGRU model matches or surpasses classical GRU models, and outperforms QRNN. This indicates that memory and forget mechanisms in quantum circuits can enhance performance, revealing the potential of quantum neural networks.

References

[1] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91:045002, Dec 2019.

[2] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, 2017.

[3] Jeffrey L Elman. Finding structure in time. *Cognitive science*, 14(2):179–211, 1990.

[4] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[5] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Phys. Rev. A*, 98:032309, Sep 2018.

[6] Maria Schuld, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. Circuit-centric quantum classifiers. *Phys. Rev. A*, 101:032308, Mar 2020.

[7] Peng-Fei Zhou, Rui Hong, and Shi-Ju Ran. Automatically differentiable quantum circuit for many-qubit state preparation. *Phys. Rev. A*, 104:042601, Oct 2021.

Probabilistic unitary and state synthesis with optimal accuracy

Seiseki Akibue[1] * Go Kato[2] † Seiichiro Tani[1] ‡

[1] *NTT Communication Science Labs., NTT Corporation.*
3–1, Morinosato-Wakamiya, Atsugi, Kanagawa 243-0198, Japan
[2] *Advanced ICT Research Institute, NICT.*
4–2–1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan

Abstract. A new synthesis approach, called probabilistic synthesis, samples a gate sequence to suppress the approximate error of a target unitary transformation or pure state. We reveal the fundamental limitations on the approximation error obtained by the optimal probabilistic synthesis compared to the conventional approach. We also construct efficient probabilistic synthesis algorithms for single-qubit unitaries and qudit pure states, rigorously estimate their time complexity, and show their quadratic error reduction. These results are based on our general lemma about the optimal convex approximation of states or transformations. As a byproduct, our lemma provides a novel method to analyze an entanglement measure.

Keywords: synthesis, convex approximation, probabilistic approximation, entanglement

1 Background

To realize information processing in a gate model quantum computer, we need to prepare an initial state and perform unitary transformations on a fixed-size system with the desired accuracy. This is possible by exploiting quantum error correction [1] or the nature of the system [2]. However, those techniques usually force us to prepare a target pure state ϕ or realize a target unitary transformation Υ by using a circuit formed from a finite gate set such as $\{H, T, CNOT\}$. As a result of the discretization, we can only prepare an approximated pure state $\hat{\phi}$ or implement an approximated unitary transformation $\hat{\Upsilon}$ in general.

To suppress the effect of decoherence or overhead caused by the fault-tolerant implementation of each gate, various *synthesis* algorithms [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15] have been proposed for minimizing the approximation error or the circuit size. Following the celebrated Solovay-Kitaev algorithm [3], the final goal of conventional synthesis algorithms is to deterministically find one of the best circuits for the approximation. Thus, the minimum approximation error obtained by such *deterministic* unitary (or state) synthesis is given by $\min_{x\in X} \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ (or $\min_{x\in X}\left\|\phi - \hat{\phi}_x\right\|_{\mathrm{tr}}$), where X is the label set of unitary transformations (or pure states) realized by circuits with a certain cost, e.g., the circuit size, depth, or the number of T gates.

While it makes sense to approximate a target unitary transformation (or pure state) by utilizing an approximated unitary (or pure state) generated by a single circuit, a recently proposed approach called *probabilistic synthesis* probabilistically samples a circuit for the approximation. Suppose that the probabilistic algorithm independently samples a circuit $\mathcal{C}_x$ implementing $\hat{\Upsilon}_x$ (or generating $\hat{\phi}_x$) in accordance with a probability distribution $p(x)$ each time Υ (or ϕ) is required in quantum information processing. Then, each realized physical transformation (or generated state) is described by $\sum_x p(x)\hat{\Upsilon}_x$ (or $\sum_x p(x)\hat{\phi}_x$). This can be interpreted as the transition from coherent errors to incoherent errors [16, 17, 18], and recent studies have experimentally demonstrated that this transition reduces the approximation error of pure states [19]. Moreover, Campbell [20] and Vadym et al. [21] constructed synthesis algorithms that reduce the approximation error of unitary transformations into $\frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond = O(\epsilon^2)$ by choosing $p(x)$ appropriately, where $\{\hat{\Upsilon}_x\}_x$ causes the worst approximation error $\epsilon := \max_\Upsilon \min_x \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ if the deterministic synthesis is used.

Despite its importance, the limitation of probabilistic synthesis, especially the minimum approximation error $\min_p \frac{1}{2}\left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_\diamond$ (or $\min_p\left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\mathrm{tr}}$), remains unknown, nor is it clear how to find the optimal probability distribution p. While a few analytical results are obtained for the case of a qubit transformation [22] or state [23, 24, 25] in the context of the optimal convex approximation of a quantum transformation or state, minimax optimization to compute the minimum approximation error makes analyses quite difficult in general. Note that the result of optimal probabilistic unitary synthesis does not contain that of state synthesis, and vice versa. This is because the generated state in state synthesis is obtained by applying a unitary transformation to a fixed input state $|0\rangle$ while the approximation error in unitary synthesis is quantified for the worst input state. Moreover, a target state could be approximated by probabilistically mixing two unitary transformations whose behaviors are totally different, except for $|0\rangle$.

2 Our contributions

Before presenting our results, we provide intuitive examples demonstrating the capability of probabilistic synthesis in Fig. 1. As a generalization of the qubit examples, we obtain the fundamental relationship between the

*seiseki.akibue@ntt.com
†go.kato@nict.go.jp
‡seiichiro.tani@ntt.com

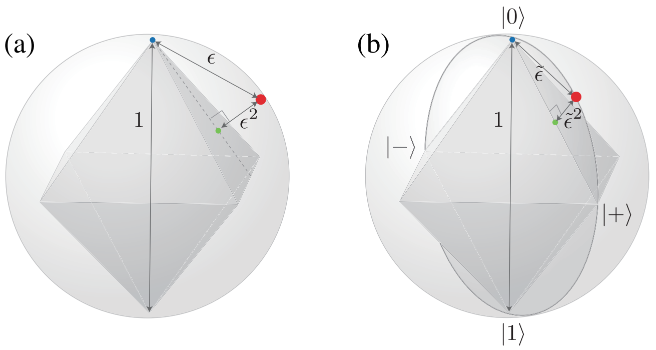


Figure 1: Quadratic reduction of the approximation error by using probabilistic synthesis. We assume that we can exactly generate an eigenstate $\hat{\phi}_x$ of the Pauli operators, represented by the six extreme points of the octahedron. We represent the Bloch sphere by a sphere with radius $\frac{1}{2}$, where the trace distance between two quantum states equals the Euclidean distance between the corresponding points. (a) We can compute $\min_p \left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\mathrm{tr}} = \epsilon^2 = \frac{1}{2\sqrt{3}}\left(\sqrt{3}-1\right)$ and $\min_x \left\|\phi - \hat{\phi}_x\right\|_{\mathrm{tr}} = \epsilon$, where ϕ is the furthest state from $\{\hat{\phi}_x\}_{x=1}^{6}$, represented as a large red point. (b) Suppose that the target state is chosen from $S_G := \{\phi : |\phi\rangle = \cos t|0\rangle + \sin t|1\rangle, t \in \mathbb{R}\}$, represented by a meridian. We can compute $\min_p \left\|\phi - \sum_x p(x)\hat{\phi}_x\right\|_{\mathrm{tr}} = \tilde{\epsilon}^2 = \frac{1}{2}\left(1 - \frac{1}{\sqrt{2}}\right)$ and $\min_x \left\|\phi - \hat{\phi}_x\right\|_{\mathrm{tr}} = \tilde{\epsilon}$, where ϕ is the furthest state in S_G from $\{\hat{\phi}_x\}_{x=1}^{6}$, represented as a large red point.

minimum approximation errors obtained by the deterministic synthesis and the probabilistic one:

Theorem 1 (simplified version) *[26, Theorem 1] Let G be a finite subgroup of unitary and antiunitary operators and $S_G := \{\phi : \forall U \in G, [U, \phi] = 0\}$ be the set of pure states invariant under the action of G. If $\{\hat{\phi}_x\}_{x\in X} = \{U\hat{\phi}_x U^\dagger\}_{x\in X}$ for all $U \in G$ with a finite set X, it holds that*

$$\max_{\phi\in S_G} \min_{p} \left\|\phi - \sum_{x\in X} p(x)\hat{\phi}_x\right\|_{tr} = \max_{\phi\in S_G} \min_{x\in X} \left\|\phi - \hat{\phi}_x\right\|_{tr}^2. \tag{1}$$

This theorem compares the *worst* approximation errors occurring when one synthesizes the target state in a subset S_G that is most difficult to approximate by using $\{\hat{\phi}_x\}_x$, which does not need to be a subset of S_G. It implies that the optimal probabilistic synthesis at least quadratically reduces the approximation error for any target state $\phi \in S_G$ compared to the worst approximation error caused by the optimal deterministic synthesis. This theorem holds for various S_G by tailoring G. For example, S_G coincides with the set of pure states when $G = \{\mathbb{I}\}$. In such a case, Theorem 1 is applicable to any $\{\hat{\phi}_x\}_x$. When $G = \{\mathbb{I}, \theta\}$, where θ represents the complex conjugation with respect to the computational basis, S_G coincides with $\{\cos t|0\rangle + \sin t|1\rangle : t \in \mathbb{R}\}$, which is generated by an axial rotation $\exp(-it\sigma_Y)$ along a fixed axis from $|0\rangle$. Such one-parameter pure states, more generally known as conjugation-invariant pure states, are often utilized in the optimal parameter estimation [27]. For the last example, S_G coincides with the set of pure states in a subspace $\mathcal{V}$ or its orthogonal complement $\mathcal{V}_\perp$ when $G = \{\mathbb{I}, 2\Pi_{\mathcal{V}} - \mathbb{I}\}$, where $\Pi_{\mathcal{V}}$ is the Hermitian projector whose range is $\mathcal{V}$. Preparing a state in a particular subspace is an extensively used subroutine in various quantum information processing tasks.

The technique used to prove Theorem 1 is also applicable to analyzing the minimum trace distance between a general mixed state ρ and a convex hull of $\{\hat{\phi}_x\}_x$, both of which are invariant under the action of G. For example, we can analyze the entanglement measure by setting G and $\{\hat{\phi}_x\}_{x\in X}$ to be a subset of non-entangling unitary operators and the set of pure product states, respectively. As a byproduct, we analytically compute $\min_{\sigma\in\mathrm{SEP}} \|\rho - \sigma\|_{\mathrm{tr}}$ when ρ is the isotropic state and the Werner state, which coincides with a conjecture numerically found in [28]. Moreover, we provide alternate succinct proof about the following coincidence between the entanglement measure and coherence measure [29].

Proposition 2 *[29, Theorem 3] For pure states $|\Phi\rangle = \sum_{i=0}^{d-1} \alpha_i |ii\rangle$ and $|\phi\rangle = \sum_{i=0}^{d-1} \alpha_i |i\rangle$, it holds that*

$$\min_{\sigma\in\mathbf{SEP}} \|\Phi - \sigma\|_{tr} = \min_{\rho\in I} \|\phi - \rho\|_{tr}, \tag{2}$$

where $I := conv\left(\{|i\rangle\langle i|\}_{i=0}^{d-1}\right)$ is called a set of incoherent states and $\{|i\rangle\}_{i=0}^{d-1}$ is an orthonormal basis.

In the case of unitary synthesis, we obtain the following theorem.

Theorem 3 (simplified version) *[30, Theorem 4.3] For an integer $d \geq 2$ specified below, let Υ and $\{\hat{\Upsilon}_x\}_x$ be a target unitary transformation and a finite set of unitary transformations on $\mathbb{C}^d$, respectively. It then holds that*

$$\frac{4\delta}{d}\left(1 - \frac{\delta}{d}\right) \leq \max_{\Upsilon} \min_{p} \frac{1}{2} \left\|\Upsilon - \sum_x p(x)\hat{\Upsilon}_x\right\|_{\diamond} \leq \epsilon^2$$
$$\text{with } \begin{cases} \delta = 1 - \sqrt{1-\epsilon^2} \quad \text{and} \\ \epsilon = \max_{\Upsilon} \min_x \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_{\diamond}. \end{cases} \tag{3}$$

This theorem provides bounds on the worst approximation error caused when one probabilistically synthesizes the target unitary that is most difficult to approximate. The gap between the upper and lower bounds exists if and only if $d \geq 3$. We can show that the gap is inevitable by constructing $\{\hat{\Upsilon}_x\}_x$ for achieving the upper bound and that for achieving the lower bound for any d and ϵ. That is, Ineq. (3) represents the fundamental relationship of the approximation error between the probabilistic and deterministic approximation that depends only on the dimension d of the system. Moreover, the two theorems reveal the distinction between unitary and state synthesis.

From a computational point of view, we show that the optimal probability distribution for approximating Υ (or ϕ) can be computed by the semidefinite program (SDP) when $\{\hat{\Upsilon}_x\}_x$ (or $\{\hat{\phi}_x\}_x$) realized by using a gate sequence is given. (This set is computable with certain synthesis algorithms.) In addition to its optimality, we can rigorously estimate the worst time complexity of our SDP due to established methods for numerically solving SDPs. As the second main result, we construct a probabilistic synthesis algorithm for single-qubit unitaries:

Theorem 4 (informal version) *[30, Theorem 5.4] For a given gate set, there exists a probabilistic synthesis algorithm for a single-qubit unitary with*

INPUT*: a target single-qubit unitary Υ and target approximation error $\epsilon \in (0,1)$*

OUTPUT*: a gate sequence implementing a single-qubit unitary $\hat{\Upsilon}_x$ sampled from a set $\{\hat{\Upsilon}_x\}_x$ in accordance with probability distribution $\hat{p}(x)$.*

such that the algorithm satisfies the following properties:

- Efficiency*: The algorithm calls a deterministic synthesis algorithm constant times and the whole running time is polylog $\left(\frac{1}{\epsilon}\right)$,*
- Quadratic improvement*: The approximation error $\frac{1}{2}\left\|\Upsilon - \sum_x \hat{p}(x)\hat{\Upsilon}_x\right\|_\diamond$ obtained with this algorithm is upper bounded by ϵ^2, whereas the error $\min_x \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond$ obtained by deterministic synthesis using the unitaries in $\{\hat{\Upsilon}_x\}_x$ is upper bounded by ϵ.*

The first property of the algorithm is desirable for fault-tolerant quantum computation (FTQC) to maintain a polynomial speedup over classical computation. Due to the second property of the algorithm, we can verify that it surpasses current algorithms [16, 20, 21] with respect to the approximation error. The second property guarantees that our algorithm improves the performance of existing deterministic synthesis algorithms with the small additional cost of classically sampling circuits

In the case of state synthesis, we construct a similar algorithm that is applicable to a general target pure state on $\mathbb{C}^d$ and satisfies the same two properties as the unitary synthesis algorithm [26, Theorem 2]. (In the case of state synthesis, the second property is measured by the trace distance.) Moreover, it can be extended into the case when a target pure state is restricted on S_G, defined in Theorem 1. This extension sometimes dramatically improves the runtime, as mentioned in the next section.

Since probabilistic state synthesis reduces the approximation error, it also reduces the size of a circuit to approximately generate a target state for a given approximation error. However, the reduction rate depends on the circuit's construction, e.g., which gate set and synthesis algorithm are used. There is a universal lower bound on the circuit size obtained by regarding a synthesized circuit as a classical description of a pure state. To analyze how probabilistic synthesis reduces this lower bound, we investigate the minimum size of classical memory to store a pure state ϕ so as to approximately reconstruct the original state and obtain the following theorem.

Theorem 5 (simplified version) *[26, Theorem 3] Let n_{det} (or n_{prob}) be the minimum size of memory to encode an arbitrary pure state ϕ on $\mathbb{C}^d$ by assigning a bit string deterministically (or probabilistically) so as to reconstruct a state $\hat{\rho}$ satisfying $\|\phi - \hat{\rho}\|_{tr} \leq \epsilon$. Then, it holds that*

$$\lim_{\epsilon\to 0} \frac{n_{prob}}{n_{det}} = \lim_{d\to\infty} \frac{n_{prob}}{n_{det}} = \frac{1}{2}. \tag{4}$$

3 Technical Outline

In the proof of Theorem 1, we analyze the minimum approximation error

$$\min_p \left\|\rho - \sum_x p(x)\hat{\rho}_x\right\|_{\mathrm{tr}} = \min_p \max_{0\leq M\leq \mathbb{I}} \mathrm{tr}\left[M(\rho - \sum_x p(x)\hat{\rho}_x)\right] \tag{5}$$

for general mixed states ρ and ρ_x, which contains minimax optimization by definition. The first tool for the analysis is the strong duality of semidefinite programming. This enables us to formulate the minimum approximation error as a semidefinite program (SDP). Such reformulation also plays a key role in the proof of Theorem 3. The second tool for the proof is the symmetrization of M in Eq. (5) by exploiting the symmetry of ρ and $\{\hat{\rho}_x\}_x$. This dramatically simplifies the optimization.

Such analysis is formulated as a general lemma [26, Lemma 2] about the optimal convex approximation of a quantum state by using a restricted subset of states. While the optimal convex approximation and state synthesis have been studied in different contexts, our lemma has demonstrated that analyzing the former problem provides not only the fundamental limitation of probabilistic synthesis but also a construction of an efficient synthesis algorithm. Furthermore, our lemma contributes to the original motivation of the studies of the optimal convex approximation [23, 24, 25], which is quantifying a resource measure in convex resource theories [31, 32, 33] such as the resource theory of entanglement, as shown in Proposition 2.

The reformulation of the minimum approximation error as an SDP enables us to efficiently compute the optimal probability distribution to achieve it. By using Theorem 3, we can verify that by solving this SDP with an ϵ-*covering* $\{\hat{\Upsilon}_x\}_{x\in X}$, i.e., $\max_\Upsilon \min_{x\in X} \frac{1}{2}\left\|\Upsilon - \hat{\Upsilon}_x\right\|_\diamond \leq \epsilon$, we obtain a probability distribution $\hat{p}$ that achieves the quadratic reduction of the approximation error. However, the size of this SDP is too large to achieve the efficiency shown in Theorem 4, since the size $|X|$ of the ϵ-covering is $\left(\frac{1}{\epsilon}\right)^{\Omega(1)}$. This problem can be resolved by proving that sampling $\hat{\Upsilon}_x \in \mathcal{B}_{2\epsilon}(\Upsilon)$ is sufficient to optimally approximate Υ, where $\mathcal{B}_{2\epsilon}(\Upsilon)$ is the 2ϵ-ball around Υ. By using a similar argument, we can construct an efficient algorithm for state synthesis. Moreover, when the target state ϕ is restricted on S_G, sampling $\hat{\phi}_x$ from an ϵ-covering of S_G and in $\mathcal{B}_{2\epsilon}(\phi)$ is sufficient to optimally approximate ϕ. For some S_G, the size of the ϵ-covering of S_G is much smaller than that of the set of pure states.

References

[1] B. M. Terhal, "Quantum error correction for quantum memories," *Rev. Mod. Phys.*, vol. 87, pp. 307–346, Apr 2015.

[2] A. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2–30, 2003.

[3] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation.* American Mathematical Society, 2002.

[4] A. W. Harrow, B. Recht, and I. L. Chuang, "Efficient discrete approximations of quantum gates," *Journal of Mathematical Physics*, vol. 43, no. 9, pp. 4445–4451, 2002.

[5] V. Kliuchnikov, D. Maslov, and M. Mosca, "Practical approximation of single-qubit unitaries by single-qubit quantum clifford and t circuits," *IEEE Transactions on Computers*, vol. 65, no. 1, pp. 161–172, 2016.

[6] V. Kliuchnikov, D. Maslov, and M. Mosca, "Asymptotically optimal approximation of single qubit unitaries by clifford and t circuits using a constant number of ancillary qubits," *Phys. Rev. Lett.*, vol. 110, p. 190502, May 2013.

[7] N. J. Ross, "Optimal ancilla-free clifford+v approximation of z-rotations," *Quantum Info. Comput.*, vol. 15, pp. 932–950, sep 2015.

[8] A. Bocharov, M. Roetteler, and K. M. Svore, "Efficient synthesis of universal repeat-until-success quantum circuits," *Phys. Rev. Lett.*, vol. 114, p. 080502, Feb 2015.

[9] A. G. Fowler, "Constructing arbitrary steane code single logical qubit fault-tolerant gates," *Quantum Info. Comput.*, vol. 11, pp. 867–873, sep 2011.

[10] A. Bouland and T. Giurgica-Tiron, "Efficient universal quantum compilation: An inverse-free solovay-kitaev algorithm." arXiv:2112.02040.

[11] N. Mahmud, A. MacGillivray, M. Chaudhary, and E. El-Araby, "Optimizing quantum circuits for arbitrary state synthesis and initialization," in *2021 IEEE 34th International System-on-Chip Conference (SOCC)*, pp. 19–24, 2021.

[12] Z. Zhang, Q. Wang, and M. Ying, "Parallel quantum algorithm for hamiltonian simulation," 2023. arXiv:2105.11889.

[13] X.-M. Zhang, M.-H. Yung, and X. Yuan, "Low-depth quantum state preparation," *Phys. Rev. Res.*, vol. 3, p. 043200, Dec 2021.

[14] K. Fukui, S. Takeda, M. Endo, W. Asavanant, J.-i. Yoshikawa, P. van Loock, and A. Furusawa, "Efficient backcasting search for optical quantum state synthesis," *Phys. Rev. Lett.*, vol. 128, p. 240503, Jun 2022.

[15] S. Ashhab, N. Yamamoto, F. Yoshihara, and K. Semba, "Numerical analysis of quantum circuits for state preparation and unitary operator synthesis," *Phys. Rev. A*, vol. 106, p. 022426, Aug 2022.

[16] M. B. Hastings, "Turning gate synthesis errors into incoherent errors," *Quantum Info. Comput.*, vol. 17, pp. 488–494, mar 2017.

[17] J. J. Wallman and J. Emerson, "Noise tailoring for scalable quantum computation via randomized compiling," *Phys. Rev. A*, vol. 94, p. 052325, Nov 2016.

[18] K. Kuroiwa and Y. O. Nakagawa, "Clifford+t-gate decomposition with limited number of t gates, its error analysis, and performance of unitary coupled cluster ansatz in pre-ftqc era," 2023. arXiv:2301.04150.

[19] 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, "Randomized compiling for scalable quantum computing on a noisy superconducting quantum processor," *Phys. Rev. X*, vol. 11, p. 041039, Nov 2021.

[20] E. Campbell, "Shorter gate sequences for quantum computing by mixing unitaries," *Phys. Rev. A*, vol. 95, p. 042306, Apr 2017.

[21] V. Kliuchnikov, K. Lauter, R. Minko, A. Paetznick, and C. Petit, "Shorter quantum circuits," 2022. arXiv:2203.10064.

[22] M. F. Sacchi and T. Sacchi, "Convex approximations of quantum channels," *Phys. Rev. A*, vol. 96, p. 032311, Sep 2017.

[23] M. F. Sacchi, "Optimal convex approximations of quantum states," *Phys. Rev. A*, vol. 96, p. 042325, Oct 2017.

[24] X.-B. Liang, B. Li, L. Huang, B.-L. Ye, S.-M. Fei, and S.-X. Huang, "Optimal approximations of available states and a triple uncertainty relation," *Phys. Rev. A*, vol. 101, p. 062106, Jun 2020.

[25] L. qiang Zhang, D. hui Yu, and C. shui Yu, "The optimal approximation of qubit states with limited quantum states," *Physics Letters A*, vol. 398, p. 127286, 2021.

[26] S. Akibue, G. Kato, and S. Tani, "Probabilistic state synthesis based on optimal convex approximation," 2023. arXiv:2303.10860.

[27] J. Miyazaki and K. Matsumoto, "Imaginarity-free quantum multiparameter estimation," *Quantum*, vol. 6, p. 665, Mar. 2022.

[28] A. Girardin, N. Brunner, and T. Kriváchy, “Building separable approximations for quantum states via neural networks,” *Phys. Rev. Res.*, vol. 4, p. 023238, Jun 2022.

[29] J. Chen, S. Grogan, N. Johnston, C.-K. Li, and S. Plosker, “Quantifying the coherence of pure quantum states,” *Phys. Rev. A*, vol. 94, p. 042313, Oct 2016.

[30] S. Akibue, G. Kato, and S. Tani, “Probabilistic unitary synthesis with optimal accuracy,” 2023. arXiv:2301.06307.

[31] M. HORODECKI and J. OPPENHEIM, “(quantumness in the context of) resource theories,” *International Journal of Modern Physics B*, vol. 27, no. 01n03, p. 1345019, 2013.

[32] F. G. S. L. Brandão and G. Gour, “Reversible framework for quantum resource theories,” *Phys. Rev. Lett.*, vol. 115, p. 070503, Aug 2015.

[33] E. Chitambar and G. Gour, “Quantum resource theories,” *Reviews of modern physics*, vol. 91, no. 2, p. 025001, 2019.

Next Generation Quantum Reservoir Computing: An Efficient Quantum Algorithm for Forecasting Quantum Dynamics

Apimuk Sornsaeng[1] * Ninnat Dangniam[2] † Thiparat Chotibut[1] ‡

[1] *Chula Intelligent and Complex Systems, Department of Physics, Faculty of Science, Chulalongkorn University, Bangkok, Thailand, 10330*

[2] *The Institute for Fundamental Study, Naresuan University, Phitsanulok, Thailand, 65000*

Abstract. Next Generation Reservoir Computing (NG-RC) is a modern class of model-free machine learning that enables an accurate forecasting of time series data generated by dynamical systems. However, adopting a classical NG-RC for many-body quantum dynamics prediction is computationally prohibitive due to the large Hilbert space of sample input data. In this work, we propose an end-to-end quantum algorithm for many-body quantum dynamics forecasting with a quantum computational speedup via the block-encoding technique. This proposal presents an efficient model-free fast-forwarding scheme to forecast quantum dynamics coherently, bypassing inductive biases incurred in a model-based approach.

Keywords: Quantum reservoir computing, Forecasting quantum dynamics

1 Motivation and Introduction

Learning quantum dynamics presents a fundamental challenge in quantum physics, to which numerous machine learning techniques have been applied [1, and references therein]. Simultaneously, quantum systems themselves can be harnessed as computational resources. However, quantum algorithms that work on a large batch of classical data typically requires extreme assumptions about data loading and readout, which remain points of contention, see for example [2, 3, 4].

The two challenges spur the proposals of quantum algorithms that can learn directly from quantum data [5, 6]. For example, there are quantum machine learning algorithms that aim to fast-forward the dynamics, that is, to obtain the learned model that evolves faster than the natural dynamics [7]. Our work adds to the literature by providing a novel quantum algorithm for learning from quantum data based on NG-RC. Importantly, this model-free approach of reservoir computing requires only time-series of quantum data, without assuming a quantum dynamical ansatz *a priori*. Such approach circumvents inductive biases that could arise in learning complex quantum dynamics.

Reservoir computing (RC). RC is a computational paradigm in machine learning that harnesses a recurrent neural network (RNN) to learn time-series data, such as the states of a dynamical system. Even when the dynamics is complicated or chaotic, well-tuned RC can accurately forecast the future states of such dynamics up to many Lyapunov times [8, 9].

Importanty, RC bypasses the training of an RNN by utilizing a fixed, randomly initialized RNN, called a *reservoir*, consists of a large number L of hidden neurons. Suppose that an input data $\boldsymbol{s}_i$ is fed into the reservoir, where i could be a label for the time step of some dynamical process. The data is represented as a feature vector $\boldsymbol{x}_i = f(\boldsymbol{s}_i) \in \mathbb{R}^L$, where f is typically a highly non-linear function that represents the dynamics of the reservoir. The feature vector is then linearly transformed in the final, trainable output layer into a prediction vector $\hat{\boldsymbol{y}}_i = W\boldsymbol{x}_i$. In particular, the input $\boldsymbol{s}_i$ at the ith time step and the output at the previous time step $\boldsymbol{y}_{i-1}$ may be fed together as inputs into the reservoir to train the feature vector $\boldsymbol{x}_i$ for the next time step , i.e., $\boldsymbol{x}_i = f(\boldsymbol{s}_i, \boldsymbol{y}_{i-1})$.

To prevent overfitting in a supervised learning via RC, the weight matrix W is obtained by optimizing $\hat{\boldsymbol{y}}_i$ in the least-square sense with respect to the desired target, which can be done via the Tikhonov regularization

$$W = YX^T(XX^T + \lambda I)^{-1}. \tag{1}$$

Here $Y = (\boldsymbol{y}_0|\cdots|\boldsymbol{y}_{T-1}) \in \mathbb{R}^{D\times T}$ is the target matrix, $X = (\boldsymbol{x}_0|\cdots|\boldsymbol{x}_{T-1}) \in \mathbb{R}^{L\times T}$ is the feature matrix, and $\lambda \geq 0$ is the regularization parameter.

Versions of quantum RC have been proposed across various scenarios, encompassing cases where the input data is quantum, where the reservoir (or "subtrate") is quantum, or scenarios where both the input data and the reservoir are quantum [10, 11].

NG-RC. Despite the fast training protocol offered by RC, the random initialization of the reservoir presents its own problem: there are overwhelmingly large number of hyperparameters to be optimized and there is no consensus on how to pick an optimal reservoir.

NG-RC is an alternative approach that takes advantage of the discovery that RC can be equivalently performed using a *linear* reservoir and a *nonlinear* trainable output layer [12, 13]. The latter is in turn equivalent to a nonlinear vector autoregression (NVAR) machine [14]. An NVAR machine predicts future observations of a time series using past observations. In particular, the underlying state space of the dynamics to be learned can be reconstructed using linear and nonlinear functions of past observations. Inspired by this correspondence, NG-RC [15] proposes taking k time-delay data $\boldsymbol{o}_i = \boldsymbol{s}_i \oplus \boldsymbol{s}_{i-\Delta} \oplus \boldsymbol{s}_{i-2\Delta} \oplus \ldots \oplus \boldsymbol{s}_{i-(k-1)\Delta}$ with step size Δ as an input, and, forgoing the need for a reservoir,

*apimuk25@hotmail.com

†ninnatdn@gmail.com

‡thiparatc@gmail.com

directly constructing a feature vector,

$$\boldsymbol{x}_i = \boldsymbol{o}_i \oplus (\boldsymbol{o}_i)^{\otimes p}, \tag{2}$$

whose nonlinearity arises from a degree-p monomial of the k-delay data, Δ, k, and p being the NG-RC hyperparameters. The feature vector is then optimized via a linear least-square regularization to predict the target dynamics.

NG-RC with $\Delta = 1$, $k = 2$, and $p = 2$ has been used to efficiently predict the dynamics of the Lorenz attractor using only small data sets [15]. Here the feature vector is a $(4D^2 + 2D)$-dimensional vector of the form

$$\boldsymbol{x}_i = \boldsymbol{s}_i \oplus \boldsymbol{s}_{i-1} \oplus [(\boldsymbol{s}_i \oplus \boldsymbol{s}_{i-1}) \otimes (\boldsymbol{s}_i \oplus \boldsymbol{s}_{i-1})]. \tag{3}$$

The primary computational bottleneck arises from matrix inversion in Eq. (1). The Tikhonov-regularized least squares requires $O(MN^2)$ matrix operations, where N and M are, respectively, the larger and the smaller dimension of the matrix. This complexity is computationally prohibitive when the data matrix is collected from many-body quantum states, since the dimension of each column vector scales exponentially with the system size. In the following section, we present an end-to-end, next generation quantum reservoir computing algorithm that does not suffer from the exponential complexity of the classical counterpart. Our quantum algorithm takes as inputs quantum data and outputs the predicted future quantum states.

2 Method and Result

Block encoding. We employ the block-encoding technique to construct the non-linear feature matrix and perform the Tikhonov regularization. Specifically, a relevant matrix A is embedded as a submatrix of a unitary gate U such that

$$\left(\langle 0|^{\otimes a} \otimes I^{\otimes s}\right) U \left(|0\rangle^{\otimes a} \otimes I^{\otimes s}\right) = A/\alpha \text{ for } U = \begin{pmatrix} A & * \\ * & * \end{pmatrix}, \tag{4}$$

where $|0\rangle^{\otimes a}$ is the fiducial state of an a-qubit ancillary system, and $\alpha \geq \|A\|$ due to the unitary constraint. Once the block-encoded matrix is in place, the quantum singular value transform (QSVT) allows us to construct a degree-q polynomial approximation of essentially any well-behaved function of the singular values of A, using the number of gates U and controlled operations that are efficient in q [16, 17]. This approach enables straightforward creation of matrix polynomials, and in particular, creation of the Moore-Penrose pseudo-inverse by inverting the singular values of A.

Input assumptions. The forecasting of quantum dynamics via NG-QRC is divided into two phases: the training phase and the prediction phase. The data are assumed to be given by the oracles

$$\begin{aligned} \mathcal{O}_0 : \quad & |0\rangle^{\otimes d} |k\rangle \mapsto |s_k\rangle |k\rangle, && (5)\\ \mathcal{O}_{-1} : \quad & |0\rangle^{\otimes d} |k\rangle \mapsto |s_{k-1}\rangle |k\rangle, && (6)\\ \tilde{\mathcal{O}} : \quad & \left|\tilde{k}\right\rangle |0\rangle^{\otimes d} \mapsto \left|\tilde{k}\right\rangle \left|\tilde{s}_{-\tilde{k}}\right\rangle, && (7) \end{aligned}$$

where $|s_i\rangle$ and $|\tilde{s}_i\rangle$ are d-qubit input data in the training phase and in the prediction phase respectively, and $k = 0, \ldots, T-1$ and $\tilde{k} = 0, 1$. This assumption is equivalent to an access to the controlled version of a unitary that generates each data point, a common assumption in the block-encoding literature. In particular, we must be able to create coherent superpositions of the form $(|0\rangle |s_0\rangle + |1\rangle |s_{-1}\rangle)/\sqrt{2}$. To express the total query complexity of the algorithm, we denote the number of calls to the oracle $\mathcal{O}$ (resp. $\tilde{\mathcal{O}}$) by $T_{\mathcal{O}}$ (resp. $T_{\tilde{\mathcal{O}}}$).

(In practice, we may only handed the data $|\tilde{s}_i\rangle$ in the prediction phase, in which case we can create coherent superpositions of such data by consuming multiple copies of each data point [18]. However, the complexity of the algorithm depends on the unknown overlap between the data points we wish to superpose.)

Training phase According to the NG-RC procedure, by utilizing regularized linear optimization, we obtain an optimal weight matrix through the regularized pseudoinverse of the feature matrix X (cf. Eq. (1)). Therefore, we begin by efficiently constructing the feature matrix X. Assuming the existence of oracles $\mathcal{O}_0$ and $\mathcal{O}_{-1}$, we can generate the linear component of the feature vector, denoted as $|o_k\rangle$, by introducing an additional qubit to entangle with these oracles. This process is depicted by the operator U^{lin}, which maps the state $|0\rangle^{\otimes d+1} |k\rangle$ to $|o_k\rangle |k\rangle$ as shown in Fig. 1 (green box). To incorporate the nonlinear component $|o_k\rangle \otimes |o_k\rangle$ into the feature vector, we must apply the operator U^{lin} twice due to the constraints imposed by the no-cloning theorem. The resulting feature vector $|x_k\rangle$ (Eq. (3)) is represented by the quantum circuit U^f, which maps the state $|0\rangle^{2d+3} |k\rangle$ to $|x_k\rangle |k\rangle$, as also illustrated in Fig. 1 (orange box). We can coherently construct the feature matrix X with the block-encoding technology using quantum gates

$$\left(I^{\otimes \max(0, 2d+3-t)} \otimes H^{\otimes t} \otimes I^{\otimes \max(2d+3, t)}\right) \cdot \mathrm{SWAP} \cdot U^f, \tag{8}$$

which is the $(\sqrt{T}, \max(2d+3, t), 0)$-block-encoding of the feature matrix X, illustrated in Fig. 1 (yellow box). According to the regularized linear optimization, we can find the optimal weight due to Eq. (1). We can construct the weight matrix W with block-encoding stated in Theorem 1.

Theorem 1 *Let $\delta_X, \delta_Y \in (0, 1]$. Suppose we have the $(\sqrt{T}, \max(2d+3, t), 0)$-block-encoding of the feature matrix X and the $(\sqrt{2}\|Y\|, \max(2d+3, t)+1, \delta_Y)$-block-encoding of training target matrix Y. Let κ_X be the conditional number of X, and*

$$\kappa = \kappa_X \sqrt{\frac{\|X\|^2 + \lambda}{\|X\|^2 + \lambda \kappa_X^2}}. \tag{9}$$

where $\lambda \geq 0$ is the regularization parameter [19].

Define $w = 2\max(2d+3, t) + 2$ is ancilla qubits used in the block-encoding of W, we can construct the $\left(\frac{2\sqrt{2}\|Y\|\kappa}{\|X\|+\sqrt{\lambda}}, w, \sqrt{2}\|Y\|\delta_X + \frac{2\kappa\delta_Y}{\|X\|+\sqrt{\lambda}}\right)$-block-encoding of the weight matrix W in

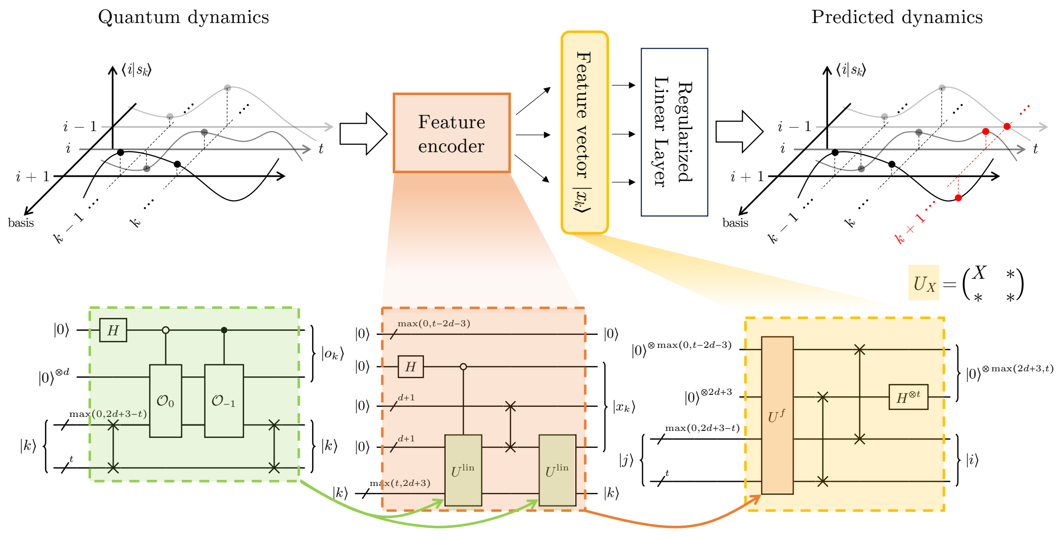


Figure 1: (Top) The schematic of the NG-QRC algorithm. The initial part of the algorithm encodes the history of quantum dynamics collected from a time series into nonlinear feature vectors, which is then processed by a regularized linear layer (whose weight matrix W training follows from Theorem 1) to predict future quantum states. (Bottom) The quantum circuits for encoding nonlinear feature vectors via the block-encoding protocol.

$T_W = O\Big(\Big(\frac{\kappa}{\|X\|+\sqrt{\lambda}}\log\Big(\frac{\kappa}{\delta_X}\Big)+\frac{1}{\|Y\|}\log\Big(\frac{\|Y\|}{\delta_Y}\Big)\Big)T_{\mathcal{O}}\sqrt{T}\Big)$ *queries, where $T_{\mathcal{O}}$ is the number of calls to the oracles.*

Prediction phase By assuming $\tilde{\mathcal{O}}$, we can create two initial states, $\{|\tilde{s}_{-1}\rangle, |\tilde{s}_0\rangle\}$. To construct the linear part of the initial feature vector $|\tilde{o}_0\rangle$, we apply a Hadamard gate to an ancilla qubit of $\tilde{\mathcal{O}}$. The feature vector $|\tilde{x}_0\rangle$ is then obtained by duplicating the operation that creates $|\tilde{o}_0\rangle$. Finally, this state is multiplied algebraically by the weight matrix W from Theorem 1 to obtain the first predicted state $|\tilde{s}_1\rangle \sim W|\tilde{x}_0\rangle/\|W|\tilde{x}_0\rangle\|$. Note that the block-encoded matrix W is a square matrix with dimension $\max(T, 4D^2+2D)\times\max(T,4D^2+2D)$ including zero padding. The resulting state $|0\rangle^{\otimes \max(t-d,d+3)}|\tilde{s}_1\rangle$ consists of $\max(t-d,d+3)$ qubits in a fiducial state that results from the zero padding, whereas $|\tilde{s}_1\rangle$ is the prediction for the quantum state after the first time step.

The query complexity of a W-operation has the cost of $O\Big(\frac{\kappa\kappa_W}{\|X\|+\sqrt{\lambda}}\frac{\|Y\|}{\|W\|}(T_W+T_{\tilde{\mathcal{O}}})\Big)\sim O(T_{\mathcal{O}}\sqrt{T}+T_{\tilde{\mathcal{O}}})$, where κ_W is the condition number of W. From W and the initial states $\{|\tilde{s}_{-1}\rangle, |\tilde{s}_0\rangle\}$, we can construct the operator

$$\tilde{U}_1 : |k\rangle\,|0\rangle^{\otimes(w'+2d+3)} \mapsto |0\rangle^{\otimes(w'+d+3)}\,|k\rangle\,|\tilde{s}_{1-k}\rangle\,, \quad (10)$$

where $k=0$ or 1 and $w' = w+\max(0,t-2d-3)$ is the number of ancilla qubits used in the block encoding of W and w is defined in Theorem 1.

To predict the next k time steps, the operator $\tilde{U}_k$ will recursively contain $\tilde{U}_{k-1}$ as a subroutine, and the register that contained $|\tilde{s}_1\rangle$ will now contain $|\tilde{s}_k\rangle$.

3 Conclusion

Drawing inspiration from the paradigm of NG-RC, we develop a novel, end-to-end quantum algorithm for predicting many-body quantum dynamics. The algorithm is purely data-driven, only requiring a time-series of quantum data and no assumption is made on the nature of the dynamics i.e. the forecasting is *model-free*. The algorithm employs block encoding to efficiently handle matrix algebra subroutines, including matrix multiplication, inversion, and regularized linear optimization. In addition to providing a quantum computational speedup and avoiding exponential resource consumption in storing many-body quantum states classically, our algorithm coherently processes and generates quantum data, thereby circumventing classical-quantum data conversion problems during encoding and readout procedures.

References

[1] Hsin-Yuan Huang, Sitan Chen, and John Preskill, *Learning to predict arbitrary quantum processes*, arXiv:2210.14894.

[2] Scott Aaronson, *Read the fine print*, Nature Physics **11**, 291–293 (2015).

[3] Ewin Tang, *A quantum-inspired classical algorithm for recommendation systems*, STOC 2019: Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, 217-228 (2019).

[4] Olivia Di Matteo, Vlad Gheorghiu, Michele Mosca, *Fault tolerant resource estimation of quantum*

random-access memories, IEEE Transactions on Quantum Engineering, **1**, 1-13 (2020).

[5] Louis Schatzki, Andrew Arrasmith, Patrick J. Coles, M. Cerezo, *Entangled Datasets for Quantum Machine Learning*, arXiv:2109.03400.

[6] Hsin-Yuan Huang *et al.*, *Quantum advantage in learning from experiments*, Science, 376(6598), 1182-1186 (2022).

[7] Cristina Cirstoiu, Zoe Holmes, Joseph Iosue, Lukasz Cincio, Patrick J. Coles, Andrew Sornborger, *Variational Fast Forwarding for Quantum Simulation Beyond the Coherence Time*, npj Quantum Information 6, 82 (2020).

[8] Jaideep Pathak, Zhixin Lu, Brian R. Hunt, Michelle Girvan, and Edward Ott, *Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data*, Chaos, **27**, 121102 (2017).

[9] Jaideep Pathak, Brian R. Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott, *Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach*, Physical Review Letters, **120**, 24102 (2018).

[10] Mujal, Pere, Rodrigo Martínez-Peña, Johannes Nokkala, Jorge García-Beni, Gian Luca Giorgi, Miguel C. Soriano, and Roberta Zambrini. *Opportunities in quantum reservoir computing and extreme learning machines*, Advanced Quantum Technologies 4, no. 8 (2021): 2100027.

[11] Fujii, Keisuke, and Kohei Nakajima. *Quantum reservoir computing: a reservoir approach toward quantum machine learning on near-term quantum devices*, Reservoir Computing: Theory, Physical Implementations, and Applications (2021): 423-450.

[12] Lukas Gono, and Juan-Pablo Ortega. *Reservoir computing universality with stochastic inputs* IEEE transactions on neural networks and learning systems 31(1), 100-112 (2019).

[13] Allen G. Hart, James L. Hook, Jonathan H. P. Dawes, *Echo State Networks trained by Tikhonov least squares are $L^2(\mu)$ approximators of ergodic dynamical systems*, Physica D: Nonlinear Phenomena, 421, 132882 (2021).

[14] Erik Bollt, *On explaining the surprising success of reservoir computing forecaster of chaos? The universal machine learning dynamical system with contrast to VAR and DMD*, Chaos 31, 013108 (2021).

[15] Daniel J. Gauthier, Erik Bollt, Aaron Griffith, and Wendson AS Barbosa, *Next generation reservoir computing*, Nature communications, **12**, 5564 (2021).

[16] András Gilyén, Yuan Su, Guang Hao Low, Nathan Wiebe, *Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics*, STOC 2019: Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, 193–204 (2019).

[17] John M. Martyn, Zane M. Rossi, Andrew K. Tan, and Isaac L. Chuang. *Grand unification of quantum algorithms.* PRX Quantum 2, no. 4 (2021): 040203.

[18] Shelby Kimmel, Cedric Yen-Yu Lin, Guang Hao Low, Maris Ozols, and Theodore J. Yoder, *Hamiltonian Simulation with Optimal Sample Complexity*, npj Quantum Information 3:13 (2017).

[19] Shantanav Chakraborty, Aditya Morolia, and Anurudh Peduri.*Quantum regularized least squares*, Quantum 7 (2023): 988.

Analysis of the effects of the two-photon temporal distinguishability on measurement-device-independent quantum key distribution

Haobo GE[1] * Akihisa TOMITA[1] † Atsushi OKAMOTO[1] ‡ Kazuhisa OGAWA[2] §

[1] *Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan*
[2] *Center for Quantum Information and Quantum Biology, Osaka University, Osaka 565-0871, Japan*

Abstract. Measurement-device-independent quantum key distribution (MDI-QKD) can remove all loopholes in the measurement devices of QKD. The arrival times of the pulses send by Alice and Bob fluctuate independently. According to the Hong-Ou-Mandel (HOM) interference occurring at Charlie's relay, time delay between two photons has the greatest effect on the distinguishability. We studied the effects of the two-photon temporal distinguishability in terms of the visibility of the HOM-dip of two photons on the final key rate of MDI-QKD with different Bell state measurement (BSM) implements. The acceptable time delay range is also estimated based on photons with Gaussian spectral amplitude functions. This study has been published in IEEE Transactions on Quantum Engineering (DOI: 10.1109/TQE.2023.3259043).

Keywords: MDI-QKD, Bell state measurement, decoy method, Hong-Ou-Mandel interference, time delay.

1 Introduction

Quantum key distribution (QKD) ensures theoretical unconditional security. However, the actual QKD system may have various security loopholes due to the inevitable errors and defects of the equipment.

For security holes caused by device imperfection, Acin et al. [1] proposed device-independent QKD (DI-QKD). It is possible to prove the unconditional security of a QKD system. However, this scheme is difficult to achieve with the current experimental technology, and its key rate is relatively low.

Lo et al. [2] proposed the measurement-device independent QKD (MDI-QKD). This scheme allows Alice and Bob to send single photon pulses to a third untrusted party Charlie for BSM. It can be immune to any detector side channel attacks. Alice and Bob can distill the secret key from public information as long as they ensure that their sources are secret and provide near-perfect state preparation.

According to Lo et al. [2], it is critical for the photons emitted by two independent lasers to be indistinguishable. Since MDI-QKD protocol is based on the photon bunching effect of two indistinguishable photons at a 50:50 beam splitter (BS), stable HOM interference [3, 4] should be observed. However, it is unclear how the imperfect HOM interference affects the security of a practical system. The relationship between the visibility of the HOM interference and the final key rate must be clarified, and methods that improve visibility must be established. So far, few research have explored this issue, with exceptions including the study by Curty et al. [5], which calculated only the effect of misalignment error in the limit of zero distance.

We explored the acceptable indistinguishability of the MDI-QKD. We used a three-intensity decoy-state MDI-QKD protocol to calculate the effect of the visibility of the two-photon interference on the key generation rate with different BSM implements. Then, we calculated the acceptable time delay of the two Gaussian pulses at a 50:50 BS.

2 Analysis of the effects of the two-photon temporal distinguishability

We consider a symmetric MDI-QKD protocol with three intensities as a photon-number channel model when the phase of pulses is fully randomized. The final key rate can be written as [5-9]

$$\begin{aligned} R \geq & p_{\mu_2} p_{\nu_2} p^z_{\mu_2} p^z_{\nu_2} \{ \mu_2 \nu_2 e^{-\mu_2 - \nu_2} s^z_{11} [1 - H(e^x_{11})] \\ & - S^z_{\mu_2 \nu_2} f H(E^z_{\mu_2 \nu_2}) \} \end{aligned} \tag{1}$$

The overall counting rate and error rate on the x basis and z basis are shown as [7, 8]

$$\begin{aligned} S^x_{\mu_i \nu_j} =& 2y^2 [1 + 2y^2 - 4y I_0(s) + I_0(s)] \\ S^x_{\mu_i \nu_j} E^x_{\mu_i \nu_j} =& e_0 S^x_{\mu_i \nu_j} - 2(e_0 - e_d) y^2 [I_0(s) - 1] \\ S^z_{\mu_i \nu_j} =& S_C + S_E \\ S^x_{\mu_i \nu_j} E^x_{\mu_i \nu_j} =& e_d S_C + (1 - e_d) S_E \end{aligned} \tag{2}$$

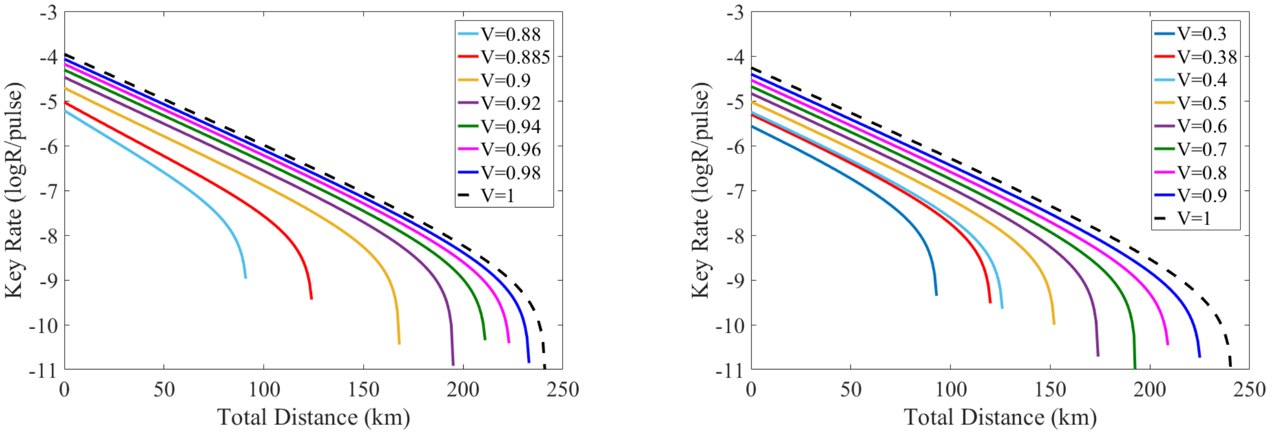
where $I_0(s)$ is the modified Bessel function of the first kind, pd is the dark count rate of the photon detector, e_0 is the error rate of the background, and e_d is the error rate due to two-photon distinguishability.

Then, we focus on the error rate ed, which is directly related to the visibility of the two-photon interference, and it can be written as

$$e_d = e^0_d + \frac{1 - V}{2} \tag{3}$$

where e^0_d is the correction parameter and is assumed to be 0. Note that different BSM implementations adopted by different protocols should cause the error rates for x-

*katsu@optnet.ist.hokudai.ac.jp
†tomita@ist.hokudai.ac.jp
‡ao@optnet.ist.hokudai.ac.jp
§k-ogawa@qiqb.osaka-u.ac.jp



(a) Complete BSM.

(b) BS+PBS BSM.

Figure 1: Key rate with different visibilities of infinite sized MDI-QKD protocol with a complete BSM(a) and BS+PBS BSM (b).

and z-basis are also different. According to the results of BSM, we divide it into three categories, namely complete BSM, BSM with a beam splitter followed by polarization beam splitters (BS+PBSs), and BSM with only a BS. Their success probabilities are 1, 1/2 and 1/4, respectively. For complete BSM and BSM with only a BS, both x- and z-basis will have errors with the probability of 1/2. However, for BSM with BS+PBSs, there are errors with the probability of 1/2 in only x-basis but no bit error in z-basis. In this case, should have no effect on z-basis in the error rate calculation in Equation (2). The visibility of the two-photon interference V can be directly estimated from the coincidence probability in the HOM interference experiment by

$$V = \frac{p_{\max} - p_{\min}}{p_{\max}} \tag{4}$$

where $p_{\max}$ and $p_{\min}$ are the maximum and minimum coincidence probabilities, respectively.

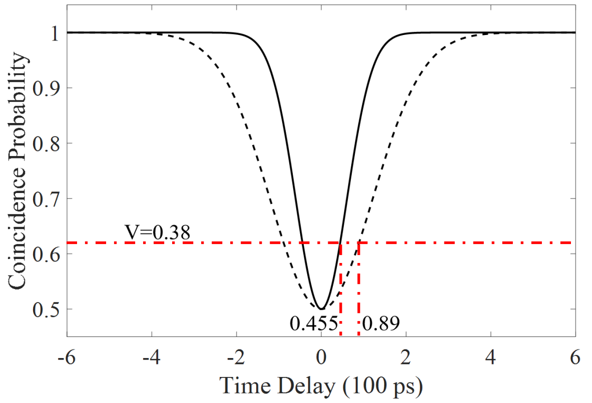


Figure 2: HOM-dip of 100 ps (black solid line) and 200 ps (black dotted line) time duration. The red dotted line of V=0.38 represents the position with the minimum coincidence probability of 0.62.

3 Simulation results

Considering the different success rates of different BSM methods, we need to multiply the key rate R in Equation (1) by a coefficient, which is 1 for complete BSM, 1/2 for BS+PBS and 1/4 for BS-only. We verify the difference between the effects of indistinguishability of complete BSM and BS+PBS BSM methods as shown in Fig.1 with the parameters given in Table 1.

The key rate of complete BSM is highest when $V = 1$, but when V is near 0.9, the key rate becomes lower than the BS+PBS method. We can also clearly see that the BS+BPS method has much higher tolerance for indistinguishability.

Due to the lack of an evaluation criterion, we tentatively decided on the definition of acceptable visibility range. First, we define the maximum communication distance where the key rate falls into zero in our simulation. Then, we define the acceptable visibility which provide the maximum communication distance more than the half of that calculated for the ideal situation ($V = 1$). The minimum visibility is 0.38 for successful infinite-sized key generation.

With the acceptable visibility we can also calculate the acceptable time delay between the two photons from Alice and Bob. We considered two Gaussian photon pulses, we can calculate the coincidence probability, p as follow

Table 1: Parameters for Simulation of MDI-QKD

Parameter	symbol	Quantity
error correction inefficiency	f	1.16
loss coefficient of fiber	α	0.2 dB/km
dark count rate	p_d	3×10^{-6}
error rate of background	e_0	0.5
detection efficiency	η_d	%14.5

$$p = 1 - \frac{1}{2} e^{-\frac{(2\ln 2)\tau^2}{\tau_L^2}} \tag{5}$$

where the time delay of Alice's and Bob's photon pulses is τ and the time duration of the photon pulse is assumed to be τ_L.

In the following, we fix the time duration to 100 and 200 ps. Because of the different key rates obtained by different decoy state calculation methods, we choose the result of the most efficient infinite-sized protocol. The HOM dips are shown in Fig.2. They show that the acceptable time delay is 45.5 ps for 100-ps width and 89.0 ps for 200-ps width.

4 Conclusion

We analyzed the decoy state MDI-QKD protocol, which allows the receiver to be protected from attacks on the measurement device. For the implementation of this protocol, the photons generated by the two independent laser sources must be indistinguishable. We calculated the final key rate determine the effects of two-photon distinguishability on the visibility of their interference. We also estimated an acceptable time delay between two photons from two independent pulse lasers.

It should be noted that the calculation results we obtained are based on the three-intensity model. It was suggested in four-intensity model [9, 10] will improve the performance for smaller number of pulses. Since the small data size is very important for practical QKD application, we should explore the improvement of the estimation with decoy method in the future. Fortunately, our conclusions are based on HOM interference, so this method is applicable to any quantum communication model that depends on two-photon interference.

This study provides quantitative conditions for timing-control accuracy, which will play an important role in improving the performance of practical MDI-QKD systems. Because synchronization is crucial to achieving high visibility of two-photon interference, we still need to improve the method to measure and control the relative time difference between photons from remote sources.

References

[1] 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.

[2] H.-K. Lo, M. Curty, and B. Qi. Measurement-Device-Independent Quantum Key Distribution. *Phys.Rev.Lett.* 108, 130503, 2012.

[3] C. K. Hong, Z. Y. Ou and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Phys.Rev.Lett.* 59, 2044, 1987.

[4] A. M. Brańczyk. Hong-Ou-Mandel Interference. arXiv: 1711. 00080v1, 2017.

[5] M. Curty, F. Xu, W. Cui, C. C. W. Lim, K. Tamaki and H.-K. Lo. Finite-key analysis for measurement-device-independent quantum key distribution. *Nat Commun.* 5, 3732, 2014.

[6] X. Ma, and M. Razavi. Alternative schemes for measurement-device-independent quantum key distribution. *Phys.Rev.A.* 86, 062319, 2012.

[7] X. Ma, C.-H. F. Fung, and M. Razavi. Statistical fluctuation analysis for measurement-device-independent quantum key distribution. *Phys.Rev.A.* 86, 052305, 2012.

[8] S. H. Sun, M. Gao, C. Y. Li, and L. M. Liang. Practical decoy-state measurement-device-independent quantum key distribution. *Phys.Rev.A.* 87, 052329, 2013.

[9] X-L. Hu, C. Jiang, Z-W. Yu, and X-B. Wang. Practical Long-Distance Measurement-Device-Independent Quantum Key Distribution By Four-Intensity Protocol. *Adv. Quantum Technol.* 4, 2100069, 2021.

[10] Y. Zhou, Z. Yu and X. Wang. Making the decoy-state measurement-device-independent quantum key distribution practically useful. *Phys.Rev.A.* 93, 042324, 2016.

Extended Abstract : Certifying beyond quantumness of locally quantum no-signaling theories through a quantum-input Bell test

Edwin Peter Lobo,[1] Sahil Gopalkrishna Naik,[1] Samrat Sen,[1]
Ram Krishna Patra,[1] Manik Banik,[1] and Mir Alimuddin[1,2]

[1]*School of Physics, IISER Thiruvananthapuram, Vithura, Kerala 695551, India.*
[2]*Department of Physics of Complex Systems,*
S.N. Bose National Center for Basic Sciences,
Block JD, Sector III, Salt Lake, Kolkata 700106, India.

Abstract : Physical theories constrained with local quantum structure and satisfying the no-signaling principle can allow beyond-quantum global states. In a standard Bell experiment, correlations obtained from any such beyondquantum bipartite state can always be reproduced by quantum states and measurements, suggesting the local quantum structure and no-signaling to be the axioms to isolate quantum correlations. In this Letter, however, we show that if the Bell experiment is generalized to allow local quantum inputs, then beyond-quantum correlations can be generated by every beyond-quantum state. This gives us a way to certify the beyond quantumness of locally quantum no-signaling theories and in turn suggests the requirement of additional information principles along with the local quantum structure and no-signaling principle to isolate quantum correlations. More importantly, our work establishes that the additional principle(s) must be sensitive to the quantum signature of local inputs. We also generalize our results to multipartite locally quantum no-signaling theories and further analyze some interesting implications.

Keywords : Quantum Correlations, Prover-Verfier, Composition of Local Quantum Systems

Reference : Phys. Rev. A **106**, L040201

Introduction: Correlations among distant events established through the violation of Bell type inequalities confirm nonlocal behavior of the physical world [1–4]. Nonseparable multipartite quantum states yielding such correlations, in Schrödinger's words, are "...the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought" [5]. The advent of quantum information science identifies the power of such nonlocal correlations in numerous device independent protocols – cryptographic key distribution [6], randomness certification [7] and amplification [8], dimension witness [9] are few canonical examples. Cirel'son's result [10], however, establishes that the nonlocal strength of quantum correlations is limited compared to the general *no-signaling* (NS) ones [11] as depicted in the celebrated Clauser-Horne-Shimony-Holt (CHSH) inequality violation [12].

In a recent work, Barnum *et al.* have shown that the set of bipartite correlations attainable from the POPT states is precisely the set of quantum correlations [18]. Consequently, their result provokes a far-reaching conclusion "... that if nonlocal correlations beyond quantum mechanics are obtained in any experiment then quantum theory would be invalidated even locally." In this work we analyze the correlations of multipartite POPT states obtained from local measurements performed on their constituent parts by considering a generalized Bell scenario as introduced in [25]. While in the standard Bell scenario spatially separated parties receive some classical inputs and accordingly generate some classical outputs by performing local measurements on their respective parts of some composite system, recently Buscemi has generalized the scenario where the parties receive quantum inputs instead of classical variables [25]. Considering this generalized scenario, here we show that not all correlations obtained from bipartite POPT states are quantum simulable. In fact, every beyond quantum POPT state produces some beyond quantum correlations in some quantum input game. On the other hand, to illustrate the limitations of the standard Bell scenario, we show that there are POPT states which produce classical-input-classical-output correlations that are not only quantum simulable, rather simulable classically. Our result shows that the *strong* claim made by the authors in [18] will not be correct anymore in this generalized Bell scenario which is allowed within the framework of local quantum theory. From a foundational perspective our study welcomes new information principles incorporating this generalized Bell type scenario to isolate quantum correlation from beyond-quantum ones. We also analyze the implication of this generalized scenario for multipartite correlations and answer an open question raised in [19].

Gleason's theorem and origin of POPT: We investigate the class of locally quantum theories studied in a series of works in the recent past [14–21]. For simplicity, we consider two parties, Alice and Bob. A more technical description for multiple parties is provided in the technical manuscript. A bipartite state is shared between Alice and Bob such that the elementary system possessed by each party has a valid quantum description. However, we do not assume that the global state is quantum. Unentangled Gleason's theorem tells us that if Alice and Bob are restricted to performing local measurements, then global states that are beyond quantum are mathematically allowed [23]. These states, denoted by W, are called POPT (positive on pure tensors) states. Mathematically, W is a Hermitian, unit trace operator on the tensor product Hilbert space of Alice and Bob. We will denote the set of POPT states as $\mathcal{W}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and the set of quantum states (density operators) by $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. $\mathcal{W}(\mathcal{H}_A \otimes \mathcal{H}_B)$ includes $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ as a proper subset. A W will be called 'beyond quantum state' (BQS) whenever $W \in \mathcal{W}(\mathcal{H}_A \otimes \mathcal{H}_B)$ but $W \notin \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. In this work, our aim is to study the correlations obtained from BQSs. But before going to our main results we briefly recall the standard input-output scenario for the Bell correlation experiment and its generalization by Buscemi.

Standard and Generalized Bell scenario: A two party Standard Bell scenario considers two distant parties Alice and Bob who receive independent random classical inputs x and y, respectively, from a Referee. Each party then produces a classical output a and b, respectively, based on which the Referee yields some payoff $\mathscr{P} : (a, b, x, y) \mapsto \mathbb{R}$. An implicit rule is that the players cannot communicate with one another once the game starts, although they can agree upon some pre-shared strategy. Each payoff function $\mathscr{P}$ describes a game in the Bell Scenario.

The Generalized Bell scenario was introduced by Buscemi to establish the nonlocal behaviour of all entangled quantum states [25]. While in the standard Bell scenario the distant parties are given classical inputs, here they are given some quantum states as the inputs. So, Alice and Bob receive independent random *quantum* inputs ψ^x and ψ^y, respectively, from a Referee. Just like in the standard Bell scenario, they are asked to produce classical outputs a and b, respectively, based on which the Referee yields some payoff $\beta : (a, b, x, y) \mapsto \mathbb{R}$.

The players can share some quantum state or BQS (say, Z_{AB}) and by making appropriate joint measurements on their respective parts of Z_{AB} and on their input states they can generate a correlation $P_{Z_{AB}} := \{p(a, b|\psi^x, \psi^y)\} \equiv \{\mathrm{Tr}[(\pi^a_{A^oA} \otimes \pi^b_{B^oB})(\psi^x_{A^o} \otimes \psi^y_{B^o} \otimes Z_{AB})]\}$ and thus obtain the expected payoff $\mathcal{I}_{\mathbb{G}_{sq}}(Z_{AB}) := \sum_{a,b,x,y} \beta\,(a, b, x, y) \times p(a, b|\psi^x, \psi^y)$ (see Fig. 1). If it turns out that for some BQS W_{AB} we have $\mathcal{I}_{\mathbb{G}_{sq}}(W_{AB}) < 0$, while $\mathcal{I}_{\mathbb{G}_{sq}}(\rho_{AB}) \geq 0$, $\forall\ \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, then the the game $\mathbb{G}_{sq}$ establishes the correlation strength of W_{AB} over quantum states. Generalization of the Bell scenario and the standard

Bell scenario to multiple parties follows in the obvious fashion.

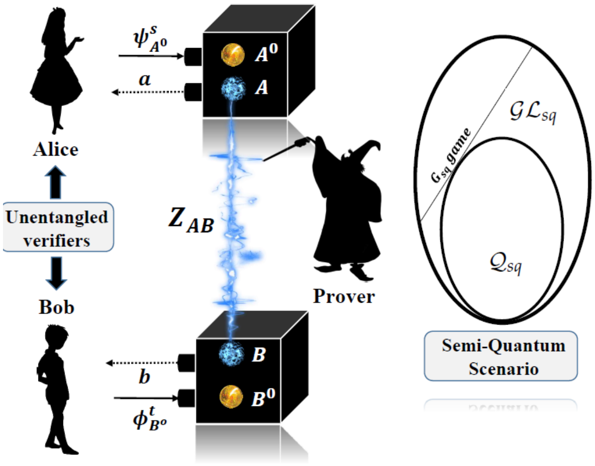


Figure 1. (Color online) A powerful but untrustworthy Prover distributes a bipartite state Z_{AB} between two distant Verifiers (Alice and Bob). The Verifiers do not have any entanglement between them, but possess their own trusted local quantum preparation device. Such limited resourceful Verifiers can verify the beyond quantumness of the state Z_{AB} provided to them (Theorem 1). The seminal Hahn-Banach separation theorem plays a crucial role in making this verification possible – the correlations produced from the bipartite quantum states form a convex-compact proper subset within the set of correlations produced from all bipartite states compatible with local quantum description and NS principle.

RESULTS

Proposition 1. *There exist beyond quantum bipartite states yielding correlations that are classically simulable in the Standard Bell Scenario.*

Theorem 1. *For every beyond quantum state $W_{AB} \in \mathcal{W}(\mathcal{H}_A \otimes \mathcal{H}_B)$ there exists a semiquantum game $\mathbb{G}_{sq}$ such that $\mathcal{I}_{\mathbb{G}_{sq}}(W_{AB}) < 0$, while $\mathcal{I}_{\mathbb{G}_{sq}}(\rho_{AB}) \geq 0$, $\forall\, \rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$.*

Theorem 2. *For every BQS $W_{A_1\cdots A_N} \in \mathcal{W}(\bigotimes_{i=1}^N \mathcal{H}_{A_i})$ there exists a semiquantum game $\mathbb{G}_{sq}$ such that $\mathcal{I}_{\mathbb{G}_{sq}}(W_{A_1\cdots A_N}) < 0$, whereas $\mathcal{I}_{\mathbb{G}_{sq}}(\rho_{A_1\cdots A_N}) \geq 0$, $\forall \rho_{A_1\cdots A_N} \in \mathcal{D}(\bigotimes_{i=1}^N \mathcal{H}_{A_i})$.*

While in the standard Bell scenario the result of Barnum *et al.* [18] ensures that the correlations obtained from any bipartite BQS is attainable in quantum theory, Theorem 1 reports beyond quantum correlations from all such BQSs in the quantum-input scenario (see Fig. 1). Naturally, it welcomes new principle(s) to isolate the quantum correlations from beyond-quantum ones in this generalized scenario.

While Theorem 1 is an existence theorem, it is not hard to see that given an arbitrary BQS there is an efficient algorithm to construct a semiquantum game. The procedure is discussed in the Appendix of the technical manuscript. It is worth mentioning that this semi quantum scenario is different from local tomography as it establishes beyond quantum nature of POPT states in a measurement device independent manner where the measurement devices used by the spatially separated parties need not to be trusted [31].

Discussion: One of the earnest research endeavours in quantum theory is to understand the limited nonlocal behaviour of quantum correlations. Apart from the foundational appeal, this question also has practical relevance as nonlocal correlations have been established as useful resources for several tasks. In the bipartite scenario the result of Barnum *et al.* [18] provides an answer to this question by assuming the description of local system to be quantum. Our work, however, points out the limitation of the scenario considered in [18]. While quantum inputs are brought into consideration, which is legitimate within the structure of local quantum description, the quantum correlations are not singled out naturally. Our Theorem 1 shows that all bipartite beyond quantum states compatible with unentangled Gleason's theorem can yield beyond quantum correlations in the quantum-input scenario, and accordingly divulges a more complex picture within the correlations zoo. Our study therefore welcomes new principles to identify the correlations in the physical world and points out that such a principle should take into consideration the type of the input-output scenario.

[1] J.S. Bell; On the Einstein Podolsky Rosen paradox, Physics **1**, 195 (1964).

[2] J. S. Bell; On the Problem of Hidden Variables in Quantum Mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[3] N. D. Mermin; Hidden variables and the two theorems of John Bell, Rev. Mod. Phys. **65**, 803 (1993).

[4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner; Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[5] E. Schrödinger; Discussion of Probability Relations Between Separated Systems, Math. Proc. Camb. Philos. Soc **31**, 555 (1935).

[6] J. Barrett, L. Hardy, and A. Kent; No Signaling and Quantum Key Distribution, Phys. Rev. Lett. **95**, 010503 (2005); A. Acín, N. Gisin, and L. Masanes; From Bell's Theorem to Secure Quantum Key Distribution, Phys. Rev. Lett. **97**, 120405 (2006); 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).

[7] S. Pironio *et al.* Random numbers certified by Bell's theorem, Nature **464**, 1021 (2010).

[8] R. Colbeck and R. Renner; Free randomness can be amplified, Nature Phys **8**, 450 (2012).

[9] N. Brunner, S. Pironio, A. Acin, N. Gisin, A. A. Méthot, and V. Scarani; Testing the Dimension of Hilbert Spaces, Phys. Rev. Lett. **100**, 210503 (2008); R. Gallego, N. Brunner, C. Hadley, and A. Acín; Device-Independent Tests of Classical and Quantum Dimensions, Phys. Rev. Lett. **105**, 230501 (2010); A. Mukherjee, A. Roy, S. S. Bhattacharya, S. Das, Md. R. Gazi, and M. Banik; Hardy's test as a device-independent dimension witness, Phys. Rev. A **92**, 022302 (2015).

[10] B. S. Cirel'son; Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).

[11] S. Popescu and D. Rohrlich; Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).

[12] 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 (1969).

[13] G. Birkhoff and J. von Neumann; The logic of quantum mechanics, Ann. Math. **37**, 823 (1936); G. W. Mackey; Mathematical Foundations of Quantum Mechanics. Benjamin, W. A. New York, 1963; Dover reprint, 2004; G. Ludwig; Attempt of an axiomatic foundation of quantum mechanics and more general theories II, III, Commun. Math. Phys. **4**, 331 (1967); Commun. Math. Phys. **9**, 1 (1968); B. Mielnik; Geometry of quantum states, Commun. Math. Phys. **9**, 55 (1968); E. Beltrametti and G. Cassinelli; The Logic of Quantum Mechanics, Addison-Wesley (1981); M. P. Solèr; Characterization of hilbert spaces by orthomodular spaces, Commun. Algebra **23**, 219 (1995); R. Haag; Local Quantum Physics: Fields, Particles, Algebras, 2nd Revised and Enlarged version, Springer (1996); R. Clifton,J. Bub, and H. Halvorson; Characterizing Quantum Theory in Terms of Information-Theoretic Constraints, Found. Phys. **33**, 1561 (2003); J. Barrett; Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007); S. Abramsky and B. Coecke; Categorical quantum mechanics, Handbook of Quantum Logic and Quantum Structures vol II, Elsevier, Amsterdam (2008); G. Chiribella, G. Mauro D'Ariano, and P. Perinotti; Informational derivation of quantum theory, Phys. Rev. A **84**, 012311 (2011).

[14] D. Foulis and C. Randall, in Interpretations and Foundations of Quantum Theory, edited by H. Neumann (Bibliographisches Institut Wissenschaftverlag, Mannheim, 1980), Vol. 5, pp. 9–20.

[15] M. Kläy, C. Randall, and D. Foulis; Tensor Products and Probability Weights, Int. J. Theor. Phys. **26**, 199 (1987).

[16] N. R. Wallach; An Unentangled Gleason's Theorem, arXiv:quant-ph/0002058.

[17] H. Barnum, C. A. Fuchs, J. M. Renes, and A. Wilce; Influence-free states on compound quantum systems, arXiv:quant-ph/0507108.

[18] H. Barnum, S. Beigi, S. Boixo, M. B. Elliott, and S. Wehner; Local Quantum Measurement and No-Signaling Imply Quantum Correlations, Phys. Rev. Lett. **104**, 140401 (2010).

[19] A. Acín, R. Augusiak, D. Cavalcanti, C. Hadley, J. K. Korbicz, M. Lewenstein, Ll. Masanes, and M. Piani; Unified Framework for Correlations in Terms of Local Quantum Observables, Phys. Rev. Lett. **104**, 140404 (2010).

[20] G. de la Torre, L. Masanes, A. J. Short, and M. P. Müller; Deriving Quantum Theory from Its Local Structure and Reversibility, Phys. Rev. Lett. **109**, 090403 (2012).

[21] M. Kleinmann, T. J. Osborne, V. B. Scholz, and A. H. Werner; Typical Local Measurements in Generalized Probabilistic Theories: Emergence of Quantum Bipartite Correlations, Phys. Rev. Lett. **110**, 040403 (2013).

[22] A. M. Gleason; Measures on the Closed Subspaces of a Hilbert Space, J. Math. Mech. **6**, 885 (1957).

[23] P. Busch; Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem, Phys. Rev. Lett. **91**, 120403 (2003).

[24] C. M. Caves, C. A. Fuchs, K. K. Manne, and J. M. Renes; Gleason-Type Derivations of the Quantum Probability Rule for Generalized Measurements, Found. Phys. **34**, 193 (2004).

[25] F. Buscemi; All Entangled Quantum States Are Nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[26] R. F. Werner; Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[27] J. Barrett; Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).

[28] A. Rai, MD. R. Gazi, M. Banik, S. Das, and S. Kunkri; Local simulation of singlet statistics for restricted set of measurement, J. Phys. A: Math. Theor. **45**, 475302 (2012).

[29] K. Kraus; States, Effects, and Operations: Fundamental Notions of Quantum Theory, Eds. A. Böhm, J. D. Dollard, and W. H. Wootters, Springer-Verlag Berlin Heidelberg (1983).

[30] O. Gühne and G. Tóth; Entanglement detection, Phys. Rep **474**, 1 (2009).

[31] C. Branciard, D. Rosset, Y-C Liang, and N. Gisin; Measurement-Device-Independent Entanglement Witnesses for All Entangled Quantum States, Phys. Rev. Lett. **110**, 060405 (2013).

[32] M. Banik; Lack of measurement independence can simulate quantum correlations even when signaling can not, Phys. Rev. A **88**, 032118 (2013).

[33] A. Chaturvedi and M. Banik; Measurement-device–independent randomness from local entangled states, EPL **112**, 30003 (2015).

[34] D. Rosset, D. Schmid, and F. Buscemi; Type-Independent Characterization of Spacelike Separated Resources, Phys. Rev. Lett. **125**, 210402 (2020).

[35] D. Schmid, D. Rosset, and F. Buscemi; The type-independent resource theory of local operations and shared randomness, Quantum **4**, 262 (2020).

[36] F. Graffitti, A. Pickston, P. Barrow, M. Proietti, D. Kundys, D. Rosset, M. Ringbauer, and A. Fedrizzi; Measurement-Device-Independent Verification of Quantum Channels, Phys. Rev. Lett. **124**, 010503 (2020).

[37] S. L. Woronowicz; Positive maps of low dimensional matrix algebras, Rep. Math. Phys. **10**, 165 (1976).

[38] M. Horodecki, P. Horodecki, and R. Horodecki; Separability of mixed states: necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[39] E. Størmer; Positive Linear Maps of Operator Algebras, Springer-Verlag Berlin Heidelberg (2013).

[40] W. van Dam; Implausible Consequences of Superstrong Nonlocality, arXiv:quant-ph/0501159 (2005); H. Buhrman, R. Cleve, S. Massar, and R. de Wolf; Nonlocality and communication complexity, Rev. Mod. Phys. **82**, 665 (2010); M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski; Information causality as a physical principle, Nature **461**, 1101 (2009); M. Navascues and H. Wunderlich; A glance beyond the quantum model, Proc. Roy. Soc. Lond. A **466**, 881 (2009); T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín; Local orthogonality as a multipartite principle for quantum correlations, Nat. Commun. **4**, 2263 (2013); A. Cabello; Simple Explanation of the Quantum Violation of a Fundamental Inequality, Phys. Rev. Lett. **110**, 060402 (2013); J. Oppenheim and S. Wehner; The uncertainty principle determines the non-locality of quantum mechanics, Science **330**, 1072 (2010); 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); M. Banik, S. S. Bhattacharya, A. Mukherjee, A. Roy, A. Ambainis, and A. Rai; Limited preparation contextuality in quantum theory and its relation to the Cirel'son bound, Phys. Rev. A **92**, 030103(R) (2015); G. Kar and M. Banik; Several foundational and information theo- retic implications of Bell's theorem; Int. J. Quantum Inform. **14**, 1640027 (2016); G. Kar, S.Ghosh, S. K. Choudhary, and M. Banik; Role of Measurement Incompatibility and Uncertainty in Determining Nonlocality, Mathematics **4**, 52 (2016); M. Banik, S. Saha, T. Guha, S. Agrawal, S. S. Bhattacharya, A. Roy, and A. S.

Majumdar; Constraining the state space in any physical theory with the principle of information symmetry, Phys. Rev. A **100**, 060101(R) (2019); S. S. Bhattacharya, S. Saha, T. Guha, and M. Banik; Nonlocality without entanglement: Quantum theory and beyond, Phys. Rev. Research **2**, 012068(R) (2020).

[41] M. Dall'Arno, S. Brandsen, A. Tosini, F. Buscemi, and V. Vedral; No-Hypersignaling Principle, Phys. Rev. Lett. **119**, 020401 (2017).

[42] S. G. Naik, E. P. Lobo, S. Sen, R. Patra, M. Alimuddin, T. Guha, S. S. Bhattacharya, and M. Banik; On composition of multipartite quantum systems: perspective from time-like paradigm,

Minimum number of experimental settings required to verify bipartite pure states and unitaries

Yunting Li[1 2 3] Haoyu Zhang[1 2 3] Zihao Li[1 2 3] Huangjun Zhu[1 2 3 *]

[1] *State Key Laboratory of Surface Physics and Department of Physics, Fudan University, Shanghai 200433, China*
[2] *Institute for Nanoelectronic Devices and Quantum Computing, Fudan University, Shanghai 200433, China*
[3] *Center for Field Theory and Particle Physics, Fudan University, Shanghai 200433, China*

Abstract. Efficient verification of quantum states and gates is crucial to the development of quantum technologies. Although the sample complexities of quantum state verification and quantum gate verification have been studied by many researchers, the number of experimental settings has received little attention and is poorly understood. In this work we study systematically quantum state verification and quantum gate verification with a focus on the number of experimental settings. We show that any bipartite pure state can be verified by only two measurement settings based on local projective measurements. Any bipartite unitary in dimension d can be verified by $2d$ experimental settings based on local operations. In addition, we introduce the concept of entanglement-free verification and clarify its connection with minimal-setting verification. Finally, we show that any two-qubit unitary can be verified with at most five experimental settings; moreover, a generic two-qubit unitary (except for a set of measure zero) can be verified by an entanglement-free protocol based on four settings. In the course of study we clarify the properties of Schmidt coefficients of two-qubit unitaries, which are of independent interest.

Keywords: quantum state verification, quantum gate verification, minimal settings, bipartite pure states and unitaries, entanglement-free verification

1 Introduction

Quantum information processing has attracted increasing attention recently due to its great potential and profound implications. To harness the power of quantum information processing, it is crucial to verify the underlying quantum states and devices efficiently based on the accessible measurements. Unfortunately, traditional tomographic approaches are notoriously inefficient since the resource overhead increases exponentially with the system size under consideration. To overcome this problem, a number of alternative approaches have been proposed recently; see Refs. [1, 2, 3, 4] for an overview.

Among alternative approaches proposed so far, *quantum state verification* (QSV) is particularly appealing because it can achieve a high efficiency based on local operations and classical communication (LOCC) [5, 6, 7, 8, 12, 10]. Notably, efficient verification protocols based on local projective measurements have been constructed for bipartite pure states [5, 11, 13, 14, 15], stabilizer states [16, 8, 17, 18, 10, 19, 20], hypergraph states [18], weighted graph states [21], and Dicke states [22, 23]. Moreover, the efficiency of QSV has been demonstrated in a number of experiments [24, 25, 26, 27]. Recently, the idea of QSV was generalized to *quantum gate verification* (QGV) [28, 29, 30] (cf. Refs. [31, 32, 33, 34, 35]), which enables efficient verification of various quantum gates and quantum circuits based on LOCC. Notably, all bipartite unitaries and Clifford unitaries can be verified with resources that are independent of the system size, while the resource required to verify the generalized controlled-NOT (CNOT) gate and generalized controlled-Z (CZ) gate grows only linearly with the system size. The efficiency of QGV has also been demonstrated in several experiments recently [36, 37].

So far most works on QSV and QGV have exclusively focused on the sample efficiency as the main figure of merit. By contrast, the number of experimental settings has received little attention, although this figure of merit is also of key interest to both theoretical study and practical applications. Even for bipartite pure states, it is still not clear how many measurement settings are required to construct a reliable verification protocol. The situation is even worse in the case of bipartite unitaries, not to mention the multipartite scenario. This problem becomes particularly important when it is difficult or slow to switch measurement settings, which is the case in many practical scenarios.

In this work we study systematically QSV and QGV with a focus on the number of experimental settings based on LOCC. We show that any bipartite pure state can be verified by two measurement settings based on nonadaptive local projective measurements. By contrast, at least d experimental settings based on local operations are required to verify each bipartite unitary in dimension d, while $2d$ settings are sufficient. In addition, we introduce the concept of entanglement-free verification, which is of special interest to both theoretical study and practical applications. Moreover, we show that any entanglement-free verification protocol can be turned into a minimal-setting protocol, and vice versa.

For each two-qubit unitary, we determine the minimum number of required experimental settings explicitly. Our study shows that any two-qubit unitary can be verified using only five experimental settings, while a generic two-qubit unitary (except for a set of measure zero) can be verified by an entanglement-free protocol based on four settings. Explicit entanglement-free protocols are constructed for CNOT, CZ, controlled-phase (C-Phase),

*zhuhuangjun@fudan.edu.cn

and SWAP gates, respectively. In the course of study we clarify the properties of Schmidt coefficients of two-qubit unitaries and their implications for studying the equivalence relation under local unitary transformations, which are of interest beyond the main focus of this work.

2 Quantum state verification and quantum gate verification

For quantum state verification, consider a quantum system associated with the Hilbert space $\mathcal{H}$. A quantum device is supposed to produce the target state $|\Psi\rangle$, but actually produces the N states $\rho_1, \rho_2, \ldots, \rho_N$ in N runs. To distinguish the two situations, we can perform a random test in each run. Each test is determined by a test operator E_l, which is associated with a two-outcome measurement of the form $\{E_l, I - E_l\}$, where I is the identity operator. To guarantee that the target state can always pass the test, the test operator E_l should satisfy the condition $\langle\Psi|E_l|\Psi\rangle = 1$, which means $E_l|\Psi\rangle = |\Psi\rangle$. If the test E_l is performed with probability p_l, then the performance of the above verification procedure is determined by the verification operator $\Omega = \sum_l p_l E_l$. Note that a positive spectral gap is necessary and sufficient for verifying the target state reliably, assuming that the total number of tests is not limited.

For quantum gate verification, consider a quantum device that is expected to perform the unitary transformation $\mathcal{U}$ associated with the unitary operator U on $\mathcal{H}$, but actually realizes an unknown quantum process Λ. In order to verify whether this quantum process is sufficiently close to the target unitary transformation, we need to construct a set $\mathscr{T} = \{|\psi_j\rangle\}_j$ of test states. In each run we randomly prepare a test state from the set $\mathscr{T}$ and apply the quantum process Λ. Then we verify whether the output state $\Lambda(\rho_j)$ is sufficiently close to the target output state $\mathcal{U}(\rho_j) = U\rho_j U^\dagger$ by virtue of QSV, where $\rho_j = |\psi_j\rangle\langle\psi_j|$ [29, 28]. By construction, the target unitary transformation can always pass each test.

A set $\mathscr{T} = \{|\psi_j\rangle\}_j$ in $\mathcal{H}$ can *identify* the unitary transformation $\mathcal{U}$ if the condition $\Lambda(|\psi_j\rangle\langle\psi_j|) = \mathcal{U}(|\psi_j\rangle\langle\psi_j|), \quad \forall j$ implies that $\Lambda = \mathcal{U}$, that is, $\Lambda(\rho) = \mathcal{U}(\rho), \quad \forall \rho \in \mathscr{D}(\mathcal{H})$, where $\mathscr{D}(\mathcal{H})$ denotes the set of all density operators on the Hilbert space $\mathcal{H}$. In this case, the set $\mathscr{T}$ is referred to as an *identification set* (IS).

3 Verification of bipartite pure states with minimal settings

Given a bipartite or multipartite pure state $|\Psi\rangle$, how many measurement settings are necessary to verify $|\Psi\rangle$ reliably? Here we focus on verification protocols based on nonadaptive local projective measurements, which are amenable to experimental realization. Although it is known that any bipartite pure state can be verified by two distinct tests based on adaptive local projective measurements [13], one test based on adaptive local projective measurements may entail many different measurement settings, so the result presented in Ref. [13] does not resolve the current problem under consideration.

Here we show that any bipartite pure state can be verified by at most two measurement settings, thereby resolving the minimal-setting problem in the bipartite scenario completely.

Theorem 1 *Every bipartite pure product state can be verified by one measurement setting. Every bipartite pure entangled state can be verified by two measurement settings, but not one measurement setting.*

4 Verification of unitary transformations with minimal settings

4.1 Minimal identification sets

Here we are particularly interested in ISs with as few elements as possible. The set $\mathscr{T}$ is a *minimal identification set* (MIS) if, in addition, any proper subset is not an IS. MISs are crucial to constructing verification protocols for unitary transformations with minimal settings. To understand the properties of ISs and MISs, we need to introduce several additional concepts and the results are shown in the following lemmas.

Lemma 2 *A set of pure states in $\mathcal{H}$ is an IS iff it is a connected spanning set.*

Lemma 3 *A set of pure states in $\mathcal{H}$ is a MIS iff it is a connected basis.*

Lemma 4 *Suppose $\mathscr{T}$ is a connected spanning set in $\mathcal{H}$. Then any maximal connected linearly independent set (CLIS) contained in $\mathscr{T}$ is a connected basis.*

Lemma 5 *Every connected spanning set in $\mathcal{H}$ contains a subset that forms a connected basis. Every set in $\mathcal{H}$ that contains a connected spanning subset is a connected spanning set.*

4.2 Minimal-setting verification and Entanglement-free verification

A verification protocol for U is *entanglement free* if all input test states and the corresponding output states (after the action of U) are product states; in addition, all measurements are based on local projective measurements. An entanglement-free protocol does not generate any entanglement in the verification procedure and hence the name. Such verification protocols are particularly appealing to both theoretical study and experimental realization. It turns out entanglement-free verification is intimately connected to minimal-setting verification.

Denote by Prod the set of pure product states; denote by $\mathrm{Prod}(U)$ the set of product states that remain product states after the action of U:

$$\mathrm{Prod}(U) = \{|\psi\rangle \in \mathrm{Prod} \mid U|\psi\rangle \in \mathrm{Prod}\}. \quad (1)$$

The dimension of the span of the set $\mathrm{Prod}(U)$ is denoted by $d_{\mathrm{Prod}}(U) = \dim \mathrm{span}(\mathrm{Prod}(U))$, which satisfies $0 \leq d_{\mathrm{Prod}}(U) \leq d$. A state $|\psi\rangle$ in $\mathcal{H}$ satisfies the *product-state constraint* associated with U if $|\psi\rangle \in \mathrm{Prod}(U)$. A set of states satisfies the product-state constraint if it is contained in $\mathrm{Prod}(U)$, so that each state satisfies the constraint.

An entanglement-free IS (EFIS) $\mathscr{T}$ for U is an IS that satisfies the product-state constraint, which implies that $\mathscr{T} \subseteq \mathrm{Prod}(U)$. Similarly, an entanglement-free MIS (EFMIS) is a MIS that satisfies the product-state constraint. Note that the definition of an EFIS (EFMIS) depends on the specific unitary transformation under consideration, although the definition of an IS (MIS) is independent of a specific unitary transformation. The unitary operator U can be verified by an entanglement-free protocol iff it admits an EFMIS, in which case $\mathrm{Prod}(U)$ contains an IS. Theorem 6 below further clarifies the connections among the product-state constraint as determined by $\mathrm{Prod}(U)$, minimal-setting verification, and entanglement-free verification.

Theorem 6 *Suppose U is a unitary operator on a composite Hilbert space $\mathcal{H}$ of dimension d. Then the following five statements are equivalent:*

1. $\mu(U) = d$.
2. $\mathrm{Prod}(U)$ *is a connected spanning set.*
3. $\mathrm{Prod}(U)$ *contains a connected basis as a subset.*
4. *U admits an EFMIS.*
5. *U can be verified by an entanglement-free protocol.*

4.3 Minimal settings for verifying bipartite unitaries

Here we focus on the verification of general bipartite unitaries and show that the minimum number of settings required to verify a generic bipartite unitary grows linearly with the total dimension.

Theorem 7 *Suppose U is a unitary operator acting on a d-dimensional bipartite Hilbert space $\mathcal{H}$. Then the minimum number of experimental settings $\mu(U)$ required to verify U satisfies $d \leq \mu(U) \leq 2d$.*

Proposition 8 *Let U be a unitary operator acting on a d-dimensional bipartite Hilbert space $\mathcal{H}$. If $d_{\mathrm{Prod}}(U) < d$, then*

$$\mu(U) = d_{\mathrm{Prod}}(U) + 2[d - d_{\mathrm{Prod}}(U)]. \tag{2}$$

In the case $d_{\mathrm{Prod}}(U) = d$, we have $\mu(U) = d$ if the set $\mathrm{Prod(U)}$ *is connected and $\mu(U) = d + 1$ otherwise.*

5 Verification of two-qubit unitaries with minimal settings

Let $\mathcal{H} = \mathcal{H}_{\mathrm{A}} \otimes \mathcal{H}_{\mathrm{B}}$ be the Hilbert space associated with a two-qubit system shared by A and B. According to Refs. [38, 39], any two-qubit unitary operator U_{AB} acting on $\mathcal{H}$ can be expressed as $U_{\mathrm{AB}} = V_{\mathrm{A}} \otimes W_{\mathrm{B}} U \tilde{V}_{\mathrm{A}} \otimes \tilde{W}_{\mathrm{B}}$, where $V_{\mathrm{A}}, W_{\mathrm{B}}, \tilde{V}_{\mathrm{A}}, \tilde{W}_{\mathrm{B}}$ are four qubit unitary operators,

$$\begin{gathered} U = U(\alpha_1, \alpha_2, \alpha_3) = \mathrm{e}^{-\mathrm{i}\sum_{k=1}^{3} \alpha_k H_k}, \\ 0 \leq |\alpha_3| \leq \alpha_2 \leq \alpha_1 \leq \pi/4, \\ H_i = \sigma_i \otimes \sigma_i, i = 1, 2, 3, \end{gathered} \tag{3}$$

and $\sigma_1, \sigma_2, \sigma_3$ are the three Pauli operators.

We determine the minimum number of experimental settings required to verify an arbitrary two-qubit unitary and derive a simple criterion for determining whether a general two-qubit unitary can be verified by an entanglement-free protocol. Our main result is summarized in the following theorem.

Theorem 9 *Suppose U is a two-qubit unitary operator with Schmidt coefficients s_0, s_1, s_2, s_3 arranged in nonincreasing order. Then*

$$\mu(U) = \begin{cases} 5 & \textit{if } s_0 > s_1 = s_2 = s_3 > 0, \\ 4 & \textit{otherwise}; \end{cases} \tag{4}$$

in addition, the unitary operator U can be verified by an entanglement-free protocol unless $s_0 > s_1 = s_2 = s_3 > 0$.

According to the relation between $\alpha_1, \alpha_2, \alpha_3$ and the Schmidt coefficients, define

$$\mathcal{S} := \Big\{(\alpha_1, \alpha_2, \alpha_3) \Big| 0 \leq \alpha_3 \leq \alpha_2 \leq \alpha_1 \leq \frac{\pi}{4}\Big\}, \tag{5}$$

$$\mathcal{S}_{\mathrm{E}} := \Big\{(\alpha, \alpha, \alpha) \Big| 0 < \alpha < \frac{\pi}{4}\Big\}, \quad \mathcal{S}_{\mathrm{EF}} := \mathcal{S} \setminus \mathcal{S}_{\mathrm{E}}. \tag{6}$$

Theorem 10 *Suppose $0 \leq \alpha_3 \leq \alpha_2 \leq \alpha_1 \leq \pi/4$. Then*

$$\mu(U(\alpha_1, \alpha_2, \alpha_3)) = \begin{cases} 4 & \textit{if } (\alpha_1, \alpha_2, \alpha_3) \in \mathcal{S}_{\mathrm{EF}}, \\ 5 & \textit{if } (\alpha_1, \alpha_2, \alpha_3) \in \mathcal{S}_{\mathrm{E}}. \end{cases} \tag{7}$$

$U(\alpha_1, \alpha_2, \alpha_3)$ can be verified by an entanglement-free protocol iff $(\alpha_1, \alpha_2, \alpha_3) \in \mathcal{S}_{\mathrm{EF}}$.

6 Summary

We studied systematically QSV and QGV with a focus on the number of experimental settings based on local operations. We showed that any bipartite pure state can be verified by only two measurement settings based on local projective measurements. The minimum number of experimental settings required to verify a bipartite unitary increases linearly with the total dimension. In addition, we introduced the concept of entanglement-free verification, which does not generate any entanglement in the verification procedure. The connection with minimal-setting verification is also clarified. Finally, we determined the minimum number of experimental settings required to verify each two-qubit unitary. It turns out any two-qubit unitary can be verified using at most five settings based on local operations, and a generic two-qubit unitary requires only four settings.

In addition, our work shows that verification protocols with minimal settings are in general not balanced and thus do not have natural analogs in QSV, which reflects a key distinction between QGV and QSV that is not recognized before. In the future it would be desirable to generalize our results to the multipartite setting.

References

[1] 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).

[2] M. Kliesch and I. Roth. Theory of quantum system certificatio. *PRX Quantum* 2, 010201 (2021).

[3] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller. Theoretical and experimental perspectives of quantum verification. *PRX Quantum* 2, 010102 (2021).

[4] X.-D. Yu, J. Shang, and O. Gühne. Statistical methods for quantum state verification and fidelity estimation. *arXiv* 2109.10805.

[5] 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).

[6] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert. Reliable quantum certification of photonic state preparations. *Nat. Commun.* 6, 8498 (2015).

[7] Y. Takeuchi and T. Morimae. Verification of Many-Qubit States. *Phys. Rev. X* 8, 021060 (2018).

[8] S. Pallister, N. Linden, and A. Montanaro. Optimal Verification of Entangled States with Local Measurements. *Phys. Rev. Lett.* 120, 170502 (2018).

[9] H. Zhu and M. Hayashi. Efficient Verification of Pure Quantum States in the Adversarial Scenario. *Phys. Rev. Lett.* 123, 260504 (2019).

[10] H. Zhu and M. Hayashi. General framework for verifying pure quantum states in the adversarial scenario. *Phys. Rev. A* 100, 062335 (2019).

[11] H. Zhu and M. Hayashi. Optimal verification and fidelity esti- mation of maximally entangled states. *Phys. Rev. A* 99, 052346 (2019).

[12] H. Zhu and M. Hayashi. Efficient Verification of Pure Quantum States in the Adversarial Scenario. *Phys. Rev. Lett.* 123, 260504 (2019).

[13] Z. Li, Y.-G. Han, and H. Zhu. Efficient verification of bipartite pure states. *Phys. Rev. A* 100, 032316 (2019).

[14] K. Wang and M. Hayashi. Optimal verification of two-qubit pure states. *Phys. Rev. A* 100, 032315 (2019).

[15] X.-D. Yu, J. Shang, and O. Gühne. Optimal verification of general bipartite pure states. *npj Quantum Inf.* 5, 112 (2019).

[16] M. Hayashi and T. Morimae. Verifiable Measurement-Only Blind Quantum Computing with Stabilizer Testing. *Phys. Rev. Lett.* 115, 220502 (2015).

[17] A. Kalev, A. Kyrillidis, and N. M. Linke. Validating and certifying stabilizer states. *Phys. Rev. A* 99, 042337 (2019).

[18] H. Zhu and M. Hayashi. Efficient Verification of Hypergraph States. *Phys. Rev. Appl.* 12, 054047 (2019).

[19] Z. Li, Y.-G. Han, and H. Zhu. Optimal Verification of Greenberger-Horne-Zeilinger States. *Phys. Rev. Appl.* 13, 054002 (2020).

[20] N. Dangniam, Y.-G. Han, and H. Zhu. Optimal verification of stabilizer states. *Phys. Rev. Res.* 2, 043323 (2020).

[21] M. Hayashi and Y. Takeuchi. Verifying commuting quantum computations via fidelity estimation of weighted graph states. *New J. Phys.* 21, 093060 (2019).

[22] Y.-C. Liu, X.-D. Yu, J. Shang, H. Zhu, and X. Zhang. Efficient Verification of Dicke States. *Phys. Rev. Appl.* 12, 044020 (2019).

[23] Z. Li, Y.-G. Han, H.-F. Sun, J. Shang, and H. Zhu. Verification of phased Dicke states. *Phys. Rev. A* 103, 022601 (2021).

[24] 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, et al.. Experimental Optimal Verification of Entangled States Using Local Measurements. *Phys. Rev. Lett.* 125, 030506 (2020).

[25] L. Lu, L. Xia, Z. Chen, L. Chen, T. Yu, T. Tao, W. Ma, Y. Pan, X. Cai, Y. Lu, et al.. Three-dimensional entanglement on a silicon chip. *npj Quantum Inf.* 6, 30 (2020).

[26] X. Jiang, K. Wang, K. Qian, Z. Chen, Z. Chen, L. Lu, L. Xia, F. Song, S. Zhu, and X. Ma. Towards the standardization of quantum state verification using optimal strategies. *npj Quantum Inf.* 6, 90 (2020).

[27] W.-H. Zhang, X. Liu, P. Yin, X.-X. Peng, G.-C. Li, X.-Y. Xu, S. Yu, Z.-B. Hou, Y.-J. Han, J.-S. Xu, et al.. Classical communication enhanced quantum state verification. *npj Quantum Inf.* 6, 103 (2020).

[28] Y.-C. Liu, J. Shang, X.-D. Yu, and X. Zhang. Efficient verification of quantum processes. *Phys. Rev. A* 101, 042315 (2020).

[29] H. Zhu and H. Zhang. Efficient verification of quantum gates with local operations. *Phys. Rev. A* 101, 042316 (2020).

[30] P. Zeng, Y. Zhou, and Z. Liu. Quantum gate verification and its application in property testing. *Phys. Rev. Res.* 2, 023306 (2020).

[31] H. F. Hofmann. Complementary Classical Fidelities as an Efficient Criterion for the Evaluation of Experimentally Realized Quantum Operations. *Phys. Rev. Lett.* 94, 160504 (2005).

[32] D. M. Reich, G. Gualdi, and C. P. Koch. Minimum number of input states required for quantum gate characterization. *Phys. Rev. A* 88, 042309 (2013).

[33] K. Mayer and E. Knill. Quantum process fidelity bounds from sets of input states. *Phys. Rev. A* 98, 052326 (2018).

[34] Y.-D. Wu and B. C. Sanders. Efficient verification of bosonic quantum channels via benchmarking. *New J. Phys.* 21, 073026 (2019).

[35] YA. Elben, B. Vermersch, R. van Bijnen, C. Kokail, T. Brydges, C. Maier, M. K. Joshi, R. Blatt, C. F. Roos, and P. Zoller. Cross-Platform Verification of Intermediate Scale Quantum Devices. *Phys. Rev. Lett.* 124, 010504 (2020).

[36] R.-Q. Zhang, Z. Hou, J.-F. Tang, J. Shang, H. Zhu, G.-Y. Xiang, C.-F. Li, and G.-C. Guo. Efficient Experimental Verification of Quantum Gates with Local Operations. *Phys. Rev. Lett.* 128, 020502 (2022).

[37] M. Luo, X. Zhang, and X. Zhou. Proof-of-principle experimental demonstration of quantum gate verification. *arXiv* 2107.13466.

[38] B. Kraus and J. I. Cirac. Optimal creation of entanglement using a two-qubit gate. *Phys. Rev. A* 63, 062309 (2001).

[39] W. Dür, G. Vidal, and J. I. Cirac. Optimal Conversion of Nonlocal Unitary Operations. *Phys. Rev. Lett.* 89, 057901 (2002).

Rate-fidelity trade-off in entanglement distribution between distant ion-cavity systems

Kazufumi Tanji[1] * Wojciech Roga[1] † Hiroki Takahashi[2] ‡ Masahiro Takeoka[1] §

[1] *Department of Electronics and Electrical Engineering, Faculty of Science and Engineering, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, Kanagawa, 223-8522 Japan*

[2] *Experimental Quantum Information Physics Unit, Okinawa Institute of Science and Technology Graduate University, 1919-1 Tancha, Onna, Kunigami, Okinawa 904-0495, Japan*

Abstract. Entanglement distribution between distant ions is necessary for realization of a large-scale quantum computer. Therefore, various experiments and theoretical research have been conducted to improve the entanglement distribution. However, the inherent problem of spontaneous decay of ions degrades the entanglement fidelity and generation rate. In this study we investigate the relationship between a waveform of a pump and spontaneous decay. First, we numerically show the trade-off between the entanglement fidelity and generation rate with various Gaussian pump pulses. Following that we propose how to determine better waveforms. Finally, we show that a Gaussian waveform is sufficient for the entanglement swapping.

Keywords: Quantum computer, trapped ion, cavity-QED, entanglement swapping

1 Introduction

Trapped ions have a range of properties [1] required for a good quantum computer such as homogeneity, long coherence time, high fidelity of gates, state preparation, and measurements, as well as good connectivity. On the other hand scalability is a challenge that needs to be solve to create a full scale distributed quantum computer[2]. The key element on the pathway to the scalable quantum computing is entanglement swapping [3] schematically shown in Fig. 1 which has been already experimentally demonstrated for instance in [4, 5, 6]. However, the performance of the entanglement swapping is inherently limited by spontaneous decays.

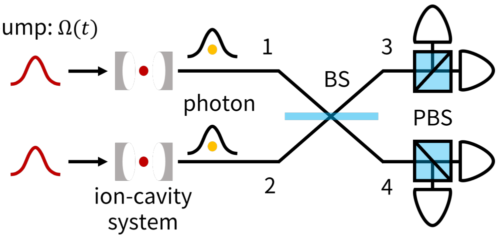


Figure 1: Entanglement swapping

The detailed analysis shows that there are two kinds of spontaneous decays relevant for the entanglement distribution, γ_g and γ_u shown in Fig 2. Here, γ_g refers to the process in which the excited state decays to the final state without emitting a communication photon. It decreases the probability of emitting photon. This decay has been well studied and the analytical form of the optimal waveform of the pump pulse which allow to mitigate the impact of this decay on entanglement sharing has been found [7].

On the other hand, in the process denoted by γ_u, the excited state decays to the initial state. The decay γ_u decreases the fidelity of ion-ion entanglement and the visibility of a photon but does not decrease the probability because the state after the decay can be excited again. This impact has been experimentally and numerically confirmed, e.g., in ref. [8, 9, 10]. In the strong coupling regime, $g \gg \kappa, \gamma$, these spontaneous decays are negligible if adiabatic excitation [11] or large detuning [12] is applied, however achieving sufficiently strong coupling is experimentally challenging. Indeed, the current maximum g [13] is not sufficient to use the techniques mentioned above.

Thus, the relationship between a waveform of a pump pulse $\Omega(t)$ and the re-excitable spontaneous decay γ_u is currently still not well understood. In this study we investigate this relation using both numerical and analytical approaches with realistic experimental conditions and give the strategy to choose a better waveform of the pump pulse.

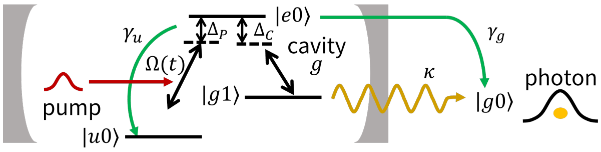


Figure 2: The ion-cavity model. The ion has the Λ-type energy level. $|\mu, n\rangle = |\mu\rangle_{\text{atom}} \otimes |n\rangle_{\text{photon}}$, $\Omega(t)$: the waveform of the pump pulse, Δ_p and Δ_c: detuning, g: the ion-cavity coupling factor, γ_u: the spontaneous decay rate $|e\rangle \to |u\rangle$, γ_g: the spontaneous decay rate $|e\rangle \to |g\rangle$, and κ: the cavity decay rate.

*kizehemu@keio.jp

†wojciech.roga@keio.jp

‡hiroki.takahashi@oist.jp

§takeoka@elec.keio.ac.jp

2 Methods

In this study, in order to focus our attention on the relationship between $\Omega(t)$ and γ_u we assume that

$$\Delta_p = \Delta_c = \gamma_g = 0. \tag{1}$$

The calculation method follows ref. [8, 14]. Assuming that the measurement unit of entanglement swapping is perfect, the entanglement fidelity is determined by the dynamics of the ion-cavity system. The dynamics is described by the following quantum master equation;

$$\begin{aligned}\frac{d}{dt}\rho(t) &= -i\left(H_{\mathrm{eff}}(t)\rho(t) - \rho(t)H_{\mathrm{eff}}^{\dagger}(t)\right)\\ &\quad + 2\gamma_u \left|u0\right\rangle\!\left\langle e0\right| \rho(t) \left|e0\right\rangle\!\left\langle u0\right|\\ &\quad + 2\kappa \left|u0\right\rangle\!\left\langle g1\right| \rho(t) \left|g1\right\rangle\!\left\langle u0\right|\\ H_{\mathrm{eff}}(t) &= (\Delta_c - i\kappa)\left|g1\right\rangle\!\left\langle g1\right| + \Delta_p \left|u0\right\rangle\!\left\langle u0\right|\\ &\quad - i\gamma_u \left|e0\right\rangle\!\left\langle e0\right|\\ &\quad + \left[\Omega(t)\left|u0\right\rangle\!\left\langle e0\right| + g\left|e0\right\rangle\!\left\langle g1\right| + h.c.\right].\end{aligned} \tag{2}$$

Considering the initial state as $\rho(0) = |u0\rangle\langle u0|$, the master equation leaves the state in the four-dimensional space. The probability of emitting photon P_{ex} is the probability of the state to be in $|g0\rangle$. Therefore,

$$P_{ex} = 2\kappa \int dt \left\langle g1\right| \rho(t) \left|g1\right\rangle. \tag{3}$$

The fidelity of the ion-ion entanglement [8] is

$$\begin{aligned}F &= \frac{1}{2}(1 + \mathrm{Re}\{J\}), \quad \text{where}\\ J &= \frac{\iint dt dt' \langle a_c^{\dagger}(t)a_c(t')\rangle\langle a_c^{\dagger}(t)a_c(t')\rangle^{*}}{P_{ex}^{2}}.\end{aligned} \tag{4}$$

The two-point correlation function $\langle a_c^{\dagger}(t)a_c(t')\rangle$ can be calculated by the quantum regression theorem [15] using the following relations,

$$\begin{aligned}&\langle a_c^{\dagger}(t)a_c(t')\rangle = \mathrm{Tr}[\Lambda(t,t')a_c],\\ &\frac{d}{dt'}\Lambda(t,t') = -i\left(H_{\mathrm{eff}}(t')\Lambda(t,t') - \Lambda(t,t')H_{\mathrm{eff}}^{\dagger}(t')\right)\\ &\qquad + 2\gamma_u \left|u0\right\rangle\!\left\langle e0\right| \Lambda(t,t') \left|e0\right\rangle\!\left\langle u0\right|\\ &\qquad + 2\kappa a_c \Lambda(t,t') a_c^{\dagger},\\ &\Lambda(t,t) = \rho(t)a_c^{\dagger}.\end{aligned} \tag{5}$$

3 Results

We calculated F and P_{ex} for different sets of parameters presented in Table 1 with various Gaussian pump pulses by python. In this proceeding, due to the limited space, we show only the results for the intermediate regime. First, in Fig. 3 we show the relationship between F and P_{ex} for various Gaussian pump pulses . We observe the trade-off between F and P_{ex}.

In Fig. 4 we classified the points from the previous figure with respect to the pulse area to investigate the trade-off further. We checked how these plots are changed with pulse width as shown in Fig. 5 (a) and (b), where we chose the pulse area as 0.7 and 7.0 respectively. The black dot shows the shortest pulse width, and the white one shows the longest pulse width. We found that the shorter pulse gives higher fidelity in the same probability when the pulse area is 0.7, and the longer pulse gives higher fidelity in the same probability when the pulse area is 7.0. We repeated the same procedure for the other pulse area plots, and we find the same characteristics as in the above examples.

Table 1: Parameters of numerical simulation

Regime	g	κ	γ_u	C
Intermediate (a)	1	1	1	1
Strong (b)	$\sqrt{10}$	1	1	10
Weak (c)	$\sqrt{10}$	10	10	0.1
Purcell (d)	10	2	50	1
(e)	10	50	2	1

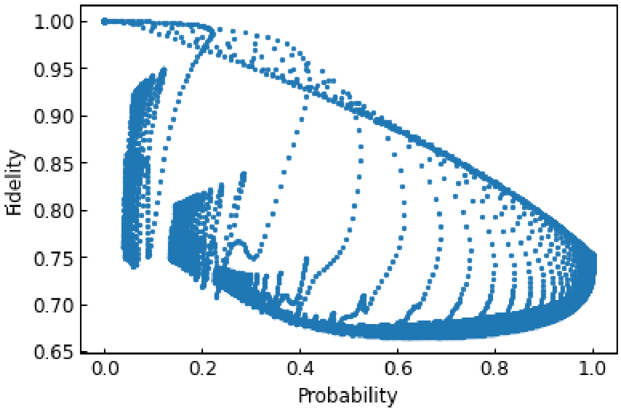


Figure 3: The trade-off between F and P_{ex}.

This result can be intuitively understood as follows. When the probability is small, to neglect the spontaneous decay, the pulse width should be much shorter than γ_u. On the other hand, when the probability is high, the adiabatic pumping is more effective [11].

We also considered the effect of non-Gaussian pumping, Fig. 6. We plot results for asymmetric-Gaussian pulses $\Omega(t) = 2\Omega_0/(1+\sqrt{r})\exp\{-(t-t_c)^2/2\sigma^2\}$, where $\sigma = \sigma_1(t \le t_c)$, $\sigma = \sigma_2(t \ge t_c)$, and r is the ratio of σ_1 and σ_2 such that the pulse area is kept constant. In Fig. 6(a), σ_1 is fixed, and σ_2 is changed. In Fig. 6(b), σ_2 is fixed, and σ_1 is changed. We can conclude that although we can see some improvement it is slight.

Finally, we considered the analytical form of the fidelity from Eq. (2)-(5) and found the following relation,

$$\begin{aligned}J &\le \frac{4C}{4C+3} - \int_0^{\infty} \frac{4}{\kappa(4C+3)}\left(\frac{d\rho_{g1,g1}(t)}{dt}\right)^2 dt\\ &\quad + \int_0^{\infty} \frac{16gC}{4C+3}(\rho_{u0,g1}(t)\,\mathrm{Im}\{\rho_{e0,u0}(t)\}\\ &\quad - \rho_{u0,u0}(t)\,\mathrm{Im}\{\rho_{e0,g1}(t)\})dt,\end{aligned} \tag{6}$$

where $\rho_{a,b} = \langle a|\rho(t)|b\rangle$, and $C = \frac{g^2}{\kappa\gamma_u}$.

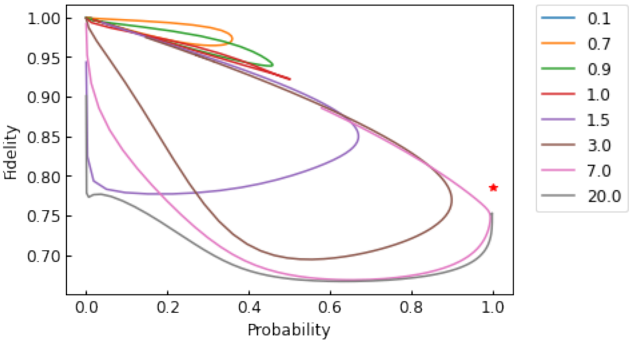
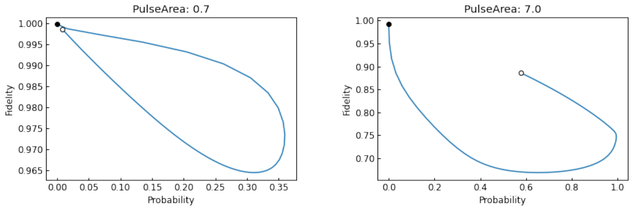


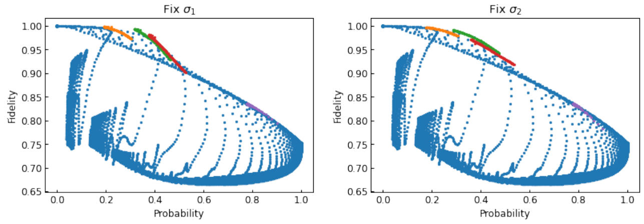
Figure 4: Classified trade-off with the pulse area. Each color shows the pulse area divided by 2π. The red star shows $1+\frac{2C}{4C+3}$ in Eq. (7)



(a)

(b)

Figure 5: The relationship of the pulse width and the trade-off curve, preserving the pulse area. Black dots show the result calculated by the shortest pulse, and White dots show the result calculated by the longest pulse. (a) Pulse area is 0.7. (b) Pulse area is 7.0.



(a)

(b)

Figure 6: The trade-off curve with asymmetric Gaussian pumping. Blue: symmetric Gaussian pulse, Orange: $\Omega_0 = 1.21$ and fixed $\sigma = 0.414$, Green: $\Omega_0 = 1.81$ and fixed $\sigma = 0.414$, Red: $\Omega_0 = 1.41$ and fixed $\sigma = 0.616$, and Purple: $\Omega_0 = 0.414$ and fixed $\sigma = 10.1$. (a) σ_1 is fixed. $\sigma_2 = 0.01\sigma_1\ 2\sigma_1$. (b) σ_2 is fixed. $\sigma_1 = 0.01\sigma_2\ 2\sigma_2$.

The first term on the right-hand side has a significant contribution, although we need the numerical calculations to determine the second and third terms. The first term is plotted as a red star in Fig. 4. The first term gives a good upper bound, even though it is a loos bound due to the derivation process. This result suggests that a Gaussian pulse is sufficient and that it is difficult to drastically improve the trade-off by optimizing the waveform further.

4 Conclusion

In this study, we consider the effect of the waveform of the pump pulse in the entanglement distribution scenario. First, we show the trade-off between F and P_{ex} for various Gaussian pulses.

Then, we find that the shorter pulse gives higher fidelity than the other pulses with the same probability when the P_{ex} is small. On the other hand, the longer pulse gives higher fidelity than the other pulses with the same probability when the P_{ex} is high.

Moreover, we show the Gaussian pump pulse is enough to distribute entanglement between distant ions by comparing the analytical and numerical results.

These results show how to determine the waveform of the pump pulse when generating entanglement between distant ions.

Finally, our results together with previous research [7, 8] help to design the optimized realistic experimental setup.

References

[1] C. D. Bruzewicz, *et al.*, *Appl. Phys. Rev.* 6, 021314, 2019.

[2] C. Monroe, *et al.*, *Phys. Rev. A* 89, 022317, 2014.

[3] L. M. Duan, *et al.*, *Phys. Rev. Lett.* 90, 90, 2003.

[4] D. Hucul, *et al.*, *Nat. Phys.* 11, 37, 2015.

[5] L. J. Stephenson, *et al.*, *Phys. Rev. Lett.* 124, 110501, 2020.

[6] V. Krutyanskiy, *et al.*, *Phys. Rev. Lett.* 130, 050803, 2023.

[7] T. Utsugi, *et al.*, *Phys. Rev. A* 106, 023712, 2022.

[8] S. Gao, Ph.D. thesis, University of Oxford, 2020.

[9] T. Walker, *et al.*, *Phys. Rev. A* 102, 032616, 2020.

[10] M. Meraner, *et al.*, *Phys. Rev. A* 102, 052614, 2020.

[11] A. Kuhn, *et al.*, *Appl. Phys. B* 69, 373, 1999.

[12] C. Di Fidio, *et al.*, *Phys. Rev. A* 65, 033825, 2002.

[13] H. Takahashi, *et al.*, *Phys. Rev. Lett.* 124, 013602, 2020.

[14] K. A. Fischer, *et al.*, *New J. Phys.* 18, 113053, 2016.

[15] C. Gardiner and P. Zoller *Quantum Noise*, Springer Berlin, Heidelberg, 2004

Quantum Agents are more energetically efficient at responding in real time

Jayne Thompson[1 2 *] Paul M. Riechers[3] Andrew J.P. Garner[3 4] Thomas J. Elliott[5] Mile Gu[3 6]

[1] *Institute of High Performance Computing, Agency for Science, Technology and Research (A*STAR), Singapore*
[2]*Horizon Quantum Computing, Singapore*
[3] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore*
[4] *Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Vienna, Austria*
[5] *Department of Mathematics, University of Manchester, Manchester, M13 9PL, UK*
[6] *Centre for Quantum Technologies, National University of Singapore*

Abstract. Agents often execute complex strategies - continually adapting their reactions to input stimuli to synergize with past actions. Here, we show there is a minimal energetic cost for classical agents to execute a given strategy, implying that they must dissipate a certain amount of heat with each decision. We prove that quantum agents can reduce this dissipation below classical limits and identify the necessary and sufficient conditions on a targeted strategy to guarantee this energetic quantum advantage. Our results point to new thermodynamic signatures of quantumness and a fundamental energetic advantage for agents leveraging quantum processing to enable complex adaptive behaviour.

Keywords: Quantum Thermodynamics, Agents

1 Introduction

Agents often enact complex strategies to survive in comptetive or resource scarce environments. From preditors chasing prey, to self driving cars, or a card counter trying to beat the house at black jack, each such agent must coordinate their actions over multiple time steps, predicating future decisions based not only on present stimuli but also on what has happened in the past. To do this an agent must track enviornmental stimuli over potentially long time frames, retaining information about the past which can be used to anticipate the consequences of future decisions and thus select rewarding future actions.

This coordination implies that agents must continuously expend energy to operate in complex environments. Consider an example of an agent Alice who is asked one of two questions at each time step: q_0 : 'Are you hungry?', or q_1 : 'Do you like electric sheep?'. Alice is required to exhibit the following behaviour: If the interrogator repeats the same question in two consecutive rounds, Alice's answers must agree; otherwise, her response to the second question must be random. It is clear that to win at this game Alice needs to retain two bits of information. The first bit being which question was asked last (so she can determine whether that question is being repeated), and the second being her previous answer. However at most one of these bits is ever reflected in Alice's future answers. Whats more if she gets asked two differing questions on consecutive rounds then both bits are discarded (as neither is relevant to generating a random answer), an operation which incurs intrinsic thermodynamic cost. Yet Alice can not play this simple game any more efficiently, there was an intrinsic cost to coordinating her actions. However this analysis assumes Alice is entirely classical.

Here we explore the potential for Alice to use quantum information processing to play such a game more efficiently than classically possible. We place fundamental bounds on the thermodynamic efficiency of any classical agent responding to a series of environmental stimuli in real time. We then isolate necessary and sufficient properties for a strategy to be executable by a quantum agent with improved energetic efficiency.

2 Strategies and Games

We describe adaptive agents as systems that interact with their environment at discrete timestepe $t \in \mathbb{Z}$. At each discrete point in time t the agent recieves and input stimuli or question $x_t \in \mathcal{X}$ from the environment and responds by generating an output response $y_t \in \mathcal{Y}$. Taking $t = 0$ as the present, we denote the past sequences of stimuli and actions as $\overleftarrow{x} := \ldots x_{-2}x_{-1}$ and $\overleftarrow{y} := \ldots y_{-2}y_{-1}$ respectively. For shorthand we denote the pair $z := (x, y)$, and similarly $\overleftarrow{z} := (\overleftarrow{x}, \overleftarrow{y})$ for the entire history.

A strategy $\mathcal{P}$ then specifies the desired statistical response to each possible future input sequence, conditioned on what has happened so far in the past. Mathematically, we adopt the framework of computational mechanics [1, 2], describing each strategy by $\mathcal{P} = \{P(Y_{0:K} = y_{0:K}|x_{0:K}, \overleftarrow{z})\}_{K>0}$, the probability the agent should answer $y_{0:K} = y_0, y_1, \ldots y_{K-1}$ when given a sequence of K future inputs $x_{0:K} = x_0, x_1, \ldots x_{K-1}$ for each natural number K and history $\overleftarrow{z}$. Note that while this definition specifies the random variable Y_t that governs each y_t, it makes no such specification on x_t. This is because a strategy constitutes a promise about the agent's response for all possible future x_t, regardless of how it is distributed. We thus adopt similar conventions to Bell tests, such that the agent cannot gain any information about future inputs based on past inputs. Therefore, we take x_t to be independent identically-distributed with Shannon entropy h_x. Let $\mathbf{X}_t$ and $\mathbf{Y}_t$ denote the physical systems that respectively encode x_t and y_t. We assume the interigator preconfigures the state of $\mathbf{X}_t$ to encode the question x_t before passing it to the agent, meanwhile the output tape is assumed to be initialized in a maximally mixed default state

[*] thompson.jayne2@gmail.com

governed by Y_{dflt} with entropy $h_{\text{dflt}} = \log_2 |\mathcal{Y}|$.

To understand the cost of this real-time response, we will also consider quasi real-time agents that process L-inputs at a time. Such agents are able to take in (length L) segments of the tapes $\mathbf{X}_{0:L} = \mathbf{X}_0 ... \mathbf{X}_{L-1}$ and $\mathbf{Y}_{0:L} = \mathbf{Y}_0 ... \mathbf{Y}_{L-1}$, and jointly transform them to generate the next L answers simultaneously. We call such agents L-stride, in analogy to a related automata in the literature of pattern machines.

To play the game over multiple rounds the agent must be able to retain relevant information about what has happened in the past in order to inform future decisions. To do this the agent must record relevant past information by configuring the initial state of a memory register $\mathbf{M}$ to some function of the past $f(\overleftarrow{z})$. We assume the agent is causal (i.e. the memory state is always a function of the past), such that someone who has access to the entire past knows as much about the future as the agent does. We also assume the agent is stationary, which is to say that if there is any internal clock dictating the behaviour of the agent then it has to be recorded in the state $f(\overleftarrow{z})$, for proper resource accounting.

Generating statistically correct behaviour at each time step then requires the L-stride agent to implement some physical process $\mathcal{T}^{(L)}$ that couples the memory $\mathbf{M}$ to $\mathbf{X}_{0:L}$ and $\mathbf{Y}_{0:L}$, while

1. Changing the state of $\mathbf{Y}_{0:L}$ to encode the output $y_{0:L}$ with probability $P(Y_{0:L}|\overleftarrow{z}, x_{0:L})$; and

2. Updating the state of $\mathbf{M}$ to one consistent with the new history $\overleftarrow{z}\, z_{0:L}$ (i.e, $f(\overleftarrow{z}\, z_{0:L})$).

3. Leaving the state of $\mathbf{X}_{0:L}$ unchanged.

Here the first two conditions ensure that the agent executes statistically correct behaviour meanwhile the last condition forbids the agent from extracting energy from the incoming tape of questions (ensuring we are characterizing the energetic cost of *responding* to environmental stimuli). These conditions remain true regaurdless of whether the agent is classical or quantum, ensuring that both counterparts are playing by the same rules.

3 Work cost of executing a strategy

To analyse the work cost of executing a complex strategy we treat $\mathcal{T}^{(L)}$ as a physical process that interacts $\mathbf{X}_{0:L}, \mathbf{Y}_{0:L}$ and $\mathbf{M}$, imprinting the final states of applying the update map onto each of these systems. This process takes place in contact with a heat bath at temperature T. We also assume the Hamiltonians governing these systems are degenerate at the beginning and end of the protocol.

Landauer's principle then puts immediate bounds on the ultimate thermal efficiency of such an agent, implying that the energetic cost of realizing the update map $\mathcal{T}^{(L)}$ is bounded below by the difference in entropy of the inputs vs. outputs of the channel. Thus an L-stride agent requires $w^{(L)}$ units of work per symbol it emits, where Landauer's principle implies

$$\frac{w^{(L)}}{kT \ln 2} \geq h_{\text{dflt}} + \frac{1}{L}[I(Z_{0:L}, M_L) - H(Y_{0:L}|X_{0:L})]. \quad (1)$$

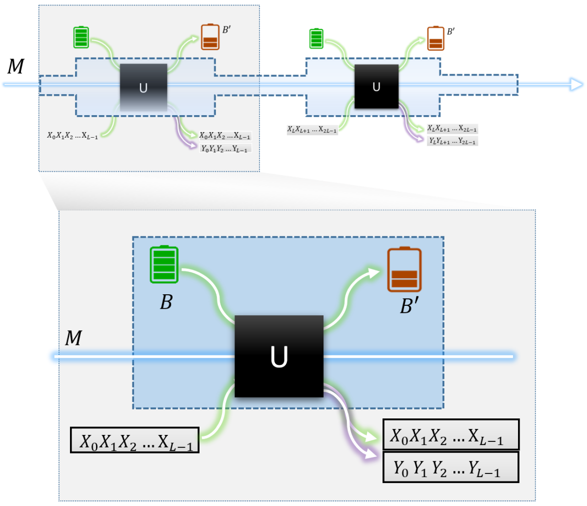

M
U
B'
$X_0X_1X_2 \ldots X_{L-1}$
$Y_0Y_1Y_2 \ldots Y_{L-1}$
U
B'
$X_LX_{L+1} \ldots X_{2L-1}$
$X_LX_{L+1} \ldots X_{2L-1}$
$Y_LY_{L+1} \ldots Y_{2L-1}$
B
B'
M
U
$X_0X_1X_2 \ldots X_{L-1}$
$X_0X_1X_2 \ldots X_{L-1}$
$Y_0\,Y_1\,Y_2 \ldots Y_{L-1}$

Figure 1: The energetic cost of a quantum agent executing a policy map $\mathcal{T}^{(L)}$ can be characterized using information batteries. To execute $\mathcal{T}^{(L)}$, an agent harnesses a information battery B comprised of $d \gg 1$ qubits of which $\lambda \gg 1$ are maximally mixed and all others are pure. U then represents a Stinespring dilation that couples B with the agent memory, and the input tape $\mathbf{X}_{t:t+L}$ (here $t = 0$ for illustrative purposes). At end of the operation, the battery will be depleted and ejected. The energy cost needed to reset this battery (while coupling it to a heat bath at some temperature T) then gives the single-shot work cost in implementing $T^{(L)}$.

Here $I(Z_{0:L}, M_L)$ is the mutual information between the joint output state of the two tapes $\mathbf{X}_{0:L}, \mathbf{Y}_{0:L}$, and the terminal state of the memory system $\mathbf{M}$ at time $t = L$, and $H(Y_{0:L}|X_{0:L})$ the conditional entropy.

4 Classical Agents

Classical agents operate with classical memory, whereby $\mathcal{T}^{(L)}$ is a classical stochastic map. It is then possible to saturate Landauer's bound using isothermal channels and changing energy landscapes [3]. Thus the energy-minimal agent should choose a memory encoding $f(\overleftarrow{z})$ that minimizes $I(Z_{0:L}, M_L)$ – that is, the encodings that store minimal information about the past.

This minimum is obtained when the memory states are associated with an encoding function ϵ such that $\epsilon(\overleftarrow{z}) = \epsilon(\overleftarrow{z}')$ if and only if $P(Y_{0:K} = y_{0:K}|x_{0:K}, \overleftarrow{z}) = P(Y_{0:K} = y_{0:K}|x_{0:K}, \overleftarrow{z}')$ for all $K \in \mathbb{N}$ and future input choices $x_{0:K}$ [2],. The resulting agent has a special significance in classical complexity science literature where it is known as the ϵ-transducer and regarded as the memory minimal automata capable of executing a given strategy $\mathcal{P}$. This gives an ultimate classical limit on the thermodynamic cost of responding to input stimuli of

$$\frac{w_c^{(L)}}{kT \ln 2} = h_{\text{dflt}} + \frac{1}{L}\left[I(Z_{0:L}, S_L) - H(Y_{0:L}|X_{0:L})\right] \quad (2)$$

where S_L is the random variable governing the memory state of the ϵ-transducer at time $t = L$.

Associated with this cost we can now derive an ultimate limit on the extra cost of having to respond to questions one at a time, by comiting the answer y_t at each time step before x_{t+1} is asked, vs. the cost of responding to all future questions at once. In particular we find that (for i.i.d. inputs x_t) the energetic cost of real time response for a classical agent is

$$\begin{aligned} \epsilon_{\mathrm{rt}} &= w_c^{(1)} - \lim_{L\to\infty} w_c^{(L)} \\ &= kT \ln 2[I(Z_0, S_1) - I(S_0, Z_0)]. \end{aligned} \tag{3}$$

In the context of our introductory example of Alice the agent who is asked one of two questions at each time step, $I(Z_0, S_1) = 2$ as the agent records 2 bits about the past time-step, but $I(S_0, Z_0) = 0.5$ as half the time one of these bits is relevant. Thus here the classical cost of real time response is $\epsilon_{\mathrm{rt}} = 1.5kT \ln 2$.

5 Quantum Agents

Quantum agents allow for a quantum memory $\mathbf{M}$ that can store quantum states [4]. Our quantum agents operate by associating each memory state of the ϵ-transducer, s_k, with some quantum memory state $|\sigma_k\rangle$.

Once configured in the appropriate memory state $\epsilon_q(\overleftarrow{z}) = |\sigma_j\rangle\langle\sigma_j|$, a L-stride quantum agent implements $\mathcal{T}^{(L)}$ by applying a unitary process U. As depicted in Fig. [1], this unitary acts jointly on an information battery system $\mathbf{B}$ initialised in some default pure state $|0\rangle$, alongside the initial memory state $|\sigma_k\rangle$ and input-output tapes. The interaction transforms the two tapes to encode the input-output response sequence $z_{0:L}$, while suitably updating the memory to record this output. This results in a joint state

$$\sum_{y_{0:L},k} \sqrt{T_{jk}^{y_{0:L}|x_{0:L}}} |\sigma_k\rangle_M |z_{0:L}\rangle_{Z_{0:L}} |\psi(z_{0:L}, i)\rangle_B \tag{4}$$

where $|\psi(z_{0:L}, i)\rangle_B$ are junk states accumulated on the battery register [4]. Thermodynamic cost is then incurred due to the discarding of the depleted battery register [5, 6].

The resulting work rate $w_q^{(L)}$ associated with discarding the information battery register can saturate Eq. (2). This implies that the energetic advantage of a quantum agent over a classical agent per time-step is

$$w_c^{(L)} - w_q^{(L)} = \frac{kT \ln 2}{L}[I(Z_{0:L}, S_L) - I(Z_{0:L}, M_L)], \tag{5}$$

where $I(Z_{0:L}, M_L)$ represents the amount of information, our quantum agent retains about the past L exchanges.

In general we show that the quantum agent is always more energetically efficient than its optimal classical counterpart such that

$$\Delta w^{(L)} = w_c^{(L)} - w_q^{(L)} > 0 \tag{6}$$

whenever there is step-wise intrinsic irreversibly in the optimal classical agent's internal map $\mathcal{T}$. Specifically this step-wise irreversability is characterized by the existence of two memory states s_i, s_j of the ϵ-transducer, such that irrespective of the sequence of future inputs $\overrightarrow{x}$ it is impossible to perfectly discriminate these two starting states based on future output responses, i.e. $\sum_{\overrightarrow{y}} P(\overrightarrow{y}|\overrightarrow{x}, s_i)P(\overrightarrow{y}|\overrightarrow{x}, s_j) > 0$.

Indeed this knocks on directly into energetic savings in the cost of real time response for a quantum agent. In particular the difference between the quantum and classical real time response costs of responding to i.i.d. x_t inputs, is

$$\Delta\epsilon_{\mathrm{rt}} = kT \ln 2[I(Z_0, S_1) - I(Z_0, M_1)]. \tag{7}$$

To see this savings in action we return to our initial example of Alice the agent under integration by an adversary who is fond of electric sheep. Quantum mechanical Alice would be capable of generating statistically correct output behaviour by encoding all past information into a single qubit. If she was previously asked q_0 she simply prepares her memory in either $|0\rangle$ or $|1\rangle$ depending on her last answer. Analogously if she was asked q_1 she configures her memory in one of the two Pauli X eigenstates $|+\rangle$ or $|-\rangle$ depending on her last answer. She thus retains $I(Z_0, M_1) = 1$ qubits of information about the past. Afterwards correct answers can always be generated by associating q_0 with a Pauli Z basis measurement, and q_1 with a measurement in the Pauli X eigenbasis. We see that already in this simple case there is a quantum advantage in the cost of real time response of $\Delta\epsilon_{\mathrm{rt}} = kT \ln 2[2 - 1] = kT \ln 2$ units of work.

6 Discussion

Agents often need to commit their answers in real time. Whether it is a black jack player who is expected to play a card each round, or a preditor chasing prey who needs to swerve immediately in response to their target changing direction, most agents need to make decisions on the spot. We have discovered that there is a quantum advantage in the amount of energy an agent must invest to be capable of making decisions in real time.

This advantage complements recent work on the thermodynamics of energy harvesting, where an agent's target is to anticipate the distribution over upcoming input strings so as to extract energy from input stimuli. In contrast here we characterize directly the thermodynamic cost of generating output responses, encountering a situation similar to quantum random access codes. In the sense that a quantum system which may be asked one of two different questions in the future but has not the capacity to record the answers of both questions, almost seems to postpone making a decision about which answer it will commit to memory. Analogously our quantum agents partially avoiding committing and potentially later discarding the irrelevant answer, leading to thermodynamic enhancement.

References

[1] Crutchfield, J.P. and Young, K. Inferring statistical complexity. Physical Review Letters, 63(2), p.105, 1989

[2] Barnett, N. and Crutchfield, J.P. omputational mechanics of input–output processes: Structured transformations and the ϵ-transducer. Journal of Statistical Physics, 161(2), pp.404-451. (2015)

[3] Boyd, A.B., Mandal, D. and Crutchfield, J.P. Thermodynamics of modularity: Structural costs beyond the Landauer bound. Physical Review X, 8(3), p.031036. (2018)

[4] Elliott, T.J., Gu, M., Garner, A.J. and Thompson, J. Quantum adaptive agents with efficient long-term memories. Physical Review X, 12(1), p.011007. (2022)

[5] Faist, P., Dupuis, F., Oppenheim, J. and Renner, R. *The minimal work cost of information processing*. Nature Communications, 6(1), p.7669. 2015

[6] Rio, L.D., Åberg, J., Renner, R., Dahlsten, O. and Vedral, V., The thermodynamic meaning of negative entropy. Nature, 474(7349), pp.61-63. (2011)

Spatiotemporal Classification of Quantum Correlations

Minjeong Song[1] * Varun Narasimhachar[1] † Bartosz Regula[2] Thomas J. Elliott[3 4]
Mile Gu[1 5 6] ‡

[1] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, 637371, Singapore*
[2] *Department of Physics, The University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*
[3] *Department of Physics & Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom*
[4] *Department of Mathematics, University of Manchester, Manchester M13 9PL, United Kingdom*
[5] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543, Singapore*
[6] *MajuLab, CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore 117543, Singapore*

Abstract. We give a spatiotemporal classification of general quantum correlations in spacetime. Based on this, we answer a fundamental question as to when the observed bipartite correlation from local measurements is (a)temporal or (a)spatial. To understand (a)temporality of correlations, we introduce atemporality—an efficiently computable real-valued property that is zero if and only if the correlation is temporal. Atemporality is asymmetric under time reversal; this includes cases of quantum correlations one way temporal but not the other. Surprisingly, we find that some entangled states are temporally replicable, although we prove that entanglement is upper-bounded to be temporally replicable.

Keywords: Quantum causality, Temporal quantum correlations, Entanglement

1 Introduction

Consider the scenario in which Alice and Bob are each in their own laboratory. In each round, they are each given a respective qubit labelled A and B. The qubits are prepared in the same way in each round. Such a preparation scheme may be

- *Spatially distributed*, such that A and B correspond to two arms of some bipartite state ρ_{AB} (see Fig. 1a).
- *Temporally distributed*, such that B is the output of A subject to some fixed quantum channel (completely-positive trace-preserving map) $\mathcal{E}$, or vice versa (see Fig. 1b).
- Neither purely spatially nor temporally distributed, such as when A evolves to B via *non-Markovian* evolution [1], or, more generally, when A and B are related by some general *process matrices* [2].

Let $\boldsymbol{\sigma}_0 = \mathbf{I}$, $\boldsymbol{\sigma}_1 = \mathbf{X}$, $\boldsymbol{\sigma}_2 = \mathbf{Y}$, $\boldsymbol{\sigma}_3 = \mathbf{Z}$ by the identity and the three standard Pauli operators. Let $\Pr(x, y|a, b)$ then denote the probability of Alice getting outcome x and Bob getting outcome y when Alice chooses to measure in basis $\boldsymbol{\sigma}_a$ and Bob in $\boldsymbol{\sigma}_b$. Alice and Bob do not perform any other interventions. By choosing appropriate Pauli measurements over a large number of rounds, Alice and Bob would then be able to determine the expected values $\langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle = \sum_{x,y} xy \Pr(x, y|a, b)$ describing how their measurement outcomes correlated in various Pauli basis. Alice and Bob then pass this information to us. From this information, what can we conclude about the causal mechanism behind the preparation of A and B?

We employ the pseudo-density operator formalism to tackle this question. Given Pauli correlations $\langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle$ for each $a, b \in \{0, 1, 2, 3\}$, we can describe the information received concisely via the pseudo-density operator (PDO)

$$\mathbf{R}_{AB} \equiv \sum_{a,b=0}^{3} \frac{\langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle}{4} \boldsymbol{\sigma}_a \otimes \boldsymbol{\sigma}_b .$$

The PDO contains all the information in $\langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle$, since the latter can be retrieved directly by noting $\langle \boldsymbol{\sigma}_a, \boldsymbol{\sigma}_b \rangle = \mathrm{tr}[\mathbf{R}_{AB}(\boldsymbol{\sigma}_a \otimes \boldsymbol{\sigma}_b)]$. Thus our capacity to infer causal mechanisms from the Pauli correlations coincides with our capacity to infer causal mechanisms from the corresponding PDO. Indeed, it is easy to see that when qubits A and B are spatially distributed, the definition of $\mathbf{R}_{AB}$ coincides with that of a standard density operator. As such, PDOs were proposed as a generalization of standard density operators to quantum measurements made in space-time [3].

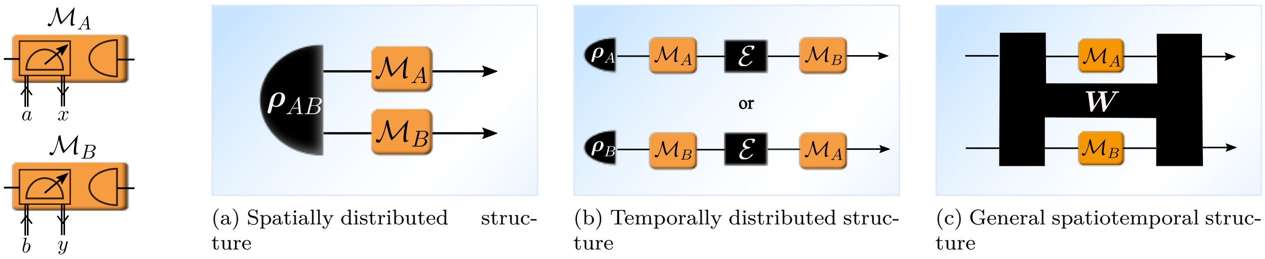
2 Main Results

Spatial-temporal compatibility. We introduce two distinct criteria on PDOs: We say that $\mathbf{R}_{AB}$ is *spatially compatible*, or belongs to $\mathcal{S}$ if its statistics can be generated via a spatial distribution mechanism. Similarly, we say that $\mathbf{R}_{AB}$ is *temporally compatible*, or belongs to $\mathcal{T}$, if its statistics can be generated via a temporal distribution mechanism [1]. PDOs that lie outside of $\mathcal{S}$ are referred to as *aspatial*, and those that lie outside of $\mathcal{T}$ are referred to as *atemporal.*

We then divide the set of all PDOs into four separate classes based on their spatial-temporal compatibility: Those that (a) lie in $\mathcal{S}$ and $\mathcal{T}$ and are thus spatial-

[*] song.at.qit@gmail.com
[†] varun.achar@gmail.com
[‡] mgu@quantumcomplexity.org

[1] We will often use the terms *spatial* and *temporal* for brevity, but we stress that they only mean *compatible with* a spatially or temporally distributed structure.



(a) Spatially distributed structure

(b) Temporally distributed structure

(c) General spatiotemporal structure

Figure 1: Alice and Bob in their respective laboratories make projective measurements $\mathcal{M}_A, \mathcal{M}_B$ on each of their quantum systems, and $\mathcal{M}_A, \mathcal{M}_B$ are depicted by the orange rounded boxes with expanded view on the left. Above three quantum circuit–like diagrams illustrate possible spatiotemporal structures. The black boxes represent background quantum processes not in the observers' control. The index $a \in \{0, 1, 2, 3\}$ denotes a measurement basis choice, with the value 0 indicating no measurement and the rest Pauli basis measurements with outcome x and post-measurement state $\mathbf{\Pi}_{x|a}$; likewise for b, y.

temporal compatible, (b) lie in $\mathcal{S}$ but not $\mathcal{T}$ and thus rule out purely temporal distribution mechanisms, (c) lie in $\mathcal{T}$ but not $\mathcal{S}$ and thus rule out purely spatial distribution mechanisms and (d) those that lie outside $\mathcal{S}$ and $\mathcal{T}$ that cannot be explained by either purely spatial or temporal distribution scheme but rather, rely on a more complicated combination of spatial and temporal mechanism. We cannot infer anything conclusive about the causal mechanism for those that lie within (a), in alignment with classical statistical theory. Quantum correlations, however, enable PDOs in each of (b), (c) and (d), where certain causal mechanisms can be ruled out.

To better understand what PDOs lie within each subclass, we need necessary and sufficient criterion for aspatiality and atemporality. Past studies of causality have focused on the former, showing that the negativity of $\mathbf{R}_{AB}$ is a necessary and sufficient for aspatiality [3]. We will derive analogous conditions for when a PDO is atemporal, and thus build a full picture of spatial-temporal compatibility. We will also provide some examples for each subclass, from which we can also show that $\mathcal{T}$ does not form a convex set.

Certifying Atemporality. Given $\mathbf{R}_{AB}$, our goal is to propose atemporality indicators that rule out its compatibility of temporal distribution mechanisms. To do this, find it useful to first consider asymmetric *atemporality*: The forward atemporality $\overrightarrow{f}$ and the reverse atemporality $\overleftarrow{f}$. $\overrightarrow{f}(\mathbf{R}_{AB})$ is zero if and only if $\mathbf{R}_{AB}$ has statistics consistent with a temporal distribution mechanism from A to B, whereas $\overleftarrow{f}$ is zero if and only if $\mathbf{R}_{AB}$ has statistics consistent with a temporal distribution mechanism from B to A. Together they naturally induce a general *atemporality* measure $f = \min(\overrightarrow{f}, \overleftarrow{f})$ that is zero if and only if $\mathbf{R}_{AB}$ lies in $\mathcal{T}$.

Our approach involves introducing *pseudo-channels*, a temporal analogue of pseudo-density operators. If a given PDO $\mathbf{R}_{AB}$ is induced from a temporal distribution mechanism from A to B with a quantum channel $\overrightarrow{\Lambda}$ and an initial state $\boldsymbol{\rho}_A$, it is known in Ref. [4, 5] that $\mathbf{R}_{AB}$ can be explicitly written as

$$\mathbf{R}_{AB} = \left(\mathcal{I}_A \otimes \overrightarrow{\Lambda}\right) \mathbf{K}_{AB} \tag{1}$$

where $\mathbf{K}_{AB} \equiv \left\{\boldsymbol{\rho}_A \otimes \frac{\mathbf{I}_B}{2}, \mathbf{S}_{AB}\right\}$ with $\mathcal{I}$ the identity channel, $\{\cdot,\cdot\}$ the anti-commutator, $\mathbf{I}$ the identity operator, and $\mathbf{S}$ the swap operator. It is noteworthy that a physical channel is represented by a completely positive (CP) and trace-preserving (TP) map. When $\mathbf{R}_{AB}$ is incompatible with such a temporal distribution mechanism, no such CPTP maps exist. However, we can drop the completepositivity (CP) requirements on $\overrightarrow{\Lambda}$, that is, $\overrightarrow{\Lambda}$ remains trace-preserving, Hermiticity-preserving and linear, but is no longer CP. This allows us to interpret *any* spatiotemporal correlations as if $\mathbf{R}_{AB}$ results from this map $\overrightarrow{\Lambda}$ acting on A to generate B. We say that $\overrightarrow{\Lambda}$ is a *forward pseudo-channel* compatible with $\mathbf{R}_{AB}$. Similarly, we have a *reverse pseudo-channel* by interchanging the roles of A and B.

Recall that the Choi–Jamiołkowski isomorphism [6] allows us to represent a linear map $\overrightarrow{\Lambda}$ by its corresponding Choi state $\boldsymbol{\chi}_{\overrightarrow{\Lambda}}$, and states that $\boldsymbol{\chi}_{\overrightarrow{\Lambda}}$ is positive semidefinite if and only if $\overrightarrow{\Lambda}$ is CP. Thus negativity $\mathcal{N}(\boldsymbol{\chi}_{\overrightarrow{\Lambda}})$ of the Choi state (sum over all absolute value of its negative eigenvalues) provides a necessary and sufficient indicator of *nonphysicality* of $\overrightarrow{\Lambda}$. These provide a natural means to define $\overrightarrow{f}$ (similarly $\overleftarrow{f}$, and thus f).

The minimal nonphysicality of such a channel then motivates our definition for *forward atemporality*:

$$\overrightarrow{f}(\mathbf{R}_{AB}) \equiv \min_{\overrightarrow{\Lambda}} \mathcal{N}(\boldsymbol{\chi}_{\overrightarrow{\Lambda}}), \tag{2}$$

where the minimization is over all forward pseudochannels $\overrightarrow{\Lambda}$ that is compatible with $\mathbf{R}_{AB}$ (i.e., those that satisfy Eq. (1)). Similarly, we define the *reverse atemporality* $\overleftarrow{f}(\mathbf{R}_{AB})$ by reverse pseudo-channels, and thus the general atemporality $f \equiv \min(\overrightarrow{f}, \overleftarrow{f})$.

That f is a necessary and sufficient conditions for atemporality is guaranteed by the following theorem:

Theorem 1. Any 2-qubit PDO $\mathbf{R}_{AB}$ has at least one compatible forward (reverse) pseudo channel. Moreover, if the marginal $\mathrm{tr}_B\,\mathbf{R}_{AB}$ (respectively, $\mathrm{tr}_A\,\mathbf{R}_{AB}$) has full rank, $\mathbf{R}_{AB}$ has a *unique* forward (respectively, reverse) pseudo channel.

Moreover, we have a systematic method to determine such compatible pseudo-channels for any given $\mathbf{R}_{AB}$. When $\mathbf{R}_{AB}$ are full rank marginals, the algorithm is particularly simple:

Algorithm 1 (Forward) Choi state of pseudo channel construction

```
Require: 2-qubit PDO $\mathbf{R}_{AB}$
  $\boldsymbol{\rho}_A \leftarrow \mathrm{tr}_B\, \mathbf{R}_{AB}$
  $\mathbf{L} \leftarrow \left(\boldsymbol{\rho}_A - \frac{\mathbf{I}}{2}\right) \otimes \mathrm{tr}_A\left[\left(\frac{1}{2}\boldsymbol{\rho}_A^{-1} \otimes \mathbf{I}\right)\mathbf{R}_{AB}\right] + \frac{\mathbf{I}}{2} \otimes \mathrm{tr}_A\left[\left(\left(\mathbf{I} - \frac{1}{2}\boldsymbol{\rho}_A^{-1}\right) \otimes \mathbf{I}\right)\mathbf{R}_{AB}\right]$
  $\boldsymbol{\mathcal{X}} \leftarrow (T \otimes \mathcal{I})(\mathbf{R}_{AB} - \mathbf{L})$   ▷ $T$ denotes the transpose map
  return $\boldsymbol{\mathcal{X}}$
```

When the marginal $\boldsymbol{\rho}_A$ is rank-deficient (i.e. some pure state $|\phi\rangle\langle\phi|$), the pseudo-channels compatible with $\mathbf{R}_{AB}$ is no longer unique. This is because any such causal interpretation corresponds to Alice initiating her system in $|\phi\rangle$, leaving us free to choose how Λ acts the the perpendicular state $|\phi^{\perp}\rangle$. We show that atemporality can be computed effectively by semidefinite programming, even in rank-deficient cases.

Atemporality and quantum correlations. While atemporality can be applied to all PDOs, it is particularly interesting to investigate the measurement in the context of correlations that are also compatible with spatial distribution mechanism i.e., those that lie $\mathcal{S}$, whose PDOs are standard density operators.

We prove that our atemporality measure coincides with the entanglement negativity [7] whenever $\mathbf{R} \in \mathcal{S}$ is pure or its marginals are given by the maximally mixed state. Given this, one may expect that atemporality and entanglement are equivalent. However, we find this to be false:

Observation 1. Atemporality and entanglement do not always coincide. There exist entangled states that are temporally compatible.

An example is the parameterized family of biased Werner states $\boldsymbol{\rho}_{p,q} \equiv (1-p)\boldsymbol{\rho}_q^{\text{Werner}} + p\,|00\rangle\langle 00|$, achieved by mixing a standard Werner state $\boldsymbol{\rho}_q^{\text{Werner}}$ [8] with the state $|00\rangle\langle 00|$. For example, $\boldsymbol{\rho}_{1/2,1/4}$ has zero atemporality, by an entanglement negativity of approximately 0.0087 (see Fig. 2). Nevertheless, we also observe that sufficiently strong entanglement does guarantee atemporality. We indeed prove the following:

Observation 2. Any temporally compatible state must have entanglement negativity of at most $\frac{1}{2}(\sqrt{2}-1)$.

Thus we see that atemporality is strongly suggestive of strong quantum correlations. Our results suggest that two concepts are heavily correlated but not the same - with atemporality looking to be a stronger notion of nonclassical correlations than entanglement.

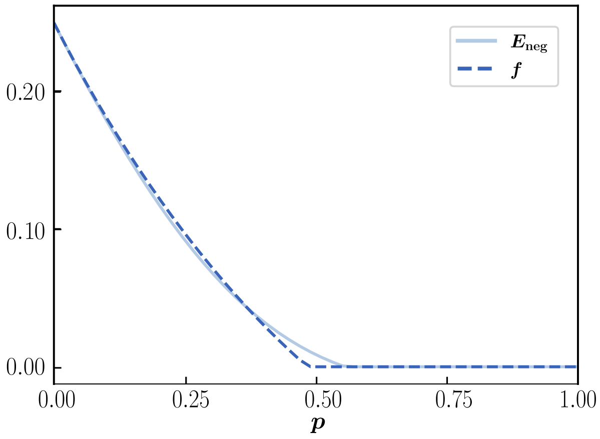


Figure 2: The graph of entanglement negativity E_{neg}, atemporality of biased Werner states $\boldsymbol{\rho}_{p,q=1/4}^{\text{B.W}}$ parameterized by $p \in [0,1]$, are plotted by the light blue solid line, the blue dashed line, respectively.

Time-reversal asymmetry. Another noteworthy property of atemporality is its asymmetry under time reversal. Some PDOs may be temporal from A to B (i.e., *forward-temporal*), but not the other way; others yet may admit a temporal interpretation only from B to A. We find the example of forward-temporal PDOs but not reverse-temporal, and Fig. 3 illustrates their non-zero reverse atemporality.

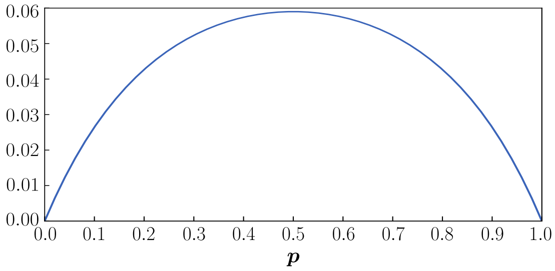


Figure 3: The graph of the reverse atemporality $\overleftarrow{f}$ for forward-temporal PDOs obtained from a forward temporal process, against the parameter p, where the temporal process is fixed as follows. Suppose A is prepared in a state $\boldsymbol{\rho}_A = \frac{1}{4}|+\rangle\langle+| + \frac{3}{4}|-\rangle\langle-|$, and after Alice's observation, evolves via a quantum channel $\mathcal{E}_p : \boldsymbol{\rho}_A \mapsto p\boldsymbol{\rho}_A + (1-p)\mathbf{Z}\boldsymbol{\rho}_A\mathbf{Z}^{\dagger}$ for $p \in [0,1]$. By construction, $\mathbf{R}_{AB} \in \overrightarrow{\mathcal{T}}$; however, for $p \in (0,1)$, $\mathbf{R}_{AB} \notin \overleftarrow{\mathcal{T}}$.

3 Discussion

Quantum correlations differ crucially from classical counterparts in that they can be fundamentally incompatible with certain underlying causal mechanisms. Prior work showed that the correlations between various Pauli measurements on two qubits A and B can be *aspatial* – such that they cannot be explained purely by a common cause. In this work, we complete this picture by introducing *atemporality*, the situation where such correlations can only be explained when common causes are involved. Overall, our worked catalyzes the development a framework to classify general quantum correlations based their compatibility to *spatial* and *temporal* causal mechanisms.

References

[1] Simon Milz and Kavan Modi. Quantum stochastic processes and quantum non-markovian phenomena. *PRX Quantum*, 2:030201, Jul 2021.

[2] Ognyan Oreshkov, Fabio Costa, and Caslav Brukner. Quantum correlations with no causal order. *Nature communications*, 3(1):1–8, 2012.

[3] Joseph F Fitzsimons, Jonathan A Jones, and Vlatko Vedral. Quantum correlations which imply causation. *Scientific reports*, 5:18281, 2015.

[4] Zhikuan Zhao, Robert Pisarczyk, Jayne Thompson, Mile Gu, Vlatko Vedral, and Joseph F Fitzsimons. Geometry of quantum correlations in space-time. *Physical Review A*, 98(5):052312, 2018.

[5] Dominic Horsman, Chris Heunen, Matthew F. Pusey, Jonathan Barrett, and Robert W. Spekkens. Can a quantum state over time resemble a quantum state at a single time? *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 473(2205):20170395, 2017.

[6] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[7] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65:032314, Feb 2002.

[8] The Werner states are a well-studied class of quantum states on bipartite systems AB with subsystems of equal dimensions. A Werner state $\boldsymbol{\rho}_{AB}$ satisfies $\boldsymbol{\rho}_{AB} = (\mathbf{U} \otimes \mathbf{U})\boldsymbol{\rho}_{AB}(\mathbf{U}^\dagger \otimes \mathbf{U}^\dagger)$ for all unitaries $\mathbf{U}$. For two qubit case, these states can be represented parametrically as $\boldsymbol{\rho}_q^{\text{Werner}} = \frac{q}{3}\mathbf{P}_s + (1-q)\mathbf{P}_a$, where $q \in [0, 1]$ and $\mathbf{P}_s, \mathbf{P}_a$ are the projectors onto the symmetric and the antisymmetric space, respectively.

Experimental demonstration of the dynamics of quantum coherence evolving under a PT-symmetric Hamiltonian on an NMR quantum processor

Akanksha Gautam[1] [*] Kavita Dorai[1] [†] Arvind[1 2] [‡]

[1] *Indian Institute of Science, Education and Research, Mohali, India*
[2] *Punjabi University, Patiala, Punjab, India*

Abstract. In this work, we theoretically investigate the dynamics of quantum coherence (total, global and local coherence) present in maximally entangled bipartite and tripartite states in both the unbroken and the broken phase of PT-symmetry and at an exceptional point when one qubit is acted upon by the local PT-symmetric Hamiltonian. Our results indicate that quantum coherence behaves differently under PT-symmetry, when the dimensionality of the quantum system is increased. The dynamics of quantum coherence present in the maximally entangled bipartite state is experimentally verified by implementing the PT-symmetric Hamiltonian on an NMR quantum processor.

Keywords: PT-Symmetric Hamiltonian, Quantum Coherence, NMR

1 Introduction

- Theoretically, we study the dynamics of quantum coherence (total coherence, global coherence and local coherence) evolving under a local PT-symmetric Hamiltonian in maximally entangled bipartite and tripartite states.

- Our results indicate that quantum coherence in the bipartite state oscillates in the unbroken phase regime of the PT-symmetric Hamiltonian. Interestingly, in the broken phase regime, while the global coherence decays exponentially, the local and total coherences enter a "freezing" regime where they attain a stable value over time.

- A similar pattern is observed for the dynamics of total and local coherences in the maximally entangled tripartite state, while the dynamics of global coherence in this state differs from that of the bipartite state.

- Experimentally, these results were validated for a maximally entangled bipartite state on a three-qubit nuclear magnetic resonance (NMR) quantum processor, with one of the qubits acting as an ancilla.

PT-Symmetric Hamiltonian: The PT-symmetric Hamiltonian for a single qubit can be written as:

$$H_{PT} = \sigma_x + ir\sigma_z \tag{1}$$

where, $r > 0$ is the amount of non-Hermicity and $\sigma_x, \sigma_y, \sigma_z$ are Pauli matrices. The energy gap of PT-symmetric Hamiltonian is $2\sqrt{1-r^2}$ and it's eigenvalues are $\pm\sqrt{1-r^2}$ where for $\mid r \mid < 1$, the eigenvalues are positive that means PT-symmetry is unbroken and for $\mid r \mid > 1$, the eigenvalues become complex that leads to broken PT-symmetry. The Hamiltonian has an exceptional point at $\mid r \mid = 1$ where both eigenvalues as well as eigenvectors coalesce.

Dynamics of quantum coherence: Different measures of basis dependent quantum coherence have been proposed such as l1-norm and relative entropy. Based on these measures, various types of quantum coherences have been defined in multipartite systems such as global coherence, local coherence and total coherence:

- **Total Quantum Coherence:** The relative entropy of total coherence is defined as

$$\begin{aligned} C_T(\rho) &= min_{\sigma\epsilon I} S(\rho \parallel \sigma) \\ &= S(\rho_d) - S(\rho) \end{aligned} \tag{2}$$

where $S(\rho) = -tr(\rho log_2 \rho)$ is the Von Neumann entropy of ρ, and ρ_d is the matrix of ρ with all off-diagonal terms set to zero in the basis $|i\rangle$.

- **Local Coherence:** Quantum coherence which is localized on each qubit of the entire system is called local coherence. It can also be defined in terms of relative entropy as

$$\begin{aligned} C_L(\rho) &= min_{\sigma\epsilon I} S(\delta(\rho) \parallel \sigma) \\ &= S(\delta_d(\rho)) - S(\delta(\rho)) \end{aligned} \tag{3}$$

where $\delta(\rho) = \rho_1 \otimes \rho_2$ ($\delta(\rho) = \rho_1 \otimes \rho_2 \otimes \rho_3$) for two qubits system (three qubits system) and $\rho_1 = Tr_2\rho_{12}$ ($\rho_1 = Tr_{23}\rho_{123}$) is the single-qubit reduced density matrix.

- **Global Coherence:** Quantum coherence that originates due to the collective nature of the whole system is called global coherence

$$C_G(\rho) = C_T(\rho) - C_L(\rho) \tag{4}$$

[*] akanksha.gautam512@gmail.com
[†] kavita@iisermohali.ac.in
[‡] arvind@iisermohali.ac.in

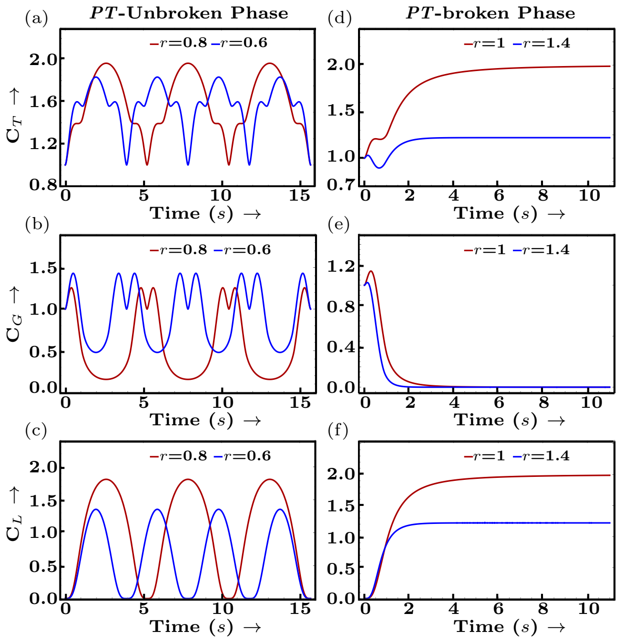


Figure 1: Plots of the evolution of quantum coherence present in the maximally entangled bipartite (Bell) state under different phases of PT-symmetry

2 Simulation Results

Two-qubit maximally entangled (Bell) state: A two-qubit Bell state defined as: $\rho_{BS} = |\psi_{BS}\rangle\langle\psi_{BS}|$ where $|\psi_{BS}\rangle = (|00\rangle + |11\rangle)/\sqrt{2})$. We study the dynamics of quantum coherence evolving under local PT-Symmetric Hamiltonian and results are shown in Figure 1

Three-qubit maximally entangled state (GHZ): We consider the three qubits maximally entangled GreenbergerHorneZeilinger (GHZ) state : $\rho_{123} = \rho_{GHZ} = |\psi_{GHZ}\rangle\langle\psi_{GHZ}|$, where $|\psi_{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$. The results of dynamics of quantum coherence evolving under local PT-Symmetric Hamiltonian are shown in Figure 2

3 Experimental Demonstration

The experimental implementation of PT-symmetric Hamiltonian on an NMR quantum processor is realized using three qubits system where three spin-1/2 nuclei (^{1}H,^{19}F and ^{13}C) of sample ^{13}C-labeled diethylfluoromalonate act as a three qubit system and the sample is dissolved in acetone-D6 . The first two nuclear spins ^{1}H and ^{19}F are used to simulate two qubit system while ^{13}C spin is utilized as the ancillary qubit.

Quantum Circuits: Quantum circuits to prepare a maximally entangled Bell state and to simulate a PT-symmetric Hamiltonian is shown in Figure 3.

4 Results and Conclusions

- Our results in Figure 4 show that in the Bell state, the total, local and global coherences oscillate with time in the unbroken phase (r = 0.6) while the amplitudes are different for various quantum coherences.

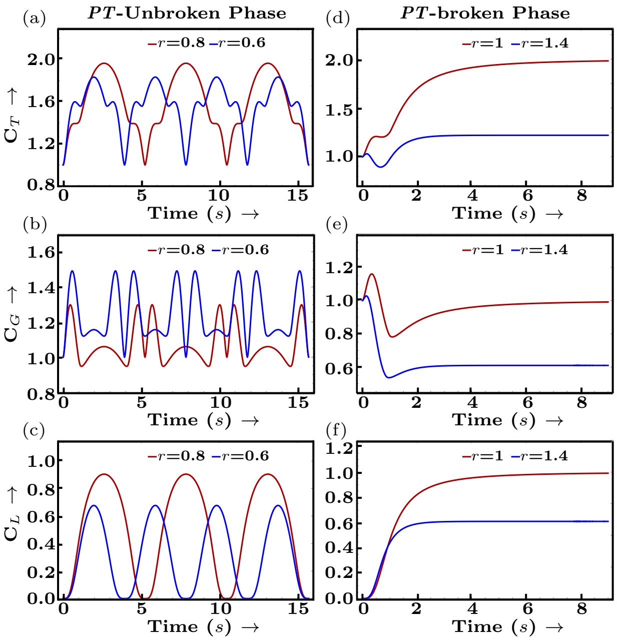


Figure 2: Plots of the evolution of quantum coherence present in a maximally entangled tripartite (GHZ) state under different phases of PT-symmetry

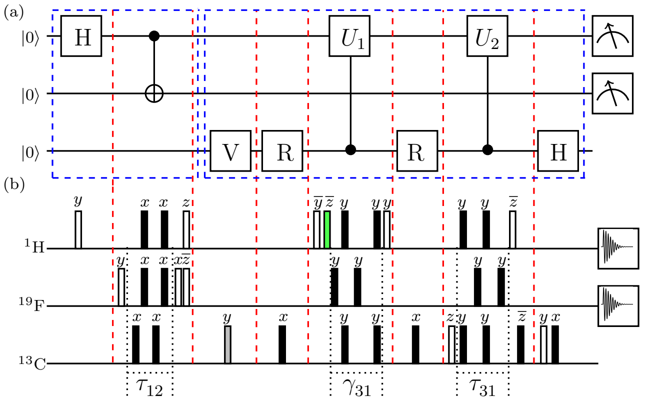


Figure 3: **(a)** Quantum circuits: The gates in the first box create a maximally entangled Bell state on the first two qubits, with the third qubit acting as an ancillary qubit. The gates in the second box simulate a PT-symmetric Hamiltonian on the first qubit. **(b)**The corresponding NMR pulse sequence, where the unfilled rectangles are $\frac{\pi}{2}$ pulses, the black rectangles denote π pulses, the gray rectangle is a θ pulse and the green rectangle denotes a ϕ pulse. Pulse phases are given above each pulse and a bar over a phase represents negative phase. Further details are given in Ref.[4]

- In the broken phase (r = 1.4), the total coherence initially increases and then freezes at later times. Similarly, our results indicate that the local coherence, which is not present initially in the Bell state, is created and increases under the broken phase of

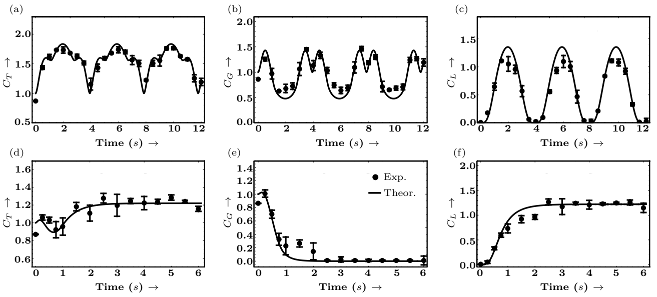


Figure 4: Dynamics of quantum coherence present in two qubit Bell State under PT-Symmetric Hamiltonian where (a), (b) and (c) represent the dynamics of total coherence (C_T), global coherence (C_G) and local coherence (C_L) respectively with time in unbroken phase of PT-Symmetry $(r = 0.6)$. (d), (e) and (f) represent the dynamics of total coherence (C_T), global coherence (C_G) and local coherence (C_L) respectively with time in broken phase of PT-Symmetry $(r = 1.4)$

PT-symmetry and freezes at later times whereas the global coherence initially increases and then exponentially decays with time.

- Our work sheds some light on the effect of the PT-symmetric Hamiltonian on quantum coherences in a multipartite system which can further help in gaining an understanding of the effects of the PT-symmetric Hamiltonian in quantum thermodynamics and in quantum information processing.

References

[1] C. M. Bender, S. Boettcher. Real spectra in non-Hermitian Hamiltonians having PT symmetry. *Phys. Rev. Lett.* 80, 5243, 1998.

[2] J. Wen, C. Zheng, X. Kong, S. Wei, T. Xin, G. Long. Experimental demonstration of a digital quantum simulation of a general PT -symmetric system. *Phys. Rev. A* 99, 062122, 2019.

[3] T. Baumgratz, M. Cramer, M. B. Plenio. Quantifying Coherence. *Phys. Rev. Lett.* 113, 140401, 2014.

[4] A. Gautam, K. Dorai, Arvind. Experimental demonstration of the dynamics of quantum coherence evolving under a PT-symmetric Hamiltonian on an NMR quantum processor. *Quant. Inf. Proc.* 21, 329, 2021.

Extended abstract:
Quantum-optimal information encoding using noisy passive linear optics

(Full version paper: arXiv:2304.12365)

Andrew Tanggara[1 2 *] Ranjith Nair[2] Syed Assad[3] Varun Narasimhachar[2 4]
Spyros Tserkis[3] Jayne Thompson[5 1] Ping Koy Lam[3 6] Mile Gu [†2 1 7]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543.*
[2] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 639673.*
[3] *Centre for Quantum Computation and Communication Technology, Department of Quantum Science, Research School of Physics and Engineering, Australian National University, Canberra ACT, Australia 2601.*
[4] *Institute of High Performance Computing, Agency for Science, Technology and Research, 1 Fusionopolis Way #20-10, Connexis North Tower, Singapore 138632.*
[5] *Horizon Quantum Computing, 05-22 Alice@Mediapolis, 29 Media Circle, Singapore 138565.*
[6] *Institute of Materials Research and Engineering, Agency for Science Technology and Research (A*STAR), 2 Fusionopolis Way, 08-03 Innovis, Singapore 138634.*
[7] *CNRS-UNS-NUS-NTU International Joint Research Unit, UMI 3654, Singapore 117543.*

Abstract. The amount of information that a noisy channel can transmit has been one of the primary subjects of interest in information theory. For a large family of optical quantum channels that can be implemented without an external energy source, we optimize the Holevo information over information encodings in the attenuations and phase-shifts that they apply to a resource state of finite energy. We show that for any given input state and environment temperature, the maximum Holevo information can be achieved by an encoding procedure that uniformly distributes the channel's phase-shift parameter. Moreover for large families of input states, any maximizing encoding schemes only involve a finite number of channel attenuation values, giving codewords that form a finite number of rings around the origin in the phase space.

Keywords: quantum channel coding, passive linear-optical channel, noisy quantum channel, continuous-variable quantum information, quantum information theory, holevo information, quantum reading

1 Thermal channel and encoding

In a classical channel coding scenario with a sender (Alice) transmitting messages to a receiver (Bob) through a channel, one quantifies the rate at which information can be reliably transmitted from Alice to Bob by the Shannon mutual information, which depends on the encoding procedure over the channel's input symbols and the channel's output statistics [1]. When a quantum channel is used, Alice's messages are instead encoded to an ensemble of quantum states that Bob can make (joint) measurements on, where the Holevo information now quantifies the amount of information that can be reliably transmitted [2]. As in the classical scenario, optimizing over the encoding procedure gives the ultimate capacity of the quantum channel.

In this work, we focus on optimizing the Holevo information of a family of quantum channels called the *thermal channels*, which are linear-optical quantum channels without external energy source where one encodes information by mixing a given finite-energy input state with the environment in some thermal state and then phase-shifting it (see Fig. 1). The thermal channel model are applicable in a number of tasks as well as fundamentally interesting in its own right. Most notably, it is directly applicable to the task of optical quantum reading tasks [3, 4, 5, 6, 7, 8] where one probes a set of quantum channels acting as "memory cells" to decode information encoded in their parameters (see Fig. 2). The thermal channels extend some of the aforementioned quantum reading results that assumed zero-temperature (i.e. vacuum) environment. Moreover, a thermal channel separates out the energy sources it utilizes, allowing one to analyze energy as a resource to perform the task at hand (e.g. for thermodynamic tasks [9]). From a fundamental perspective, the thermal channel framework also allows one to analyze its information capacity with a more restrictive *peak* energy constraint as opposed to the more well understood average energy constraint. As it is known in the latter case that the information capacity of large classes

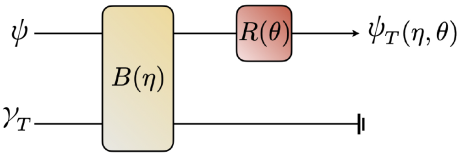


Figure 1: Thermal channel. Input "resource" state ψ with energy E is mixed with an "environment" in a thermal state γ_T by a beamsplitter $B(\eta)$ of transmittance η and then undergoes a θ phase-rotation operation $R(\theta)$, giving an output codeword state $\psi_T(\eta,\theta)$.

[*] andrew.tanggara@gmail.com
[†] mgu@quantumcomplexity.org

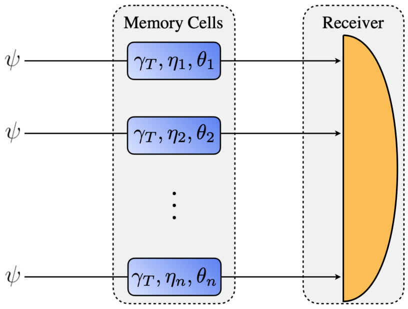

Memory Cells
Receiver
ψ
$\gamma_T, \eta_1, \theta_1$
ψ
$\gamma_T, \eta_2, \theta_2$
$\vdots$
ψ
$\gamma_T, \eta_n, \theta_n$

Figure 2: Application of thermal channel in quantum memory reading. Each memory cell is a thermal channel with temperature T where information is encoded in its attenuation η_j and phase θ_j. One decodes the information by sending a probe state ψ through and measuring the output.

of Bosonic Gaussian channels are achieved by Gaussian encodings [10, 11, 12], this is not always the case with a peak energy constraint as we will later show. Lastly, a peak energy constraint is also relevant to model many practical scenarios, such as: channels with limited energy tolerance, technological limitations on generating Fock states of large occupation number, and satellite-based laser communication system where energy is scarce.

A thermal channel is formally modeled as a beamsplitter operation that mixes a given resource state ψ with energy E with the thermal state of the environment with a given temperature T, followed by a phase shift operation (see Fig. 1). The procedure that encodes information by modulating the thermal channel's attenuation parameter $\eta \in [0,1]$ and phase parameter $\theta \in [-\pi,\pi]$ is described by a joint cumulative distribution function $F(\eta,\theta)$ called a *thermal encoding*. This procedure generates the codeword ensemble $\{F(\eta,\theta),\psi_T(\eta,\theta)\}_{\eta,\theta}$ that defines the average codeword state $\psi_{\text{ave}} = \int dF(\eta,\theta)\psi_T(\eta,\theta)$. In this work, we characterize thermal encodings F that maximizes the Holevo information for a given resource state ψ and environment temperature T given by,

$$\chi[F] = S(\psi_{\text{ave}}) - \int dF(\eta,\theta)\; S(\psi_T(\eta,\theta))\;, \qquad (1)$$

where S is the von Neumann entropy.

2 Optimal thermal encoding

Our first result on characterizing the optimal thermal encodings indicates that instead of maximizing χ over all possible encoding F, we only need to consider a class of encodings that distributes the phase parameter uniformly, called the *circularly symmetric encodings*. This is because for any given thermal encoding, there exist a circularly symmetric encoding that is at least as good.

Proposition 1 *Given a resource state ψ and environment temperature T and an arbitrary thermal encoding F, there exists a circularly symmetric encoding F' such that $\chi[F'] \geq \chi[F]$.*

Optimizing over the circularly symmetric encoding is advantageous as one only needs to maximize the Holevo information over the encodings of the attenuation parameter, both practically and theoretically simplifying the optimization. Additionally, the codeword ensemble of a circularly symmetric encoding can be nicely visualized in the phase-space as shown in Fig.3 for coherent state resources, where for each fixed attenuation η one obtains a ring structure representing the uniformly distributed phase θ. For any given circularly symmetric encoding F, we can further use a measure of information called the *marginal information density* $i[\eta,F]$ for each ring at attenuation point η, which average gives us the Holevo information

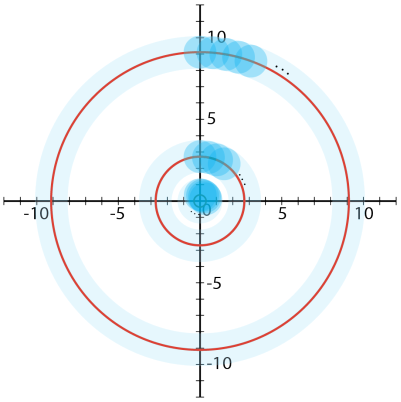


Figure 3: Phase-space visualization of codewords for an optimal encoding given coherent state resource with energy $E \sim 9.2$ and a zero-temperature environment. Radius of each red ring indicates energy of the attenuated and phase-shifted coherent state codewords (small blue circles).

$$\chi[F] = \int dF(\eta)\; i[\eta,F]\;. \qquad (2)$$

Using this relation, we show that all circularly symmetric encoding that maximizes χ necessarily and sufficiently satisfy the following conditions.

Proposition 2 *For any given resource state ψ, a circularly symmetric encoding F^* is uniquely optimal if and only if*

$$i[\eta,F^*] \leq \chi[F^*] \quad \text{for all } \eta \in [0,1] \qquad (3a)$$

$$i[\eta,F^*] = \chi[F^*] \quad \text{if and only if } F^*(\eta) \text{ is increasing}\;. \qquad (3b)$$

These conditions for an optimal circularly symmetric encoding F^* shows it only contains the attenuation points η where most amount of information can be encoded. Beside being a technical tool to show further characterizations of optimal thermal encodings, the optimality conditions in Proposition 2 are practically useful in numerical optimization of thermal encodings. As one might note that there might be uncountably many attenuation points η in an optimal encoding, we further show that this is not the case for large families of resource states, particularly for coherent state resource in temperature $T = 0$ and for thermal state resource. In fact, an optimal circularly symmetric encoding for those

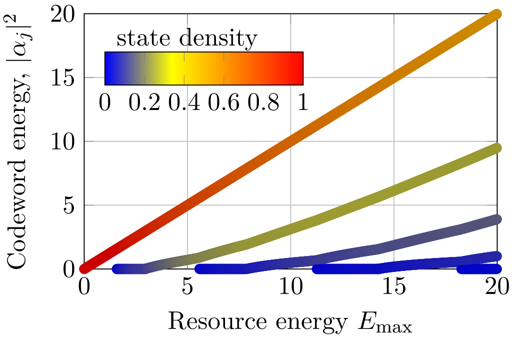


Figure 4: Optimal circularly symmetric distributions F^* for coherent state resource with energy $E_{\max}$. For a fixed $E_{\max}$, there are a finite number of codeword energies $\{\eta_j E_{\max}\}_j$ each corresponds to a ring (as shown in Fig. 3 for $E_{\max} \sim 9.2$) with its corresopnding attenuation probability $dF(\eta_j)$ color encoded.

resource states involves only a (countably) finite number of attenuation points.

Proposition 3 *For any coherent-state resource $\psi = |\alpha\rangle\langle\alpha|$ in zero-temperature environment and for a thermal-state resource $\psi = \rho_{\text{th}}(n_{\text{res}})$ in any environment temperature, the attenuation-coefficients η of the optimal circularly-symmetric encoding takes a finite number of values in* $[0, 1]$.

Proposition 3 states that the codewords from these optimal circularly symmetric encoding is a finite mixture of ring states, visualized as discrete rings in its phase-space representation (see Fig. 3). The number of rings however, varies as the resource state energy changes (see Fig. 4).

Although we are able to show this property rigorously only for these coherent state and thermal state resource cases, numerical evidence suggests that it also holds for the optimal encoding of larger classes of resource states. Numerical optimization suggests that the encodings satisfying the conditions in Proposition 2. The resulting Holevo information from the numerical optimization for given a displaced thermal state resource at zero-temperature environment can be seen in Fig 5, whereas the results for given coherent state resource at non-zero temperature environment is shown in Fig. 6.

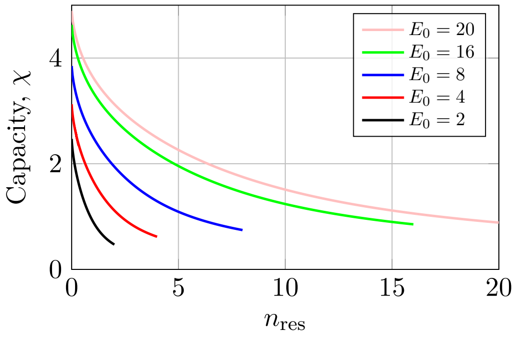


Figure 5: Capacity for a displaced thermal-state resource with energy E_0 and mean photon number n_{res} at $T = 0$. For each line, the total energy of the resource is fixed at E_0. The left end of each line (with $n_{\text{res}} = 0$) corresponds to a pure coherent state resource, while the right end is a thermal state resource with mean photon number $n_{\text{res}} = E_0$. For each E_0, the capacity χ decreases as the resource mean photon number n_{res} increases.

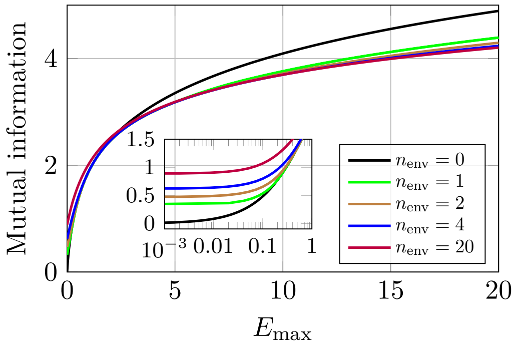


Figure 6: Thermal-encoding capacity for coherent state resource $|\alpha_{\max}\rangle$, where the Holevo information varies with environment mean photon number n_{env}.

References

[1] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[2] Alexander Semenovich Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Problemy Peredachi Informatsii*, 9(3):3–11, 1973.

[3] Hsu, Delaubert, Bowen, Fabre, Bachor, and Koy Lam. A quantum study of multibit phase coding for optical storage. *IEEE Journal of Quantum Electronics*, 42(10):1001–1007, Oct 2006.

[4] Stefano Pirandola. Quantum reading of a classical digital memory. *Physical Review Letters*, 106(9), Mar 2011.

[5] Stefano Pirandola, Cosmo Lupo, Vittorio Giovannetti, Stefano Mancini, and Samuel L Braunstein. Quantum reading capacity. *New Journal of Physics*, 13(11):113012, Nov 2011.

[6] Saikat Guha, Zachary Dutton, Ranjith Nair, Jeffrey H. Shapiro, and Brent J. Yen. Information capacity of quantum reading. *Frontiers in Optics 2011/Laser Science XXVII*, 2011.

[7] Mark M Wilde, Saikat Guha, Si-Hui Tan, and Seth Lloyd. Explicit capacity-achieving receivers for optical communication and quantum reading. In *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, pages 551–555. IEEE, 2012.

[8] Saikat Guha and Jeffrey H Shapiro. Reading boundless error-free bits using a single photon. *Physical Review A*, 87(6):062306, 2013.

[9] Varun Narasimhachar, Syed Assad, Felix C. Binder, Jayne Thompson, Benjamin Yadin, and Mile Gu. Thermodynamic resources in continuous-variable quantum systems, September 2019.

[10] V. Giovannetti, R. García-Patrón, N. J. Cerf, and A. S. Holevo. Ultimate classical communication rates of quantum optical channels. *Nature Photonics*, 8(10):796–800, Sep 2014.

[11] Vittorio Giovannetti, Saikat Guha, Seth Lloyd, Lorenzo Maccone, Jeffrey H Shapiro, and Horace P Yuen. Classical capacity of the lossy bosonic channel: The exact solution. *Physical review letters*, 92(2):027902, 2004.

[12] Alexander S Holevo. Quantum systems, channels, information. In *Quantum Systems, Channels, Information*. de Gruyter, 2019.

Virtual distillation with noise dilution

Yong Siah Teo[1] * Seongwook Shin[1] Hyukgun Kwon[1] Seok-Hyung Lee[1]
Hyunseok Jeong[1] †

[1] *Department of Physics and Astronomy, Seoul National University, 08826 Seoul, South Korea*

Abstract. Virtual distillation is an error-mitigation technique for arbitrary noisy environments. Considering isolatable noise from peripherals of a quantum circuit, such as delay lines, we find that virtual-distillation performance improves if the peripheral is uniformly distributed across the circuit. For multi-qubit loss and Pauli noise channels of a fixed overall error rate, error-mitigated circuit-output states improve in quality monotonically as the peripheral is split (diluted) into more layers. We show that second-order distillation is sufficient for near-optimal mitigation. These results are applied to quantum-computing clusters, where detectors are limited and delay lines are necessary to queue output qubits from multiple circuits.

Keywords: variational quantum algorithm, noisy intermediate-scale quantum (NISQ), error mitigation, virtual distillation, two-designs, delay lines, noise dilution, Pauli channel, photon loss

1 Virtual distillation

In the current era of noisy intermediate-scale quantum (NISQ) devices, their exploitation for any viable advantage in computation should accompany the mitigation of errors arising from noisy environments. Virtual distillation [1–3] is a technique that can mitigate noise of small error rates for arbitrary noise models.

Suppose that an effective noise channel Φ_ϵ, of some error rate ϵ, acts on a pure state $\rho = |\ \rangle\langle\ |$ of Hilbert-space dimension d. The resulting noisy state

$$\rho' = \Phi_\epsilon[\rho] = |\lambda_0(\epsilon)\rangle\lambda_0(\epsilon)\langle\lambda_0(\epsilon)| + \sum_{k=1}^{d-1} |\lambda_k(\epsilon)\rangle\lambda_k(\epsilon)\langle\lambda_k(\epsilon)| \tag{1}$$

possesses a spectral decomposition with ordered eigenvalues $\lambda_0 > \lambda_1 \geq \ldots \geq \lambda_{d-1}$, where for sufficiently small ϵ, the eigenstate $|\lambda_0(\epsilon)\rangle\langle\lambda_0(\epsilon)| \cong \rho$.

Virtual distillation utilizes a basic linear-algebraic principle. If the *dominant* eigenvalue $\lambda_0(\epsilon) > \lambda_{k>0}(\epsilon)$ is *nondegenerate*, which is typical for noisy environments,

$$\lim_{M\to\infty} \frac{\rho'^M}{\mathrm{tr}\left\{\rho'^M\right\}} = |\lambda_0(\epsilon)\rangle\langle\lambda_0(\epsilon)| \cong \rho\,. \tag{2}$$

Thus, a sufficiently large *distillation order* M, followed by a trace normalization, amplifies the singly-dominant eigenstate that shall approximately be the target ρ.

In a variational quantum algorithm (VQA) setting [4–10], one is generally interested in measuring the expectation value $\langle O\rangle$ of a Hermitian observable O with respect to some target $\rho = |\ \rangle\langle\ |$. Virtual distillation correspondingly entails the measurement of $\mathrm{tr}\left\{\rho'^M O\right\}/\mathrm{tr}\left\{\rho'^M\right\}$, which can be done either with a correcting circuit [2] or shadow tomography [11].

2 Concept of noise dilution

Consider a circuit peripheral that contributes to isolatable noise, where the map $\Phi^{\mathrm{i.i.d.}}_\epsilon$ comprises independent and identically distributed (i.i.d.) noisy channels. This leads to the concept of *noise dilution* [12].

Suppose that this peripheral is used in conjunction with an n-qubit quantum circuit described by the unitary operator U, which can be split into several smaller portions. The corresponding noiseless target output $\rho = U|\mathbf{0}\rangle\langle\mathbf{0}|U^\dagger$, where $|\mathbf{0}\rangle\langle\mathbf{0}|$ is some fixed initialized pure state. Given the possible decomposition $U = W_1W_2\ldots W_{L_\mathrm{err}}$ in terms of unitary operators W_1, W_2, ..., W_{L_err}, the observer may choose to distribute the peripheral across U, so that the noisy output state is

$$\begin{aligned}\rho'_{L_\mathrm{err}} &= \Phi^{\mathrm{i.i.d.}}_{\frac{\epsilon}{L_\mathrm{err}}}\left[W_{L_\mathrm{err}}\ldots\Phi^{\mathrm{i.i.d.}}_{\frac{\epsilon}{L_\mathrm{err}}}\left[W_2\,\Phi^{\mathrm{i.i.d.}}_{\frac{\epsilon}{L_\mathrm{err}}}\left[W_1|\mathbf{0}\rangle\langle\mathbf{0}|W_1^\dagger\right]W_2^\dagger\right]\right.\\ &\qquad\left.\ldots W^\dagger_{L_\mathrm{err}}\right]\\ &= \left(1-\frac{\epsilon}{L_\mathrm{err}}\right)^{L_\mathrm{err}n}\rho + \left[1-\left(1-\frac{\epsilon}{L_\mathrm{err}}\right)^{L_\mathrm{err}n}\right]\rho_\mathrm{err}^{(L_\mathrm{err})}\,,\end{aligned} \tag{3}$$

where each of the L_err peripheral layers acts with an error rate of ϵ/L_err.

A relevant situation is the arrangement of delay lines in the practical implementation of the NISQ clusters, where the total error rate ϵ of these lines is small. Supposing that the delay line of a certain noise decay rate γ for which the error rate is $\epsilon = 1 - \mathrm{e}^{-\gamma\tau}$ after some delay time period τ. For a small τ, we find that $\epsilon \cong \gamma\tau$, so that the L_err-layered noise dilution scheme outlined here is equivalent to splitting the delay line into L_err equal delay times τ/L_err and distributing them evenly throughout the quantum circuit, whilst preserving the total delay time τ. The error rate of each portioned peripheral is then ϵ/L_err. In this case, one may either use delay lines after the computation with U or distribute them uniformly across U as in Eq. (3). Figure 1 illustrates noise dilution in a four-qubit circuit.

The mean squared-error (MSE, $\mathcal{D}$) is used to measure the performance of error mitigation:

$$\mathcal{D} = \left\langle \mathrm{tr}\left\{\left(\rho - \frac{\rho'^M}{\mathrm{tr}\left\{\rho'^M\right\}}\right)^2\right\}\right\rangle\,. \tag{4}$$

*yong.siah.teo@gmail.com
†h.jeong37@gmail.com

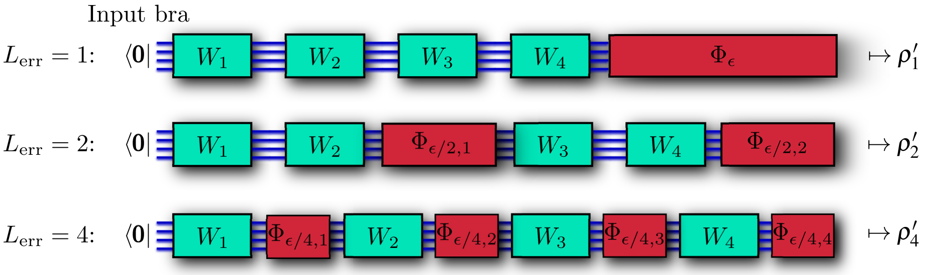

Input bra
$L_{\rm err}=1$: $\langle\mathbf{0}|$ W_1 W_2 W_3 W_4 Φ_ϵ $\mapsto\rho'_1$
$L_{\rm err}=2$: $\langle\mathbf{0}|$ W_1 W_2 $\Phi_{\epsilon/2,1}$ W_3 W_4 $\Phi_{\epsilon/2,2}$ $\mapsto\rho'_2$
$L_{\rm err}=4$: $\langle\mathbf{0}|$ W_1 $\Phi_{\epsilon/4,1}$ W_2 $\Phi_{\epsilon/4,2}$ W_3 $\Phi_{\epsilon/4,3}$ W_4 $\Phi_{\epsilon/4,4}$ $\mapsto\rho'_4$

Figure 1: Dilution of a peripheral ($\Phi^{\rm i.i.d.}_\epsilon$) for a four-qubit quantum circuit into $L_{\rm err}=1$ (no dilution), 2 and 4 layers.

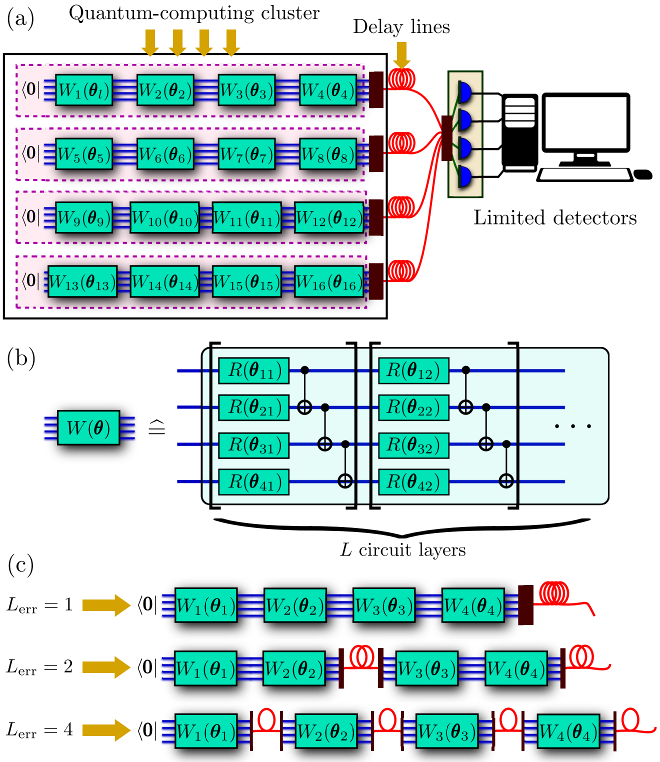

(a)
Quantum-computing cluster
Delay lines
Limited detectors
$W_1(\boldsymbol{\theta}_1)$ $W_2(\boldsymbol{\theta}_2)$ $W_3(\boldsymbol{\theta}_3)$ $W_4(\boldsymbol{\theta}_4)$
$W_5(\boldsymbol{\theta}_5)$ $W_6(\boldsymbol{\theta}_6)$ $W_7(\boldsymbol{\theta}_7)$ $W_8(\boldsymbol{\theta}_8)$
$W_9(\boldsymbol{\theta}_9)$ $W_{10}(\boldsymbol{\theta}_{10})$ $W_{11}(\boldsymbol{\theta}_{11})$ $W_{12}(\boldsymbol{\theta}_{12})$
$W_{13}(\boldsymbol{\theta}_{13})$ $W_{14}(\boldsymbol{\theta}_{14})$ $W_{15}(\boldsymbol{\theta}_{15})$ $W_{16}(\boldsymbol{\theta}_{16})$
(b)
$W(\boldsymbol{\theta})$ $\widehat{=}$
$R(\boldsymbol{\theta}_{11})$ $R(\boldsymbol{\theta}_{21})$ $R(\boldsymbol{\theta}_{31})$ $R(\boldsymbol{\theta}_{41})$
$R(\boldsymbol{\theta}_{12})$ $R(\boldsymbol{\theta}_{22})$ $R(\boldsymbol{\theta}_{32})$ $R(\boldsymbol{\theta}_{42})$
L circuit layers
(c)
$L_{\rm err}=1$
$L_{\rm err}=2$
$L_{\rm err}=4$

Figure 2: Noise dilution on a cluster of NISQ circuits. Each circuit is parametrized by layers of single-qubit gates $R(\boldsymbol{\theta})$ and controlled-NOT (CNOT) gates, where delay lines are used to queue output qubits after quantum computation.

3 Main results with NISQ clusters

3.1 Premise

We illustrate the main results with an example of a quantum-computing cluster (see Fig. 2) that houses multiple n-qubit quantum circuits that are accessible by the public domain. A user may login to the cluster and use one such circuit to perform variational computation. We additionally assume a limited number of detectors, so that the output qubits are to be queued with delay lines for the final measurement whenever necessary. These delay lines are the main sources of isolatable errors that can be rearranged to optimize error-mitigation performance.

In a typical NISQ cluster, each unitary operator W_l describes a circuit comprising L layers of single-qubit and CNOT gates. From [13], randomized circuits of this kind are approximately two-designs if $L=O(\mathrm{poly}(n))$.

3.2 Noise channels

Two single-qubit noise channels, each of error rate ϵ, are considered: the loss channel that mixes a qubit state $\rho_{\rm qubit}$ with the vacuum state $|\textsc{vac}\rangle\langle\textsc{vac}|$, and the Pauli channel that introduces rotation errors by the standard Pauli operators X, Y and Z. Succinctly,

$$\begin{aligned}\Phi^{\rm loss}_\epsilon[\rho_{\rm qubit}] &= (1-\epsilon)\rho_{\rm qubit}+|\textsc{vac}\rangle\epsilon\langle\textsc{vac}|\,,\\ \Phi^{\rm Pauli}_\epsilon[\rho_{\rm qubit}] &= (1-\epsilon)\,\rho_{\rm qubit}+\epsilon_1 X\,\rho_{\rm qubit}\,X\\ &\quad+\epsilon_2 Y\,\rho_{\rm qubit}\,Y+\epsilon_3 Z\,\rho_{\rm qubit}\,Z\,. \end{aligned}\tag{5}$$

For i.i.d. loss channels, since $|\textsc{vac}\rangle\langle\textsc{vac}|$ is orthogonal to any $\rho_{\rm qubit}$, the complete MSE up to $O(\epsilon^{2M})$ is

$$\begin{aligned}\mathcal{D}^{\rm i.i.d.\,loss}_{M,L_{\rm err}} = \left(\frac{\epsilon}{L_{\rm err}}\right)^{2M}\Bigg[n\left\langle\mathrm{tr}\left\{\mathrm{tr}_1\left\{|\ \rangle\langle\ |\right\}^{2M}\right\}\right\rangle\\ +\left\langle\left(\sum_{j=1}^{n}\mathrm{tr}\left\{\mathrm{tr}_j\left\{|\ \rangle\langle\ |\right\}^{M}\right\}\right)^2\right\rangle\Bigg]\,,\end{aligned}\tag{6}$$

I.i.d. Pauli channels are trickier to handle for their non-commuting error terms. This gives rise to *persistent errors* that no longer vanish according to $O((\epsilon_l\epsilon_{l'})^M)$. The respective MSEs for $M=1$ and $M\geq2$ up to $O(\epsilon_l\epsilon_{l'})$ are

$$\begin{aligned}\mathcal{D}^{\rm i.i.d.\,Pauli}_{M=1,L_{\rm err}} &= (n\epsilon)^2\Bigg[\frac{d^3}{(d+1)(d^2-1)}\\ &\quad-\frac{d}{L_{\rm err}(d+1)(d^2-1)}\Bigg]+\frac{n\,d}{L_{\rm err}(d+1)}\sum_{l=1}^{3}\epsilon_l^2\,,\\ \mathcal{D}^{\rm i.i.d.\,Pauli}_{M\geq2,L_{\rm err}} &= 2\sum_{l,l'=1}^{3}\frac{\epsilon_l\epsilon_{l'}}{L_{\rm err}^2}\sum_{j,j'=1}^{n}\Big[\left\langle\mathrm{tr}\left\{\rho\,T_j^{(l)}T_{j'}^{(l')}\right\}\right\rangle\\ &\quad-\left\langle\mathrm{tr}\left\{\rho\,T_j^{(l)}\right\}\mathrm{tr}\left\{\rho\,T_{j'}^{(l')}\right\}\right\rangle\Big]\,,\\ T_j^{(l)} &= P_j^{(l)}\rho P_j^{(l)}+W_{L_{\rm err}}P_j^{(l)}\rho_{L_{\rm err}-1}P_j^{(l)}W^\dagger_{L_{\rm err}}+\ldots\\ &\quad+W_{L_{\rm err}}W_{L_{\rm err}-1}\ldots W_2P_j^{(l)}\rho_1P_j^{(l)}W_2^\dagger\ldots W^\dagger_{L_{\rm err}-1}W^\dagger_{L_{\rm err}}\,,\end{aligned}\tag{7}$$

where $P_j^{(l)}$ is a single-qubit Pauli operator of error rate ϵ_l.

We investigate the use of delay lines, that give rise to these two isolatable noise channels. The first scenario is when photon loss is dominant, which could be the case

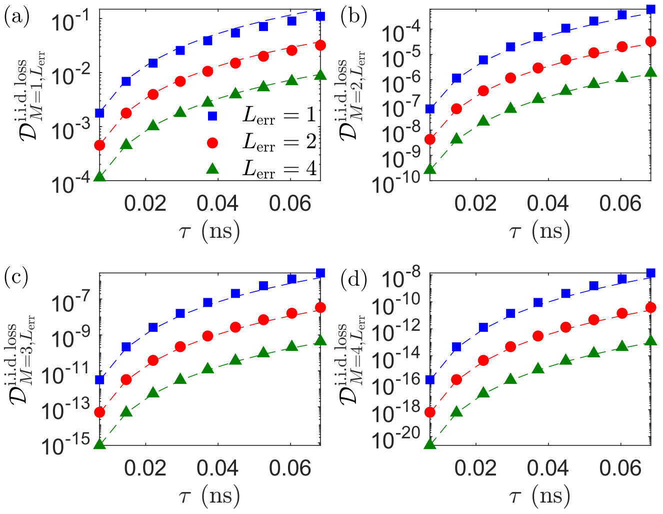


Figure 3: Virtual-distillation MSE curves for four-qubit hardware-efficient circuits ($L = 2$) with respect to the total delay time τ under the i.i.d. loss channel. Performances for various M and L_{err} are illustrated.

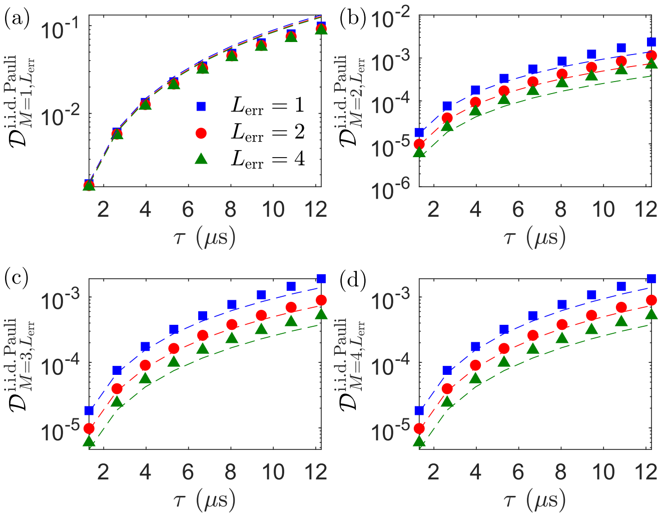


Figure 4: Virtual-distillation MSE curves for four-qubit hardware-efficient circuits ($L = 2$) with respect to the total delay time τ under the i.i.d. depolarizing channel ($\epsilon_1 = \epsilon_2 = \epsilon_3 = \epsilon/3$).

when the cluster is integrated into a photonic chip. We take the decay rate $\gamma_{\text{loss}} = 0.2$ dB cm^{-1} [14] or equivalently $\gamma_{\text{loss}} = 6$ dB ns^{-1}, where $\epsilon_{\text{loss}} = 1 - 10^{-\gamma_{\text{loss}}\tau/10}$ is related to the delay time τ (in ns). The second scenario is when polarization drifts occur more frequently in regular optical-fiber-based delay lines [15–17], such that the depolarizing channel serves as an appropriate noise model. We take the depolarizing rate to be $\gamma_{\text{depol}} = 1.3 \times 10^{-4}$ s^{-1} [18], which is equivalent to an error rate of about 0.01 for a 400-meter optical fiber. In other words, $\epsilon_{\text{depol}} = 1 - \mathrm{e}^{-\gamma_{\text{depol}}\tau}$.

Figures 3 and 4 plot error-mitigation performances based on four-qubit hardware-efficient circuits that are *shallow*, where the number of single-qubit-CNOT layers $L = 2$, such that a total of eight circuit layers is considered for each entire computation circuit. Hence, $L_{\text{err}} = 2$, for instance, implies that one of the diluted noise layers is sandwiched between two unitary subcircuits, each with $2L = 4$ circuit layers [see Fig. 2(c)].

3.3 Result 1: Error-mitigation improves with increasing L_{err}

For a fixed total error rate, the more diluted the peripheral is across the quantum circuit, the higher the error-mitigative power with virtual distillation. This finding is substantiated by the analytical answers in (6) and (7), as well as Figs. 3 and 4. In particular, (6) and (7) show that $\mathcal{D}$ drops monotonically with increasing L_{err} for sufficiently small error rates. This entails that the MSEs are well approximated with leading orders of the error rates.

Hence, for qubit delay lines of a fixed common length each, rather than delaying the qubits after first running them through the entire NISQ circuit, delaying the qubits uniformly across the circuit can help reduce the noise influence, and therefore better facilitate error mitigation for the same order virtual-distillation order M.

Interestingly, such a noise dilution technique may be employed to further reduce the effects of persistent noise channels (such as the Pauli channel), where virtual distillation beyond the second order brings no additional advantages.

3.4 Result 2: $M = 2$ is sufficient for small error rates

For the i.i.d. photon-loss channel, since the error term (namely the vacuum state $|\text{VAC}\rangle\langle\text{VAC}|$) commutes with the ideal target state, the MSE $\mathcal{D}^{\text{i.i.d. loss}}_{M,L_{\text{err}}} = O(\epsilon^{2M}/L_{\text{err}})$ drops exponentially with M. Figure 3 confirms this with up to $M = 4$ distillation order. Even when noise dilution is not applied, $M = 2$ is already sufficient in driving down the MSE by about three orders of magnitude. Significant enhancement is achievable with noise dilution.

On the other hand, the i.i.d. Pauli channel, as mentioned in Sec. 3.2, introduces error terms that result in order $\mathcal{D}^{\text{i.i.d. Pauli}}_{M,L_{\text{err}}} = O(\epsilon_l\epsilon_{l'}/L_{\text{err}})$ regardless of the distillation order $M \geq 2$. Figure 4 illustrates such an error persistence for the depolarizing channel.

It is clear that for both of these channels, $M = 2$ is sufficient to achieve near-optimal error-mitigation when the error rate is small.

References

[1] B. Koczor, “Exponential error suppression for near-term quantum devices,” *Phys. Rev. X*, vol. 11, p. 031057, Sep 2021.

[2] 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*, vol. 11, p. 041036, Nov 2021.

[3] K. Yamamoto, S. Endo, H. Hakoshima, Y. Matsuzaki, and Y. Tokunaga, "Error-mitigated quantum metrology via virtual purification," 2021.

[4] J. Biamonte, "Universal variational quantum computation," *Phys. Rev. A*, vol. 103, p. L030401, Mar 2021.

[5] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, pp. 625–644, Sep 2021.

[6] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik, "Quantum Chemistry in the Age of Quantum Computing," *Chemical Reviews*, vol. 119, no. 19, pp. 10856–10915, 2019. PMID: 31469277.

[7] 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*, vol. 90, p. 032001, Mar 2021.

[8] S. McArdle, S. Endo, A. Aspuru-Guzik, S. C. Benjamin, and X. Yuan, "Quantum computational chemistry," *Rev. Mod. Phys.*, vol. 92, p. 015003, Mar 2020.

[9] S. Shin, Y. S. Teo, and H. Jeong, "Exponential data encoding for quantum supervised learning," *Phys. Rev. A*, vol. 107, p. 012422, Jan 2023.

[10] Y. S. Teo, "Optimized numerical gradient and hessian estimation for variational quantum algorithms," *Phys. Rev. A*, vol. 107, p. 042421, Apr 2023.

[11] A. Seif, Z.-P. Cian, S. Zhou, S. Chen, and L. Jiang, "Shadow distillation: Quantum error mitigation with classical shadows for near-term quantum processors," *PRX Quantum*, vol. 4, p. 010303, Jan 2023.

[12] Y. S. Teo, S. Shin, H. Kwon, S.-H. Lee, and H. Jeong, "Virtual distillation with noise dilution," *Phys. Rev. A*, vol. 107, p. 022608, Feb 2023.

[13] A. W. Harrow and R. A. Low, "Random Quantum Circuits are Approximate 2-designs," *Communications in Mathematical Physics*, vol. 291, pp. 257–302, Oct. 2009.

[14] J. M. Arrazola, V. Bergholm, K. Brádler, T. R. Bromley, M. J. Collins, I. Dhand, A. Fumagalli, T. Gerrits, A. Goussev, L. G. Helt, J. Hundal, T. Isacsson, R. B. Israel, J. Izaac, S. Jahangiri, R. Janik, N. Killoran, S. P. Kumar, J. Lavoie, A. E. Lita, D. H. Mahler, M. Menotti, B. Morrison, S. W. Nam, L. Neuhaus, H. Y. Qi, N. Quesada, A. Repingon, K. K. Sabapathy, M. Schuld, D. Su, J. Swinarton, A. Száva, K. Tan, P. Tan, V. D. Vaidya, Z. Vernon, Z. Zabaneh, and Y. Zhang, "Quantum circuits with many photons on a programmable nanophotonic chip," *Nature*, vol. 591, pp. 54–60, Mar 2021.

[15] A. Dragan and K. Wódkiewicz, "Depolarization channels with zero-bandwidth noises," *Phys. Rev. A*, vol. 71, p. 012322, Jan 2005.

[16] A. Bayat, V. Karimipour, and I. Marvian, "Threshold distances for transmission of epr pairs through pauli channels," *Physics Letters A*, vol. 355, no. 2, pp. 81–86, 2006.

[17] M. Karpiński, C. Radzewicz, and K. Banaszek, "Fiber-optic realization of anisotropic depolarizing quantum channels," *J. Opt. Soc. Am. B*, vol. 25, pp. 668–673, Apr 2008.

[18] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, "Violation of Bell's Inequality under Strict Einstein Locality Conditions," *Phys. Rev. Lett.*, vol. 81, pp. 5039–5043, Dec 1998.

Catalysis always degrades external quantum correlations

Seok Hyung Lie[1] and Nelly H.Y. Ng[1]
[1] *School of Physical and Mathematical Sciences, Nanyang Technological University, 21 Nanyang Link, Singapore, 637371*

(Dated: March 7, 2023)

Catalysts used in quantum resource theories need not be in isolation and therefore are possibly correlated with external systems, which the agent does not have access to. Do such correlations help or hinder catalysis, and does the classicality or quantumness of such correlations matter? To answer this question, we first focus on the existence of a non-invasively measurable observable that yields the same outcomes for repeated measurements, since this signifies macro-realism, a key property distinguishing classical systems from quantum systems. We show that a system quantumly correlated with an external system so that the joint state is necessarily perturbed by any repeatable quantum measurement, also has the same property against general quantum channels. Our full characterization of such systems called *totally quantum systems*, solves the open problem of characterizing tomographically sensitive systems raised in [Lie and Jeong, Phys. Rev. Lett. 130, 020802 (2023)]. An immediate consequence is that a totally quantum system cannot catalyze any quantum process, even when a measure of correlation with its environment is arbitrarily low. It generalizes to a stronger result, that the mutual information of totally quantum systems cannot be used as a catalyst either. These results culminate in the conclusion that, out of the correlations that a generic quantum catalyst has with its environment, only classical correlations allow for catalysis, and therefore using a correlated catalyst is equivalent to using an ensemble of uncorrelated catalysts.

Introduction.— Catalysis in quantum resource theory, a concept inspired by catalysis in chemistry, is a paradigm that utilizes some quantum resources without altering or deteriorating it, while expanding the set of accessible quantum states (or channel) transformations [1]. Initially, catalysis was often studied under the condition of being uncorrelated with the final state of the system [2–7]. In recent years, however, a notable trend is the investigation of how the power of catalysis, in enabling state transitions, can be increased by allowing correlations to persist between system and catalyst after the process [8–18]. The relaxation of this constraint simplifies the conditions for state transition significantly, and often leads to a characterization of von Neumann quantities (e.g. entropy) – which have a strong operational significance previously only in the asymptotic i.i.d. regime – in one-shot settings. It is argued that in certain scenarios, the resultant correlation can be ignored, which assumes that only the marginal state of catalyst is relevant when catalyst is separated from the system.

However, this line of thought clashes with an often-used concept of 'catalyst bank' [19–21], a hypothetical entity that lends quantum resource catalysts to (possibly many) agents and retrieves thereafter (See FIG. 1.) A catalyst could be only a part of a large collection of quantum systems possessed by the bank. In this case, it is operationally natural for the bank to require the agent who borrows the catalyst to return it in a fashion that the whole quantum system stays in the same state, even though the agent used only a small portion. Even if the catalyst is prepared uncorrelated with other systems, after a single round of correlated catalysis by some user, the catalyst will form correlation. When the next user borrows the catalyst, again, it is natural to require the correlations to be preserved, as nothing forbids the same user to borrow the same catalyst twice, and not wasting any resource not possessed by oneself is the prime premise of resource catalysis.

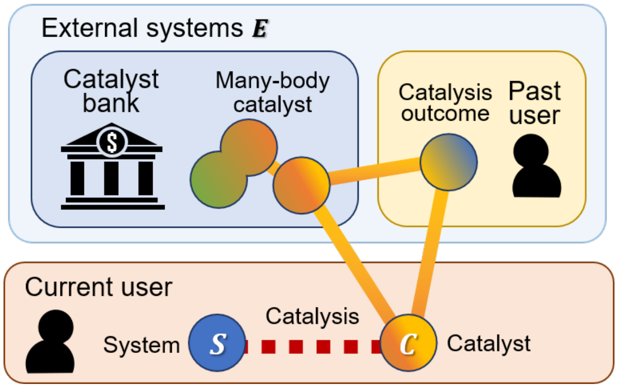


FIG. 1. A quantum resource catalyst could be correlated with external systems inaccessible to a user in many plausible scenarios. First, a catalyst can be a part of multipartite collection of catalysts of the bank. In this case, it is natural for the bank to demand the multipartite state to remain intact after each catalysis. Second, even for an initially uncorrelated catalyst, after a round of correlation-forming catalysis, it remains correlated with its previous user. The same user can borrow the catalyst again, and it is natural to expect the relation with the catalyst to remain the same as the previous round of catalysis.

Some studies on correlating catalysis deal with such potential problems, by showing that the amount of correlation formed in catalysis can be made arbitrarily small, e.g. [9, 14]. However, we will show that whenever the correlation between the catalyst and the external systems is of a quantum-mechanical nature (which we formally specify later), even *arbitrarily small* correlations forbid catalysis, in the sense that the joint state cannot be left unperturbed (Theorem 5).

In doing so, we fully characterize multi-partite states that cannot be used as a catalyst when only local access is allowed. We show that this characterization coincides with the description of quantum states that have no local classical observable (Theorem 3). Here, classical observables mean those that can be measured without perturbing the global quantum state, in other words, they have non-invasive measurability of deterministically distinguishable states, i.e., obey macro-realism [22–25]. It turns out that characterizing catalyst with local access is equivalent to characterizing the property known as (tomographical) sensitivity, which was an open problem in the previous work [26].

The observation that quantum correlation that a catalyst has with other systems only hinders catalysis leads to the conclusion (Theorem 9) that only classical correlation in correlated catalyst allows for meaningful catalysis. It yields a rather surprising consequence that utilizing a catalyst correlated with an external system is functionally equivalent to using an ensemble of uncorrelated catalysts. In other words, considering correlated catalyst does not introduce new types of non-trivial catalytic transformations, but it only induces probabilistic mixtures of conventional catalytic transformations.

Background.— Recall that a bipartite state ρ_{AB} is said to be classical-quantum (C-Q) when there exists an orthonormal basis $\{|i\rangle\}$ of A such that the following expression is possible:

$$\rho_{AB} = \sum_i p_i \, |i\rangle\!\langle i|_A \otimes \rho_B^{(i)}, \tag{1}$$

with some probability distribution (p_i) and a set of states $\{\rho_B^{(i)}\}$ on B. It is equivalent to the existence of a rank-1 projective measurement $\{|i\rangle\!\langle i|\}$ on A that does not disturb the global state ρ_{AB} after the measurement, i.e.,

$$\rho_{AB} = \sum_i (|i\rangle\!\langle i|_A \otimes \mathbb{1}_B)\rho_{AB}(|i\rangle\!\langle i|_A \otimes \mathbb{1}_B). \tag{2}$$

One can generalize this definition where the projective measurement need not be rank-1 anymore, leading to the following definition of partial classicality.

Definition 1. A bipartite state ρ_{AB} is said to be partially classical-quantum (PC-Q) when there exists a projective measurement $\{\Pi_k\}_{k=1}^n$ with $n > 1$ on A that preserves ρ_{AB}, i.e.,

$$\rho_{AB} = \sum_{k=1}^n (\Pi_k \otimes \mathbb{1}_B)\rho_{AB}(\Pi_k \otimes \mathbb{1}_B). \tag{3}$$

We will sometimes say that a single system is PC when it is implicitly assumed to be correlated with another system and they are in a PC-Q state. When a system is not PC, then we will say that it is *totally quantum* (TQ) [18], so a non PC-Q state is a TQ-Q state. Note that the correlations in a TQ-Q state can be in general significantly weaker compared to entanglement. For example, the following evidently separable state is a TQ-Q state:

$$\rho_{AB} = \frac{1}{2}\left(\sum_{i=0}^{d-1} \lambda_i \, |i\rangle\!\langle i|_A\right) \otimes |0\rangle\!\langle 0|_B + \frac{1}{2}\,|+\rangle\!\langle +|_A \otimes |1\rangle\!\langle 1|_B\,, \tag{4}$$

where $\sum_{i=0}^{d-1} \lambda_i \, |i\rangle\!\langle i|_A$ is a nondegenerate quantum state on A and $|+\rangle_A = d^{-1/2}\sum_{i=0}^{d-1}|i\rangle_A$ is a maximally coherent state on A. This definition of classicality can be further generalized to input or output systems of quantum channels or completely positive (CP) maps: we say that the output system of a quantum channel $\mathcal{N}$ is partially classical (PC) when there exists a projective measurement $\mathcal{P} := \sum_k \Pi_k(\cdot)\Pi_k$ that fixes $\mathcal{N}$, i.e., $\mathcal{P}\circ\mathcal{N} = \mathcal{N}$. Similarly we say that the input system of $\mathcal{N}$ is PC when a projective measurement $\mathcal{Q}$ exists such that $\mathcal{N}\circ\mathcal{Q} = \mathcal{N}$. Likewise, we say that the input or output system is TQ when it is non-PC.

Totally quantumness and sensitivity.— One could question the generality of the notion of totally quantumness, since allowing weaker measurements such as positive operator valued measures (POVM) instead of projective measurements in Definition 1 may give rise to a qualitatively different characterization. Perhaps the most natural definition of totally quantumness could be as follows: A of ρ_{AB} is said to be *totally quantum** (TQ*), or ρ_{AB} is said to be a TQ*-Q state, when any non-trivial quantum measurement on A necessarily perturbs ρ_{AB}. However, by noting that every quantum channel with Kraus operators $\{K_i\}$ can be considered an implementation of the POVM $\{K_i^\dagger K_i\}$, we observe that totally quantumness* defined above is equivalent to the concept of (tomographical) sensitivity introduced in Ref. [26], which characterizes a state's ability to detect the action of *any* non-trivial local channel.

Definition 2. A bipartite state ρ_{AB} is sensitive on A to a set of quantum operations $\mathfrak{Q}$ with $\mathrm{id} \in \mathfrak{Q}$ when for every $\mathcal{S} \in \mathfrak{Q}$

$$(\mathcal{S}_A \otimes \mathrm{id}_B)(\rho_{AB}) = \rho_{AB} \implies \mathcal{S} = \mathrm{id}_A. \tag{5}$$

When a quantum state is sensitive to the set of all quantum channels, we simply say that it is sensitive, or equivalently TQ*-Q. Similarly, through the Choi-Jamiołkowski isomorphism, we say that a linear map $\mathcal{N}$ is sensitive to $\mathfrak{Q}$ when for every $\mathcal{S} \in \mathfrak{Q}$

$$\mathcal{S}\circ\mathcal{N} = \mathcal{N} \implies \mathcal{S} = \mathrm{id}. \tag{6}$$

Our first main result shows that actually the more general definition of totally quantumness is equivalent to the weaker one. This result solves the open problem questioned in Ref. [18].

Theorem 3 (TQ*=TQ)**.** A quantum channel is sensitive if and only if its output system is TQ. Similarly, a bipartite state ρ_{AB} is sensitive on A if and only if it is a TQ-Q state.

Theorem 3 can be shown using the structure result of fixed points of quantum channels [27], which in turn follows from the Artin-Wedderburn theorem [28, 29]. Self-contained elementary proofs of these results are given in the Supplemental Materials. Theorem 3 says that there is no intermediate level of classicality when it comes to non-invasive measurability; in other words, sensitivity to projective measurements automatically implies sensitivity to general quantum channels. This implies that no classical value can be read from a non-PC system without perturbing it globally, even through weak measurements [30–32].

Theorem 3 yields an interesting property of totally quantumness that it is *contagious*; if any system prepared in a pure state unitarily interacts with a totally quantum system, then it also becomes totally quantum.

Proposition 4. Let ρ_{AB} be a TQ-Q state. For any isometry $V : A \to AK$ such that the marginal state τ_E of $\tau_{KAB} = (V \otimes \mathbb{1}_B)\rho_{AB}(V^\dagger \otimes \mathbb{1}_B)$ is full-rank, τ_{KAB} is a TQ-Q state with respect to the bipartition $K|AB$.

An alternative interpretation of Prop. 4 is that any subsystem of a totally quantum system is also totally quantum. This result is analogous to that of Ref. [33], where a local projective measurement on a Q-Q state inevitably forms entanglement with a measurement device. Prop. 4 shows that this holds similarly even when one considers the more general case of POVMs, and TQ-Q states (proof in Supplemental Material).

Proposition 4 shows that it is impossible to circumvent Theorem 3 by unitarily extracting a macro-realistic part K from system A invasively. It means that there could be a non-trivial quantum channel action $\mathcal{R}$ on A as a back-action of the invasive measurement, and one can interpret V as the Stinespring dilation of the quantum channel $\mathcal{R}$. It provides additional motivation for the nomenclature *totally quantum* system for non-PC systems, as it has no classical property even in the weakest sense, i.e. when classicality means macro-realism and non-invasive measurability.

Theorem 3 offers another intuitive explanation of why quantum key distribution (QKD) is secure. A typical example of TQ-Q state can be found in the BB84 protocol [34]. When Alice wants to send her random bit (say) 0 to Bob through a quantum channel that could be eavesdropped, she encodes that bit in either of two random bases and record it in her memory M,

$$\rho_{AM} = \frac{1}{2}\left|0\middle\rangle\middle\langle 0\right|_A \otimes \left|0\middle\rangle\middle\langle 0\right|_M + \frac{1}{2}\left|+\middle\rangle\middle\langle +\right|_A \otimes \left|1\middle\rangle\middle\langle 1\right|_M . \quad (7)$$

Since ρ_{AM} is a special case of (4), by Theorem 3, any eavesdropper interacting with the qubit A in a non-trivial fashion must alter the global state ρ_{AB}, which results in detectable statistical difference in the later steps of the protocol.

Local catalysis of bipartite state.— Theorem 3 has a significant consequence about catalysis utilizing correlated states. In resource theories, conventionally catalytic transformations mean processes described as

$$\rho_S \to \sigma_S := \mathrm{Tr}_C[\Lambda(\rho_S \otimes \tau_C)], \quad (8)$$

with the catalytic constraint requiring that the catalyst remains in its original state in the process:

$$\mathrm{Tr}_S[\Lambda(\rho_S \otimes \tau_C)] = \tau_C, \quad (9)$$

where C is called the catalyst and Λ is a free operation on SC.

However, we do not limit ourselves to state transitions between two fixed quantum states. In this work, a catalytic transformation means a general quantum channel $\Phi(\rho)$ given as $\Phi(\rho) := \mathrm{Tr}_C[\Lambda(\rho_S \otimes \tau_C)]$ regardless of whether the initial state ρ_S is fixed or not. Note that sometimes it is required that the final state of joint system SC has arbitrarily weak correlation, i.e. $\|\Lambda(\rho_S \otimes \tau_C) - \sigma_S \otimes \tau_C\|_1 < \epsilon$. However, we do not make such an assumption here for generality.

Now, using a correlated system CE in state τ_{CE} as a catalyst when only access to C is given means the transformation of the form in (8) (where τ_S is interpreted as $\mathrm{Tr}_E\, \tau_{CE}$) with the modified constraint

$$\mathrm{Tr}_S[\Lambda_{SC} \otimes \mathrm{id}_E(\rho_S \otimes \tau_{CE})] = \tau_{CE}. \quad (10)$$

A typical example of correlated catalyst τ_{CE} is the product of a previous catalysis, i.e. $\tau_{CE} = \Lambda(\rho_E \otimes \tau_C)$, here Λ is same with that in (8) but acts on CE (See FIG. 1.) In other words, catalyst τ_C is 'borrowed' by E for catalysis Λ, formed correlation with E, and returned to be borrowed by S again for another round of catalysis. As discussed in Introduction, the bipartite state τ_{CE} could be used as a resource whenever two systems C and E are combined again, and any change of τ_{CE} by S can alter its resourceful nature. Our second main result then shows that the catalysis constraint Eq. (10) severely limits the usability of correlated catalyst.

Theorem 5. A TQ-Q state τ_{CE} cannot be used to catalytically implement a non-trivial transformation when only access to C is given.

Proof. We focus on the fact that once the initial state ρ_S and the interaction channel Λ on SC is fixed as in (8), then the following channel on C is induced.

$$\Gamma(\eta_C) := \mathrm{Tr}_S[\Lambda(\rho_S \otimes \eta_C)]. \quad (11)$$

It follows that a catalyst τ_{CE} must be a fixed point of $\Gamma_C \otimes \mathrm{id}_E$. However, by Theorem 3, the only channel on

C that can fix a TQ-Q τ_{CE} is the identity channel id_C. It follows that the channel Λ on SC must be factorized into $\Lambda_{SC} = \Xi_S \otimes \mathrm{id}_C$. It implies that there is no interaction between S and C, therefore the catalysis is trivial. □

We remark that our proofs did not assume that Λ is a free operation, hence the results are applicable to a framework of catalysis much more general than the conventional one where the interaction between system and catalyst should be a free operation. If we assume that Λ_{SC} is a free operation, Ξ_S must be free too, if feeding into Λ_{SC} a free state on C and discarding it are all free operations. Therefore, the whole transformation is of the form $\rho_S \to \Xi_S(\rho_S)$, which is simply a transformation through a free operation.

In principle, catalysts can be also used for non-free operations to reduce the cost (or boost the rate) of transformation, but Theorem 5 tells us that not even such generalized catalysis is possible with totally quantum catalysts.

Mutual information catalysis.— If preserving the whole state of τ_{CE} is too severe a constraint, one might want to preserve only one measure of its correlation, the mutual information $I(C:E)_\tau$. In other words, one might want to catalytically implement the transformation $\rho_S \to \sigma_S := \mathrm{Tr}_{CE}[\Lambda_{SC} \otimes \mathrm{id}_E(\rho_S \otimes \tau_{CE})]$ with the constraint that $I(C:E)_\tau = I(C:E)_\eta$, where

$$\eta_{CE} := \mathrm{Tr}_S[\Lambda_{SC} \otimes \mathrm{id}_E(\rho_S \otimes \tau_{CE})]. \tag{12}$$

If the above holds, then we say that the transformation $\rho \to \sigma$ is MI(mutual information)-catalytically implemented.

For this purpose, we first prove the following lemma: if one is required to preserve the mutual information of a TQ-Q state ρ_{AB}, then the only actions one can locally apply on the system A are unitary operations.

Lemma 6. If ρ_{AB} is a TQ-Q state, then for any quantum channel $\mathcal{N}$ on A with $\sigma_{AB} := (\mathcal{N}_A \otimes \mathrm{id}_B)(\rho_{AB})$ satisfying $I(A:B)_\rho = I(A:B)_\sigma$ must be a unitary operation.

Proof. By the data processing inequality, we have $I(A:B)_\rho \geq I(A:B)_\sigma$. By the saturation condition of the data processing inequality, there exists a recovery channel $\mathcal{R}_\mathcal{N}$ acting on A such that [35]

$$((\mathcal{R}_\mathcal{N} \circ \mathcal{N})_A \otimes \mathrm{id}_B)(\rho_{AB}) = \rho_{AB}. \tag{13}$$

Since ρ_{AB} is a TQ-Q state, it implies that $\mathcal{R}_\mathcal{N} \circ \mathcal{N} = \mathrm{id}_A$. As the dimensions of input and output systems of $\mathcal{N}_A$ are same, it follows that $\mathcal{N}_A$ is a unitary operation [36]. □

By using a similar proof to that of Theorem 5, but substituting the usage of Theorem 3 with Lemma 6, we can show that this type of catalysis grants us no additional power either.

Corollary 7. A TQ-Q state τ_{CE} cannot be used to MI-catalytically implement a non-trivial transformation when only access to C is given.

This technique provides an answer to the following question: What if different parties try to utilize a multipartite state as a catalyst at the same time? One might wonder if it is possible for two local actions at different sites can cancel each other to enable the recovery of the mutual information. The following result shows that it is nevertheless impossible. In other words, Corollary 7 explicitly shows that indeed quantum correlation would be a hidden resource; whenever a catalyst is quantumly correlated with an environment, no catalysis is possible without destroying such quantum correlations, as quantified by the mutual information.

Proposition 8. For any TQ-Q state ρ_{AB} and two channels $\mathcal{N}_A$ and $\mathcal{M}_B$, if $(\mathcal{N}_A \otimes \mathcal{M}_B)(\rho_{AB}) = \rho_{AB}$, then $\mathcal{N}_A$ is a unitary operation.

Proof. Let $\tau_{AB} := (\mathcal{N}_A \otimes \mathrm{id}_B)(\rho_{AB})$ and $\sigma_{AB} := (\mathcal{N}_A \otimes \mathcal{M}_B)(\rho_{AB})$. By the data processing inequality, we have $I(A:B)_\rho \geq I(A:B)_\tau \geq I(A:B)_\sigma$. However, as $\sigma_{AB} = \rho_{AB}$, we have $I(A:B)_\rho = I(A:B)_\tau$. By Lemma 6, it follows that $\mathcal{N}$ is a unitary operation. □

General correlated catalysts.— One may get the impression that the results above only have implications for a restricted class of bipartite states that are TQ-Q. However, any non TQ-Q state τ_{CE} is a PC-Q state. In particular, as a consequence of the Koashi-Imoto theorem [37], any bipartite state τ_{CE} can be decomposed into the following form,

$$\tau_{CE} = \sum_i p_i \tau_{C_i^L} \otimes \tau_{C_i^R E}, \tag{14}$$

where $C = \bigoplus_i C_i$ is a direct sum of subspaces $C_i := C_i^L \otimes C_i^R$ and each $\tau_{C_i^R E}$ is either a TQ-Q state, or $|C_i^R| = 1$ (in which case $\tau_{C_i^R E}$ is uncorrelated, but we include it for completeness). See Supplemental Materials for a self-contained and elementary proof of the Koashi-Imoto theorem based on that of Ref. [38].

This observation tells us how restricted the usage of a general correlated catalyst with local access is. The subspaces C_i of the decomposition above can be interpreted as the 'classical degrees of freedom' for C that can be read out without disturbing τ_{CE}, and any quantum channel on C preserving τ_{CE} also preserve these sectors. It naturally leads us to the following conclusion. (See Supplemental Materials for a detailed discussion.)

Theorem 9. If a transformation can be catalytically achieved by using the catalyst $\tau_{CE} = \sum_i p_i \tau_{C_i^L} \otimes \tau_{C_i^R E}$ with access to C, such that τ_{CE} is preserved, then the same transformation can be achieved by an ensemble $\left\{p_i, \tau_{C_i^L}\right\}$ of local catalysts.

Here, by the ensemble $\left\{p_i, \tau_{C_i^L}\right\}$, we mean the probabilistic mixture of quantum states with classical handle of index i, distinguished from the mixed state $\sum_i p_i \tau_{C_i^L}$. Intuitively, using τ_{CE} on C goes as follows: First, one measures which subspace C_i it is supported on without disturbing τ_{CE}. Since no action can act on C_i^R without disturbing $\tau_{C_i^R E}$, only $\tau_{C_i^L}$ can be utilized as a catalyst. Naturally, it is equivalent to having a local catalyst $\tau_{C_i^L}$ with probability p_i. In summary, quantum correlations in correlated catalyst only hinder catalysis, and only the classical part of correlation allows for catalysis because the property of non-invasive measurability, which is essential for recovering the catalyst state.

Conclusion. — We inspect the extent to which external correlations that a catalyst might have with its environment (inaccessible to agent) would be affected, when the catalyst is used to facilitate a process. We find that correlations of a quantum-mechanical nature, i.e. contained in TQ-Q states, necessarily degrade in the process of utilizing the catalyst to perform non-trivial transformations. This cautions against potential embezzlement via the consumption of correlations as resources unaccounted for. Alternatively, from a more constructive viewpoint, our work also shows that there is no advantage of a catalyst bank in creating quantumly correlated states and loaning parts of the states out for catalytic purposes – they might as well prepare classical ensembles of various independent ancillas.

We emphasize again that the 'classicality' here means the non-invasive measurability which means that one can measure a system without altering the state. Hence, our results are not in conflict with previous results on quantum catalysis, where 'classicality' indicates other properties, e.g., non-entangled [39], non-coherent [40], non-imaginary [41], etc. In other words, even when quantum properties of catalyst are utilized, the *state* should be classically known to its user.

Acknowledgments. SHL thanks Chae-Yeun Park for helpful discussions on the mathematical proofs. This work was supported by the start-up grant of the Nanyang Assistant Professorship of Nanyang Technological University, Singapore.

SUPPLEMENTAL MATERIALS

Notations

Throughout this Supplemental Materials we will use the following notations. First, every Hilbert space associated with a quantum system is assumed to be finite-dimensional and $|\mathcal{H}|$ denotes the dimension of a vector space $\mathcal{H}$. The operator space over a Hilbert space $\mathcal{H}$ is denoted by $B(\mathcal{H})$. We slightly abuse the notation and denote the identity map on $B(\mathcal{H})$ by $\text{id}_{\mathcal{H}}$. Also, we will sometimes say that a linear map Φ defined on $B(\mathcal{H})$ is a linear map on $\mathcal{H}$, in the sense we identify the operator space associated with a physical system with the system itself. The identity operator in $B(\mathcal{H})$, on the other hand, is denoted by $\mathbb{1}_{\mathcal{H}}$. A linear map Φ on $B(\mathcal{H})$ is called completely positive (CP) when $\Phi \otimes \text{id}_{\mathcal{K}}$ is positive for any Hilbert space $\mathcal{K}$. A linear map Φ on $B(\mathcal{H})$ is unital if $\Phi(\mathbb{1}_{\mathcal{H}}) = \mathbb{1}_{\mathcal{H}}$.

For any subspace $\mathcal{K}$ of a Hilbert space $\mathcal{H}$ and a linear map Φ on $B(\mathcal{H})$, we define its limitation $\Phi|_{\mathcal{K}}$ as Φ whose domain is limited to $B(\mathcal{K})$ without limiting its image. A 'quantum channel on $\mathcal{K}$' means a quantum channeldefined on $B(\mathcal{K})$ whose image is also in $B(\mathcal{K})$. We will also identify matrices with operators; the term 'matrix' will be used to emphasize its algebraic properties. Especially, a 'full matrix algebra' is a full operator set over a finite-dimensional Hilbert space, emphasizing that operation composition, i.e. matrix multiplication is well-defined. For any $M \in B(\mathcal{H})$, the linear map Ad_M is given as $\text{Ad}_M(\rho) := M\rho M^\dagger$. For any $Q \in B(\mathcal{H})$ such that $Q \geq 0$, $\text{supp}\,(Q)$ is used to denote the support of Q, the sum of eigenspaces corresponding to strictly positive eigenvalues of Q.

Structure theorem for fixed points of quantum channel

This section contains the proof of our first main result, i.e. Theorem 3 in the main text. We first provide the proof of a technical result, known in literature [27] as the structure theorem, and use it subsequently to prove Theorem 3.

Block-diagonal structure

Of central importance for this section is the structure theorem for the fixed point set of a quantum channel, which we state as S.Theorem 10. Many proofs of this theorem relies on the Artin-Wedderburn theorem, where the proof often requires mathematically advanced tools such as ring theory [42, 43], functional analysis [27, 38], or lengthy linear algebraic arguments [37]. In this section, we give an elementary and self-contained proof, by focusing on the more concrete case of our interest.

We begin with the following.

S.Lemma 1. For a unital CP map Φ, if $\Phi(\rho) = \rho$ when ρ is a Hermitian operator with the spectral decomposition $\rho = \sum_{i=1}^n \lambda_i \Pi_i$, we have $\Phi(\Pi_i) = \Pi_i$ for all i.

Proof. Without loss of generality, we let $\lambda_i > \lambda_{i+1}$ for all i. Let $r_i := \text{Tr}[\Pi_i]$ and $\pi_i := \Pi_i / r_i$, then we get that $q_i := \text{Tr}[\pi_n \Phi(\Pi_i)]$ is a probability distribution. With respect to this distribution,

$$\langle \lambda_i \rangle = \sum_i q_i \lambda_i = \mathrm{Tr}[\pi_n \Phi(\rho)] = \mathrm{Tr}[\pi_n \rho] = \lambda_n. \tag{15}$$

Since the left hand side is a mean of $\{\lambda_i\}$ and the right hand side is the smallest element of the set being averaged, it follows that the probability distribution $\{q_i\}$ must satisfy that $q_n = 1$ and $q_i = 0$ for any $1 \le i \le n-1$. It implies that $\Phi(\Pi_n) = \Pi_n$, however since Φ is unital, we also have that

$$\Phi\left(\sum_{i=1}^{n-1} \Pi_i\right) = \sum_{i=1}^{n-1} \Pi_i. \tag{16}$$

By limiting Φ onto the support of $\sum_{i=1}^{n-1} \Pi_i$, we can repeat the same argument and conclude that $\Phi(\Pi_{n-1}) = \Pi_{n-1}$. Repeating this process n times gives us the desired result. □

S.Lemma 2. For a unital CP map Φ with Kraus operators $\{K_i\}$, Φ fixes ρ if and only if $[K_i, \rho] = \left[K_i^\dagger, \rho\right] = 0$ for all i.

Proof. Showing the "if" part is trivial; $\Phi(\rho) = \sum_i K_i \rho K_i^\dagger = \rho \sum_i K_i K_i^\dagger = \rho$. For the other direction, first let us assume that ρ is Hermitian with the spectral decomposition $\rho = \sum_j \lambda_j \Pi_j$. If $\Phi(\rho) = \rho$, then by S.Lemma 1 we get that $\Phi(\Pi_j) = \Pi_j$ for all j. Next, by conjugating Π_j on $\Phi(\Pi_k) = \sum_i K_i \Pi_k K_i^\dagger = \Pi_k$, we get

$$\sum_i T^i_{jk} T^{i\dagger}_{jk} = \delta_{jk} \Pi_k \tag{17}$$

where $T^i_{jk} := \Pi_j K_i \Pi_k$. From the positivity of each $T^i_{jk} T^{i\dagger}_{jk}$, we get that $T^i_{jk} = \Pi_j K_i \Pi_k = 0$ for all i, whenever $j \neq k$. It means that every K_i is block-diagonal with respect to $\{\Pi_j\}$;

$$K_i \Pi_j = K_i(\mathbb{1} - \sum_{k\neq j} \Pi_k) = K_i - \sum_{k \neq j} \Pi_k K_i \Pi_k = \Pi_j K_i. \tag{18}$$

Hence, $[K_i, \Pi_j] = 0$ for all i and j. Finally, this also then implies that $[K_i, \rho] = 0$. By using the fact that a general matrix can be decomposed into Hermitian and anti-Hermitian parts, and a similar analysis for the anti-Hermitian part, we get the desired result. □

As a side note to complement S.Lemma 2, we remark that for any set of matrices $\{K_i\}$, the centralizer of $\left\{K_i, K_i^\dagger\right\}$ is the set of fixed points of a unital CP map. This can be shown by downscaling $K_i \to cK_i$ by some constant c such that $\sum_i K_i K_i^\dagger \le \mathbb{1}$ holds, and using $\{K_i\} \cup \left\{\sqrt{\mathbb{1} - \sum_i K_i K_i^\dagger}\right\}$ as Kraus operators to construct a unital CP map.

Our main point of interest for this section involves the set of fix points for a unital map Φ, for which we denote as F_Φ. By S.Lemma 2, we know that if $\rho_1, \rho_2 \in F_\Phi$, then their product $\rho_1 \rho_2 \in F_\Phi$ also. We remark that because of Hermitian-preserving property of Φ, if $X \in F_\Phi$ then $X^\dagger \in F_\Phi$. Furthermore, let us also define $\min F_\Phi$ as the set of *minimal, non-zero central projectors* in F_Φ. Here, a minimal projector means that it is not a sum of two nonzero projectors, and being central means that it commutes with every element in F_Φ.

We start with a quick observation on the relation between minimal projectors. Note that for any minimal projector S, $SMS = cS$ for some complex number c (we assume that M is Hermitian without loss of generality). Otherwise SMS has a nontrivial spectral decomposition, and a projector onto an eigenspace of SMS will be smaller than S, which violates that S is minimal.

S.Lemma 3. If two minimal projectors S and T commute, either $S = T$ or $ST = 0$.

Proof. This follows from the fact that if $ST \neq 0$ and $S \neq T$, then ST is a projector smaller than S, which contradicts that S is minimal. □

The next theorem, known as a variant of the Artin-Wedderburn theorem, states that for unital CP maps Φ, the structure of F_Φ always admits a particular decomposition, i.e. a direct sum according to projections unto $\min F_\Phi$. This particular approach focusing on finding matrix basis elements is inspired by Ref. [43].

S.Theorem 4 (Artin-Wedderburn [28])**.** For any unital CP map Φ,

$$F_\Phi = \bigoplus_{P \in \min F_\Phi} F_\Phi P. \tag{19}$$

Moreover, there exists a tensor product structure for each $\mathrm{supp}\,(P) = \mathcal{H}_P \otimes \mathcal{L}_P$ and $F_\Phi P$ factorizes into $B(\mathcal{H}_P) \otimes \mathbb{1}_{\mathcal{L}_P}$, where $B(\mathcal{H}_P)$ is the full matrix algebra on $\mathcal{H}_P$.

Proof. We start by noticing that every two different projectors $P_1, P_2 \in \min F_\Phi$ are orthogonal to each other by S.Lemma 3. Otherwise, P_1 cannot be minimal since $P_1 = P_1 P_2 + P_1 P_2^\perp$ as both $P_1 P_2$ and $P_1 P_2^\perp$ are central projectors. Next, one can observe that by definition, every central projector in F_Φ is a sum of minimal central projections. Especially, the identity operator, which is in F_Φ since Φ, if it is non-minimal, is then also the sum of all projectors in $\min F_\Phi$. Finally, each $F_\Phi P$ is an algebra with P as its unity.

Now, let us focus on each algebra $F_P := F_\Phi P$. Although P is a minimal central projector, there could in general be minimal projectors Q in F_P such that $Q < P$, even though they are not central, otherwise P would not be in $\min F_\Phi$. In particular, there exists a decomposition of P into minimal projectors

$$P = \sum_i Q_i. \tag{20}$$

By S.Lemma 3, we also know that all the Q_i in the decomposition of P are mutually orthogonal.

Consider the following relation between Q_i: $Q_i \sim Q_j$ if there exists $X \in F_P$ such that $Q_i X Q_j \neq 0$. We claim that this relation is an equivalence relation – the fact that it is reflexive and symmetric is straightforward, while transitivity is also true: if there exists X and Y in F_P such that $Q_i X Q_j \neq 0$ and $Q_j Y Q_k \neq 0$, then

$$Z := Q_i X Q_j Y Q_k \neq 0. \tag{21}$$

This follows from the fact that

$$Q_j Y Q_k Y^\dagger Q_j = c_{jk} Q_j, \tag{22}$$

with $c_{jk} = \|Q_j Y Q_k\|_2^2 / \operatorname{Tr}[Q_j]$ which is non-zero and similarly, $Q_i X Q_j X^\dagger Q_i = c_{ij} Q_i$ with $c_{ij} = \|Q_i X Q_j\|_2^2 / \operatorname{Tr}[Q_i] \neq 0$ so that $ZZ^\dagger = c_{ij} c_{jk} Q_i \neq 0$. Therefore the equivalence relation $\sim$ splits $\{Q_i\}$ into equivalence classes. Moreover, there exists only one equivalence class; Suppose we have two distinct equivalent classes $\mathcal{I}$ and $\mathcal{J}$ with $Q_{\mathcal{I}} := \sum \mathcal{I} = \sum_{i : Q_i \in \mathcal{I}} Q_i$ (similarly for $\mathcal{J}$), we then have

$$Q_{\mathcal{I}} F_P Q_{\mathcal{J}} = 0. \tag{23}$$

It follows that for any $X \in F_P$, $Q_{\mathcal{I}} X (P - Q_{\mathcal{I}}) = 0$, hence $Q_{\mathcal{I}}$ becomes central in F_P, which contradicts that P is a minimal central projector. As a result, for any two Q_i and Q_j, we have $Q_i \sim Q_j$, i.e., there exists $X \in F_P$ such that $Q_i X Q_j \neq 0$.

Now, we let $E_{11} := Q_1$ and $E_{1j} := Q_1 X_j Q_j$ for some X_j, which is guaranteed to exist, such that $\|Q_{1j}\|_2 = \operatorname{Tr}\left[Q_1 X_j Q_j X_j^\dagger\right] = \operatorname{Tr}[Q_1]$. By letting $X_1 := Q_1$, we can interpret E_{11} as a special case of E_{1i}. Then we define $E_{i1} := E_{1i}^\dagger = Q_i X_j^\dagger Q_1$ and $E_{ij} := E_{i1} E_{1j} = Q_i X_{ij} Q_j$ where $X_{ij} := X_i^\dagger Q_1 X_j$ for all $i, j > 1$. Because of the property

$$Q_i X Q_i = (\operatorname{Tr}[Q_i X] / \operatorname{Tr}[Q_i]) Q_i, \tag{24}$$

for all i and $X \in F_P$, we have $E_{1i} E_{j1} = Q_1 X_i Q_i Q_j X_j^\dagger Q_1 = \delta_{ij} \left(\operatorname{Tr}\left[Q_1 X_i Q_i X_i^\dagger\right] / \operatorname{Tr}[Q_1] \right) Q_1 = \delta_{ij} Q_1$. Therefore,

$$E_{ij} E_{kl} = E_{i1} (E_{1j} E_{k1}) E_{1l} = \delta_{jk} E_{il}, \tag{25}$$

and that $\operatorname{Tr}\left[E_{ij}^\dagger E_{kl}\right] = \delta_{ik} \delta jl$ so that $\|E_{ij}\|_2^2 = \operatorname{Tr}[Q_1]$ for all i and j. Especially, $E_{ii}^2 = E_{ii}$. Because $E_{ii} = Q_i (X_i^\dagger Q_1 X_i) Q_i = r Q_i$ for some positive number r, we have $r = 1$ so that $E_{ii} = Q_i$. It follows that $\operatorname{Tr}[Q_i] = \operatorname{Tr}[E_{i1} E_{1i}] = \operatorname{Tr}[E_{1i} E_{i1}] = \operatorname{Tr}[Q_1]$ for all i, i.e., every minimal projector Q_i has the same rank.

Next, note that for any $Z \in F_P$, we have $Q_i Z Q_j E_{ji} = Q_i (Z Q_j X_{ji}) Q_i = c Q_i$ for some complex number c and by taking the trace of the both hands we get $c = (\operatorname{Tr}[E_{ji} Z] / \operatorname{Tr}[Q_1])$ because $Q_j E_{ji} Q_i = E_{ji}$ and $\operatorname{Tr}[Q_i] = \operatorname{Tr}[Q_1]$. Therefore,

$$Q_i Z Q_j = \left(\operatorname{Tr}\left[E_{ij}^\dagger Z\right] / \operatorname{Tr}[Q_1] \right) E_{ij}, \tag{26}$$

and it follows that for any $Z \in F_P$, $Z = PZP = \sum_{i,j} Q_i Z Q_j = \sum_{i,j} \left(\operatorname{Tr}\left[E_{ij}^\dagger Z\right] / \operatorname{Tr}[Q_1] \right) E_{ij}$, so $\{E_{ij}\}$ is an orthonormal basis of F_P.

Now, we define a linear map Ψ given as for the basis elements $\{E_{ij}\}$

$$\Psi(E_{ij}) := |i\rangle\langle j|_{\mathcal{H}_P} \otimes \mathbb{1}_{\mathcal{L}_P}, \tag{27}$$

with some Hilbert spaces $\mathcal{H}_P$ and $\mathcal{L}_P$ such that $|\mathcal{L}_P| = \operatorname{Tr}[Q_1]$. We can see that it is an isomorphism from the fact that $\Psi(E_{ij}) \Psi(E_{kl}) = \delta_{jk} \Psi(E_{il})$ and $\operatorname{Tr}\left[\Psi(E_{kl})^\dagger \Psi(E_{ij})\right] = \delta_{ik} \delta_{jl} \operatorname{Tr}[Q_1] = \operatorname{Tr}\left[E_{kl}^\dagger E_{ij}\right]$. Therefore, one can see that $F_P = F_\Phi P$ is isomorphic to $B(\mathcal{H}_P) \otimes \mathbb{1}_{\mathcal{L}_P}$. □

For any $\rho_{AB} \geq 0$, if $\operatorname{Tr}_B(\rho_{AB}) = c \, |\psi\rangle\langle\psi|_A$ for some pure state $|\psi\rangle_A$ on A, then we necessarily have $\rho_{AB} = |\psi\rangle\langle\psi|_A \otimes \rho_B$. Thus, by considering the Choi matrix of each limitation $\Phi|_{\operatorname{supp}(P)}$, we have the following result.

S.Corollary 5. For a unital CP map Φ with the fixed point set F_Φ with the decomposition (19), each limitation $\Phi|_{\operatorname{supp}(P)}$ decomposes into

$$\Phi|_{\operatorname{supp}(P)} = \mathrm{id}_{\mathcal{H}_P} \otimes \Phi_{\mathcal{L}_P}, \tag{28}$$

for some unital CP map $\Phi_{\mathcal{L}_P}$ on $\mathcal{L}_P$ with $\mathbb{1}_{\mathcal{L}_P}$ as the unique fixed point.

Proof. The set of fixed points of the limitation $\Phi|_{\operatorname{supp}}(P)$ is $M(\mathcal{H}_P) \otimes \mathbb{1}_{\mathcal{L}_P}$. Therefore every Kraus operator of $\Phi|_{\operatorname{supp}(P)}$ commutes with every matrix of the form $A_{\mathcal{H}_P} \otimes \mathbb{1}_{\mathcal{L}_P}$, which implies that every Kraus operator is in the form of $\mathbb{1}_{\mathcal{H}_P} \otimes K_i$, hence $\Phi|_{\operatorname{supp}(P)} = \mathrm{id}_{\mathcal{H}_P} \otimes \Phi_{\mathcal{L}_P}$. □

S.Corollary 5 is a key tool that is necessary for our goal of proving the structure theorem. Additionally, we need a few technical lemmata. We present S.Lemma 7 and 9 from Ref. [44] with an alternative proof based on elementary linear algebra. The proofs of S.Lemma 6 and 8 are taken from Ref. [27] and presented here for completeness.

S.Lemma 6 (Proposition 6.8, [27])**.** If a quantum channel preserves an operator, then it also preserves the Hermitian and anti-Hermitian part of the operator. Moreover, the channel also preserves the positive and negative parts of the (anti-)Hermitian part.

Proof. Let Φ be a quantum channel and $\Phi(X) = X$. If $X = H + iA$ where both H and A are Hermitian, then $\Phi(H) - H = i(A - \Phi(A))$. Since the only operator that is both Hermitian and anti-Hermitian is 0, we have $\Phi(H) = H$ and $\Phi(A) = A$. Moreover, if $H = P - N$ where $P, N \geq 0$ and $PN = 0$, then by letting Π_P be the projector onto $\operatorname{supp}(P)$ and similarly for N, we have

$$\begin{aligned}\operatorname{Tr}[P] &= \operatorname{Tr}[\Pi_P(P-N)] = \operatorname{Tr}[\Pi_P\Phi(P-N)]\\ &\leq \operatorname{Tr}[\Phi(P)] = \operatorname{Tr}[P].\end{aligned} \tag{29}$$

□

S.Lemma 7. If λ is an eigenvalue of $A \in B(\mathcal{H})$, then the complex conjugate λ^* is an eigenvalue of $A^\dagger$. Moreover, the geometric multiplicity of λ for A is same with that of λ^* for $A^\dagger$.

Proof. The first part immediately follows from that $\det[A - \lambda\mathbb{1}] = 0$ is equivalent to $\det\left[A^\dagger - \lambda^*\mathbb{1}\right] = 0$. For the second part, recall that the geometric multiplicity of λ for A is equal to $|\mathcal{H}| - r(A - \lambda\mathbb{1})$ where $r(X)$ is the rank of an operator X. Because rank is invariant under the adjoint transformation, we have

$$|\mathcal{H}| - r(A - \lambda\mathbb{1}) = |\mathcal{H}| - r(A^\dagger - \lambda^*\mathbb{1}), \tag{30}$$

and thus we get the wanted result. □

S.Lemma 8 (Proposition 6.10, [27])**.** For any fixed point $\rho \geq 0$ of a quantum channel Φ, if Q is the projector onto the support of ρ, then $\operatorname{Tr}[(\mathbb{1} - Q)\Phi(Q)] = 0$ and

$$\sigma \leq Q \implies \Phi(\sigma) \leq Q. \tag{31}$$

Moreover, for projectors Q, the condition (31) is equivalent to

$$\Phi^\dagger(Q) \geq Q. \tag{32}$$

A similar argument can be given for N, too.

Proof. For minimal and maximal positive eigenvalues λ_m and λ_M of ρ, we have $\lambda_m Q \leq \rho \leq \lambda_M Q$, hence

$$\begin{aligned}0 &\leq \lambda_m \operatorname{Tr}[(\mathbb{1} - Q)\Phi(Q)]\\ &\leq \operatorname{Tr}[(\mathbb{1} - Q)\Phi(\rho)]\\ &= \operatorname{Tr}[(\mathbb{1} - Q)\rho]\\ &\leq \lambda_M \operatorname{Tr}[(\mathbb{1} - Q)Q] = 0.\end{aligned}$$

For the second part:
($\implies$) Let $Q^\perp := \mathbb{1} - Q$. Then,

$$\operatorname{Tr}\left[Q^\perp\Phi(Q)\right] = \operatorname{Tr}\left[\Phi^\dagger(Q^\perp)Q\right] = 0, \tag{33}$$

hence $Q^\perp\Phi^\dagger(Q^\perp)Q^\perp = \Phi^\dagger(Q^\perp)$. Using $\Phi(\mathbb{1}) = \mathbb{1}$, we get

$$\Phi^\dagger(Q) = Q + Q^\perp\Phi^\dagger(Q)Q^\perp \geq Q. \tag{34}$$

($\impliedby$) For any $\sigma \leq Q$, we have that

$$\begin{aligned}\operatorname{Tr}[\sigma] = \operatorname{Tr}[\sigma Q] &= \operatorname{Tr}\left[\sigma\Phi^\dagger(Q)\right]\\ &= \operatorname{Tr}[\Phi(\sigma)Q]\\ &\leq \operatorname{Tr}[\Phi(\sigma)] = \operatorname{Tr}[\sigma].\end{aligned}$$

Therefore, $\operatorname{Tr}[\Phi(\sigma)Q] \leq \operatorname{Tr}[\Phi(\sigma)]$, which implies that $\Phi(\sigma) \leq Q$. □

S.Lemma 9 (Theorem 2, [44])**.** For any quantum channel Φ on A, if every fixed point of $\Phi^\dagger$ is proportional to $\mathbb{1}_A$, then Φ also has a unique fixed density matrix ρ.

Proof. As a linear map on $B(A)$ (by identifying A with its associated Hilbert space), fixed points of Φ are equivalent to eigenvectors corresponding to the eigenvalue 1. If every fixed point of $\Phi^\dagger$ is proportional to $\mathbb{1}_A$, then it means that the geometric multiplicity, or the dimension of the eigenspace, of eigenvalue 1 of $\Phi^\dagger$ is 1. By S.Lemma 7, a linear map and its adjoint have the same eigenvalues and the same geometric multiplicities, the geometric multiplicity of 1 as an eigenvalue of Φ is also 1. It means that there is a unique operator ρ in $B(A)$ such that $\Phi(\rho) = \rho$. This ρ has to be Hermitian because of the Hermitian preserving property of Φ and moreover, $\rho \geq 0$ because if a Hermitian operator is fixed by a quantum channel, then both of its positive and negative parts should be fixed by the channel by S.Lemma 6, which makes the geometric multiplicity of 1 larger than 1. After the normalization, it follows that there is only a single quantum state fixed by Φ. □

S.Theorem 10 (Structure theorem for fixed points of quantum channel, [27])**.** For any quantum channel Φ on A, the set F_Φ of all fixed points of Φ have the decomposition of the form

$$F_\Phi = \bigoplus_i B(\mathcal{H}_i) \otimes \rho_i. \tag{35}$$

Here, for any vector space $\mathcal{K}$, $\mathcal{K} \otimes \rho_i := \{v \otimes \rho_i : v \in \mathcal{K}\}$.

Proof. By the previous lemmata, $F_{\Phi^\dagger}$ has the decomposition Eq. (19) and for any $P \in \min F_{\Phi^\dagger}$, $\Phi^\dagger|_{\operatorname{supp}(P)} = \operatorname{id}_{\mathcal{H}_P} \otimes \Phi^\dagger_{\mathcal{L}_P}$. Moreover, since $\Phi^\dagger(P) = P$, by S.Lemma 8, the image of $\Phi|_{\operatorname{supp}(P)}$ is also contained in $B(\operatorname{supp}(P))$, so that $\left(\Phi|_{\operatorname{supp}(P)}\right)^\dagger = \Phi^\dagger|_{\operatorname{supp}(P)}$. It follows that each $\Phi_{\mathcal{L}_P} := \Phi^{\dagger\dagger}_{\mathcal{L}_P}$ is a quantum channel on $\mathcal{L}_P$ and has the unique fixed point, say, ρ_P, by S.Corollary 5 and S.Lemma 9. It follows that, for any $X \in B(\mathcal{H}_P)$,

$$\begin{aligned}\Phi(X_{\mathcal{H}_P} \otimes \rho_P) &= \Phi|_{\mathcal{H}_P \otimes \mathcal{L}_P}(X_{\mathcal{H}_P} \otimes \rho_P)\\ =&\operatorname{id}_{\mathcal{H}_P}(X) \otimes \Phi_{\mathcal{L}_P}(\rho_P) = X_{\mathcal{H}_P} \otimes \rho_P.\end{aligned}$$

In other words, $B(\mathcal{H}_P) \otimes \rho_P \subseteq F_\Phi$. Therefore, the following direct sum

$$G_\Phi := \bigoplus_{P \in \min F_\Phi^\dagger} B(\mathcal{H}_P) \otimes \rho_P, \tag{36}$$

is a subspace of F_Φ as a vector space over the complex number field, because each summand is fixed by Φ.

Now, we count the dimension of each space. Recall that the dimension of a direct sum of vector spaces is the sum of the dimension of all the individual summand. Because $|B(\mathcal{H}_P)\otimes\rho_P| = |B(\mathcal{H})||\operatorname{span}\{\rho_P\}| = |B(\mathcal{H})|$ as ρ_P is understood as a single-point set and the span of it is 1-dimensional, we have

$$|G_\Phi| = \sum_{P\in\min F_{\Phi^\dagger}} |B(\mathcal{H}_P)| = \sum_{P\in\min F_{\Phi^\dagger}} |\mathcal{H}_P|^2. \tag{37}$$

On the other hand, similarly $|B(\mathcal{H}_P)\otimes\mathbb{1}_{\mathcal{L}_P}| = |B(\mathcal{H}_P)|$, therefore it follows that

$$|F_{\Phi^\dagger}| = \sum_{P\in\min F_{\Phi^\dagger}} |B(\mathcal{H}_P)| = \sum_{P\in\min F_{\Phi^\dagger}} |\mathcal{H}_P|^2. \tag{38}$$

Finally, note that $|F_\Phi| = |F_{\Phi^\dagger}|$ by S.Lemma 7, because they correspond to the geometric multiplicities of 1 for Φ and $\Phi^\dagger$, respectively. It follows that $G_\Phi = F_\Phi$ because their dimensions are the same. □

Proof of Theorem 3

Proof. The first part of the theorem is straightforward by looking at the contrapositive: if a quantum channel $\mathcal{N}$ has an output system that is not totally quantum, i.e. PC, then by Eq. (6) there exists a non-trivial projective measurement, which is a quantum channel that fixes $\mathcal{N}$, hence $\mathcal{N}$ cannot be sensitive.

We show the other direction. Let the output system of $\mathcal{N}$ be TQ and let us say that a quantum channel $\mathcal{S}$ fixes $\mathcal{N}$, i.e., $\mathcal{S}\circ\mathcal{N} = \mathcal{N}$. This means that the image of $\mathcal{N}$ is a subset of fixed points of $\mathcal{S}$. By S.Theorem 10, the space of all fixed points $F_\mathcal{S}$ of $\mathcal{S}$ must have the following unique decomposition,

$$F_\mathcal{S} = \bigoplus_i \mathcal{M}_{d_i}\otimes\sigma_i, \tag{39}$$

with respect to an appropriate basis and for some fixed full-rank quantum states σ_i and the full matrix algebras $\mathcal{M}_{d_i}$.

If there is more than one term in the direct sum of Eq. (39), we will call the set of projectors onto their supports $\{\Pi_i\}$. Note that the projective measurement described by $\{\Pi_i\}$ fixes all the states in $F_\mathcal{S}$, and therefore all the output states of $\mathcal{N}$. This contradicts the assumption that the output system of $\mathcal{N}$ is TQ. Hence, there should be only one term in the direct sum of Eq. (39) so that $F_\mathcal{S}$ should have the form of $F_\mathcal{S} = \mathcal{M}_d\otimes\sigma$ for some full matrix algebra $\mathcal{M}_d$ and a fixed quantum state σ. However, suppose that there are more than one distinct eigenvalue for σ so that there are at least two terms in the spectral decomposition of $\sigma = \sum_j \lambda_j P_j$, where P_j is the projector onto the eigenspace corresponding to the eigenvalue λ_j. Then, we can now identify a projective measurement that fixes $F_\mathcal{S}$, i.e. $\{\mathbb{1}_d\otimes P_j\}$ where $\mathbb{1}_d$ is the identity matrix in $\mathcal{M}_d$. It also violates the output system of $\mathcal{N}$ being non-PC, hence σ should be a 1-dimensional state so that $F_\mathcal{S}$ is isomorphic to a full matrix algebra, i.e., $\mathcal{S}$ fixes *every* operator. This means that $\mathcal{S} = \mathrm{id}$ and hence $\mathcal{N}$ is sensitive.

The statement for bipartite states follows the proof above by applying the Choi-Jamiołkowski isomorphism to $\mathcal{S}$. □

Proof of Proposition 4

Proof. Let us start by noting that for any isometry $V : A\to AK$, there exists a unitary operator U on AK such that $V = U(|\phi\rangle_K\otimes\mathbb{1}_A)$ for some state $|\phi\rangle_K$. Additionally, consider a measurement operation $\mathcal{P}_K(\cdot) = \sum_i \Pi_i\cdot\Pi_i$ with more than one mutually orthogonal projectors $\{\Pi_i\}$ on K that fixes τ_{KAB}. Consider an induced quantum channel on A defined as

$$\mathcal{Q}(\cdot) := \mathrm{Tr}_K[\mathcal{U}^\dagger\circ\mathcal{P}_K\circ\mathcal{U}(|\phi\rangle\langle\phi|_K\otimes\cdot_A)] \tag{40}$$

where $\mathcal{U}$ is the unitary channel describing the action of U on AK. Since $\mathcal{P}_E$ fixes τ_{KAB}, $\mathcal{Q}_A$ fixes ρ_{AB}. Since it is a TQ-Q state, it implies that $\mathcal{Q}$ is the identity channel. Since unitary operations cannot form correlation with other systems, (or by the Schrödinger-HJW theorem [45, 46]) that $\mathcal{U}^\dagger\circ\mathcal{P}_K\circ\mathcal{U}(|\phi\rangle\langle\phi|_K\otimes\mathrm{id}_A) = \sigma_K\otimes\mathrm{id}_A$ for some state σ_K. Applying this channel to ρ_{AB} on A, we get

$$\mathcal{U}^\dagger_{KA}\circ\mathcal{P}_E\circ\mathcal{U}_{KA}(|\phi\rangle\langle\phi|_K\otimes\rho_{AB}) = \sigma_K\otimes\rho_{AB}. \tag{41}$$

Note that $\mathcal{U}_{KA}(|\phi\rangle\langle\phi|_K\otimes\rho_{AB}) = \tau_{KAB}$ and $\mathcal{P}_K$ does not alter τ_{KAB}. Hence, the left hand side is $\mathcal{U}^\dagger_{KA}\circ\mathcal{U}_{KA}(|\phi\rangle\langle\phi|_K\otimes\rho_{AB}) = |\phi\rangle\langle\phi|_K\otimes\rho_{AB}$. It follows that $\sigma_K = |\phi\rangle\langle\phi|_K$.

Now, a quantum state η is fixed by measurement operation $\mathcal{P}(\cdot) = \sum_i \Pi_i\cdot\Pi_i$ if and only if $\Pi_i\eta\Pi_i = \eta$ for only one i. Moreover, $\mathcal{P}_K\otimes\mathrm{id}_A$ is also a measurement operation. Considering the Choi matrix, it follows that there exists a unique Π_i such that $(\Pi_i\otimes\mathbb{1}_A)U(|\phi\rangle_K\otimes\mathbb{1}_A) = U(|\phi\rangle_K\otimes\mathbb{1}_A)$ and $(\Pi_j\otimes\mathbb{1}_A)U(|\phi\rangle_K\otimes\mathbb{1}_A) = 0$ for any other Π_j. By conjugating the operator above to ρ_{AB} on A and tracing out B, we get that $\Pi_j\tau_K\Pi_j = 0$, which contradicts τ_A being full-rank. □

Proof of Theorem 9

The Koashi-Imoto Theorem and Decomposition of PC-Q state

Here, we prove that arbitrary bipartite state τ_{CE} can be decomposed into the from

$$\tau_{CE} = \bigoplus_i p_i \tau_{C_i^L} \otimes \tau_{C_i^R E} \tag{42}$$

where $C = \bigoplus_i C_i^L \otimes C_i^R$ and each $\tau_{C_i^R E}$ is a TQ-Q state. Equivalently, through the Choi-Jamiołkowski isomorphism, for any quantum channel $\mathcal{T} : E \to C$, one has the following unique decomposition.

$$\mathcal{T}(\rho) = \bigoplus_i \tau_{C_i^L} \otimes \mathcal{T}_{C_i^R}(\rho). \tag{43}$$

with respect to the same C above for any input state ρ, where $\tau_{C_i^L}$ is a quantum state on C_i^L and $\mathcal{T}_{C_i^R} : E \to C_i^R$ is a sensitive, trace non-increasing [47] CP map (or a trivial map with $|C_i^R| = 1$, which we simply count as a special case of sensitive map for simplicity). This result can be attained by using the Koashi-Imoto theorem [37] as a lemma. Here, we provide a concise statement and proof of the Koashi-Imoto theorem that mainly follows that of Ref. [38] without using the result of Ref. [48] directly, by using the tools that facilitated the structure theorem (S.Theorem 10) we developed earlier.

S.Lemma 11 (Koashi-Imoto [37, 38])**.** For any set of quantum states $\{\rho_k\}$ on a Hilbert space $\mathcal{H}$, there exists a unique decomposition of $\mathcal{H} = \bigoplus_i \mathcal{H}_{i^L} \otimes \mathcal{H}_{i^R}$ that satisfies the following:

(i) Each ρ_k decomposes as

$$\rho_k = \bigoplus_i q_{i|k}\, \omega_{i^L} \otimes \rho_{i^R|k}, \tag{44}$$

where $(q_{i|k})$ is a probability distribution over i and $\rho_{i^R|k}$ is a quantum state on $\mathcal{H}_{i^R}$ depending on k, while ω_{i^L} is a quantum state on $\mathcal{H}_{i^L}$ independent of k.

(ii) Any quantum channel $\mathcal{C}$ on $\mathcal{H}$ that fixes all ρ_k is a quantum channel on each subspace $\mathcal{H}_{i^L} \otimes \mathcal{H}_{i^R}$ and

$$\mathcal{C}|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}} = \mathcal{C}_{\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}, \qquad \forall i, \tag{45}$$

where each $\mathcal{C}_{\mathcal{H}_{i^L}}$ fixes ω_{i^L}, i.e., $\mathcal{C}_{\mathcal{H}_{i^L}}(\omega_{i^L}) = \omega_{i^L}$.

Proof. Let $\mathbf{F} := \{\mathcal{F} : \mathcal{F}(\rho_k) = \rho_k, \forall k\}$ be the set of all quantum channels that fixes each $\{\rho_k\}$. Then, when $F_{\mathcal{C}}$ is the set of all fixed points of a linear map $\mathcal{C}$, we let

$$F_0 := \bigcap_{\mathcal{F}\in\mathbf{F}} F_{\mathcal{F}^\dagger}. \tag{46}$$

Since every $F_{\mathcal{F}^\dagger}$ is finite dimensional, F_0 can be actually expressed as a finite intersection:

$$F_0 = F_{\mathcal{F}_1^\dagger} \cap F_{\mathcal{F}_2^\dagger} \cap \cdots \cap F_{\mathcal{F}_M^\dagger}, \tag{47}$$

for some $\mathcal{F}_1, \cdots, \mathcal{F}_M \in \mathbf{F}$. To see this, observe that intersecting one more $F_{\mathcal{F}}$ can never increase the dimension of the intersection, hence becasue of the finite dimensionality, only a finite number of $F_{\mathcal{F}_i^\dagger}$ nontrivially affect the interaction $\bigcap_{\mathcal{F}\in\mathbf{F}} F_{\mathcal{F}^\dagger}$. Let us consider the quantum channel $\mathcal{F}_0$ given as

$$\mathcal{F}_0 := \frac{1}{M}\sum_{n=1}^{M} \mathcal{F}_n. \tag{48}$$

The Kraus operators of $\mathcal{F}_0^\dagger$ are simply the union of scalar multiples of those of $\mathcal{F}_i^\dagger$. Suppose now, that we have a state $\rho \in F_{\mathcal{F}_0^\dagger}$. Then, by S.Lemma 2, ρ commutes with all the Kraus operators of each $\mathcal{F}_i^\dagger$. Therefore, again by S.Lemma 2, ρ is fixed by all of $\mathcal{F}_i^\dagger$, i.e., $\rho \in F_{\mathcal{F}_1^\dagger} \cap F_{\mathcal{F}_2^\dagger} \cap \cdots \cap F_{\mathcal{F}_M^\dagger}$. Hence $\rho \in F_0$. It shows that $F_{\mathcal{F}_0^\dagger} \subseteq F_0$, and therefore $F_0 = F_{\mathcal{F}_0^\dagger}$. It follows that there exists the decomposition of the form (19)

$$F_{\mathcal{F}_0^\dagger} = \bigoplus_i \mathbb{1}_{\mathcal{H}_{i^L}} \otimes B(\mathcal{H}_{i^R}), \tag{49}$$

with respect to a decomposition $\mathcal{H} = \bigoplus_i \mathcal{H}_{i^L} \otimes \mathcal{H}_{i^R}$. Define $Q_i := \mathbb{1}_{\mathcal{H}_{i^L}} \otimes \mathbb{1}_{\mathcal{H}_{i^R}}$ to be the projector onto the subspace $\mathcal{H}_{i^L} \otimes \mathcal{H}_{i^R}$ of $\mathcal{H}$. Then, we can observe from the block-diagonal structure (49) that

$$Q_i \in F_{\mathcal{F}_0^\dagger} = \bigcap_{\mathcal{G}\in\mathbf{F}} F_{\mathcal{G}^\dagger} \subseteq F_{\mathcal{F}^\dagger} \tag{50}$$

for any $\mathcal{F} \in \mathbf{F}$. It means that $\mathcal{F}^\dagger(Q_i) = Q_i$ and thus $\mathcal{F}^\dagger(Q_i) \geq Q_i$ and $\mathcal{F}^\dagger(\mathbb{1} - Q_i) \geq \mathbb{1} - Q_i$ for any $\mathcal{F} \in \mathbf{F}$ and i because $\mathcal{F}^\dagger(\mathbb{1}) = \mathbb{1}$. By S.Lemma 8, it follows that, for all i and $\mathcal{F} \in \mathbf{F}$, $\mathcal{F}|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}}$ is a quantum channel on $\mathcal{H}_{i^L} \otimes \mathcal{H}_{i^R}$ and $\left(\mathcal{F}|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}}\right)^\dagger = \mathcal{F}^\dagger|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}}$.

Especially for $\mathcal{F}_0$, from (49) it follows that

$$\mathcal{F}_0^\dagger|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}} = \mathcal{G}_{0\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}, \tag{51}$$

with some unital CP map $\mathcal{G}_{0\mathcal{H}_{i^L}}$ on $\mathcal{H}_{i^L}$ with the one-dimensional fixed point set $\{c\mathbb{1}\}$. By S.Lemma 9,

$$\mathcal{F}_0|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}} = \mathcal{F}_{0\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}, \tag{52}$$

with some quantum channel $\mathcal{F}_{0\mathcal{H}_{i^L}}$ on $\mathcal{H}_{i^L}$ with a unique fixed quantum state, say, ω_{i^L}. Therefore,

$$F_{\mathcal{F}_0} = \bigoplus_i \omega_{i^L} \otimes B(\mathcal{H}_{i^R}). \tag{53}$$

It follows that every ρ_k, as a fixed point of $\mathcal{F}_0$, indeed has the expression of the form (44).

Now we prove (ii). For any $\mathcal{F} \in \mathbf{F}$, again, $\mathcal{F}^\dagger|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}} = \mathcal{F}^\dagger_{\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}$ for some quantum channel $\mathcal{F}_{\mathcal{H}_{i^L}}$ on $\mathcal{H}_{i^L}$, because $F_0 = F_{\mathcal{F}_0^\dagger} \subseteq F_{\mathcal{F}^\dagger}$, so $\mathcal{F}|_{\mathcal{H}_{i^L}\otimes\mathcal{H}_{i^R}} = \mathcal{F}_{\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}$. Since $\{\rho_k\} \subseteq F_{\mathcal{F}}$, it must be that $\omega_{i^L} \otimes \rho_{i^R|k}$ is a fixed point of $\mathcal{F}_{\mathcal{H}_{i^L}} \otimes \mathrm{id}_{\mathcal{H}_{i^R}}$ for all i. Thus, every $\mathcal{F}_{\mathcal{H}_{i^L}}$ must have ω_{i^L} as a fixed state. It proves (ii).

□

We apply this result on the image of the quantum channel $\mathcal{T}$, $\{\mathcal{T}(\rho)\}$ so that the input state ρ functions as the index k in the statement of the Koashi-Imoto theorem above. Hence, according to (i) above, for any ρ, $\mathcal{T}(\rho)$ decomposes as

$$\mathcal{T}(\rho) = \bigoplus_i \omega_{i^L} \otimes \mathcal{T}_{i^R}(\rho). \tag{54}$$

Since ω_{i^L} is independent of ρ, it follows that each $\mathcal{T}_{i^R}$ is linear in ρ. Moreover, it is immediate that they are CP. Now we claim that each $\mathcal{T}_{i^R}$ is sensitive, because, otherwise, there exists a non-identity quantum channel $\mathcal{N}_{i^R}$ on $\mathcal{H}_{i^R}$ such that $\mathcal{N}_{i^R} \circ \mathcal{T}_{i^R} = \mathcal{T}_{i^R}$ and it contradicts (ii) of Lemma 11.

Once we have the decomposition of arbitrary quantum channel (43), by using the Choi-Jamiołkowski isomorphism again, for each CP map $\mathcal{T}_{i^R}$ one can find the corresponding bipartite quantum state $\rho_{C_i^R E}$ and normalization factor $q_i := \mathrm{Tr}[\mathcal{T}_{i^R}(\mathbb{1}_{\mathcal{H}})]/|\mathcal{H}|$ to get (42), which should sum up to 1 due to the trace preserving property of $\mathcal{T}$. In this case, for any quantum channel $\mathcal{C}$ on C such that $\mathcal{C}_C \otimes \mathrm{id}_E(\rho_{CE}) = \rho_{CE}$, $\mathcal{C}$ should satisfy (45).

Proof of Theorem 9

We follow the logic of the proof of Theorem 3 above. Consider any interaction Λ on SC that implements the catalytic transformation of a state ρ_S on S using τ_{CE} with access to C. We consider the channel $\mathcal{S}$ on S, defined as

$$\mathcal{S}(\rho) := \mathrm{Tr}_{CE}[\Lambda_{SC} \otimes \mathrm{id}_E(\rho_S \otimes \tau_{CE})]. \tag{55}$$

Recall the statement of the theorem presumes the following decomposition of the catalyst with its environment:

$$\tau_{CE} = \sum_i p_i \tau_{C_i^L} \otimes \tau_{C_i^R E}. \tag{56}$$

Let $\mathcal{P} = \sum_j \Pi_j \cdot \Pi_j$ be the pinching map on C where each Π_j is the projector onto the subspace $C_j^L \otimes C_j^R$. This models the process that acts locally on the system after a projective measurement processes. Note that $\mathcal{P}_C \otimes \mathrm{id}_E$ fixes τ_{CE}, i.e.

$$\mathcal{P}_C \otimes \mathrm{id}_E(\tau_{CE}) = \tau_{CE}. \tag{57}$$

Thus, one can perform the pinching operation before applying Λ_{SC} without changing $\mathcal{S}$:

$$\begin{aligned} &\mathrm{Tr}_{CE}[\Lambda_{SC} \otimes \mathrm{id}_E(\rho_S \otimes \tau_{CE})] \\ &= \mathrm{Tr}_{CE}[\Lambda_{SC} \circ (\mathrm{id}_{SE} \otimes \mathcal{P}_C)(\rho_S \otimes \tau_{CE})] \\ &= \mathrm{Tr}_C[\Lambda_{SC} \circ (\mathrm{id}_S \otimes \mathcal{P}_C)(\rho_S \otimes \tau_C)]. \end{aligned} \tag{58}$$

Hence, from the discussion of the previous section, we get the factorization

$$\Lambda_{SC}(\rho_S \otimes \Pi_j \tau_C \Pi_j) = p_j \Lambda_{SC_j^L}(\rho_S \otimes \tau_{C_j^L}) \otimes \tau_{C_j^R} \tag{59}$$

with the limitation $\Lambda_{SC_j^L}$ of Λ_{SC} onto SC_j^L which is by itself a quantum channel on SC_j^L. The resultant transformation of ρ_S is therefore of the form

$$\mathcal{S}(\rho) = \sum_j p_j \,\mathrm{Tr}_C[\Lambda_{SC_j^L}(\rho_S \otimes \tau_{C_j^L})] \tag{60}$$

Therefore, one can see that $\mathcal{S}$ can be implemented with the ensemble $\left\{p_i, \tau_{C_i^L}\right\}$.

Conversely, any catalysis possible with the ensemble $\left\{p_i, \tau_{C_i^L}\right\}$ is also possible with τ_{CE} through the protocol explained in the main text after Theorem 9.

[1] C. Datta, T. Varun Kondra, M. Miller, and A. Streltsov, arXiv e-prints , arXiv (2022).

[2] D. Jonathan and M. B. Plenio, Physical Review Letters **83**, 3566 (1999).

[3] S. Daftuar and M. Klimesh, Physical Review A **64**, 042314 (2001).

[4] M. Klimesh, arXiv preprint arXiv:0709.3680 (2007).

[5] G. Aubrun and I. Nechita, Communications in Mathematical Physics **278**, 133 (2008).

[6] E. T. Campbell, Physical Review A **83**, 032317 (2011).

[7] F. Brandao, M. Horodecki, N. Ng, J. Oppenheim, and S. Wehner, Proceedings of the National Academy of Sciences **112**, 3275 (2015).

[8] J. Åberg, Physical Review Letters **113**, 150402 (2014).

[9] M. P. Müller, Physical Review X **8**, 041051 (2018).

[10] N. Shiraishi and T. Sagawa, Physical Review Letters **126**, 150502 (2021).

[11] H. Wilming, R. Gallego, and J. Eisert, Entropy **19**, 241 (2017).

[12] P. Boes, J. Eisert, R. Gallego, M. P. Müller, and H. Wilming, Physical Review Letters **122**, 210402 (2019).

[13] S. Rethinasamy and M. M. Wilde, Physical Review Research **2**, 033455 (2020).

[14] T. V. Kondra, C. Datta, and A. Streltsov, arXiv preprint arXiv:2102.11136 (2021).

[15] S. H. Lie and H. Jeong, Physical Review A **101**, 052322 (2020).

[16] S. H. Lie and H. Jeong, Physical Review Research **3**, 013218 (2021).

[17] S. H. Lie and H. Jeong, arXiv preprint arXiv:2104.00300 (2021).

[18] S. H. Lie and H. Jeong, arXiv preprint arXiv:2206.11469 (2022).

[19] C. Sparaciari, *Multi-resource theories and applications to quantum thermodynamics*, Ph.D. thesis, UCL (University College London) (2018).

[20] Y. Feng, R. Duan, and M. Ying, IEEE transactions on information theory **51**, 1090 (2005).

[21] T. Fritz, Mathematical Structures in Computer Science **27**, 850 (2017).

[22] M. M. Wilde and A. Mizel, Foundations of Physics **42**, 256 (2012).

[23] A. J. Leggett and A. Garg, Physical Review Letters **54**, 857 (1985).

[24] L. Clemente and J. Kofler, Physical Review Letters **116**, 150401 (2016).

[25] A. Pan, Physical Review A **102**, 032206 (2020).

[26] S. H. Lie and H. Jeong, Physical Review Letters **130**, 020802 (2023).

[27] M. M. Wolf, Lecture notes available at http://www-m5. ma. tum. de/foswiki/pub M **5** (2012).

[28] E. Artin, in *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, Vol. 5 (Springer, 1927) pp. 251–260.

[29] J. Wedderburn, Proceedings of the London Mathematical Society **2**, 77 (1908).

[30] T. A. Brun, American Journal of Physics **70**, 719 (2002).

[31] S. Gudder, Fuzzy sets and systems **155**, 18 (2005).

[32] A. Winter, IEEE Transactions on Information Theory **45**, 2481 (1999).

[33] A. Streltsov, H. Kampermann, and D. Bruß, Physical Review Letters **106**, 160401 (2011).

[34] C. H. Bennett and G. Brassard, Theoretical Computer Science **560**, 7 (2014).

[35] M. M. Wilde, *Quantum information theory* (Cambridge University Press, 2013).

[36] A. Nayak and P. Sen, Quantum Information & Computation **7**, 103 (2007).

[37] M. Koashi and N. Imoto, Physical Review A **66**, 022318 (2002).

[38] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Communications in mathematical physics **246**, 359 (2004).

[39] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Reviews of modern physics **81**, 865 (2009).

[40] A. Streltsov, G. Adesso, and M. B. Plenio, Reviews of Modern Physics **89**, 041003 (2017).

[41] K.-D. Wu, T. V. Kondra, S. Rana, C. M. Scandolo, G.-Y. Xiang, C.-F. Li, G.-C. Guo, and A. Streltsov, Physical Review A **103**, 032401 (2021).

[42] W. K. Nicholson, New Zealand J. Math **22**, 83 (1993).

[43] M. Brešar, Expositiones Mathematicae **28**, 79 (2010).

[44] D. Burgarth, G. Chiribella, V. Giovannetti, P. Perinotti, and K. Yuasa, New Journal of Physics **15**, 073045 (2013).

[45] E. Schrödinger, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 31 (Cambridge University Press, 1935) pp. 555–563.

[46] L. P. Hughston, R. Jozsa, and W. K. Wootters, Physics Letters A **183**, 14 (1993).

[47] This is also why the probabilities in Eq. (42) were omitted in Eq. (43), since w.l.o.g. they can be absorbed into $\mathcal{T}_{C_i^R}$.

[48] M. Takesaki *et al.*, *Theory of operator algebras I* (Springer New York, NY, 1979).

Quantum Simulation of Open Quantum Systems on NISQ Devises: Benchmarking Regularisation Approaches to Quantum Process Tomography

Ian Joel David[1] * Ilya Sinayskiy[1 2] † Francesco Petruccione[2 3] ‡

[1] *School of Chemistry and Physics, University of KwaZulu-Natal, Durban 4001, South Africa*
[2] *National Institute for Theoretical and Computational Sciences (NITheCS), Stellenbosch, South Africa*
[3] *School of Data Science and Computational Thinking, Stellenbosch University, Stellenbosch 7604, South Africa*

Abstract. Quantum Process Tomography (QPT) is essential for characterizing and validating quantum devices and algorithms, but unavoidable device noises can lead to non-physical quantum channels. This work compares regularization methods to tackle shot noise in QPT using commonly used single-qubit quantum channels. The methods' effectiveness is evaluated using the minimum eigenvalue of the Choi matrix and fidelity. Our findings demonstrate spectral transformations as the most effective solution for finite sampling issues in quantum channel reconstruction in the Noisy Intermediate-Scale Quantum (NISQ) era.

Keywords: Quantum Channel, Quantum Process Tomography, Optimisation, Spectral Transformations

Simulating large, complex quantum systems with classical computers is a computationally challenging problem and quickly becomes intractable with growing system size. For this reason, Quantum Simulation, the simulation of quantum systems with other well-controlled quantum systems (quantum computers), has been proposed [1, 2]. Several algorithms for quantum simulation have been developed. However, most of these algorithms are best suited for fault-tolerant settings. In these settings, quantum computers would provide a clear advantage over classical computers in simulating quantum systems and other tasks. Despite the advancements made in quantum hardware, large-scale fault-tolerant quantum computers are unlikely to appear in the nearest decade. This, and the growing accessibility of Noisy Intermediate Scale Quantum (NISQ) computers via cloud services, naturally lead to the development of the quantum simulation approaches in the NISQ era. Most NISQ-era quantum simulation algorithms are focused on Hamiltonian simulation or simulation of the unitary evolution.

It is well known that realistic quantum systems are always in unavoidable contact with the thermal environment, which leads to decoherence and dissipation [3]. It has been shown that one can use NISQ devices to perform a quantum simulation of simple open quantum systems [4].

The master equation describes the evolution of an open quantum system. In the simplest case, when one neglects environmental memory effects, the master equation has Gorini-Kossakowski-Sudarshan-Lindblad (GKSL) form [3]. The solution of the master equation is called a quantum channel. Our previous work focused on the NISQ quantum simulation of the non-convex geometry of CP-divisible and non-divisible channels (a particular class of open quantum evolutions with memory) [5].

However, the reconstructed quantum channels are usually non-physical due to finite sampling and noise in the NISQ devices. The process of quantum channel reconstruction is called quantum process tomography (QPT). In this contribution, we will benchmark various approaches to making the results of QPT physical (so-called regularisations of the process matrix). For this comparison, we simulate some common single-qubit quantum channels. We use two metrics, the minimum eigenvalue of the Choi matrix and the fidelity, to compare the effectiveness of these methods. Our results show that the spectral transformations perform the best overall in dealing with finite sampling present in reconstructing the quantum channel in the NISQ era [6].

References

[1] R. P. Feynman, Simulating physics with computers, Int. J. Theor. Phys, vol. 21, no. 6/7, 1982.

[2] Y. Manin, Computable and uncomputable, Sovetskoye Radio, Moscow, vol. 128, 1980.

[3] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems.* Oxford University Press, 2002.

[4] G. García-Pérez, M. A. C. Rossi, and S. Maniscalco, IBM Q Experience as a versatile experimental testbed for simulating open quantum systems. npj Quantum Inf 6, 1, 2020.

[5] I. J. David, I. Sinayskiy, F. Petruccione, Digital Simulation of Convex Mixtures of Markovian and Non-Markovian Single Qubit Channels on NISQ Devices, arXiv:2108.11343.

[6] I. J. David, I. Sinayskiy, F. Petruccione, Benchmarking Regularisation Methods for Quantum Process Tomography On NISQ Devices, 2023.

* ianjoeldavid290614@gmail.com
† sinayskiy@ukzn.ac.za
‡ petruccione@sun.ac.za

Classical Simulation of Two-Qubit Entangled States with One Bit of Communication

Peter Sidajaya[1] * Aloysius Dewen Lim[2] Baichu Yu[1 3 4] Valerio Scarani[1 2]

[1] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[2] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*
[3] *Shenzhen Institute for Quantum Science and Engineering, Southern University of Science and Technology, Nanshan District, Shenzhen, 518055, China*
[4] *International Quantum Academy (SIQA), Shenzhen 518048, China*

Abstract. Bell's theorem states that local hidden variables cannot fully explain the behavior of an entangled quantum state. Consequently, many have asked how much supplementary classical communication would be needed to simulate them. Our study aims to investigate the question of whether a partially entangled two-qubit state can be simulated with only one bit of communication. We present a semianalytical LHV model generated by a neural network, which approximates the behavior of partially entangled states. We characterize the error values in terms of a hypothesis testing scenario.

Keywords: nonlocality, entanglement, communication complexity, machine learning, neural network

1 Introduction to the problem

It is well established that Local Hidden Variables (LHVs) are inadequate to describe the behaviours of entangled quantum states. More recently, some have asked how much supplementary resources, especially classical communication, do we need to simulate entangled states [1, 2, 3]. The problem was originally posed as a means to gain a more intuitive understanding of the power of entanglement. However, since Toner and Bacon's protocol to simulate a maximally entangled two-qubit state, progress in the field has been slower [3]. There has been some works focusing on finding quantum behaviour that is unsimulatable with just one bit of communication [4, 5, 6], but none has been found as of yet. On the other hand, Renner et al. has shown a protocol to simulate some partially entangled two-qubit states with just a single bit of communication [7]. The question that we are trying to answer in this work is then: *Can the other partially entangled two-qubit states be simulated using a single bit of communication?*

2 Using Machine Learning to generate a protocol

In this work, we tried to answer the question for the rest of the two-qubit entangled states by using a neural network to generate numerical protocols that try to simulate their behaviours. There have been numerous uses of neural networks in nonlocality and entanglement [8, 9, 10], but our approach was inspired by the work of Krivachy et al. [11] where a neural network which was built with locality constraints in its architecture was used as an oracle to test whether a distribution is local. We modify the design of the network in order to use it to generate local strategies that can be done with one bit of supplementary communication.

The locality of the network is done by having separate networks represent the different parties and routing the different inputs according to what each party should receive. Communication is added to the model by first looking at one bit communication as a power for one party to choose between two options for both of them. This way, we can actually model one bit of communication by having two local models (which are neural networks in themselves) and a third one that takes in Alice's inputs and outputs a number between 0 and 1 which denotes the probability of Alice choosing the first strategy, and thus sending the bit 0 to Bob. The final output is then, a convex combination of the two local models averaged over the LHVs. The architecture is illustrated in Fig. 1.

3 A semianalytical approximation to the protocol

We first tested our model on the maximally entangled state and reobtained Toner and Bacon's model, with slight modifications. We then proceeded to train our neural network to simulate the behaviours of partially entangled states. The performance of our models, measured by the average Kullback-Leibler divergence of its behaviour with the quantum behaviour, is shown in Fig. 2. We then wrote down the functions that approximate the behaviours of the neural network, which we will call our *semianalytical* protocols.

The outputs of Alice are of the form of

$$P(A_1 = +1 \mid \hat{a}) = \frac{1}{2}(1 - \mathrm{sgn}(\hat{a} \cdot \vec{\lambda}_{a1} + b_{a1})),$$

where $\hat{\lambda}_{a1} = u_{a1}\vec{\lambda}_1 + \vec{\lambda}_2 + v_{a1}\hat{z}$ decides the hemisphere direction and $b_{a1} = w_{a1} + x_{a1}\vec{\lambda}_1 \cdot \hat{z} + y_{a1}\vec{\lambda}_2 \cdot \hat{z}$ decides the size of the hemisphere. Similarly,

$$P(A_2 = +1 \mid \hat{a}) = \frac{1}{2}(1 + \mathrm{sgn}(\hat{a} \cdot \vec{\lambda}_{a2} + b_{a2})),$$

[*] `peter.sidajaya@u.nus.edu`

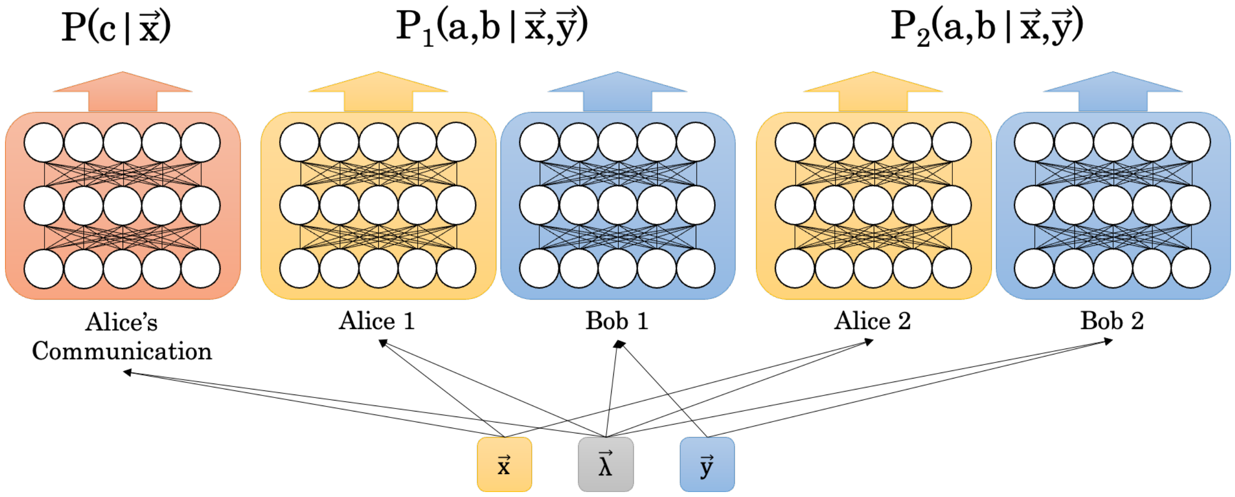


Figure 1: The architecture of the Artificial Neural Network (ANN). The model consists of two local distributions and a communication network and in each distributions the two parties are constrained by locality by routing the input accordingly. The communication network outputs a value between 0 and 1, and represents the probability of Alice sending a certain bit to Bob. The output for a particular round is then simply the convex combination of the two local distributions.

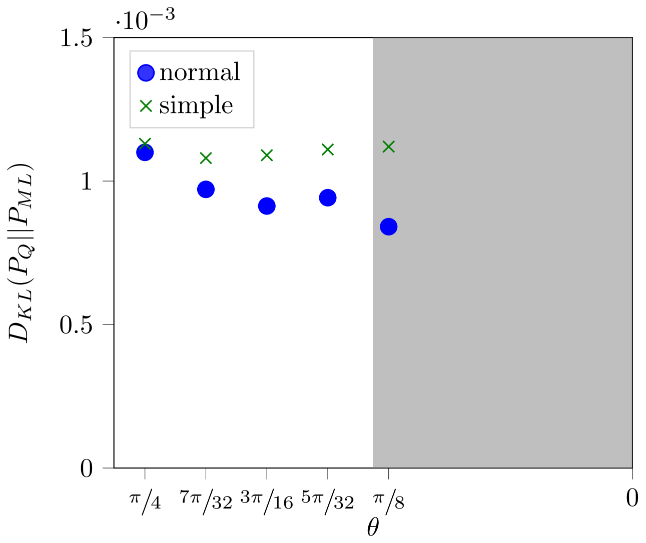


Figure 2: The relative error between the neural network models' behaviours and the quantum behaviours. The blue dots are for the original model described, while the green crosses are a simplified model where we reduces the perceptrons in the network, which we then used for the semianalytical approximations. The grey shaded region is the region in which Renner's LHV model works.

$$P(B_1 = +1 \mid \hat{b}) = \frac{1}{2}(1 + \mathrm{sgn}(\hat{b} \cdot \vec{\lambda}_{b1} + b_{b1})),$$

$$P(B_2 = +1 \mid \hat{b}) = \frac{1}{2}(1 - \mathrm{sgn}(\hat{b} \cdot \vec{\lambda}_{b2} + b_{b2})).$$

The bit of communication is given by

$$P(C = +1 \mid \hat{a}) = \frac{1}{2}(1 - \mathrm{clip}[c, -1, 1]),$$

where

$$\begin{aligned} c = {} & \Theta(\hat{a} \cdot \vec{\lambda}_1 + b_c)\Theta(\hat{a} \cdot \vec{\lambda}_2 + b_c) \\ & + \Theta(-\hat{a} \cdot \vec{\lambda}_1 + b_c)\Theta(-\hat{a} \cdot \vec{\lambda}_2 + b_c) \\ & - \Theta(-\hat{a} \cdot \vec{\lambda}_1 - b_c)\Theta(\hat{a} \cdot \vec{\lambda}_2 - b_c) \\ & - \Theta(\hat{a} \cdot \vec{\lambda}_1 - b_c)\Theta(-\hat{a} \cdot \vec{\lambda}_2 - b_c), \end{aligned}$$

with $b_c = u_c + v_c(\vec{\lambda}_2 \cdot \hat{z})(1 - \vec{\lambda}_1 \cdot \hat{z})$. Θ is the step function and clip is the clipping function.

The parameters u, v, w, x, y for all the parties are constants which depend on the states. We obtained the values of these parameters, which are listed in the manuscript, by using numerical methods.

We benchmarked our *semianalytical* protocol by comparing it with the original quantum behaviours and measuring its Kullback-Leibler divergences. To get a better intuition, consider a hypothesis testing scenario where we have a sample of n length generated by the same measurement done to n identical systems, but we only know that all the systems are quantum systems (P_Q), or our semianalytical LHV models (P_{LHV}), but we do not know which. If we take P_{LHV} as the null hypothesis, the probability for all possible decision making protocol of mistakenly rejecting a true null hypothesis (thus meaning that the sample comes from the LHV protocol, but we instead guess that it comes from an actual quantum system) is lower bounded by

$$a \geq e^{-nD_{KL}(P_Q||P_{LHV})}.$$

This can then be used to find the minimum number of sample size n needed such that we can be confident that we would be able to distinguish the two behaviours.

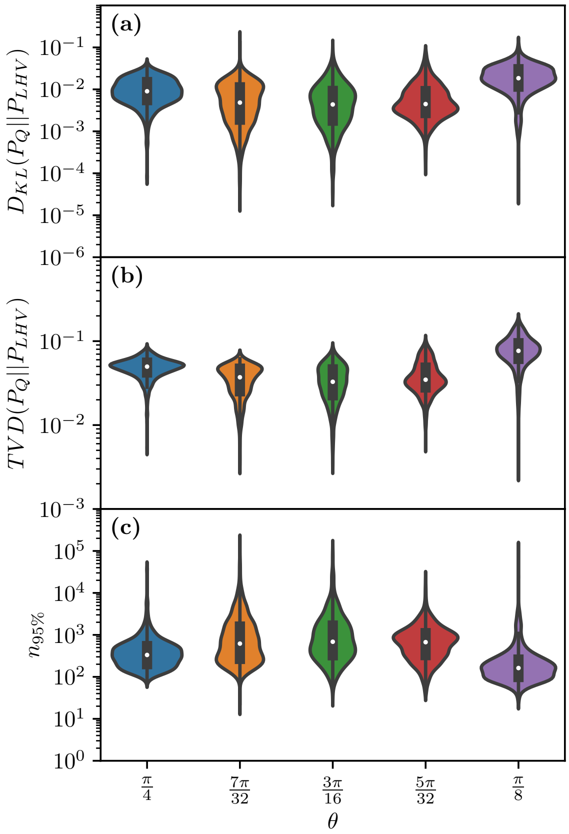


Figure 3: Violin plots describing the following values: **(a)** The Kullback-Leibler divergence between our semianalytical protocols and the quantum behaviours. **(b)** The Total Variational Distance between our semianalytical protocols and the quantum behaviours. **(c)** The minimum sample size needed to have at least 95% confidence in distinguishing the two behaviours as described in the hypothesis testing scenario. In all three, the distribution is over the different projective measurements on the two qubits. The white dots are the averages.

These benchmarks for our semianalytical protocols are shown in Fig. 3.

4 Conclusions

While the question of *exactly* simulating partially entangled states with one bit of communication remains unanswered, our works suggest that producing approximations of the quantum behaviours quite closely is possible. Our protocols requires, on average, hundreds of measurements before it could be distinguished from the real quantum behaviours, discounting other noises in an actual quantum system. Taking into account that some of the two-qubit states can already been simulated by an exact protocol, these evidences suggests that all two-qubit states can be simulated with just a single bit of communication. We still believe that an exact protocol simulating the states does exist, and we hope that we, or someone else, would be able to find it and close the problem for good.

As a side note, the question of the power of entanglement compared to classical communication continues to surprise us even when going to higher dimension. In the second part of our manuscript, we tried to find a quantum behaviour in higher dimensions that is unsimulatable by one bit of communication by verifying their membership in the communication polytope. However, we were unable to find any as most, if not all, behaviours of two-qutrit states with up to four measurement settings on each side can be simulated by a single bit of communication.

References

[1] M. Steiner, "Towards quantifying non-local information transfer: finite-bit non-locality," *Physics Letters A*, vol. 270, no. 5, pp. 239–244, 2000.

[2] J. A. Csirik, "Cost of exactly simulating a bell pair using classical communication," *Phys. Rev. A*, vol. 66, p. 014302, Jul 2002.

[3] B. F. Toner and D. Bacon, "Communication cost of simulating bell correlations," *Phys. Rev. Lett.*, vol. 91, p. 187904, Oct 2003.

[4] D. Bacon and B. F. Toner, "Bell inequalities with auxiliary communication," *Phys. Rev. Lett.*, vol. 90, p. 157904, Apr 2003.

[5] K. Maxwell and E. Chitambar, "Bell inequalities with communication assistance," *Phys. Rev. A*, vol. 89, p. 042108, Apr 2014.

[6] E. Zambrini Cruzeiro and N. Gisin, "Bell inequalities with one bit of communication," *Entropy*, vol. 21, no. 2, p. 171, 2019.

[7] M. J. Renner and M. T. Quintino, "The minimal communication cost for simulating entangled qubits," *arXiv preprint arXiv:2207.12457*, 2022.

[8] D.-L. Deng, "Machine learning detection of bell nonlocality in quantum many-body systems," *Physical review letters*, vol. 120, no. 24, p. 240402, 2018.

[9] A. Canabarro, S. Brito, and R. Chaves, "Machine learning nonlocal correlations," *Physical review letters*, vol. 122, no. 20, p. 200401, 2019.

[10] Y.-C. Ma and M.-H. Yung, "Transforming bell's inequalities into state classifiers with machine learning," *npj Quantum Information*, vol. 4, no. 1, p. 34, 2018.

[11] T. Kriváchy, Y. Cai, D. Cavalcanti, A. Tavakoli, N. Gisin, and N. Brunner, "A neural network oracle for quantum nonlocality problems in networks," *npj Quantum Information*, vol. 6, no. 1, pp. 1–7, 2020.

Security of differential phase shifted QKD against explicit individual attacks

Valliamai Ramanathan[1] * Anil Prabhakar[1] Prabha Mandayam[1]

[1] *Indian Institute of Technology, Madras*

Abstract. Quantum key distribution (QKD) is known to be unconditionally secure in principle, but quantifying the security of QKD protocols from a practical standpoint continues to remain an important challenge. Here, we focus on phase-based QKD protocols and characterize the security of the 3 and n-pulse differential-phase-shifted (DPS) quantum key distribution protocols against individual attacks. In particular, we focus on the minimum error discrimination (MED) and cloning attacks and obtain the corresponding bit error rates and the collision probability in the presence of these attacks. We compare the secure key rates thus obtained with the known theoretical lower bounds derived considering a general individual attack. In a departure from the theoretical lower bounds which have no explicit attack strategies, our work provides a practical assessment of the security of these phase-based protocols based on attacks with known implementations.

Keywords: Quantum Key Distribution, Differential Phase, Minimum error discrimination, Cloning.

1 Introduction

QKD offers the promise of secure communication over public networks, ideally with *unconditional security* [1, 2]. In practice, device imperfections and detector efficiencies compromise this notion of unconditional security. Furthermore, different QKD implementations are susceptible to different kinds of eavesdropping attacks, ranging from individual attacks like intercept-resend [3] to photon number splitting (PNS) attacks [4, 5] and collective attacks [6]. We refer to review articles [7, 8] for a comprehensive survey of the known secure key rate estimates, both in the ideal case as well as in the imperfect scenarios.

In this article, we focus on the specific class of differential-phase-shifted QKD protocols [9, 10, 11]. This class of phase-based protocols is known for its simplicity and efficient key generation. DPS QKD does not require basis reconciliation as in the case of the BB84 protocol and thus every bit that is detected contributes to the key. The ease of implementing this protocol makes it an practically relevant. Several variations of this protocol including round-robin DPS [12], the small-number-random DPS protocol [13], and measurement-device-independent DPS (MDI-DPS) [14] protocols have been studied in the recent past.

While the DPS protocol was originally proven to be secure against the basic individual attacks such as the intercept-resend and beam splitter attacks, it was subsequently shown to be secure against more general individual attacks [15]. Unconditional security was then proved under general coherent attacks in [16] for single photon DPS QKD. In proving unconditional security, the eavesdropper is considered to be capable of entangling ancillary systems to blocks of pulses and performing operations that are dependent on previous measurement results. Other attacks studied prior to the unconditional security bound [16], include specific collective attacks [17].

We study two individual attacks, namely, the minimum error discrimination and cloning attacks, whose effects on the secure key rate of the DPS QKD protocol have hitherto not been quantified. Specifically, we obtain the secure key rates for both the single-photon and weak-coherent-source 3-pulse DPS QKD protocols. We compare the secure key rates thus obtained with the known theoretical lower bounds arising from proofs of unconditional security. In what follows, we present a brief summary of our results and refer to [18] for detailed calculations and security analysis.

2 Preliminaries

We begin with a brief review of the n-pulse DPS QKD protocol [9, 10], schematically shown for the $n = 3$ case in Fig. 1. This schematic can be easily extended to n pulses by simply introducing more delay lines at the source, or equivalently, by phase modulating a single weak coherent pulse [19]. The sender (Alice) prepares a photon in an equal superposition of $n \geq 3$ pulses and encodes her bit values $\{0, 1\}$ via the relative phases $\{0, \pi\}$. The states sent by Alice in the $n = 3$ case can be written as,

$$|\psi\rangle_{\pm,\pm} = \frac{1}{\sqrt{3}}(|100\rangle \pm |010\rangle \pm |001\rangle). \tag{1}$$

The receiver (Bob) uses an unbalanced Mach Zehnder interferometer (MZI) for decoding, detects the relative phase between the pulses and hence obtains the bit value sent by Alice. Detection in any one of the $n-1$ time slots except the first and the last time bin gives the key bits, wherein the bit-value is assigned based on which detector (constructive or destructive) clicks.

The security of any QKD implementation can be quantified via the secure key rate. Assuming that the eavesdropper is restricted to only individual attacks, the asymptotic secure key rate can be calculated for the DPS QKD protocol, as [20],

$$R_{\text{sk}} = s\gamma p_{\text{click}} \; \{\tau - f(e)\,[h(e)]\}\,. \tag{2}$$

*ee21d201@smail.iitm.ac.in

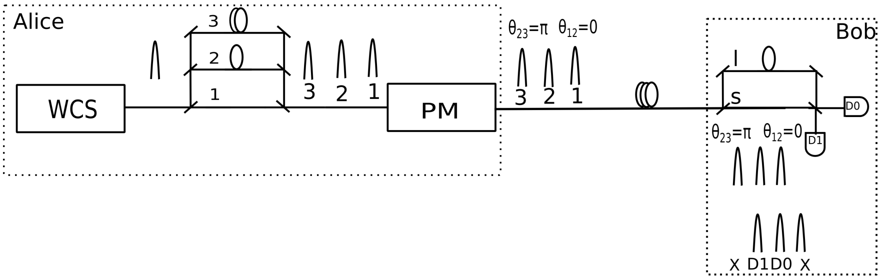


Figure 1: The 3-pulse DPS QKD protocol. PM - Phase Modulator, θ_{12} and θ_{23} are the phase introduced between the pulses by the phase modulator. $\theta_{12}, \theta_{23} \in \{0, \pi\}$. D0 and D1 are single photon detectors. Detection in the second and the third time slot gives θ_{12} and θ_{23} respectively.

Here, $R_{\text{sifted}} = s\gamma p_{\text{click}}$, where s is the sifting parameter, with $s = 2/3$ for 3-pulse case. γ is the repetition rate of the transmission, p_{click} is the probability of Bob's detection after taking into account detector inefficiencies, e is the bit error rate, and $f(e)$ characterises the performance of the classical error correction scheme. The parameter τ represents the shrinking factor due to privacy amplification and is calculated from the average collision probability, while $h(e)$ is the Shannon binary entropy [20]. The shrinking factor τ is a function of the average collision probability p_c.

$$\tau = -\log_2 p_c. \tag{3}$$

The average collision probability quantifies Eve's mutual information with Alice and Bob and can be evaluated in terms of the individual collision probability of each bit, denoted as p_{co}. In fact, for individual attacks, the overall collision probability is simply the product of the individual collision probabilities. In general, the collision probability is obtained as[15]

$$p_{\text{co}} = \sum_{x,z} p^2(X_i = x|Z_i = z)p(Z_i = z) \tag{4}$$

where X and Z are Alice's and Eve's key bit strings respectively.

For the most general individual attack, the collision probability for each bit can be bound as [15],

$$p_{\text{co}} \leq 1 - e^2 - (1 - 6e)^2/2 \tag{5}$$

Using this value in Eq. (2) gives a lower bound on the secure key rate obtainable against individual attacks.

3 Minimum error discrimination attack

In the MED attack, the eavesdropper (Eve) aims to identify the optimal quantum measurement to discriminate the set of signal states being used by the sender. The minimum error condition is enforced by maximizing the success probability of the state discrimination. Since the ideal (single-photon) n-pulse DPS states are linearly dependent for $n \geq 3$, the corresponding protocols are naturally secure against Unambiguous State Discrimination (USD) type attacks.

The ideal three-pulse single-photon DPS QKD scheme shown in Fig.1 encodes the logical bits in the relative phase. The states sent by Alice can be written as in Eq. (1). We assume that the states are chosen with equal probability, namely, 1/4. Thus to do a minimum error discrimination of these four states, Eve has to set up a positive operator valued measure (POVM) comprising of a set of four elements $\{P_1, P_2, P_3, P_4\}$, with each POVM element identifying one state perfectly. Minimum error discrimination can be stated as a semidefinite program (SDP), as described in [21]. The SDP corresponding to the MED of the DPS states in Eq. (1) is,

$$\begin{array}{rl} \text{maximize:} & \frac{1}{4}\langle \rho_i, P_i \rangle \\ \text{subject to:} & \sum_i P_i = I \\ & P_i > 0 \end{array} \tag{6}$$

Here, ρ_i is the density matrix formed from the kets corresponding to the DPS states. The SDP stated in Eq. (6) is run using the cvx solver [22, 23]. We find that each state has a 75% chance of being rightly identified by Eve. In [10] an intercept and resend attack was considered where Eve uses the same apparatus as Bob, and Eve introduces errors in 33% of the intercepted bits. In an intercepting attack using MED, Eve introduces only a 25% error. We now compare the key rate in the presence of these two attacks. The collision probability is found to be 0.72 for this attack and Eve can intercept up to $4e$ fraction of bits to go unnoticed and maintain the error rate at Bob.

The key rate is plotted in Fig.2 and it can be seen that the length up to which the protocol is secure reduces in the presence of the intercept and resend attack with MED as compared to that with intercept and resend using an MZI.

4 Cloning attack

The cloning attack is described by a quantum channel Φ, which takes the state $\rho \in \mathrm{D}(\mathcal{X})$ to the state $\Phi(\rho) \in \mathrm{D}(\mathcal{Y} \otimes \mathcal{Z})$. The channel Φ must correspond to a completely positive and trace-preserving linear mapping of the form $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{Z})$. Suppose the state that is sent is indexed k, the overall success probability of the cloning attack is,

$$\sum_{k=1}^{N} p_k \langle \psi_k \otimes \psi_k | \Phi (|\psi_k\rangle \langle \psi_k|) | \psi_k \otimes \psi_k \rangle \tag{7}$$

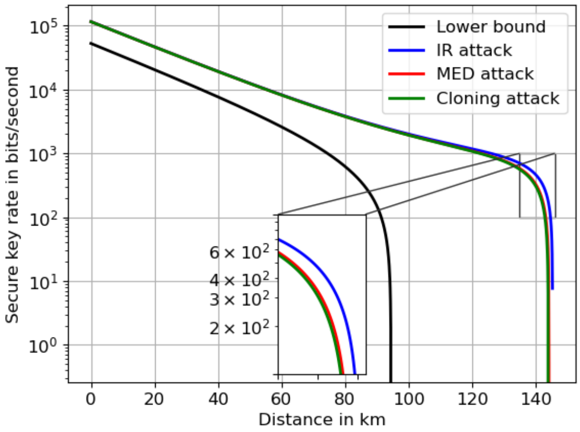


Figure 2: Secure Key rate vs channel length in the presence of intercept resend, Minimum error discrimination, cloning attack and the lower bound from[15]

The optimal success probability of a cloning strategy, which is represented by the supremum of the probability in Eq. (7) over all valid channels $\Phi : \mathrm{L}(\mathcal{X}) \to \mathrm{L}(\mathcal{Y} \otimes \mathcal{Z})$, may be represented by a semidefinite program[24].

The formulation makes use of the Choi-Jamiołkowski representation $J(\Phi) \equiv X$ of a given channel Φ. For finding a cloning machine for the DPS states, we consider the DPS three pulse state to be the states of a qutrit with basis states $\{|1\rangle, |2\rangle, |3\rangle\}$. The states are then written as,

$$|\psi\rangle_{\pm,\pm} = \frac{1}{\sqrt{3}}(|1\rangle \pm |2\rangle \pm |3\rangle) \tag{8}$$

The SDP solution for the probability of right cloning of the qutrit state gives a value of 0.78. We can obtain the map acting on Bob's state from the Choi matrix. The fidelity $\mathcal{F}$ between the original state and the cloned state is 0.81. This fidelity is obtained for all the states $|\psi(\pm,\pm)\rangle$.

The bit error rate at Bob's side due to the application of a cloning machine is found out using the probability of detecting in the constructive D1 (destructive D2) port when destructive (constructive) port is supposed to click. We write Bob's mixed state density matrix after the cloning machine as a convex combination in an appropriate basis, and then find out the probability of wrong detection.

To quantify the maximum information available to an adversary, we let Eve do a minimum error discrimination of the states reaching her after the cloning machine. Since the states reaching Eve and Bob are mixed, the MED success probability is relatively less than in the case where Eve does a MED directly on the states sent by Alice. Nevertheless, this strategy does marginally better since the error introduced at Bob due to cloning is less than the direct MED thus allowing for more bits to be attacked. Finally, we plot the secure key rate vs the distance for all the attacks discussed so far in Fig. 2.

5 Conclusion

We quantify the secure key rate for the 3-pulse DPS QKD protocol in the presence of minimum error discrimination and cloning attack. For both these attacks, we have identified the optimal strategy for the eavesdropper, for the specific set of signal states used in the 3-pulse DPS protocol. We further note that these attack strategies can be implemented by realising specific maps and are realistic in the sense that they do not require quantum memory.

Our results indicate that these sophisticated attacks are in fact comparable to the simple intercept and resend attack in terms of the information gained by the adversary. An eavesdropper gets only as much information as a simple ”Bob-like” intercept and resend strategy even if she uses the minimum error discrimination or cloning attack as described here. These explicit attack strategies give much higher secure key rates than the lower bound predicted using inequalities such as the Cauchy inequality.

Although unconditional security of DPS QKD is known under general, collective attacks, it is useful to benchmark the effectiveness of specific individual attacks so as to be able to assess their relative strengths. The fact that the secure key rates in the presence of the individual attacks studied here are still much higher than the lower bound on the secure key rate also suggests that these rates are of more practical relevance. Essentially, the secure key rates under such explicit attacks give a better sense of the realistic key rates and distances for QKD protocols.

References

[1] C. H. Bennett and G. Brassard, “Quantum cryptography: Public key distribution and coin tossing,” in *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, p. 175–179, December 1984.

[2] P. W. Shor and J. Preskill, “Simple proof of security of the BB84 quantum key distribution protocol,” *Phys. Rev. Lett.*, vol. 85, pp. 441–444, Jul 2000.

[3] M. Curty and N. Lütkenhaus, “Intercept-resend attacks in the Bennett-Brassard 1984 quantum-key-distribution protocol with weak coherent pulses,” *Phys. Rev. A*, vol. 71, p. 062301, Jun 2005.

[4] B. Huttner, N. Imoto, N. Gisin, and T. Mor, “Quantum cryptography with coherent states,” *Phys. Rev. A*, vol. 51, pp. 1863–1869, Mar 1995.

[5] N. Lütkenhaus and M. Jahma, “Quantum key distribution with realistic states: photon-number statistics in the photon-number splitting attack,” *New Journal of Physics*, vol. 4, no. 1, p. 44, 2002.

[6] E. Biham, M. Boyer, G. Brassard, J. Van De Graaf, and T. Mor, “Security of quantum key distribution against all collective attacks,” *Algorithmica*, vol. 34, no. 4, pp. 372–388, 2002.

[7] V. Scarani, H. Bechmann-Pasquinucci, N. J. Cerf, M. Dušek, N. Lütkenhaus, and M. Peev, "The security of practical quantum key distribution," *Rev. Mod. Phys.*, vol. 81, pp. 1301–1350, Sep 2009.

[8] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, "Secure quantum key distribution with realistic devices," *Rev. Mod. Phys.*, vol. 92, p. 025002, May 2020.

[9] K. Inoue, E. Waks, and Y. Yamamoto, "Differential phase shift quantum key distribution," *Phys. Rev. Lett.*, vol. 89, p. 037902, Jun 2002.

[10] S. Ranu, G. Shaw, A. Prabhakar, and P. Mandayam, "Security with 3-pulse differential phase shift quantum key distribution," *2017 IEEE Workshop on Recent Advances in Photonics (WRAP)*.

[11] K. Inoue, "Differential phase-shift quantum key distribution systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 21, no. 3, pp. 109–115, 2015.

[12] T. Sasaki, Y. Yamamoto, and M. Koashi, "Practical quantum key distribution protocol without monitoring signal disturbance," *Nature*, vol. 509, pp. 475–478, May 2014.

[13] Y. Hatakeyama, A. Mizutani, G. Kato, N. Imoto, and K. Tamaki, "Differential-phase-shift quantum-key-distribution protocol with a small number of random delays," *Phys. Rev. A*, vol. 95, p. 042301, Apr 2017.

[14] S. K. Ranu, A. Prabhakar, and P. Mandayam, "Differential phase encoded measurement-device-independent quantum key distribution," *Quantum Information Processing, Volume 20, Article no. 67*, Feb 2021.

[15] E. Waks, H. Takesue, and Y. Yamamoto, "Security of differential-phase-shift quantum key distribution against individual attacks," *Phys. Rev. A*, vol. 73, p. 012344, Jan 2006.

[16] K. Wen, K. Tamaki, and Y. Yamamoto, "Unconditional security of single-photon differential phase shift quantum key distribution," *Phys. Rev. Lett.*, vol. 103, p. 170503, Oct 2009.

[17] C. Branciard, N. Gisin, and V. Scarani, "Upper bounds for the security of two distributed-phase reference protocols of quantum cryptography," *New Journal of Physics*, vol. 10, p. 013031, jan 2008.

[18] V. Ramanathan, A. Prabhakar, and P. Mandayam, "Security of differential phase shifted QKD against explicit individual attacks," *arXiv quant-ph 2305.11822*, 2023.

[19] G. Shaw, S. Sridharan, S. Ranu, F. Shingala, P. Mandayam, and A. Prabhakar, "Time-bin superposition methods for DPS-QKD," *IEEE Photonics Journal*, vol. 14, no. 5, pp. 1–7, 2022.

[20] E. Diamanti, "Security and implementation of differential phase shift quantum key distribution systems," *ProQuest Dissertations and Theses*, p. 166, 2006.

[21] J. Watrous, "Lecture notes on quantum information theory - lecture 1,2,7,8." https://cs.uwaterloo.ca/ watrous/QC-notes/.

[22] C. R. Inc.", "Cvx: Matlab software for disciplined convex programming," 2011.

[23] M. Grant and S. Boyd, *Graph implementations for nonsmooth convex programs*. Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, 2008.

[24] A. Molina, T. Vidick, and J. Watrous, "Optimal counterfeiting attacks and generalizations for Wiesner's quantum money," *arXiv quant-ph 1202.4010*, 2012.

Maximal secret randomness of a quantum state

Shuyang Meng[1] * Fionnuala Curran[2] Máté Farkas[2] Gabriel Senno[2]
Victoria J. Wright[2] Valerio Scarani[1 3] † Antonio Acín [2 4] ‡

[1] *Department of Physics, National University of Singapore, 2 Science Drive 3, Singapore 117542*
[2] *ICFO-Institut de Ciències Fotòniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860 Castelldefels (Barcelona), Spain*
[3] *Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543*
[4] *ICREA - Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, Spain.*

Abstract. One of the most counter intuitive aspects of quantum theory is its claim that there is “intrinsic” randomness in the physical world. Quantum information science has greatly progressed in the study of randomness in the last decade. With a lot of emphasis on device-independent and semi-device-independent bounds, one of the most basic question has escaped attention: how much secret randomness can be extracted from a given state ρ, and what measurement would achieve that bound. We solve the min-max problem of finding the measurement that minimizes the maximal probability of an eavesdropper. The result is that one can guarantee an amount of randomness $H_{\text{min}} = -\log P_{\text{guess}}$ with $P^*_{\text{guess}}(\rho) = \frac{1}{d}\left(\text{tr}\sqrt{\rho}\right)^2$ by performing suitable projective measurements (in particular, a measurement in a basis that is completely unbiased with the eigenbasis of ρ).

Keywords: randomness, quantum measurement, quantum cryptography, quantum resource

1 INTRODUCTION

Randomness is one of the intrinsic properties that lies in the heart of quantum theory, which not only has fundamental implications for our world view, but is also attractive for practical use. The amount of randomness is naturally captured by the guessing probability P_{guess} of a potential eavesdropper Eve. This intuitive characterisation was found to have operational meaning: the *min-entropy* $H_{\text{min}} = -\log P_{\text{guess}}$ quantifies (informally) the fraction of perfect coin tosses than can be extracted from a string generated by the available source. There has been an explosion of designs and implementations of QRNGs certifiable under various assumptions[1, 2, 3], from device-independent[4], to semi-device-independent[5] or fully characterised devices[6].

However, one of the most basis questions was left out: how much secret randomness can be extracted from a known state ρ. We show that the answer is $H^*_{\text{min}} = -\log P^*_{\text{guess}}$ with

$$P^*_{\text{guess}}(\rho) = \frac{1}{d}\left(\text{tr}\sqrt{\rho}\right)^2 \tag{1}$$

and discuss the measurements that lead to this result.

2 RESULT

Alice holds a quantum state ρ from a Hilbert space of dimension d, to which she applies a measurement $\mathcal{M} = \{M_i\}$. We want to determine how random, that is, how unpredictable the measurement outcome is to a potential eavesdropper Eve, who has a more detailed knowledge than Alice about the process, but she cannot actively influence it (she is “outside the lab”). Concretely, in every round, Eve knows the true state ρ_c produced by the source. Given this knowledge, she guesses the most likely outcome $i = i(c)$ for that round. The state seen by Alice must be related to the states known by Eve through $\rho = \sum_c p_c\rho_c$. Without loss of generality, we can regroup all Eve's states that lead to the same guessed outcome (Eve does not gain anything in treating them as distinct). Eve's average guessing probability is therefore given by

$$P_{guess}(\{\rho_i, p_i\}, \mathcal{M}) = \sum_i p_i \,\text{Tr}(M_i\rho_i)\,. \tag{2}$$

Since we don't know the true states, we need to consider the worst case scenario, i.e. the decomposition that maximizes Eve's guessing probability:

$$\begin{aligned} P_{guess}(\rho, \mathcal{M}) &= \max_{\{\rho_i\}} \sum_i \text{Tr}(M_i\rho_i) \\ s.t. &\quad \rho_i \geq 0\,,\ \sum_i \rho_i = \rho \end{aligned} \tag{3}$$

where we redefined the density matrices so that $\text{Tr}(\rho_i) = p_i$. Actually, as long as $\mathcal{M}$ is predetermined, this value can be efficiently computed using positive semi-definite programming(SDP). However, in this context our focus is **how much randomness can be extracted from a known state** ρ. This requires optimizing Alice's measurement, i.e. computing

$$P^*_{guess}(\rho) = \min_{\mathcal{M}} P_{guess}(\rho, \mathcal{M})\,. \tag{4}$$

We solve this optimisation under the **assumption** that $\mathcal{M}$ is a rank-1 protective measurement: $i \in \{1, ..., d\}$ and $M_i = |q_i\rangle\langle q_i|$. The optimal guessing probability is (1). The core of the proof consists in rephrasing (3) as distance-based measurement

$$P_{guess}(\rho, \mathcal{M}) = \max_{\{\sigma\in\mathcal{I}\}} F(\rho, \sigma) \tag{5}$$

*e0572880@u.nus.edu
†valerio.scarani@gmail.com
‡antonio.acin@icfo.eu

where F is Uhlmann's fidelity and $\mathcal{I}$ is the set of states that are diagonal about the measurement basis $\{|q_i\rangle\}$. The invariance under unitary transforms of Uhlmann's fidelity lead to (1).

In summary, our main results are shown below. Theorem 1 gives the value of the maximal amount of randomness we can generate from ρ, and Theorem 2 gives the measurements that make this value achievable.

Theorem 1. *Under the assumption that Eve's best measurement is a projective measurement with rank 1 elements $M_i = |q_i\rangle\langle q_i|$, there is*

$$P^*_{guess}(\rho) = F(\rho, \mathbb{I}/d) = \frac{1}{d}(\text{Tr}\,\sqrt{\rho})^2\,, \tag{6}$$

Theorem 2. *When the measurement basis $\mathcal{M} = \{M_c\}$ satisfy:*

$$Tr(M_c\rho^{1/2}) \text{ are equal for } c = 1, 2, ...d. \tag{7}$$

it generates maximal randomness from ρ as in Theorem 1. Meanwhile, Eve's knowledge on this state in the worst case scenario corresponds to the decompostion

$$\rho_c = \rho^{1/2}M_c\rho^{1/2}. \tag{8}$$

which means the measurement basis is the "pretty good measurement" [7] according to the ensemble $\{p_x, \rho_x\}$, and 7 is saying that $Tr(M_c\rho_c)$ should be equal for all c's – the guessing probability is always the same no matter what the outcome turns out to be.

The most typical example of such measurement basis $\mathcal{M}$ is one that is formed by a basis that is maximally unbiased with the basis that diagonalizes ρ. This property of the measurement guarantees a trivial fulfilment of (7), and is therefore one of the desired measurement basis. We will call it "unbiased measurement" in the following text. For instance, applying this measurement to a qutrit(therefore $d = 3$), we have

$$P^*_{\text{guess}}(\rho) = \frac{1}{3}(\text{Tr}\,\sqrt{\rho})^2\,, \tag{9}$$

the solution of (3) using SDP coincides exactly with the expression (9), which verifies our proof(see Figure 1). It is important to note that unbiased measurement is not the only measurement that satisfies (7) and generates the maximal randomness.

3 Discussion

We answer the question of how much randomness one can generate from a fixed quantum state ρ (Theorem 1) and how to generate this amount of randomness (Theorem 2). One of the open questions remained is how to quantify the randomness if we allow Alice to use a positive operator-valued measure (POVM) and taking into consideration the information leaking from the measurement device[8].

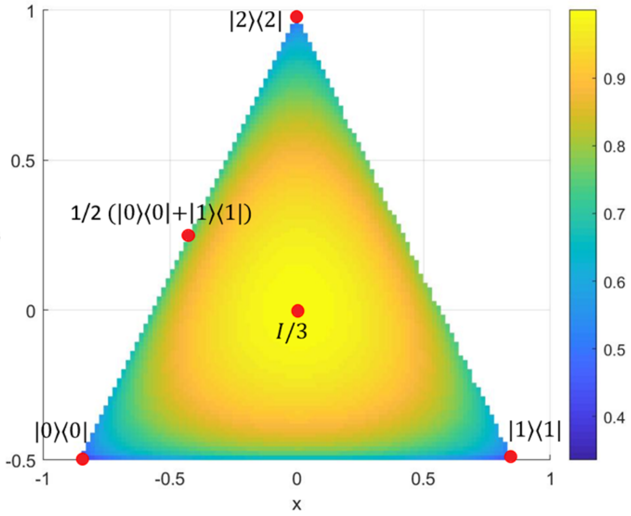


Figure 1: The color picture of $P_{\text{guess}}(\rho, \mathcal{M})$, where $\mathcal{M}$ is a measurement unbiased about the computational basis $|0\rangle$, $|1\rangle$, $|2\rangle$. ρ is a qutrit that can be regard as a mixture of state $|0\rangle$, $|1\rangle$, $|2\rangle$, where p_0, p_1, p_2 range from 0 to 1, forming a convex hull in the shape of an equilateral triangle. According to Theorem 2, it is as well a color picture of $P^*_{\text{guess}}(\rho)$.

References

[1] M. Herrero-Collantes and J. C. Garcia-Escartin, "Quantum random number generators," *Reviews of Modern Physics*, vol. 89, no. 1, p. 015004, 2017.

[2] M. N. Bera, A. Acín, M. Kuś, M. W. Mitchell, and M. Lewenstein, "Randomness in quantum mechanics: philosophy, physics and technology," *Reports on Progress in Physics*, vol. 80, no. 12, p. 124001, 2017.

[3] X. Ma, X. Yuan, Z. Cao, B. Qi, and Z. Zhang, "Quantum random number generation," *npj Quantum Information*, vol. 2, no. 1, pp. 1–9, 2016.

[4] Y. Liu, X. Yuan, M.-H. Li, W. Zhang, Q. Zhao, J. Zhong, Y. Cao, Y.-H. Li, L.-K. Chen, H. Li *et al.*, "High-speed device-independent quantum random number generation without a detection loophole," *Physical review letters*, vol. 120, no. 1, p. 010503, 2018.

[5] M. Avesani, D. G. Marangon, G. Vallone, and P. Villoresi, "Source-device-independent heterodyne-based quantum random number generator at 17 gbps," *Nature communications*, vol. 9, no. 1, p. 5365, 2018.

[6] T. Gehring, C. Lupo, A. Kordts, D. Solar Nikolic, N. Jain, T. Rydberg, T. B. Pedersen, S. Pirandola, and U. L. Andersen, "Homodyne-based quantum random number generator at 2.9 gbps secure against quantum side-information," *Nature Communications*, vol. 12, no. 1, p. 605, 2021.

[7] P. Hausladen and W. K. Wootters, “A ‘pretty good’measurement for distinguishing quantum states,” *Journal of Modern Optics*, vol. 41, no. 12, pp. 2385–2390, 1994.

[8] G. Senno, T. Strohm, and A. Acín, “Quantifying the intrinsic randomness of quantum measurements,” *arXiv preprint arXiv:2211.03581*, 2022.

Logical entanglement distribution between distant 2D array qubits

Yuya Maeda[1] * Yasunari Suzuki[2] † Toshiki Kobayashi[1 3] Yuuki Tokunaga[2]
Takashi Yamamoto[1 3] Keisuke Fujii[1 3]

[1] *Graduate School of Engineering Science, Osaka University*
[2] *NTT Computer and Data Science Laboratories*
[3] *Center for Quantum Information and Quantum Biology, Osaka university*

Abstract. Sharing logical entanglement pairs between distant quantum nodes is a key process for achieving fault-tolerant quantum computation and communication. Here, we propose an efficient logical entanglement distribution protocol based on surface codes for two distant 2D-array qubits with nearest-neighbor interaction. Our protocol allows tuning the trade-off relation between the success probability of this protocol and the infidelity of logical entanglements using post-selection. We numerically evaluated the performance of our protocol and the trade-off relations and showed that our protocol can prepare logical entangled states with improving fidelity.

Keywords: Logical entanglement distribution, Distributed fault-tolerant quantum computing, Surface code

1 Overview

A quantum network enables various quantum communication protocols, such as quantum cryptography [1] and distributed quantum computing [2, 3]. These applications can be achieved by preparing entangled qubit pairs between distant quantum devices via qubit transmission. While the qubit transmission suffers from loss and imperfection of communication channels, we can retrieve the fidelity of entangled pairs with post-selection of Bell measurement and entanglement distillation protocols [4, 5, 6].

In practice, even if we successfully generate high-fidelity entangled pairs with physical qubits, their fidelity would be immediately damaged due to their finite lifetimes. Thus, they must be kept encoded in the quantum error-correcting codes, with which we can project quantum states to the logical code space via stabilizer Pauli measurements and estimate Pauli errors from obtained values. Recently, the implementation of state-of-the-art error correcting codes, surface codes, has been demonstrated with superconducting circuits [7] or neutral atoms [8]. These devices enable experimental integration of several tens or hundreds of physical qubits on a 2D lattice and two-qubit gates acting on the nearest neighboring qubit pairs. The communication between two distant nodes is also experimentally demonstrated. Thus, the fault-tolerant communication of encoded logical qubits will be achieved in the near future. Evaluating the practical performance of logical entanglement distribution is also vital for exploring the design of distributed fault-tolerant quantum computing and quantum communications.

There have been massive efforts to explore the relation between high-performance entanglement distillation and quantum error correction [5, 6, 9, 10]. For example, Ref. [5] proved one-way entanglement distillation has the same ability of distilling entanglement to quantum error correcting code. Ref. [6] proposed a method to transform a stabilizer code into a two-way entanglement distillation protocol, enhancing its capability to distill entangled states. The protocol involves post-selection based on stabilizer measurements. However, the existing protocols do not consider logical entanglement distribution under realistic architectures, i.e., recent 2D integrated devices, and do not consider how we can leverage probabilistic generations of logical entanglement in these devices. Thus, an efficient entanglement distribution protocol tailored for current state-of-the-art devices is lacking.

In this work, we propose an efficient logical entanglement distribution protocol tailored for 2D-integrated quantum devices connected by noisy quantum channels. In this protocol, we assume that each node can share entangled states between physical qubits in parallel and can perform arbitrary two-qubit gates on nearest-neighboring pairs and single-qubit measurements. The protocol consists of three steps: position-wise entanglement generation, synchronized qubit rearrangement, and quantum error correction. First, we perform a probabilistic generation of physical entanglements. Then, we rearrange randomly distributed successful pairs only with nearest-neighbor interactions. Finally, we perform stabilizer measurements to project the state to the code space. In this step, we perform post-selection in terms of syndrome values, which enables us to tune the fidelity of logical entangled states at the cost of the success rate. It provides tunability in the design of quantum communication protocols, i.e., enables balancing between the rate and reliability of communication. We numerically analyzed the performance of our protocol and revealed the trade-off relation between the success rates and the fidelity of logical entanglement. Our results indicate that the protocol is feasible with current achievable experimental parameters. For example, when SWAP-gate fidelity is 0.995, error rates of logical entangled states can reach about 8.0×10^{-3} with a success probability larger than 0.3. Our results can be used as a baseline of practical logical entanglement distribution and can be used for estimating the bandwidth between fault-tolerant quantum comput-

*maeda-yuya@qi.mp.es.osaka-u.ac.jp
†yasunari.suzuki@ntt.com

ing nodes.

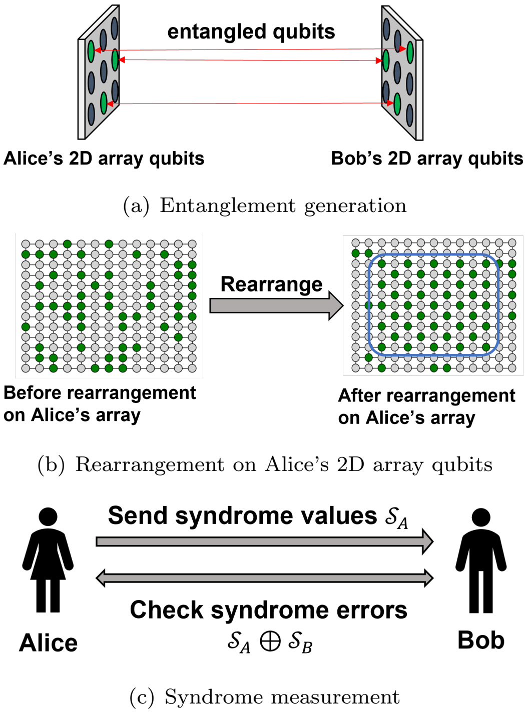
2 Protocol

Our protocol aims to share a logical entangled pair by sharing several physical entangled pairs and encoding them into the code space of surface codes. We show the three steps of this protocol and explain the detail of each step after that.

1. **Entanglement generation** Generate entangled states between two distant physical qubits on the 2D square lattice.

2. **Qubit rearrangement** Rearrange the locations of generated entangled pairs so that entangled pairs constitute 2D lattices. Note that the rearrangement process of Alice and Bob is the same since entangled pairs are located in the same positions.

3. **Syndrome measurement** Alice and Bob perform the syndrome measurements, Alice sends obtained syndrome values to Bob, and Bob performs error correction according to obtained values. This process projects the state of rearranged entangled pairs into the logical space of surface codes, which results in an entangled logical pair. If Bob finds that the number of disagreed syndrome values exceeds a threshold, they abort and restart the process.



(a) Entanglement generation

(b) Rearrangement on Alice's 2D array qubits

(c) Syndrome measurement

Figure 1: Sequence of the proposed protocol

Entanglement generation We assume that Alice and Bob have physical qubits aligned on the $L \times L$ 2D square lattice. They generate the entangled pairs between two qubits at the same coordinate in parallel. Fig. 1(a) is a conceptual diagram of entanglement generation. Green qubit pairs located at the same positions in each node represent successfully entangled qubit pairs. We denote the success probability of each generation protocol as p_{gen}. Note that this probability can be increased by repeating trials while the time for the generation process increases and the error rate of generated entangled states e_{init} is reduced. After the generation process, the entangled qubit pairs at each node are placed on the random sites in the lattice, but the arrangements of entangled pairs on two nodes are always the same since we perform entanglement generations for qubit pairs at the same position.

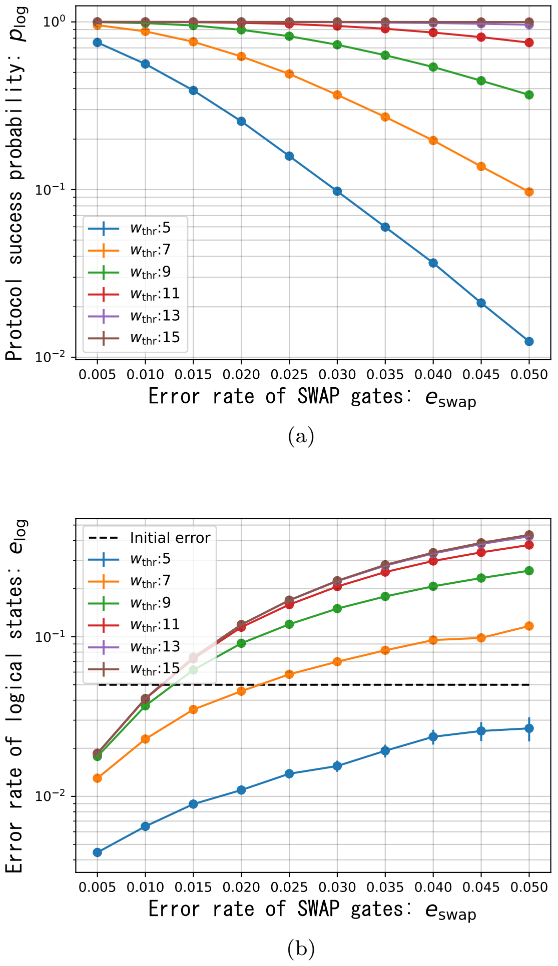
Qubit rearrangement In the rearrangement step, the generated entangled qubits are assembled to the data-qubit locations of surface codes as shown in Fig. 1(b). The distance of surface codes is chosen according to the number of successfully generated entangled states. SWAP gates on neighboring pairs are used for moving the positions of entangled qubits. We assumed SWAP gates suffer from uniform depolarizing noise, and the depolarizing rate is denoted by e_{swap}. We minimized the number of SWAP gates by converting the optimization problem into a matching problem between entangled qubit positions and target positions and numerically solved it.

Syndrome measurement Alice and Bob perform stabilizer Pauli measurements of surface codes to generate a logical entangled state. Then, Alice sends the syndrome values to Bob, and Bob marks the syndrome values that do not coincide with those observed in the same position at Alice's node as anomalous. If the number of the anomalous syndrome is no more than the acceptance threshold w_{thr}, Bob estimates errors and corrects them. If the estimation is correct, the state is projected to an expected logical entangled state. Otherwise, Bob informs the failure of the trial to Alice and requests to restart. Note that if w_{thr} is no less than the number of all the syndrome values, this protocol is never aborted, and Bob does not need to inform whether the trial is a failure or success to Alice, i.e., this protocol becomes one-way.

Our interest is to check the feasibility of the protocol in realistic parameter regions and explore the trade-off relationship between the parameters, such as the generation rate and fidelity of logical entangled states. The performance of this protocol depends on five parameters: the generation rate of entangled states p_{gen}, size of the 2D-array qubits L, initial error rate of the physical entangled states e_{init}, error rate of the SWAP gates e_{swap}, and acceptance threshold w_{thr}. Here, we can tune p_{gen} and e_{init} under a trade-off relation, and we can choose any threshold w_{thr} to obtain a desired performance.

3 Numerical evaluation

We examined the performance of our protocol with numerical simulation. The efficiency of this protocol is assessed based on the probability of post-selection p_{log}, which we call protocol success probability, and the error



(a)

(b)

Figure 2: The performance of our protocol is plotted as the function of error rates of SWAP gates e_{SWAP} for several acceptance thresholds w_{thr}. The protocol success probability $p_{\log}$ and the logical error rate $e_{\log}$ are plotted in (a) and (b), respectively. The initial error rate of physical entangled states e_{init} is shown as a black dotted line in (b).

rate of the logical states in the successful events $e_{\log}$. For simplicity, we evaluated the performance with fixed parameters in this extended abstract as $p_{\text{gen}} = 0.3$, $L = 19$, $e_{\text{init}} = 0.05$. The performance is evaluated with Monte-Carlo sampling, and we repeated 10^5 trials to evaluate each configuration. Note that the trials in this setting typically constitute $d = 7$ surface codes with few exceptional cases. So, we show the values calculated only from the samples with $d = 7$.

First, we plotted the protocol success probability $p_{\log}$ and the logical error rate of the post-selected state $e_{\log}$ as the function of e_{swap} in Fig. 2(a) and Fig. 2(b), respectively. Here, we varied e_{swap} from 0.05 to 0.005 and evaluated performances for several threshold w_{thr}. As expected, as the acceptance threshold becomes large, i.e., allowing large discrepancy of syndrome values, the post-selection probability increases, but the infidelity of resultant states is reduced. When we choose the acceptance threshold $w_{\text{thr}} = 5$, we can prepare the logical entanglement with reducing the error rates from physical ones if e_{swap} is below 0.03, which corresponds to a fidelity of about $F_{\text{CNOT}} = 0.9899$ in terms of CNOT gate conversion. Referring to the current experimental progress [7, 8, 11, 12], we can expect that this setting is a feasible example.

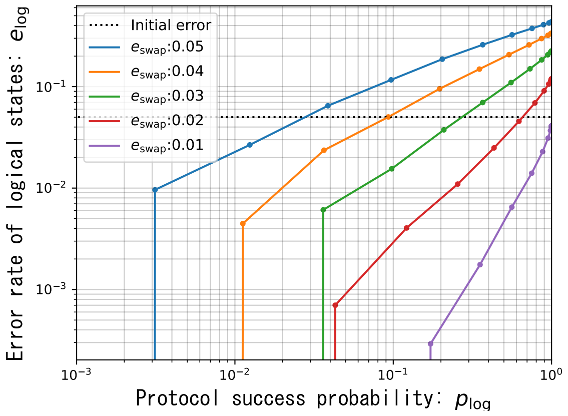


Figure 3: Trade-off relationship between success probability of the protocol $p_{\log}$ and the error rate of resultant logical entanglement $e_{\log}$.

The balance of the post-selection probability and infidelity can be tuned with threshold w_{thr}. Thus, we plotted their trade-off relation under several e_{swap} in Fig. 3. This figure illustrates the trade-off relations between quality and speed of logical entanglement generations between two distant nodes. These would be crucial parameters for estimating the performance of fault-tolerant quantum communication and distributed computation. If the achieved logical error rate is not sufficient for expected fault tolerance, we can fault-tolerantly perform subsequent entanglement distillation using several logical entanglements and lattice surgery operations. The detailed parameter exploration for optimizing practical quantum algorithms and protocols is left as future work.

4 Conclusion

In this study, we proposed a logical entanglement distribution protocol using surface codes on 2D-array qubits and evaluated its performance through numerical simulations. To encode randomly generated entangled states on 2D-array qubits using nearest-neighbor interactions, we need error-prone rearrangement of physical qubits. We proposed a method to mitigate errors in the rearrangement by optimizing the rearrangement sequence and performing post-selection based on the number of syndrome errors during encoding. This results in the trade-off relation between the generation rate of logical entangled states and their fidelity after post-selection. We conducted numerical simulations of the proposed protocol and identified the trade-off relationship between generation rate and logical error rates under various parameters.

References

[1] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[2] Daniele Cuomo, Marcello Caleffi, et al. Towards a distributed quantum computing ecosystem. *IET Quantum Communication*, 1(1):3–8, 2020.

[3] Austin G. Fowler, David S. Wang, et al. Surface code quantum communication. *Phys. Rev. Lett.*, 104:180503, May 2010.

[4] Charles H. Bennett, Gilles Brassard, et al. Purification of noisy entanglement and faithful teleportation via noisy channels. *Phys. Rev. Lett.*, 76:722–725, Jan 1996.

[5] Charles H. Bennett, David P. DiVincenzo, et al. Mixed-state entanglement and quantum error correction. *Phys. Rev. A*, 54:3824–3851, Nov 1996.

[6] Ryutaroh Matsumoto. Conversion of a general quantum stabilizer code to an entanglement distillation protocol. *Journal of Physics A: Mathematical and General*, 36(29):8113, jul 2003.

[7] Rajeev Acharya, Igor Aleiner, et al. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614(7949):676–681, Feb 2023.

[8] Dolev Bluvstein, Harry Levine, et al. A quantum processor based on coherent transport of entangled atom arrays. *Nature*, 604(7906):451–456, Apr 2022.

[9] W. Dür, H.-J. Briegel, et al. Quantum repeaters based on entanglement purification. *Phys. Rev. A*, 59:169–181, Jan 1999.

[10] Keisuke Fujii, Takashi Yamamoto, et al. A distributed architecture for scalable quantum computation with realistically noisy devices. *arXiv preprint arXiv:1202.6588*, 2012.

[11] Ivaylo S. Madjarov, Jacob P. Covey, et al. High-fidelity entanglement and detection of alkaline-earth rydberg atoms. *Nature Physics*, 16(8):857–861, Aug 2020.

[12] Youwei Zhao, Yangsen Ye, et al. Realization of an error-correcting surface code with superconducting qubits. *Phys. Rev. Lett.*, 129:030501, Jul 2022.

Extended abstract for AQIS23

Zixuan Liu[1 2 *] Ming Yang[3 †] Giulio Chiribella[1 2 4 5 ‡]

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

[2] *HKU-Oxford Joint Laboratory for Quantum Information and Computation*

[3] *Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom*

[4] *Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford, UK*

[5] *Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada*

Abstract. Certain quantum devices, such as half-wave plates and quarter-wave plates in quantum optics, are bidirectional, meaning that the roles of their input and output ports can be exchanged. Bidirectional devices can be used in a forward mode and a backward mode, corresponding to two opposite choices of the input-output direction. They can also be used in a coherent superposition of the forward and backward modes, giving rise to new operations with indefinite input-output direction. In this work we explore the potential of input-output indefiniteness for the transfer of classical and quantum information through noisy channels. We first formulate a model of communication from a sender to a receiver via a noisy channel used in indefinite input-output direction. Then, we show that indefiniteness of the input-output direction yields advantages over standard communication protocols in which the given noisy channel is used in a fixed input-output direction. These advantages range from a general reduction of noise in bidirectional processes, to heralded noiseless transmission of quantum states, and, in some special cases, to a complete noise removal. The noise reduction due to input-output indefiniteness can be experimentally demonstrated with current photonic technologies, providing a way to investigate the operational consequences of exotic scenarios characterised by coherent quantum superpositions of forward-time and backward-time processes.

Keywords: quantum communication, indefinite input-output direction, quantum resource theories

Recently, there has been an interest in exploring communication protocols where the configuration of the devices is itself quantum. For example, information could travel along different routes from the sender to the receiver, with each route traversing a different quantum device. The choice of route could be controlled by a quantum system, thus giving rise to interference of the alternative quantum evolutions [1, 2, 3, 4, 5, 6]. Similarly, the time of transmission could be controlled by a quantum system, giving rise to interference across multiple time bins [7]. A further example includes the use of different communication devices in different orders, with the choice of order controlled by a quantum system [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18].

Indefinite time direction is a new foundational concept [19]. This concept, originally motivated by foundational questions about the arrow of time, applies more generally to all quantum processes for which the roles of the input and the output can be exchanged. Examples of such processes, called *bidirectional* [19], are provided by half-wave plates, quarter-wave plates, and other optical crystals that rotate the polarization of single photons. Any such crystal can be traversed in two opposite directions, giving rise to two different quantum processes, conventionally called the forward and backward process. Forward and backward processes can be also probed in a coherent quantum superposition, by controlling a photon's trajectory through the crystal. As a result, the role of the inputs and outputs of the devices can become indefinite. A general mathematical framework for describing the use of a bidirectional device in an indefinite input-output direction was provided in Ref. [19].

A first operational consequence of input-output indefiniteness was shown in Ref. [19], which provided a quantum game in which the player can win with certainty only if the input-output direction is indefinite. Recently, this game was demonstrated experimentally with photons [20, 21]. Another consequence can be deduced from a related work on the quantum superposition of thermodynamic evo-

[*] zixuanliu@connect.hku.hk

[†] my365@cam.ac.uk

[‡] giulio.chiribella@cs.ox.ac.uk

lutions with opposite time arrows [22]. Besides these works, however, the operational consequences of indefinite input-output direction are still largely unexplored, especially in comparison to those of indefinite causal order, which have been extensively studied in the past decade [23, 24, 25, 26, 27, 28, 29, 16, 30, 31, 32, 33, 34, 35, 36, 37].

In this paper, we explore the consequences of input-output indefiniteness for the transfer of quantum information. To this purpose, we formulate a communication model that uses bidirectional devices in a coherent superposition of the forward and backward mode. We show that that the ability to coherently control the direction of a single particle travelling through a single quantum device can enhance the transmission of both classical and quantum information. For dephasing noise, corresponding e.g. to rotations of a single photon's polarization about a fixed direction, we show that a perfect, deterministic transmission of quantum bits becomes possible even if the original dephasing channel was entanglement-breaking and therefore could not transmit any quantum information. For more radical types of noise, such as depolarising noise, we show advantages both in the transmission of classical and quantum communication, observing nonzero capacities even in parameter regimes where no information can be transmitted by using the device in a fixed direction. For example, a completely depolarising qubit channel used in an indefinite input-output direction gives rise to a classical communication capacity of 0.3113 bits per channel use, and even permits a noiseless heralded transmission of qubit states with a success probability of 25%.

Our results highlight some similarities, as well as some differences, between input-output indefiniteness and the related notions of indefinite causal order and indefinite trajectories. Regarding the similarities, the communication advantages observed in all these scenarios are consequences of the ability to coherently control the configuration of noisy channels, as first observed by Gisin *et al* [3] and recently elaborated in Refs. [4, 38]. On the other hand, our result indicate that coherent control over the input-output direction can offer larger advantages. For example, putting two completely depolarising qubit channels on two alternative paths, and controlling the path of a quantum particle gives a classical capacity of at most 0.16 bits per channel use, while using the two channels in a coherently controlled order yields a capacity of at most 0.049 bits per channel use. Both of these values are strictly smaller than the classical capacity of 0.3113 bits per channel use achievable with a *single* depolarising channel used in a coherent superposition of the forward and backward mode.

By exploring the communication advantages of indefinite input-output direction, one can pin down a number of operational consequences of hypothetical scenarios in which the arrow of time is in a quantum superposition, thereby allowing agents to access quantum processes in a coherent superposition of the forward time direction, from the past to the future, and of the backward time direction, from the future to the past. While the physical realization of these scenarios is currently an open problem, the mathematical framework of quantum operations is with indefinite input-output direction provides a rigorous way to explore them a conceptual level. At the same time, the superposition of time directions can be simulated in table top experiments by coherently controlling the path of single photons, as envisaged in [19] and recently demonstrated in two experiments [20, 21]. Notably, all the communication enhancements identified in this work can already be demonstrated with a simple adaptation of these setups. The realization of these experiments is expected to contribute to the development of a toolbox for quantum control over the configuration of multiple quantum devices, which may prove technologically useful in a longer term.

In the shorter term, our work provides the starting point for a number of foundational explorations. First, an interesting direction is the investigation of scenarios where both the input-output direction and the causal order of multiple processes are subject to quantum indefiniteness. A mathematical framework for these more general operations was recently established in [19], but very little is currently known about their information-theoretic potential. The angle of quantum communication, explored in this paper, represents a promising approach to explore the capabilities of new operations with indefinite order and direction. Another interesting area of future research is the study of quantum thermodynamic tasks assisted by indefinite input-output direction. Recently, the communication advantages of indefinite causal order stimulated new research in quantum thermodynamics [29, 39, 40]. Similarly, the communication advantages provided in this paper suggest new thermodynamic protocols where the input-output direction of one or more processes is indefinite.

Communication protocols exploiting input-output indefiniteness are also interesting outside the context of quantum communication. They can be regarded as a new type of error correction, boosted by coherent quantum control over the input-output direction of the noisy processes. These protocols have some similarity with dynamical decoupling protocols in which an unknown noisy process is alternated with its inverse [41]. In our protocols, however, the advantages do not come from the alternation, but rather from the quantum interference between a process and its inverse, which facilitates the detection of errors and their subsequent correction. The working principle of these advantages appears to be the ability to distinguish between errors represented by symmetric matrices and errors represented by anti-symmetric matrices, which in particular allows one to correct Pauli Y errors separatly from the other types of single-qubit errors.

References

[1] Y. Aharonov, J. Anandan, S. Popescu, and L. Vaidman, "Superpositions of time evolutions of a quantum system and a quantum time-translation machine," *Physical review letters*, vol. 64, no. 25, p. 2965, 1990.

[2] D. K. Oi, "Interference of quantum channels," *Physical Review Letters*, vol. 91, no. 6, p. 067902, 2003.

[3] N. Gisin, N. Linden, S. Massar, and S. Popescu, "Error filtration and entanglement purification for quantum communication," *Physical Review A*, vol. 72, no. 1, p. 012338, 2005.

[4] A. A. Abbott, J. Wechs, D. Horsman, M. Mhalla, and C. Branciard, "Communication through coherent control of quantum channels," *Quantum*, vol. 4, p. 333, 2020.

[5] G. Chiribella and H. Kristjánsson, "Quantum shannon theory with superpositions of trajectories," *Proceedings of the Royal Society A*, vol. 475, no. 2225, p. 20180903, 2019.

[6] Q. Dong, S. Nakayama, A. Soeda, and M. Murao, "Controlled quantum operations and combs, and their applications to universal controllization of divisible unitary operations," *arXiv preprint arXiv:1911.01645*, 2019.

[7] H. Kristjánsson, W. Mao, G. Chiribella, *et al.*, "Witnessing latent time correlations with a single quantum particle," *Physical Review Research*, vol. 3, no. 4, p. 043147, 2021.

[8] D. Ebler, S. Salek, and G. Chiribella, "Enhanced communication with the assistance of indefinite causal order," *Physical Review Letters*, vol. 120, no. 12, p. 120502, 2018.

[9] S. Salek, D. Ebler, and G. Chiribella, "Quantum communication in a superposition of causal orders," *arXiv preprint arXiv:1809.06655*, 2018.

[10] G. Chiribella, M. Banik, S. S. Bhattacharya, T. Guha, M. Alimuddin, A. Roy, S. Saha, S. Agrawal, and G. Kar, "Indefinite causal order enables perfect quantum communication with zero capacity channels," *New Journal of Physics*, vol. 23, no. 3, p. 033039, 2021.

[11] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, "Communication enhancement through quantum coherent control of n channels in an indefinite causal-order scenario," *Entropy*, vol. 21, no. 10, p. 1012, 2019.

[12] L. M. Procopio, F. Delgado, M. Enríquez, N. Belabas, and J. A. Levenson, "Sending classical information via three noisy channels in superposition of causal orders," *Physical Review A*, vol. 101, no. 1, p. 012346, 2020.

[13] N. Loizeau and A. Grinbaum, "Channel capacity enhancement with indefinite causal order," *Physical Review A*, vol. 101, no. 1, p. 012340, 2020.

[14] M. Caleffi and A. S. Cacciapuoti, "Quantum switch for the quantum internet: Noiseless communications through noisy channels," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 3, pp. 575–588, 2020.

[15] S. Sazim, M. Sedlak, K. Singh, and A. K. Pati, "Classical communication with indefinite causal order for n completely depolarizing channels," *Physical Review A*, vol. 103, no. 6, p. 062610, 2021.

[16] G. Chiribella, M. Wilson, and H. Chau, "Quantum and classical data transmission through completely depolarizing channels in a superposition of cyclic orders," *Physical review letters*, vol. 127, no. 19, p. 190502, 2021.

[17] S. S. Bhattacharya, A. G. Maity, T. Guha, G. Chiribella, and M. Banik, "Random-receiver quantum communication," *PRX Quantum*, vol. 2, no. 2, p. 020350, 2021.

[18] T. Guha, S. Roy, and G. Chiribella, "Quantum networks boosted by entanglement with a control system," *arXiv preprint arXiv:2206.05247*, 2022.

[19] G. Chiribella and Z. Liu, "Quantum operations with indefinite time direction," *Communications Physics*, vol. 5, no. 1, pp. 1–8, 2022.

[20] Y. Guo, Z. Liu, H. Tang, X.-M. Hu, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, "Experimental demonstration of input-output indefiniteness in a single quantum device," *arXiv preprint arXiv:2210.17046*, 2022.

[21] T. Strömberg, P. Schiansky, M. T. Quintino, M. Antesberger, L. Rozema, I. Agresti, Č. Brukner, and P. Walther, "Experimental superposition of time directions," *arXiv preprint arXiv:2211.01283*, 2022.

[22] G. Rubino, G. Manzano, and Č. Brukner, "Quantum superposition of thermodynamic evolutions with opposing time's arrows," *Communications Physics*, vol. 4, no. 1, pp. 1–10, 2021.

[23] G. Chiribella, "Perfect discrimination of no-signalling channels via quantum superposition of causal structures," *Physical Review A*, vol. 86, no. 4, p. 040301, 2012.

[24] G. Chiribella, G. M. D'Ariano, P. Perinotti, and B. Valiron, "Quantum computations without definite causal structure," *Physical Review A*, vol. 88, no. 2, p. 022318, 2013.

[25] M. Araújo, F. Costa, and Č. Brukner, "Computational advantage from quantum-controlled ordering of gates," *Physical review letters*, vol. 113, no. 25, p. 250402, 2014.

[26] P. A. Guérin, A. Feix, M. Araújo, and Č. Brukner, "Exponential communication complexity advantage from quantum superposition of the direction of communication," *Physical review letters*, vol. 117, no. 10, p. 100502, 2016.

[27] D. Ebler, S. Salek, and G. Chiribella, "Enhanced communication with the assistance of indefinite causal order," *Physical review letters*, vol. 120, no. 12, p. 120502, 2018.

[28] X. Zhao, Y. Yang, and G. Chiribella, "Quantum metrology with indefinite causal order," *Physical Review Letters*, vol. 124, no. 19, p. 190503, 2020.

[29] D. Felce and V. Vedral, "Quantum refrigeration with indefinite causal order," *Physical review letters*, vol. 125, no. 7, p. 070603, 2020.

[30] L. M. Procopio, A. Moqanaki, M. Araújo, F. Costa, I. Alonso Calafell, E. G. Dowd, D. R. Hamel, L. A. Rozema, Č. Brukner, and P. Walther, "Experimental superposition of orders of quantum gates," *Nature communications*, vol. 6, no. 1, pp. 1–6, 2015.

[31] G. Rubino, L. A. Rozema, A. Feix, M. Araújo, J. M. Zeuner, L. M. Procopio, Č. Brukner, and P. Walther, "Experimental verification of an indefinite causal order," *Science advances*, vol. 3, no. 3, p. e1602589, 2017.

[32] K. Goswami, C. Giarmatzi, M. Kewming, F. Costa, C. Branciard, J. Romero, and A. G. White, "Indefinite causal order in a quantum switch," *Physical review letters*, vol. 121, no. 9, p. 090503, 2018.

[33] K. Wei, N. Tischler, S.-R. Zhao, Y.-H. Li, J. M. Arrazola, Y. Liu, W. Zhang, H. Li, L. You, Z. Wang, *et al.*, "Experimental quantum switching for exponentially superior quantum communication complexity," *Physical review letters*, vol. 122, no. 12, p. 120504, 2019.

[34] Y. Guo, X.-M. Hu, Z.-B. Hou, H. Cao, J.-M. Cui, B.-H. Liu, Y.-F. Huang, C.-F. Li, G.-C. Guo, and G. Chiribella, "Experimental transmission of quantum information using a superposition of causal orders," *Physical review letters*, vol. 124, no. 3, p. 030502, 2020.

[35] G. Rubino, L. A. Rozema, D. Ebler, H. Kristjánsson, S. Salek, P. A. Guérin, A. A. Abbott, C. Branciard, Č. Brukner, G. Chiribella, *et al.*, "Experimental quantum communication enhancement by superposing trajectories," *Physical Review Research*, vol. 3, no. 1, p. 013093, 2021.

[36] H. Cao, N.-N. Wang, Z. Jia, C. Zhang, Y. Guo, B.-H. Liu, Y.-F. Huang, C.-F. Li, and G.-C. Guo, "Quantum simulation of indefinite causal

order induced quantum refrigeration," *Physical Review Research*, vol. 4, no. 3, p. L032029, 2022.

[37] X. Nie, X. Zhu, K. Huang, K. Tang, X. Long, Z. Lin, Y. Tian, C. Qiu, C. Xi, X. Yang, *et al.*, "Experimental realization of a quantum refrigerator driven by indefinite causal orders," *Physical Review Letters*, vol. 129, no. 10, p. 100603, 2022.

[38] H. Kristjánsson, G. Chiribella, S. Salek, D. Ebler, and M. Wilson, "Resource theories of communication," *New Journal of Physics*, vol. 22, no. 7, p. 073014, 2020.

[39] T. Guha, M. Alimuddin, and P. Parashar, "Thermodynamic advancement in the causally inseparable occurrence of thermal maps," *Phys. Rev. A*, vol. 102, p. 032215, Sep 2020.

[40] K. Simonov, G. Francica, G. Guarnieri, and M. Paternostro, "Work extraction from coherently activated maps via quantum switch," *Physical Review A*, vol. 105, no. 3, p. 032217, 2022.

[41] L. Viola and S. Lloyd, "Dynamical suppression of decoherence in two-state quantum systems," *Physical Review A*, vol. 58, no. 4, p. 2733, 1998.

T-depth Reduction Method by Decomposing MCT Gates Considering the Cancellation

Koki Hirono[1] * Shigeru Yamashita[1] †

[1] *Graduate School of Information Science and Engineering*
Ritsumeikan University

Abstract. This paper proposes a method to reduce T-depth in a quantum circuit by decomposing MCT (Multi Controlled Toffoli) gates considering the cancellation of operations when multiple MCT gates are present. The method aims to reduce T-depth by decomposing two MCT gates with common control bits in a way that their operations cancel each other. Our experimental results show that the proposed method can reduce T-depth up to approximately 43.17% compared to the method of Niemann et al..

Keywords: MCT gate, T-depth, Quantum circuit

1 Introduction

General quantum algorithms have a part that calculates logical functions [1]. Quantum circuits that calculate this logical function need to be designed for each given logical function. In quantum circuit design, any logical function can be realized by combining multiple MCT (Multi Controlled Toffoli) gates. To realize the operation of MCT gates, it is necessary to decompose them into a group of gate called Clifford+T that can be executed directly [2]. Among the Clifford+T gate group, there is a gate called T gate. The number of T gates that cannot be executed simultaneously in a quantum circuit is called T-depth [3]. Since the operation time of T gates is longer than that of other gates in the Clifford+T gate group [4], T-depth is used as an index of the cost of quantum circuit design. For this reason, research has been conducted on MCT gate decomposition methods that reduce T-depth [5] [6].

Niemann et al.'s method [6] has been proposed as a method for reducing T-depth when decomposing MCT gates. This method only considers the decomposition of single MCT gate, so it may not be optimal when decomposing multiple MCT gates. Niemann et al.'s method reduces T-depth by decomposing single MCT gate into multiple MCT gates using ancilla bits. However, Niemann et al.'s method does not consider the cancellation of operations with other MCT gates in this decomposition. Therefore, this paper proposes a method for reducing T-depth by performing decomposition considering the cancellation of operations with other MCT gates.

The proposed method and Niemann et al.'s method were implemented and evaluated for T-depth. Then, the reduction rate of T-depth was calculated by comparing the proposed method and Niemann et al's method. As a result, it was confirmed that the proposed method can reduce T-depth by up to about 43.17% compared to Niemann et al.'s method.

2 Preliminaries and Previous Work

A quantum circuit is a graphical representation of quantum computation using quantum gates and qubits. Qubits and quantum gates correspond to bits and logical gates in classical computing. Quantum gates, like conventional logical gates, have inputs and outputs. By combining multiple quantum gates, desired operations can be realized. Quantum circuits must always be reversible, unlike classical circuits, and they have a different property in that the input can be uniquely determined from the output.

MCT (Multi Controlled Toffoli) gate [7] is a quantum gate composed of multiple control bits and one target bit. MCT gate applies NOT gate to the target bit when all control bits have a value of 1. Since MCT gate cannot be executed directly, it must be decomposed into a group of gates that can be directly executed, such as Clifford+T gates.

There is a decomposition method for the MCT gate proposed by Abdesaied et al. [5]. In Abdesaied et al.'s method, MCT gate with $c \geq 3$ control bits can be decomposed by using $(c-2)$ dirty ancilla bits, achieving MCT gate decomposition with T-depth of $4(c-1)$.

There is a T-depth reduction method during MCT gate decomposition proposed by Niemann et al. [6]. In Niemann et al.'s method, MCT gate with c control bits is decomposed using $k \geq 1$ clean ancilla bits. In Niemann's method, MCT gate is divided into several stages, and Abdesaied et al.'s method is applied to each stage.

In Niemann et al.'s method, when $1 \leq k \leq \frac{c}{2}$, MCT gate is decomposed into three stages. Figure 1 shows an example of MCT gate that is applied in Niemann et al.'s method when the number of clean ancilla bits is 2.

Let k clean ancillas be denoted as $a_1, ..., a_k$, and let the target bit of MCT that is decomposed gate be denoted as t. The control lines of c-controlled MCT gate are divided into groups $C_1, ..., C_{k+1}$ of $k+1$ control lines. The number of control lines of these divided groups of control lines is represented as $|C_1|, ..., |C_{k+1}|$. The arrangement of the three-stage MCT gates is described as follows:

1 In the first stage, we place in parallel k MCT gates each of which has one of $C1, \ldots, C_k$ as control bits,

*is0521kv@ed.ritsumei.ac.jp
†ger@cs.ritsumei.ac.jp

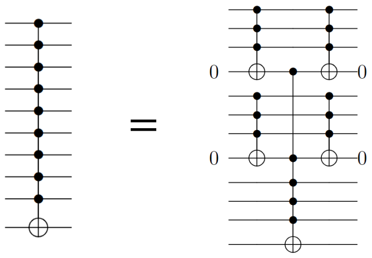


Figure 1: Niemann et al.'s method

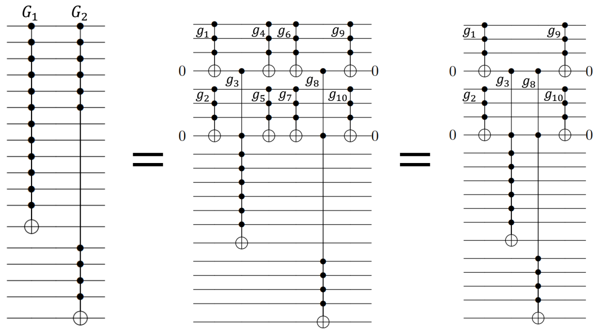


Figure 2: Example of decomposing MCT gates considering the cancellation of operations

and one of $a_1, \dots, a_k$ as target bit.

2 In the second stage, we place an MCT gate with C_{k+1} and $a_1, \dots, a_k$ as control bits and t as target bit.

3 In the third stage, we place a copy of the first stage to restore the values on the ancilla bits.

Let d be the number of additional dirty ancilla bits. To apply Abdesaied et al.'s method to each MCT gate in this decomposition, it is necessary to follow Equation 1.

$$\frac{c+2-k+d}{2} \geq |C_{k+1}| \geq \frac{c-2k-d-1}{2} \qquad (1)$$

In the method of Niemann et al., the control bits are placed so that the overall T-depth is small.

For $k > \frac{c}{2}$, the decomposition is applied recursively to the second-stage MCT gate.

3 The Proposed Method

Niemann et al.'s T-depth reduction method for MCT gates considers only the decomposition of a single MCT gate. In this section, we propose a method for reducing T-depth by considering the cancellation of MCT gates when decomposing multiple MCT gates using k clean ancilla bits. When two MCT gates have the same bit as their control bits, MCT gates can be applied decomposition considering the cancelation.

An example of decomposing MCT gates considering the cancellations using two clean ancilla bits is shown in Figure 2. Two adjacent MCT gates, G_1 and G_2, are decomposed as shown in the central circuit of Figure 2. In this circuit, the each of two MCT gates, (g_4, g_6) and (g_5, g_7) have the same control bits and target bits on the same bit, resulting in the cancellation of operations. Then, the circuit transforms into the right circuit in Figure 2. By decomposing two MCT gates into three stages, arranging them such that they have the controls and targets on the same bits, and achieving the cancellation of MCT gate operations, the proposed method can reduce T-depth in the circuit.

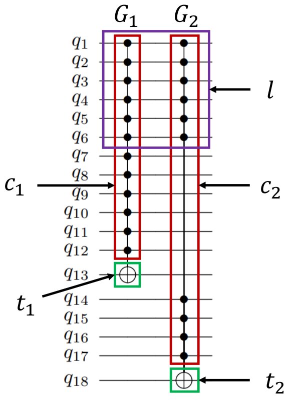


Figure 3: Example of two MCT gates have common control bits

Figure 3 shows two MCT gates with common control bits. The proposed method finds the number of common control bits, denoted as m, that can be used when applying the decomposition considering cancellation for the two MCT gates. Let n be the number of quantum bits in the circuit, k be the number of clean ancilla bits and d be the number of additional dirty ancilla bits. Let c_1 and c_2 be the number of control bits for the two MCT gates G_1 and G_2, and let l be the number of common control bits between the two MCT gates. By using a value of m that satisfies Equation 2 and Equation 3, we can decompose MCT gates considering cancellation using the common control bits.

$$m \leq \min\left\{l, \frac{n+2k+d}{2}\right\} \qquad (2)$$

$$\max\left\{2k, \frac{2c_1+k-d-n-1}{2}, \frac{2c_2+k-d-n-1}{2}\right\} \leq m \qquad (3)$$

Furthermore, T-depth of two MCT gates is obtained when applying the decomposition considering the cancellations for the two MCT gates. T-depth is categorized based on the value of k.

When $k = 1$, T-depth is as follows:

- $4(c_1 - 2) + 4(c_2 - 2) + 2$ when $m = 2$
- $4(c_1 - 1) + 4(c_2 - 1)$ when $m > 2$

When $k \geq 2$, T-depth is as follows:

- $4(c_1 - k - 1) + 4(c_2 - k - 1) + 4$ when $m = 2k$
- $4(c_1 - m + k - 1) + 4(c_2 - m + k - 1) + 8(\lceil \frac{m}{k} \rceil)$ when $m > 2k$

In the proposed method, the combination that achieves the maximum reduction in T-depth compared to applying Niemann's method is determined by applying the decomposition considering cancellations, starting from the first MCT gate in the set of sortable MCT gates. Based

Table 1: Experimental results on benchmark circuits for $k = 2, 3$

benchmarks		$k = 2$			$k = 3$		
circuit name	number of gates	Niemann's	Proposed	$\Delta T(\%)$	Niemann's	Proposed	$\Delta T(\%)$
tial_265	1041	19355	13974	27.80	14889	12779	14.17
misex3c_244	1721	42077	26838	36.22	31952	24878	22.14
cordic_218	2533	78556	44645	43.17	57527	36539	36.48
ex1010_230	2611	58220	36385	37.50	42211	35028	17.02
alu_319	15764	676098	582528	13.84	498225	445792	10.52

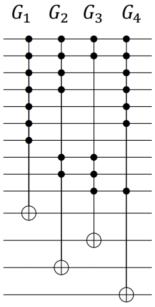


Figure 4: Example of sortable MCT gates

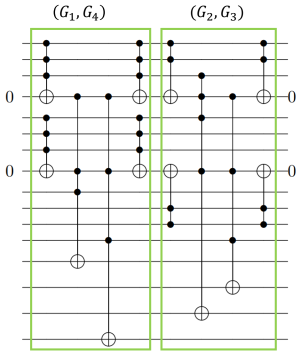


Figure 5: Example of sorting MCT gates

on the determined combination, MCT gates are sorted, and the decomposition considering the cancellations is applied to reduce the T-depth of the quantum circuit.

Example 1 *Figure 4 shows a set of four sortable MCT gates. An example of the proposed method is demonstrated when applying the decomposition considering cancellations with two clean ancilla bits for this set of MCT gates. The combinations are determined starting from G_1 in Figure 4. First, MCT gate G_4 is chosen as it provides the largest reduction in T-depth compared to Niemann et al.'s method when applying the decomposition considering cancellations with G1. Next, the combination for G_2 is determined. MCT gate G_3 provides the largest reduction in T-depth when applying the decomposition considering cancellations with G_2. Then, MCT gates (G_1, G_4) and (G_2, G_3) are arranged adjacent to apply decomposition considering the cancellations, as shown in Figure 5. In this case, the reduction in T-depth from Niemann et al.'s method is 20.*

4 Experimental Result and Conclusion

To evaluate the proposed method, we implemented both Niemann's method and the proposed method. Table 1 shows the experimental results using RevLib benchmark circuits [8].

- benchmarks
  - circuit name: Name of RevLib benchmark circuit
  - number of gates: Number of gates in RevLib benchmark circuit
- k : number of clean ancilla bits given to the circuit
  - Niemann's: T-depth of quantum circuits applying the method of Niemann et al.
  - Proposed: T-depth of quantum circuits applying the proposed method
  - $\Delta T(\%)$: T-depth reduction ratio of the proposed method from the method of Niemann et al. (%)

The $\Delta T(\%)$ values in Table 1 are all positive, indicating that the proposed method achieves a smaller T-depth compared to Niemann's method for all benchmark circuits.

This paper proposes the T-depth reduction method by decomposing MCT gates considering the cancellation. Future work includes a T-depth reduction method using the remaining common control bits after applying the decomposition with considering the cancelation.

References

[1] Shigeru Yamashita, Shin-ichi Minato, and D Michael Miller. Ddmf: An efficient decision diagram structure for design verification of quantum circuits under a practical restriction. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 91(12):3793–3802, 2008.

[2] Xinlan Zhou, Debbie W Leung, and Isaac L Chuang. Methodology for quantum logic gate construction. *Physical Review A*, 62(5):052316, 2000.

[3] Matthew Amy, Dmitri Maslov, Michele Mosca, and Martin Roetteler. A meet-in-the-middle algorithm for fast synthesis of depth-optimal quantum circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 32(6):818–830, 2013.

[4] Austin G Fowler. Time-optimal quantum computation. *arXiv preprint arXiv:1210.4626*, 2012.

[5] Nabila Abdessaied, Matthew Amy, Mathias Soeken, and Rolf Drechsler. Technology mapping of reversible circuits to clifford+ t quantum circuits. In *2016 IEEE 46th international symposium on multiple-valued logic (ISMVL)*, pages 150–155. IEEE, 2016.

[6] Philipp Niemann, Anshu Gupta, and Rolf Drechsler. T-depth optimization for fault-tolerant quantum circuits. In *2019 IEEE 49th International Symposium on Multiple-Valued Logic (ISMVL)*, pages 108–113. IEEE, 2019.

[7] Adriano Barenco, Charles H Bennett, Richard Cleve, David P DiVincenzo, Norman Margolus, Peter Shor, Tycho Sleator, John A Smolin, and Harald Weinfurter. Elementary gates for quantum computation. *Physical review A*, 52(5):3457, 1995.

[8] Robert Wille, Daniel Große, Lisa Teuber, Gerhard W Dueck, and Rolf Drechsler. Revlib: An online resource for reversible functions and reversible circuits. In *38th International Symposium on Multiple Valued Logic (ismvl 2008)*, pages 220–225. IEEE, 2008.

Quantum Circuit Synthesis Method of VQE for Traveling Salesman Problem Considering W States

Kohei Ogino[1] * Atsushi Matsuo[2] † Shigeru Yamashita[1] ‡

[1] *College of Information Science and Engineering, Ritsumeikan University*
Nojihigashi1-1-1, Kusatsushi-Siga, 525-8577
[2] *IBM Quantum, IBM Research - Tokyo*
Hakozaki-cho19-21,Nihonbashi, Chuo-ku-Tokyo, 103-8510

Abstract. VQE (Variational Quantum Eigensolver), a quantum algorithm, can be used to solve combinatorial optimization problems. A quantum circuit used in VQE is called a PQC (Parameterized Quantum Circuit). There is a PQC synthesis method for the Traveling Salesman Problem. In the existing method, only one type of PQC is used so the PQC may not be optimal for some problems. In this paper, we propose a method for synthesizing PQC for the Traveling Salesman Problem considering quantum states generated by PQC. The proposed method reduces the total cost in the Traveling Salesman Problem by 10.43% on average.

Keywords: Variational Quantum Eigensolver, Parameterized Quantum Circuit, W states

1 Introduction

VQE (Variational Quantum Eigensolver) [1] is an algorithm for finding the minimum eigenvalue of a given Hamiltonian based on the variational approach. By setting the given Hamiltonian appropriately, combinatorial optimization problems such as the max-cut problem and TSP (Traveling Salesman Problem) can be solved [2].

VQE is an algorithm that uses a quantum computer with quantum circuits and a classical computer with a classical optimizer.The quantum circuit is called a PQC (Parameterized Quantum Circuit) , and research on its synthesis has been active [3–5]. There is a PQC synthesis method for the TSP [6]. The existing method requires the generation of a quantum state called the W state, and the quantum circuit for generationg it greatly affects the performance of the algorithm. However, the circuit for generating the W state has a problem that the W state varies greatly depending on the arrangement of the quantum gates. Therefore, if the optimal W state cannot be generated, there is a possibility that an optimal solution cannot be obtained. Even if an optimal solution can be obtained, it may require a large amount of computation time.

In this paper, we propose a method for synthesizing PQC for the TSP that takes into account the quantum states generated by PQC. The proposed method synthesizes PQC by arranging quantum gates in a different way from the existing method [6] to reduce the effect of the arrangement of qubits on the W state generated. In addition, considering the W state to be generated, initial values of appropriate circuit parameters are set to the PQC. The proposed method increases the number of problems for which an optimal solution can be obtained and reduces the computation time.

2 Background

2.1 Formulation of TSP

This section describes the formulation of TSP. With a complete graph $G = (V, E)$ with N vertices, the cost of any two vertices $(u, v) \in E$ in G is defined as $W_{u,\,v}$. Formulating the TSP as a linear programming problem, it can be defined as Eq. (1), (2), (3) . The N^2 variables $x_{v,\,p}$ $(1 \leq v \leq N,\ 1 \leq p \leq N)$ take the value 1 or 0, meaning that the city v is visited for the pth time when $x_{v,\,p} = 1$ and the city v is not visited for the pth time when $x_{v,\,p} = 0$. Eq. (1) implies the total cost and takes $x_{u,\,p}$ and $x_{v,\,p+1}$ as variables. The value of $x_{u,\,p}$ represents the route with the minimum total cost. Eq. (2) means that the same city is visited only once, and Eq. (3) means that only one city is visited at a time.

$$\mathit{Minimize} \quad \sum_{(u,\,v)\in E} W_{(u,\,v)} \sum_{p=1}^{N} x_{u,\,p}\ x_{v,\,p+1} \tag{1}$$

$$\mathit{Constraints}\ 1 \quad \sum_{p=1}^{N} x_{v,\,p} = 1,\ v = 1, 2, \ldots N \tag{2}$$

$$\mathit{Constraints}\ 2 \quad \sum_{v=1}^{N} x_{v,\,p} = 1,\ p = 1, 2, \ldots N \tag{3}$$

2.2 The method of solving TSP using VQE

The method of solving TSP using VQE is shown below.

1. Set the Hamiltonian H so that $\langle\psi_\theta|\, H\, |\psi_\theta\rangle$ corresponds to the function in Eq. (1).

2. Assign N^2 variables $x_{v,\,p}$ to PQC input $|\psi\rangle = |\psi_1\rangle, \ldots, |\psi_{N^2}\rangle$.The output of PQC is $|\psi(\theta)\rangle = |\psi_{1(\theta)}\rangle, \ldots, |\psi_{N^2(\theta)}\rangle$. The details of PQC used here are explained in Section 2.3. Then, calculating the expected value $\langle\psi_\theta|\, H\, |\psi_\theta\rangle$ of H

3. Update θ so that $\langle\psi_\theta|\, H\, |\psi_\theta\rangle$ is smaller by using a classical computer. It means that the PQC parameter θ is updated so that the total cost is minimized.

*dos@ngc.is.ritsumei.ac.jp
†matsuoa@jp.ibm.com
‡ger@cs.ritsumei.ac.jp

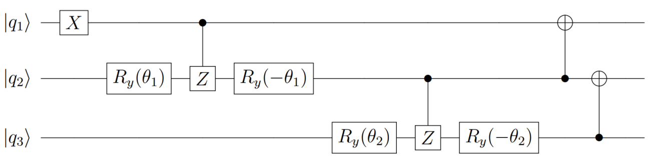


Figure 1: A quantum circuit for 3-qubit W state

4. Repeat the above two operations until the expected value $\langle\psi_\theta| H |\psi_\theta\rangle$ converge.
5. When the values of $\langle\psi_\theta| H |\psi_\theta\rangle$ converge afther repeating these operations, $|\psi_\theta\rangle$ is the path with the minimum total cost.

2.3 The PQC for TSP

By using the existing method [6], only bit strings that satisfy the Eq. (2), are obtained at the output of PQC. The method for synthesizing circuits that satisfy the constraints of the Eq. (2) for TSP is shown below. The circuit that satisfies the constraints $\sum_{p=1}^{N} x_{1,\,p} = 1$ at $v = 1$ in Eq. (2) is synthesized by the following procedures.

1. Initialize the N-bit input to $|q_1\, q_2 \ldots q_N\rangle = |00\ldots0\rangle$.
2. Apply a X gate to q_1
3. Repeat the following three operations from $i = 1$ to $i = N - 1$.
   - Apply a $R_y(\theta)$ gate to q_{i+1}
   - Apply a CZ gate to q_i and q_{i+1}
   - Apply a $R_y(-\theta)$ gate to q_{i+1}
4. Repeat the following operation from $i = 1$ to $i = N - 1$.
   - Apply a CX gate to q_{i+1} and q_i

An example for $N = 3$ is shown in Figure. 1. It generates 3-qubit W state. W state is quantum state defined as Eq. (4). Observing the N-qubit W state, only a bit sequence in which only 1 bit of any N bits is $|1\rangle$ and all other $N - 1$ bits are $|0\rangle$ is obtained. By applying this circuit to $q_4\ q_5\ q_6$ and $q_7\ q_8\ q_9$, the PQC which satisfies the constraints of the Eq. (2) can be synthesized.

$$|\psi\rangle = \sum_{i=1}^{N} \alpha_i(\phi)\, |\psi_i\rangle \quad \sum_{i=1}^{N} |\alpha_i(\phi)|^2 = 1 \tag{4}$$

$|\psi_i\rangle \in \{\ |10\ldots00\rangle\,,\ |01\ldots00\rangle\,,\ |00\ldots10\rangle\,,\ |00\ldots01\rangle\ \}$

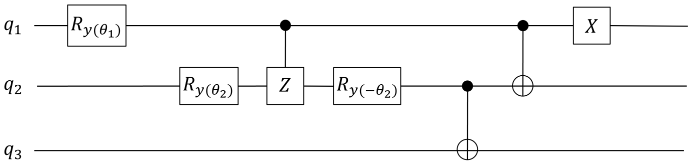


Figure 2: A quantum circuit for 3-qubit W state synthesized by proposed method

3 The proposed PQC considering W States

There are two problems with the existing method [6]. The first one is that PQC is synthesized using only one type of circuit that generates W state, despite the fact that there are multiple synthesis methods for circuits that generate a W state [7]. Second, the initial values of PQC parameters are not taken into account. The initial values of the parameters are determined by the classical optimizer during the execution of VQE. It is possible to set the initial values of the parameters in advance, but the existing method [6], does not do this. Since the probability of observing the W state generated by the parameters varies widely, the initial values of the parameters determined by the classical optimizer may not be optimal.

To solve these problems, we propose a new circuit to generate W state and use it to generate a PQC for TSP. The initial values of the PQC parameters are set to values considering the W state to be generated, in addition to the values determined by the classical optimizer. The method for synthesizing circuits we propose is shown below.

1. Initialize the N-bit input to $|q_1\, q_2 \ldots q_N\rangle = |00\ldots0\rangle$.
2. Apply a $R_y(\theta)$ gate to q_1
3. Repeat the following three operations from $i = 1$ to $i = N - 2$.
   - Apply a $R_y(\theta)$ gate to q_{i+1}
   - Apply a CZ gate to q_i and q_{i+1}
   - Apply a $R_y(-\theta)$ gate to q_{i+1}
4. Repeat the following operation from $i = 1$ to $i = N - 1$.
   - Apply a CX gate to q_{N-i} and q_{N-i+1}
5. Apply a X gate to q_1

An example for $N = 3$ is shown in Figure. 2.

4 Experimental Results

4.1 Evaluation Method

To evaluate the proposed method, we implemented the PQC for the existing method [6] and the proposed PQC in Python. In the proposed method, we set three kinds of

Table 1: Results of solving TSP

	Existing method [6]		Proposed method	
Number of cities	Total cost	Time (s)	Total cost	Time (s)
4	211.5	4.41	199.5	5.33
5	276.3	267.00	248.0	373.92
6	321.9	466.74	271.9	631.87
7	374.0	2623.82	332.1	2843.11

Table 2: The number of times each initial value was selected

Types of initial values	Selected times
1	8
2	9
3	23

initial values: (1) the initial value set by the classical optimizer, (2) the initial value such that the probability of observing each W state is equal, and (3) the initial value such that only $|10\dots0\rangle$, where only the most significant bit among N-bit W state is 1, is observed. In the experiment, we randomly prepared 10 TSP questions from 4 to 7 cities each, and used VQE with each PQC to obtain the answers. The part of VQE that uses a quantum computer was computed using a noiseless simulator. For the classical optimizer, we used COBYLA (Constrained Optimization BY Linear Approximation) [8].

4.2 Results and Consideration

The average of the experimental results for each city is shown in Table. 1. Experimental results show that the proposed method reduces the total cost by 10.43% on average compared to the existing method [6]. As for the initial values of the parameters, (3) is suitable, indicating that the initial values of the parameters in the proposed method have a significant impact on the results. In terms of computation time, the proposed method increased 27.98% on average compared to the existing method [6]. The R_y gates, where the parameters of PQC are set, are all placed on both sides of the CZ gates in the existing method, while the proposed method places them not only on both sides of the CZ gates but also on the q_1 bit. This difference in the placement of the R_y gates affects the time required for optimization with the classical optimizer, and is assumed to have increased the computation time for the proposed method. This is presumably because the classical optimizer sets the initial values of the parameters for the existing method [6], whereas the proposed method takes more time when optimizing with the classical optimizer because the parameters are set considering the generated W state.

5 Conclusions

In this paper, we proposed the method for synthesizing PQC for TSP that takes into account the W states to be generated and the new initial values of its parameters. We synthesize PQC for TSP using a different circuit that generates W state from the existing method [6]. For the initial values of the PQC parameters, we set two types of values that take into account the generated W states, in addition to the values set by the classical optimizer, which were used in the existing method [6]. Experimental results show that the method proposed in this paper reduces the total cost by 10.43% on average compared to the existing method [6]. The optimal initial values of parameters for the proposed method were also found. Computation time increased compared to the existing method due to the R_y gate arrangement in the proposed method.

The method proposed in this paper requires N^2 qubits to solve TSP for N cities. Currently, the maximum number of qubit available on a quantum computer is 127 bits [9], so it is not possible to solve the traveling salesman problem for more than 11 cities. Future work is to improve the proposed method so that multiple quantum computers can run VQE in parallel, and to support a larger number of cities. In the experiments to evaluate the proposed method, the quantum computer part used a noiseless simulator. In the actual quantum computer, errors occur due to the influence of noise [10]. It is a future task to investigate the relationship between the error due to noise and the proposed PQC by running the quantum computer part of the proposed method on a real quantum computer.

References

[1] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, Jeremy LO' brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, Vol. 5, No. 1, pp. 1–7, 2014.

[2] Nikolaj Moll, Panagiotis Barkoutsos, Lev S Bishop, Jerry M Chow, Andrew Cross, Daniel J Egger, Stefan Filipp, Andreas Fuhrer, Jay M Gambetta, Marc Ganzhorn, et al. Quantum optimization using variational algorithms on near-term quantum devices. *Quantum Science and Technology*, Vol. 3, No. 3, p. 030503, 2018.

[3] Sukin Sim, Peter D Johnson, and Alán Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid

quantum-classical algorithms. *Advanced Quantum Technologies*, Vol. 2, No. 12, p. 1900070, 2019.

[4] Atchade Parfait Adelomou, Elisabet Golobardes Ribe, and Xavier Vilasis Cardona. Using the parameterized quantum circuit combined with variational-quantum-eigensolver (vqe) to create an intelligent social workers' schedule problem solver. *arXiv preprint arXiv:2010.05863*, 2020.

[5] Hanrui Wang, Yongshan Ding, Jiaqi Gu, Yujun Lin, David Z Pan, Frederic T Chong, and Song Han. Quantumnas: Noise-adaptive search for robust quantum circuits. In *2022 IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, pp. 692–708. IEEE, 2022.

[6] Atsushi Matsuo, Yudai Suzuki, and Shigeru Yamashita. Problem-specific parameterized quantum circuits of the vqe algorithm for optimization problems. *arXiv preprint arXiv:2006.05643*, 2020.

[7] Wolfgang Dür, Guifre Vidal, and J Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, Vol. 62, No. 6, p. 062314, 2000.

[8] Michael JD Powell. A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical analysis*, pp. 144–157. Springer, 1978.

[9] IBM Quantum Compute Resources. https://quantum-computing.ibm.com/services/resources. (Accessed 2023/05/19).

[10] Andrew Steane. Quantum computing. *Reports on Progress in Physics*, Vol. 61, No. 2, p. 117, 1998.

Detecting Information Backflow via Temporal Quantum Correlations

Shrikant Utagi[*]

Department of Physics, Indian Institute of Technology Madras, Chennai 600036, India.

There are two prominent criteria of non-Markovianity, namely, one due to divisibility and the one due to distinguishability or information backflow. We show that the two recently introduced measures of non-Markovianity, those based on temporal quantum correlations, sit somewhere between these two. We show, by example, that since both causality measure and temporal steerable weight detect information backflow, they can certify a process to be non-Markovian where trace distance may fail.

I. INTRODUCTION

No quantum system is perfectly isolated from its surroundings and it may inevitably undergo the process of decoherence [1], and such a system is called *open* quantum system, who's dynamics is governed by quantum master equations or equivalently represented by a quantum channel – completely positive trace preserving (CPTP) map that takes density operators to density operators [2].

Non-Markovian open system dynamics has been of interest for many years from both physics and quantum information theory point of view [1, 3–8]. Two broad notions of non-Markovianity are based on divisibility, which relies on the concept of entanglement [9] and distinguishability which relies on the concept of distance, for instance trace distance, between two states under a channel [10]. Despite the fact that a plethora of non-Markovianity witnesses and measures exist, they are often dependent on the type of dynamics. It is known that these are non-equivalent identifiers of a given physical process as non-Markovian [11–14]. However, for certain class of dynamical maps, namely the image non-increasing maps, they are indeed equivalent [15]. Interestingly, it has been established that the trace distance as an identifier of information backflow essentially fails in two cases: (i) eternally non-Markovian channel [16] (ii) when non-unital part of the channel is solely responsible for the information backflow [17–19]. That is, P-indivisibility as an equivalent notion of information backflow has to be dealt with some care, since it might originate from non-unital part alone. In certain situations, non-unitality can be harnessed for quantum information tasks [20–22]. Therefore, studying the relationship between non-Markovianity and non-unitality may be a crucial factor for harnessing both these aspects of a quantum channel.

Quantum correlations, such as entanglement and steering, are those that cannot be realized by any local realistic theory, hence they are proven to be resources for quantum information and quantum computation tasks [2, 23, 24]. Spatial correlations are generally known to be the correlations with a common cause [25, 26]. Recent developments reveal that quantum mechanics allows for quantum correlations with direct cause dubbed temporal quantum correlations [27]. It has been shown [28] that temporal quantum correlations form a hierarchy, namely temporal non-separability, temporal steering, and temporal nonlocality, and that temporally non-separable correlations in pseudo-density operator are a form of correlations with stronger quantum direct cause while temporal steerable correlations can be interpreted to that of weaker form of quantum direct cause. To mention in the passing, note that temporal steering finds its application in quantum cryptography [29]. Recently, temporal steerable weight and causality measure were used to define measures of non-Markovianity [18, 30]. One may wonder if these quantities are equivalent indicators of information backflow, since in the case of spatial correlations two different quantifiers of entanglement may be in-equivalent in detecting correlation backflow [31], while there may exist a measure that detects "all most" all non-Markovian dynamics [32]. In this work, we address the question of the ability of temporal steering and temporal non-separable correlations to detect non-Markovianity of a generic quantum channel and establish their place in the divisibility hierarchy [12, 13]. Before answering the question, we shall review the well-known divisibility hierarchy.

II. PRELIMINARIES

A. Divisibility hierarchy

Let $\rho \in \mathcal{B}(\mathcal{H})$, where $\mathcal{B}(\mathcal{H})$ is the bounded operator space and $\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ be a quantum channel, with the operator-sum form $\Lambda[\rho] = \sum_j K_j \rho K_j^\dagger$, where K_j are called the Kraus operators with $\sum_j K_j^\dagger K_j = \mathbb{I}$. The map

[*] shrikant.phys@gmail.com

$\Lambda(t)$ is positive when it outputs a positive semidefinite state i.e., $\Lambda(t)[\rho] \geq 0$, for all $\rho \in \mathcal{B}(\mathcal{H})$. The map $\Lambda(t)$ is the said to be completely positive (CP) if the matrix

$$\chi(t) = (I \otimes \Lambda)[|\Psi\rangle\langle\Psi|], \tag{1}$$

called the Choi matrix, is positive semidefinite i.e., $\chi(t) \geq 0$, where I is the identity map and $|\Psi\rangle = |00\rangle + |11\rangle \in \mathcal{H} \otimes \mathcal{H}$ is an unnormalized maximally entangled state. The divisibility property for a map is given by concatenation of intermediate maps [9]:

$$\Lambda(t+\epsilon, 0) = \Lambda(t+\epsilon, t)\Lambda(t, 0), \tag{2}$$

for all $t + \epsilon \geq t \geq 0$. Assuming that some form of inverse Λ^{-1} exists, one can define the intermediate map

$$\Lambda(t, s) = \Lambda(t+\epsilon, 0)\Lambda^{-1}(t, 0). \tag{3}$$

Then the following definitions hold.

Definition 1 *A process is said to be CP-divisible if the intermediate map $\Lambda(t,s)$ is CP.*

Definition 2 *A process is said to be P-divisible if the intermediate map $\Lambda(t,s)$ is positive but need not be CP.*

Definition (2) is equivalent to

$$\frac{d}{dt}\|\Lambda(t)[\rho_1 - \rho_2]\|_1 \geq 0 \tag{4}$$

for any pair of $\rho_1, \rho_2 \in \mathcal{B}(\mathcal{H})$, where $\|A\|_1 = \mathrm{Tr}\sqrt{A^\dagger A}$ is the trace norm. This measure was originall proposed by Breuer-Laine-Piilo [10, 33].

We will describe next that the violation of Eq. (4) is only *sufficient* but not necessary for P-indivisibility of a given map $\Lambda(t)$.

B. Non-unital non-Markovianity

Any d-dimensional dynamical map $\Lambda(t)$ may be given a matrix representation as

$$F(t) = \left(\begin{array}{c|c} 1 & \mathbf{0}_{1\times(d^2-1)} \\ \hline \boldsymbol{\tau} & M \end{array}\right), \tag{5}$$

where $\boldsymbol{\tau} \in \mathbb{R}^{d^2-1}$ and M is a $d^2 - 1 \times d^2 - 1$ real matrix. Given the state $\rho = \frac{1}{d}(\mathbb{I} + \sum_{i=1}^{d^2-1} r_i G_i)$, where G_i are traceless orthonormal basis given by the generalized Gell-Mann matrices with $G_0 = \frac{\mathbb{I}}{\sqrt{d}}$ and the Hermitian G_i where $i \in \{1, ...d^2 - 1\}$, then the map $F_{ij}(t) := \mathrm{Tr}(G_i\Lambda[G_j])$, which transforms a Bloch vector $\boldsymbol{r}$ as

$$\boldsymbol{r} \to \boldsymbol{r}'(t) = M(t)\boldsymbol{r}(0) + \boldsymbol{\tau}. \tag{6}$$

The components of the vector $\boldsymbol{\tau}$ are obtained as $\tau_i = \mathrm{Tr}(G_i\Lambda[\mathbb{I}])$. Therefore, a map is unital if $\tau_i = 0$, i.e., $\Lambda[\mathbb{I}] = \mathbb{I}$. If $\boldsymbol{r_1}$ and $\boldsymbol{r_2}$ are two Bloch vectors corresponding to the states ρ_1 and ρ_2, then from Eq. (5) and (6), the transformation $\Lambda(t)[\rho_1 - \rho_2]$ corresponds to $M(t)[\boldsymbol{r_1} - \boldsymbol{r_2}]$. Therefore, the trace distance fails to witness non-Markovianity originating solely from $\boldsymbol{\tau}$. However, the full non-Markovianity, in the sense of P-indivisibility is encoded in the map $F(t)$. Interestingly, the divisibility property of $\Lambda(t)$ carried over as $F(t) = F(t,s)F(s)$ leads to a non-trivial relationship between unital and non-unital parts of the channel [13]:

$$M(t) = M(t,s)M(s) \text{ and } \boldsymbol{\tau}(t) = \boldsymbol{\tau}(t,s) + M(t,s)\boldsymbol{\tau}(s), \tag{7}$$

where $\boldsymbol{\tau}(t,s)$ and $M(t,s)$ parameterize $F(t,s)$.

It is known that non-unitality of a channel is necessary for increase of purity $\mathcal{P} := \mathrm{Tr}(\rho^2)$. For instance, it can be shown that when $\Lambda(t)[\mathbb{I}/2] \neq \mathbb{I}/2$, that is when $\boldsymbol{\tau} = \mathbf{0}$, then $\mathrm{Tr}(\Lambda(t)[\rho])^2 \leq \mathrm{Tr}(\rho^2)$ for all t, even when the map $\Lambda(t)$ is P-indivisible according to the Eq. (4). Therefore, the full P-indivisibility is captured in $F(t)$ and one would require a distance measure that is sensitive to non-unital part. For simplicity, the feature of non-monotonicity originating solely from the non-unital part of the channel may be called as "non-unital non-Markovianity".

An example of a purely non-unital non-Markovian channel is given in Ref. [17]. Any qubit channel, up to a unitary transformation, is equivalently described by substituting $\boldsymbol{\tau} = (0, 0, \tau_3)^T$ and $M_{ij} = \{\lambda_1, \lambda_2, \lambda_3\}$ with $M_{i\neq j} = 0$ in Eq. (5). For a qubit, the Pauli operators σ_i, $i \in \{0, 1, 2, 3)$, with $\sigma_0 = \mathbb{I}_2$, form the orthonormal Hilbert-Schmidt basis. In Ref. [17], a non-Markovian qubit generalized amplitude damping (GAD) channel $\Lambda(t)[\rho] = \sum_i A_i\rho A_i^\dagger$ was proposed given by the Kraus operators

$$A_1 = \sqrt{1-p}\begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\eta} \end{bmatrix}; \ A_2 = \sqrt{1-p}\begin{bmatrix} 0 & \sqrt{\eta} \\ 0 & 0 \end{bmatrix};$$

$$A_3 = \sqrt{p}\begin{bmatrix} \sqrt{1-\eta} & 0 \\ 0 & 1 \end{bmatrix}; \ A_4 = \sqrt{p}\begin{bmatrix} 0 & 0 \\ \sqrt{\eta} & 0 \end{bmatrix},$$

$$\text{with } \eta(t) := 1 - e^{-t} \quad \text{and} \quad p(t) := \sin^2(\omega t), \tag{8}$$

where ω is some real number. In terms of the map $F(t)$, this corresponds to the parameters $\lambda_1 = \lambda_2 = \sqrt{1-\eta}$, $\lambda_3 = 1 - \eta$ and $\tau_3 = (1 - 2p)\eta$. Clearly, trace distance measure in Eq.(4) fails to witness non-Markovianity as the non-monotonicity originates from $p(t)$ rather than $\eta(t)$.

III. TEMPORAL QUANTUM CORRELATIONS AND NON-MARKOVIANITY

In this section we show that the so-called temporal correlations, namely temporal non-separability and temporal steering correlations faithfully witness non-unital non-Markovianity. However, we show that temporal steerable correlations fail to quantify enteral non-Markovianity, which PDO can modified to define a measure of non-Markoviantiy such that it does indeed quantify eternal non-Markovianity.

A. Pseudo-density operator

Pseudo-density operator (PDO) construction captures spatial and temporal correlations in an equal footing [27, 34]. There are a number of other such frameworks that address the problem of unification of spatiotemporal correlations namely the process matrix [35], process tensor [36] that are derived from the framework of quantum combs [37] which represent most general space-time processes, and superdensity operator [38], the spatiotemporal doubled density operator [39], to name a few others. It has been shown that process matrix can be mapped to PDO in a number of ways [40].

A two-point qubit PDO can be written as

$$R = I \otimes \Lambda(t) \left[\left\{ \rho \otimes \frac{\mathbb{I}_2}{2}, \frac{1}{2} \sum_{i=0}^{3} \sigma_i \otimes \sigma_i \right\} \right], \tag{9}$$

where σ_i are Pauli operators with $\sigma_0 = \mathbb{I}_2$, $\{A, B\} = AB + BA$, I is the identity map and ρ is the input to the PDO under the quantum channel $\Lambda(t)$. Fitzsimons et. al. introduced [27] a measure of "temporalness" of correlations, dubbed causality measure as the negativity of PDO, given by $f = \|R\|_1 - 1$. Later, it was generalized to a log negativity of PDO as $F = \log \|R\|_1$ in order to preserve additivity property, thereby connecting it to the quantum capacity of a given channel [34]. We call F simply the causality measure (CM). F is nonincreasing under CPTP map, therefore a CP-divisible (or, Markovian) map yields $F_{\Lambda(t)[\rho]} \geq F_{\Lambda(t+\tau)[\rho]}$, violation of which implies information backflow. Figure (1) shows this. A measure of non-Markovianity has been defined [18] using F, for a given channel Λ and input state ρ, as the integral over positive slope of F

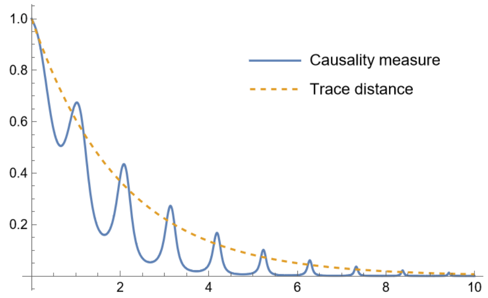


FIG. 1. (Color online.) Breakdown of monotonicity of causality measure $F(t)$ for $\omega = 3$ (bold, blue curve) and the behavior of trace distance (dashed, orange curve) under non-Markovian GAD given in Eq.(8). The x-axis is time.

$$\mathcal{N} := \max_{\rho} \int_{\frac{dF}{dt}>0} dt \, \frac{dF(\rho, \Lambda, t)}{dt}. \tag{10}$$

Interestingly, $\mathcal{N}^{\text{causality}}$ is equivalent to the entanglement negativity measure due to Rivas-Huelga-Plenio [9] if $\rho = \frac{\mathbb{I}}{2}$. This makes PDO an advantage over the Choi matrix based method since one neither requires an entangled state or ancillary systems to identify correlation backflows under non-Markovian evolution. Additionally, since PDM takes in a single density operator as input, it is free from optimization problem for the initial pair of inputs states involved in distinguishability measures.

B. Temporal steering

Similar to the steering in space with a given spatially entangled state, one may steer a state in time by making a measurement on the input state and sending it via a quantum channel followed by a complete quantum state tomography of the output state at the end of the channel. Now, we shall introduce the notion of temporal steerable weight. Alice performs a positive operator valued measure (POVM) measurement on an input state ρ at $t = 0$ transforming it into

$$\rho_{a|x} = \frac{\Pi_{a|x} \rho \Pi_{a|x}^{\dagger}}{p(a|x)}, \tag{11}$$

where $p(a|x) = \text{Tr}[\Pi_{a|x} \rho \Pi_{a|x}^{\dagger}]$ is the probability that an outcome a occurs given that Alice preforms a measurement in the basis x. Now the state $\rho_{a|x}$ is sent to Bob down a noisy quantum channel $\Lambda(t)$ for a time t. When Bob receives the state at t he performs a quantum state tomography to get the state $\sigma_{a|x}(t) = \Lambda(t)[\sigma_{a|x}(0)$. We may call the set of states $\sigma_{a|x}(t)$ as temporal assemblages, and let the unnormalized assemblage be $\sigma_{a|x}(t) \equiv p(a|x)\sigma_{a|x}$. Now, by assumption, Bob doesn't trust Alice nor her devices, and he would want to distinguish the correlations due to Alice's measurements from the correlations that might have originated from a hidden variable λ, making the correlations to satisfy locality in time and realism. Therefore, we may represent the correlations that might be produced by such classical origins as

$$\sigma^{US}_{a|x}(t) = \sum_{\lambda} P(\lambda)P(a|x,\lambda)\sigma_{\lambda}, \tag{12}$$

where $\sigma^{US}_{a|x}(t)$ is the unsteerable assemblage and $P(a|x,\lambda)$ is the probability that an outcomes a occurs given that Alice makes a measurement x, and λ the hidden variable that might have influenced the outcome, in which case Bob will be able to write down his assemblage in the form (12), and when he can't, then he is sure that the state is prepared by Alice's measurement. Now, a measure of temporal steering was introduced by [30] called temporal steerable weight (TSW). In order to define TSW consider a convex mixture

$$\sigma_{a|x}(t) = w\sigma^{US}_{a|x}(t) + (1-w)\sigma^{S}_{a|x}(t) \quad \forall a, x. \tag{13}$$

Clearly, $\sigma_{a|x}(t)$ is an assemblage which might contain both unsteerable and steerable correlations, with the constraint $0 \leq w \leq 1$. The TSW for a given assemblage $\sigma_{a|x}(t)$ is defined by

$$W^{\text{TS}} = 1 - w', \tag{14}$$

where w' is the maximum value of w. TSW may be interpreted as the minimal steerable resources required to reproduce temporal steerable assemblage. That is, $W^{\text{TS}} = 0$ and 1 for minimal and maximal steerability, respectively. w' may be obtained by semi-definite programming:

$$\begin{aligned} \text{Find} \quad & w' = \max \operatorname{Tr} \sum_{\lambda} w\sigma_{\lambda}, \\ \text{subject to} \quad & \left(\sigma_{a|x}(t) - \sum_{\lambda} q_{\lambda}(a|x) w\sigma_{\lambda}\right) \geq 0 \quad \forall a, x \\ & w\sigma_{\lambda} \geq 0 \qquad \forall \lambda, \end{aligned} \tag{15}$$

where $q_{\lambda}(a|x)$ are the extremal values of $P_{\lambda}(a|x)$.

Now, under the noisy quantum channel these correlations deteriorate and [30] have shown that W^{TS} is non-increasing under local operations. Therefore, we have the monotonicity condition

$$W^{\text{TS}}_{\rho} \geq W^{\text{TS}}_{\Lambda[\rho]}. \tag{16}$$

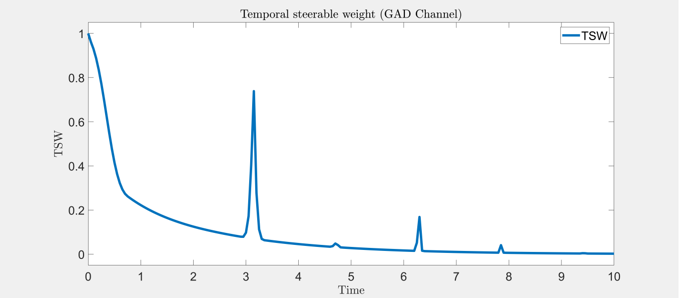


FIG. 2. (Color online.) Breakdown of monotonicity of TSW for $\omega = 3$, for the channel given in Eq.(8).

A Markov process satisfies the above condition, and a non-Markovian process violates it. This is shown in Fig. (2). Given this fact, a measure of non-Markovianity is nothing but the area under the positive slope of of $W^{\text{TS}}_{\Lambda[\rho]}$:

$$\mathcal{N}_{\text{TSW}} = \int_{t=0\,;\,\frac{dW^{\text{TS}}}{dt}>0}^{t} \frac{dW^{\text{TS}}}{dt} dt, \tag{17}$$

which by the factor of $\frac{1}{2}$ is equivalent to

$$N := \int_{t_0}^{t_{\max}} \left| \frac{dW^{\text{TS}}}{dt} \right| dt + (W^{\text{TS}}_{t_{\max}} - W^{\text{TS}}_{t_0}). \tag{18}$$

It is important to mention that $\mathcal{N}_{\text{TSW}}$ is only a sufficient and not a necessary condition for non-Markovianity of Λ. There may be channels that will be detected as Markovian by this measure while other measures may detect them as non-Markovian. Breakdown of monotonicity of TSW may be interpreted as information back-flow from the environment to the system, hence this measure captures the range of memory effects that BLP does.

IV. CONCLUSION AND A NOTE

It is known [17, 41] that trace distance as an indicator of information backflow may fail for non-Markovian channel in which non-unital part is solely responsible for information backflow. Trace distance and other correlation measures such as mutual information and entanglement negativity fail to witness "eternal non-Markovianity" [16]. The question of the equivalence between temporal steering and the temporal non-separable correlations in PDO as indicators of non-Markovianity was left open in Ref.[8]. We fill this gap in this work and we have shown by example that these two measures sit somewhere between distinguishability (such as trace distance) and divisibility-based indicators. They faithfully witness non-unital non-Markovianity, however it can be shown that they (the ones presented in the current version of the paper) fail to witness eternal non-Markoviantiy.

Note: This work is in progress and more results are awaited, and the full version shall be posted on arXiv very soon.

ACKNOWLEDGMENTS

I thank IIT Madras for the support through the Institute Postdoctoral Fellowship. I also thank Prateek Chawla for the kind help with the semidefinite program for temporal steerable weight written in MATLAB.

[1] Heinz-Peter Breuer and Francesco Petruccione. *The theory of open quantum systems*. Oxford University Press, 2002.

[2] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.

[3] Angel Rivas, Susana F Huelga, and Martin B Plenio. Quantum non-markovianity: characterization, quantification and detection. *Rep. Prog. Phys*, 77(9):094001, 2014.

[4] Heinz-Peter Breuer, Elsi-Mari Laine, Jyrki Piilo, and Bassano Vacchini. Colloquium: Non-markovian dynamics in open quantum systems. *Rev. Mod. Phys*, 88(2):021002, 2016.

[5] Inés de Vega and Daniel Alonso. Dynamics of non-markovian open quantum systems. *Rev. Mod. Phys.*, 89:015001, Jan 2017.

[6] Li Li, Michael J.W. Hall, and Howard M. Wiseman. Concepts of quantum non-markovianity: A hierarchy. *Physics Reports*, 759:1 – 51, 2018.

[7] Dariusz Chruściński. Dynamical maps beyond markovian regime. *Physics Reports*, 992:1–85, 2022.

[8] U. Shrikant and Prabha Mandayam. Quantum non-markovianity: Overview and recent developments. *Frontiers in Quantum Science and Technology*, 2, 2023.

[9] Ángel Rivas, Susana F Huelga, and Martin B Plenio. Entanglement and non-markovianity of quantum evolutions. *Phys. Rev. Lett*, 105(5):050403, 2010.

[10] Heinz-Peter Breuer, Elsi-Mari Laine, and Jyrki Piilo. Measure for the degree of non-markovian behavior of quantum processes in open systems. *Phys. Rev. Lett*, 103(21):210401, 2009.

[11] Dariusz Chruściński, Andrzej Kossakowski, and Ángel Rivas. Measures of non-markovianity: Divisibility versus backflow of information. *Physical Review A*, 83(5):052128, 2011.

[12] Dariusz Chruściński and Sabrina Maniscalco. Degree of non-markovianity of quantum evolution. *Physical review letters*, 112(12):120404, 2014.

[13] Dariusz Chruściński, Chiara Macchiavello, and Sabrina Maniscalco. Detecting non-markovianity of quantum evolution via spectra of dynamical maps. *Physical review letters*, 118(8):080404, 2017.

[14] Dariusz Chruściński, Ángel Rivas, and Erling Størmer. Divisibility and information flow notions of quantum markovianity for noninvertible dynamical maps. *Phys. Rev. Lett.*, 121:080407, Aug 2018.

[15] Sagnik Chakraborty and Dariusz Chruściński. Information flow versus divisibility for qubit evolution. *Physical Review A*, 99(4):042105, 2019.

[16] Michael J. W. Hall, James D. Cresser, Li Li, and Erika Andersson. Canonical form of master equations and characterization of non-markovianity. *Phys. Rev. A*, 89:042120, Apr 2014.

[17] Jing Liu, Xiao-Ming Lu, and Xiaoguang Wang. Nonunital non-markovianity of quantum dynamics. *Phys. Rev. A*, 87:042103, Apr 2013.

[18] Shrikant Utagi. Quantum causal correlations and non-markovianity of quantum evolution. *Physics Letters A*, 386:126983, 2021.

[19] Nina Megier, Andrea Smirne, and Bassano Vacchini. Entropic bounds on information backflow. *Physical Review Letters*, 127(3):030401, 2021.

[20] Alexander Streltsov, Hermann Kampermann, and Dagmar Bruß. Behavior of quantum correlations under local noise. *Physical review letters*, 107(17):170502, 2011.

[21] Shrikant Utagi, R Srikanth, and Subhashish Banerjee. Ping-pong quantum key distribution with trusted noise: non-markovian advantage. *Quantum Information Processing*, 19(10):1–12, 2020.

[22] Katarzyna Siudzińska. Improving classical capacity of qubit dynamical maps through stationary state manipulation. *Journal of Physics A: Mathematical and Theoretical*, 56(23):235301, 2023.

[23] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Reviews of modern physics*, 81(2):865, 2009.

[24] Roope Uola, Ana CS Costa, H Chau Nguyen, and Otfried Gühne. Quantum steering. *Reviews of Modern Physics*, 92(1):015001, 2020.

[25] Adrien Feix and Časlav Brukner. Quantum superpositions of common-causeand direct-causecausal structures. *New Journal of Physics*, 19(12):123028, 2017.

[26] Mariami Gachechiladze, Nikolai Miklin, and Rafael Chaves. Quantifying causal influences in the presence of a quantum common cause. *Physical Review Letters*,

125(23):230401, 2020.

[27] Joseph F Fitzsimons, Jonathan A Jones, and Vlatko Vedral. Quantum correlations which imply causation. *Scientific reports*, 5:18281, 2015.

[28] Huan-Yu Ku, Shin-Liang Chen, Neill Lambert, Yueh-Nan Chen, and Franco Nori. Hierarchy in temporal quantum correlations. *Physical Review A*, 98(2):022104, 2018.

[29] Karol Bartkiewicz, Antonín Černoch, Karel Lemr, Adam Miranowicz, and Franco Nori. Temporal steering and security of quantum key distribution with mutually unbiased bases against individual attacks. *Physical Review A*, 93(6):062345, 2016.

[30] Shin-Liang Chen, Neill Lambert, Che-Ming Li, Adam Miranowicz, Yueh-Nan Chen, and Franco Nori. Quantifying non-markovianity with temporal steering. *Physical review letters*, 116(2):020503, 2016.

[31] Alaor Cervati Neto, Göktuğ Karpat, and Felipe Fernandes Fanchini. Inequivalence of correlation-based measures of non-markovianity. *Physical Review A*, 94(3):032105, 2016.

[32] Dario De Santis, Markus Johansson, Bogna Bylicka, Nadja K Bernardes, and Antonio Acín. Correlation measure detecting almost all non-markovian evolutions. *Physical Review A*, 99(1):012303, 2019.

[33] Heinz-Peter Breuer, Elsi-Mari Laine, Jyrki Piilo, and Bassano Vacchini. Colloquium: Non-markovian dynamics in open quantum systems. *Rev. Mod. Phys.*, 88:021002, Apr 2016.

[34] Robert Pisarczyk, Zhikuan Zhao, Yingkai Ouyang, Vlatko Vedral, and Joseph F Fitzsimons. Causal limit on quantum communication. *Physical review letters*, 123(15):150502, 2019.

[35] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature communications*, 3(1):1–8, 2012.

[36] Felix A. Pollock, César Rodríguez-Rosario, Thomas Frauenheim, Mauro Paternostro, and Kavan Modi. Non-markovian quantum processes: Complete framework and efficient characterization. *Phys. Rev. A*, 97:012127, Jan 2018.

[37] Giulio Chiribella, Giacomo Mauro DAriano, and Paolo Perinotti. Theoretical framework for quantum networks. *Physical Review A*, 80(2):022339, 2009.

[38] Jordan Cotler, Chao-Ming Jian, Xiao-Liang Qi, and Frank Wilczek. Superdensity operators for spacetime quantum mechanics. *Journal of High Energy Physics*, 2018(9):1–57, 2018.

[39] Zhian Jia and Dagomir Kaszlikowski. The spatiotemporal doubled density operator: a unified framework for analyzing spatial and temporal quantum processes. *arXiv preprint arXiv:2305.15649*, 2023.

[40] Tian Zhang, Oscar Dahlsten, and Vlatko Vedral. Quantum correlations in time. *arXiv preprint arXiv:2002.10448*, 2020.

[41] Harri Mäkelä. Bounds for the divisibility-based and distinguishability-based non-markovianity measures. *Physical Review A*, 91(1):012108, 2015.

Variational Decoupling of Quantum Dynamics

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*

Abstract. Decoupling of systems enables the decomposition of complex problems into simpler ones. However, partitioning nature can be intricate, especially in the quantum world, where intuition may generally lose its sensation. While the analytical solutions are infeasible, variational quantum algorithms (VQA) provide a promising way to tackle this issue. In this work, we propose a variational algorithm to decouple quantum evolutions, optimizing the variational circuit to find a good approximation of the original evolution as independent evolutions on subsystems. Our algorithm can help simplify more complicated quantum systems, which helps to reveal underlying structures and accelerate variational quantum compilation. The efficient state-preparation and measurement procedures make our algorithm near-term quantum computer compatible.

Keywords: Variational Quantum Algorithms, Quantum Simulation, Quantum Compilation, Decoupling

1 Introduction

The partitioning of systems plays a crucial role in the study of nature. The separability of a system enables the decomposition of a complex problem into simpler ones. For instance, consider how the decomposition of forces simplifies the description of the free-falling objects revealing the breaking of symmetry due to gravity. The separation of variables has contributed greatly to the success of modern physics. Indeed, the attempt to decouple physical systems is not only for theoretical interest, but also of practical importance. The ability of decoupling systems is reflected in the simplification of computation. By decouple a system into smaller systems, it enables the possibility of parallelization. The parallelization breaks a task for large computers into tasks for smaller computers, which can be solved simultaneously, and save the time of the overall computation.

Fruitful results in simplifying physical systems have been achieved by decoupling them. Take how the orbits of Hydrogen atoms are solved via the separation of variables as an example. However, though the ability to identify integrated components is usually taken for granted by human beings, partitioning nature itself can be convoluted in general. This issue is further exacerbated in the quantum world, where intuition may generally lose its sensation. As the complexity of quantum systems increases, the ability to decouple quantum dynamics becomes more challenging.

Despite the difficulty, the problems on large quantum systems need to be solved for deeper studies. While the analytical methods lose their power, it is natural to ask if the decoupling of quantum systems can be automated. Variational quantum algorithms (VQA) [1] attract more and more attention as the noise-intermediate-scale-quantum (NISQ) [10] era approaching, as they can be achieved with shallow quantum circuits and are robust to some noises. VQA utilizes the techniques from classical machine learning to optimize the parameters of quantum circuits, which aims to accomplish tasks with limited analytical understanding of the tasks. This fast-developing method provides a promising way to tackle many more tasks, which includes the problem of decoupling quantum dynamics.

In this work, we propose a variational algorithm to decouple quantum evolutions. An efficiently measurable quantity is introduced to quantify the degree of decoupling, which is used to guide the optimization of the variational circuit. Based on such quantity, we can find shallow pre- and post-processing quantum circuits by optimizing the variational circuits, where the original evolution can be approximated as independent evolutions on subsystems after the processing. With this algorithm, the variational quantum compilation [6] can be accelerated and the parallelization of quantum computations can be achieved. By pre-train the quantum circuit with our method, it saves the number of total measurements for the training process. Under the scope of quantum simulation, our algorithm can also help to find the decoupling of the Hamiltonian of a system by tuning its basis. We expect our method can help to simplify more complicated quantum systems and reveal more underlying structures. As the state preparation and measurements are efficient, our method can be near-term quantum computer compatible.

2 Framework

Parallel quantum computation presents a promising avenue for accelerating quantum computations by distributing workloads across multiple quantum systems, thereby minimizing the rounds of qubit interactions within a quantum computer, which are considered challenging to implement [4, 13, 14, 16]. However, a generic unitary operator U_{AB} acting on two-party quantum systems A and B generates entanglement between the two systems, rendering the local implementation of the uni-

*ximing001@e.ntu.edu.sg

†cr.yang@nus.edu.sg

‡ceptryn@gmail.com

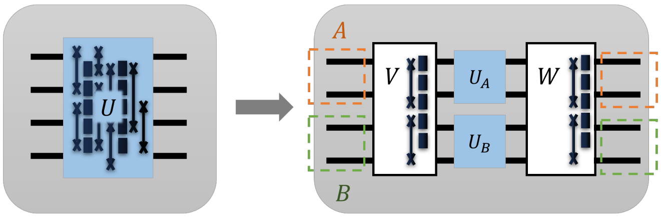


Figure 1: The task of decoupling evolutions is to find the pre-processing operation V and post-processing operation W, such that the target unitary operation U can be decomposed as $W(U_A \otimes U_B)V$, where U_A and U_B are unitary operators acting on subsystems A and B respectively.

tary operator on each subsystem impractical. Consequently, this raises the question: can we identify a simple pre-processing V and a simple post-processing W that decouples the unitary operator U_{AB} into two unitary operators U_A and U_B acting on subsystems A and B respectively? This decoupling idea is illustrated in Figure 1.

Although analytically decoupling a general evolution may be challenging, we develop a variational hybrid quantum-classical algorithms [9] by optimizing the parameters of the pre- and post-processing unitary operators. Indeed, we propose a cost function to measure the separability of the unitary operator U_{AB}, which lies at the heart of the success of the algorithm. We propose an efficient method to evaluate the cost function, leading to efficiently computing the gradient of the cost function in conjunction with the parameter shift rules. We try to find two parameterized pre- and post- transformations $W(\vec{\theta})$ and $V(\vec{\varphi})$, employing which the target unitary operator U_{AB} can be decomposed as

$$W^{\dagger} U_{AB} V^{\dagger} = U_A \otimes U_B \ . \tag{1}$$

We define our cost function C^U_{sep} as a measure of the separability of the unitary operator $VU_{AB}W$, which takes the average over the separability of quantum states after applying $VU_{AB}W$ to the product states, where

$$C^U_{\text{sep}} = \iint C_{\text{sep}}(\mathcal{U}(|\psi\rangle\langle\psi|^A \otimes |\phi\rangle\langle\phi|^B))\mathrm{d}\,|\psi\rangle\,\mathrm{d}\,|\phi\rangle \ \ , \tag{2}$$

and the $\mathcal{U}$ is a shorthand for the unitary channel corresponding to the quantum circuit. The function C_{sep} is defined for a bipartite quantum state $\rho^{A,B}$ as

$$C_{\text{sep}}(\rho^{A,B}) = 2 - \mathrm{Tr}\left[\left(\rho^A\right)^2\right] - \mathrm{Tr}\left[\left(\rho^B\right)^2\right] \ ,$$

where $\rho^{A,B}, \rho^A, \rho^B$ are the density matrices of the entire system, subsystem A, and subsystem B, respectively. We show that the cost function can be efficiently measured through quantum circuits of constant depth and a linear time classical processing, which does not require ancilla qubits or controlled operations (theorem 1). As the cost function is evaluated efficiently, we can also use the parameter shift rule to obtain the gradients efficiently. The parameters are then updated with conventional gradient descending methods, such as ADAM [7]

Theorem 1 *Our cost function can be efficiently measured by preparing a bipartite initial state $\tau_A \otimes \tau_B$ following a swap test with an observable O,*

$$C^U_{sep}(\mathcal{U}) = \mathrm{tr}\left[\mathcal{U}^{\otimes 2}_{\vec{\theta},\vec{\varphi}}(\tau_A \otimes \tau_B)O\right] \ , \tag{3}$$

where

1. *the initial state $\tau_A \otimes \tau_B$ can be prepared with a constant depth quantum circuit and a linear time classical preprocessing;*
2. *the observable O can be measured with a constant depth quantum circuit and a linear time classical postprocessing;*

This cost function is designed as a measure of the separability of the unitary operator. A quantum evolution is decoupled if and only if all separable states remain separable after applying the quantum evolution. Therefore, we define the cost function to be the average separability of all quantum states after applying the quantum operator to the separable states. Thus, for the efficiency of measurements, we consider the Tsallis-2 entropy of the Schmidt coefficients of the quantum state $|\psi_{AB}\rangle$ as the measure of entanglement, which is also known as the linear entropy. The linear entropy of the subsystem A is 0 when the entire quantum state has 0 entanglement, namely separable. Thus, we use the deviance of the linear entropy from 0 to define our quantum state separability [11]

3 Numerical Results

We demonstrate the effectiveness of our algorithm in two settings: (1) quantum compilation that aims to decompose a general unitary operator into a circuit consisting of elementary gates, thus having a broad application in quantum computation and (2) quantum simulation that mimics the behavior of another quantum system. In the first setting, our algorithm provides an alternative method to compile a quantum unitary operator into a quantum circuit with a tree structure. We apply our algorithm to compile both 2-qubit and 4-qubit unitary operators. Our numerical results show that our algorithm is faster than the conventional compilation method that uses fidelity as the cost function. In the second setting, we apply our algorithm to find a basis of a Hamiltonian in which the Hamiltonian is separable. Our numerical results further show that our cost function is quadratically correlated to the trace distance between the Hamiltonian and the decoupled Hamiltonian.

3.1 Quantum Compilation

The quantum compilation is the process of converting a generic unitary operator into a set of elementary gates that can be executed on a specific quantum computing hardware. Various variational algorithms have been proposed to discover approximate quantum circuits for a given target unitary operator [2, 5, 12, 6]. But minimizing these variational algorithms' cost function, the average fidelity $\bar{F}$ between target unitary operator U and the ansatz unitary operator V, is still challenging for large quantum circuits due to the vanishing of both fidelity $\bar{F}$ and its derivatives [8].

Our decoupling method tackles the compilation problem by breaking down learning the entire approximate quantum circuits into discovering a sequence of smaller quantum circuits. Specifically, we first apply our decoupling algorithm to find a post-processing unitary operator V that $UV^\dagger$ is approximately separable, namely $UV^\dagger \approx U_A \otimes U_B$, rather than training a circuit $V(\vec{\theta})$ to approximate U such that $UV^\dagger(\vec{\theta}) \approx \mathbb{I}$. Enlarging the optimum set of the optimization problem from $\mathbb{I} = \mathbb{I}_A \otimes \mathbb{I}_B$ to $U_A \otimes U_B$ reduces the complexity of training.

We demonstrate the effectiveness of our decoupling method in compiling 2-qubit and 4-qubit unitary operators. The target 2-qubit unitary operator U is uniformly random generated while the ansatz is a circuit that allows compiling any 2-qubit unitary operator. A similar test is also performed on a 4-qubit unitary, where the target unitary is generated by one layer of the tree ansatz with randomly generated parameters. Both methods are trained with ADAM [7] with the same hyper-parameters. In both cases, the two methods are compared by the number of gradients evaluated, which is proportional to the number of measurements made during the training. The 4-qubit case is shown in figure 2, our decoupling method converges to a better result with fewer measurements in general. Note that the average fidelities are not optimized in the first step (orange region) and the intermediate step (red region), as we are decoupling the unitary. However, the average fidelities are optimized faster in the last step (pink region), as we only need to train the local gates to approximate the single-qubit unitaries.

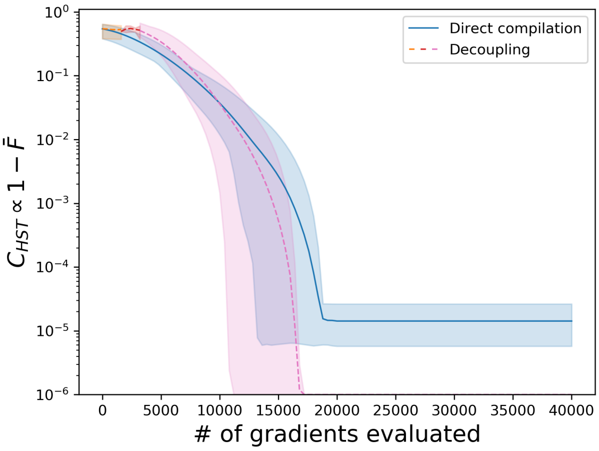


Figure 2: Compile a 4-qubit unitary with ADAM optimization. Each line is the average over 10 independent training processes, while the regions between the worst and best cases are shaded with the corresponding color. The task is divided into three steps here. The first step (orange region) decouples the unitary to 2 two-qubit systems; the second step (red region) decouples each two-qubit system to 2 single-qubit systems; and the last step minimizes the fidelities between each single-qubit system.

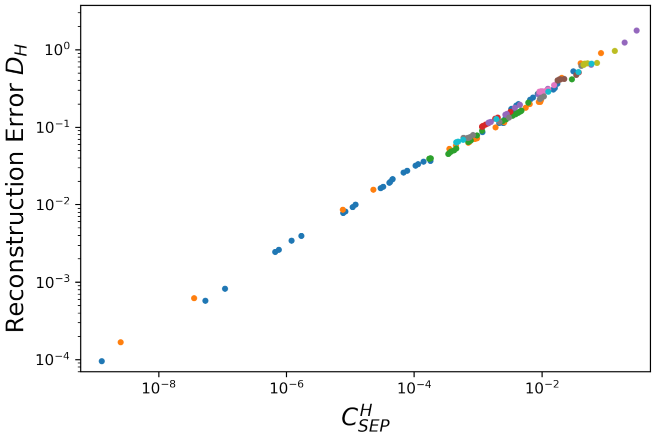


Figure 3: The relations between the reconstruction error and the cost function C^U_{sep}. While the cost function is minimized during the training process, the reconstruction errors are also non-increasing.

3.2 Quantum Simulation

Quantum simulation predicts material properties by calculating ground and excited states using qubit-based operators, which can be implemented by quantum algorithms, with variational methods enhancing results and expanding simulatable evolutions [3, 15]. For a certain given Hamiltonian H, it is important to know if its evolution e^{-itH} can be decoupled into two independent parts:

$$Ve^{-itH}V^\dagger = e^{-itVHV^\dagger} \approx e^{-itH_A} \otimes e^{-itH_B} \ , \qquad (4)$$

where H_A and H_B are some local Hamiltonians acting on subsystems A and B respectively. Thus, we apply our decoupling method to Hamiltonian decoupling by specifying the post-processing unitary V to be the inverse of the pre-processing unitary operator. We hope to first train a shallow circuit $V(\vec{\phi})$, so that $V(\vec{\phi})HV^\dagger(\vec{\phi}) \approx H_A + H_B$ if possible, where H_A and H_B only act on each subsystem A and B, respectively. This can be done with our decoupling method by minimizing the cost function C^U_{sep} (2) on e^{-itH} for some random t, which is equivalent to minimizing the trace distance between the target Hamiltonian and the decoupled Hamiltonian:

$$D_H := \frac{1}{2}\left\| H - V^\dagger(H_A + H_B)V \right\|_1 \ . \qquad (5)$$

To demonstrate how the Hamiltonians can be decoupled, we test the algorithm on 10 randomly generated 3-tuples (H_A, H_B, V) defined on two qubits, where H_A and H_B are Hamiltonians acting on a single qubit, and V is a unitary operator that acts on the entire system. For each 3-tuple, we add an interaction term with different strengths. The reconstruction errors (5) are plotted against the cost function C^H_{sep} for parameter update during the training procedure, as shown in Figure. 3. The figure shows that our cost function is nearly a quadratic function of the reconstruction error, indicating that our cost function (2) captures the reconstruction errors well for the purpose of decoupling the dynamics.

References

[1] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles. Variational Quantum Algorithms. *Nat Rev Phys*, 3(9):625–644, Aug. 2021.

[2] F. T. Chong, D. Franklin, and M. Martonosi. Programming languages and compiler design for realistic quantum hardware. *Nature*, 549(7671):180–187, Sept. 2017.

[3] I. M. Georgescu, S. Ashhab, and F. Nori. Quantum Simulation. *Rev. Mod. Phys.*, 86(1):153–185, Mar. 2014.

[4] Y. He, S. K. Gorman, D. Keith, L. Kranz, J. G. Keizer, and M. Y. Simmons. A two-qubit gate between phosphorus donor electrons in silicon. *Nature*, 571(7765):371–375, July 2019.

[5] K. Heya, Y. Suzuki, Y. Nakamura, and K. Fujii. Variational Quantum Gate Optimization, Oct. 2018.

[6] S. Khatri, R. LaRose, A. Poremba, L. Cincio, A. T. Sornborger, and P. J. Coles. Quantum-assisted quantum compiling. *Quantum*, 3:140, May 2019.

[7] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization, Jan. 2017.

[8] 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, Nov. 2018.

[9] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.*, 18(2):023023, Feb. 2016.

[10] J. Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, Aug. 2018.

[11] E. Santos and M. Ferrero. Linear entropy and Bell inequalities. *Phys. Rev. A*, 62(2):024101, July 2000.

[12] K. Sharma, S. Khatri, M. Cerezo, and P. J. Coles. Noise resilience of variational quantum compiling. *New Journal of Physics*, 22(4):undefined–undefined, 2020.

[13] V. V. Shende, I. L. Markov, and S. S. Bullock. Minimal Universal Two-qubit Quantum Circuits. *Phys. Rev. A*, 69(6):062321, June 2004.

[14] N. Yu, R. Duan, and M. Ying. Five Two-Qubit Gates Are Necessary for Implementing Toffoli Gate. *Phys. Rev. A*, 88(1):010304, July 2013.

[15] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. Benjamin. Theory of variational quantum simulation. *Quantum*, 3:191, Oct. 2019.

[16] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley. Minimum construction of two-qubit quantum operations. *Phys. Rev. Lett.*, 93(2):020502, July 2004.

Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality

Daniel Ebler[1] * Marco Túlio Quintino[2] †

[1] *Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong*
[2] *Sorbonne Université, CNRS, LIP6, F-75005 Paris, France*

Abstract. We analyse the performance of quantum circuits and general processes to transform k uses of an arbitrary unitary operation U into another unitary operation $f(U)$. When the desired function f a homomorphism, i.e., $f(UV) = f(U)f(V)$, it is known that optimal average fidelity is attainable by parallel circuits and indefinite causality does not provide any advantage. Here we show that the situation changes dramatically when considering anti-homomorphisms, i.e., $f(UV) = f(V)f(U)$. In particular, we prove that when f is an anti-homomorphism, sequential circuits exponentially outperform parallel ones and processes with indefinite causal order could outperform sequential ones. We presented explicit constructions on how to obtain such advantages for the unitary inversion task $f(U) = U^{-1}$ and the unitary transposition task $f(U) = U^T$. We also stablish a one-to-one connection between three apparently different problems: unitary estimation, parallel unitary transposition, and parallel unitary inversion, allowing one to easily import results from one field to the other. Finally, we apply our results to several concrete problem instances and present a method based on computer-assisted proofs to show optimality.

This submission is mainly based in Ref. [1], Quantum 6, 679 (2022).

Keywords: Quantum higher-order operations, Quantum Circuits, Quantum Supermaps, Quantum Combs, Quantum Indefinite causality

1 Transforming unitary quantum operations

We now present the main task analysed in this paper. Let $f : \mathcal{SU}(d) \to \mathcal{SU}(d')$ be a function which transforms unitary operators to unitary operators and $\mathcal{SU}(d)$ is the group of unitary d-dimensional operators with determinant one (special unitary group of dimension d). We consider a scenario where one has access to k uses of an arbitrary d-dimensional unitary quantum operation described by an operator $U \in \mathcal{SU}(d)$. Our goal is to design a universal quantum circuit or a quantum process which approximates the transformation $U^{\otimes k} \mapsto f(U)$ for any $U \in \mathcal{SU}(d)$.

Main Problem 1 *Given a function $f : \mathcal{SU}(d) \to \mathcal{SU}(d')$, find the optimal parallel/sequential quantum circuit or general superchannel which approximates the transformation*

$$U^{\otimes k} \mapsto f(U), \quad \forall U \in \mathcal{SU}(d) \tag{1}$$

As a figure of merit, we quantify our approximations by means of average fidelity, which for several cases considered in this work, coincides with the worst-case fidelity and with the optimal white noise robustness. This work complements previous research which studied probabilistic but exact transformations between unitary quantum operations [2, 3], scenario which has similarities and differences when compared to the deterministic non-exact analysed here.

2 Parallel strategies are optimal for the homomorphic case

In Ref. [4] the authors show that when f is a homomorphism, every superchannel admits a parallel implementation without decreasing the average fidelity. In other words, sequential strategies and even general indefinite causal order strategies cannot outperform parallel ones.

3 Sequential strategies provide exponential advantage in the anti-homomorphic case

We now consider the anti-homomorphic case, where the function f respects $f(UV) = f(V)f(U)$, case which covers unitary inversion and unitary transposition. We start by showing an interesting one-to-one connection between parallel unitary inversion, parallel unitary transposition, and the problem of estimating unitary operations uniformly sampled in $\mathcal{SU}(d)$.

Theorem 1 *For any dimension d, the optimal average fidelity for parallel unitary inversion, parallel unitary transposition, and unitary estimation are equivalent and respect the upper bound*

$$\langle F \rangle_{par} \leq 1 - \frac{1}{(k+3)^2}. \tag{2}$$

For qubits, the optimal protocol attains

$$\langle F \rangle_{par}^{d=2} = \cos^2\left(\frac{\pi}{k+3}\right). \tag{3}$$

*Ebler.Dan@gmail.com
†Marco.Quintino@lip6.fr

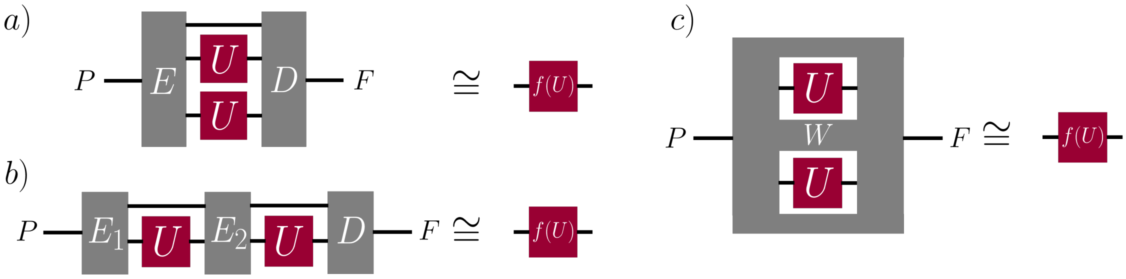


Figure 1: Three different strategies to transform $k = 2$ uses of an unknown unitary operation U into $f(U)$. $a)$ parallel circuit, E and D stand for fixed operations (encoder and decoder); $b)$ sequential circuit with multiple encoder operations; and $c)$ general processes acting on U may not have a definite causal order. Here, we analyse the performance of these strategies for different functions f.

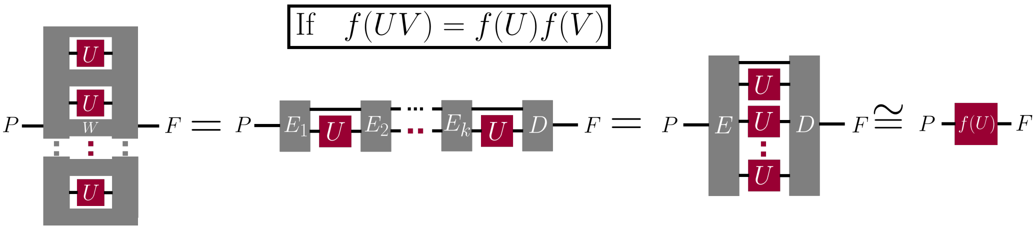


Figure 2: If the function f respects $f(UV) = f(U)f(V)$, every protocol that transform k copies of U into $f(U)$ can be made by a parallel circuit with the same average fidelity [4]. Examples of functions respecting this property are unitary complex conjugation $f(U) = U^*$ and unitary cloning $f(U) = U \otimes U$.

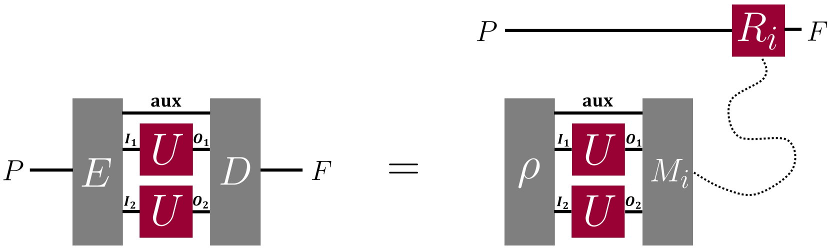
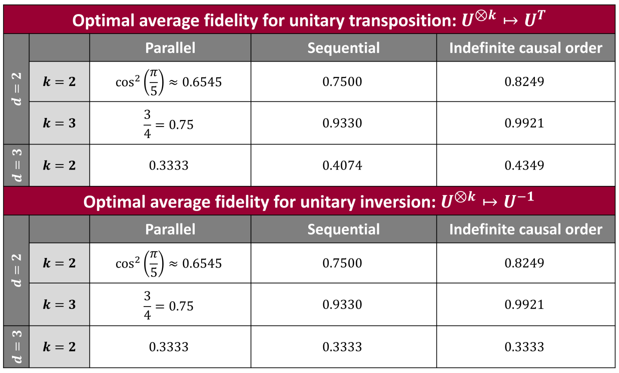


Figure 3: Every parallel strategy for unitary transposition and unitary inversion can be implemented by a prepare-and-measure strategy without changing its average fidelity performance. Additionally, the performance of parallel unitary transposition and parallel unitary inversion are both equivalent to the performance of estimating unitary quantum operations uniformly sampled in $\mathcal{SU}(d)$.



Optimal average fidelity for unitary transposition: $U^{\otimes k} \mapsto U^T$

		Parallel	Sequential	Indefinite causal order
$d = 2$	$k = 2$	$\cos^2\left(\frac{\pi}{5}\right) \approx 0.6545$	0.7500	0.8249
	$k = 3$	$\frac{3}{4} = 0.75$	0.9330	0.9921
$d = 3$	$k = 2$	0.3333	0.4074	0.4349

Optimal average fidelity for unitary inversion: $U^{\otimes k} \mapsto U^{-1}$

		Parallel	Sequential	Indefinite causal order
$d = 2$	$k = 2$	$\cos^2\left(\frac{\pi}{5}\right) \approx 0.6545$	0.7500	0.8249
	$k = 3$	$\frac{3}{4} = 0.75$	0.9330	0.9921
$d = 3$	$k = 2$	0.3333	0.3333	0.3333

Figure 4: Optimal average fidelity for deterministic protocols transforming k uses of U into its transpose and into its inverse.The values for $k = 2$ were obtained via numerical SDP optimisation and rigorously certified up to the fourth decimal digit with a computer assisted proof.

We now show that there exists a sequential quantum circuit in which the average fidelity approaches one exponentially in the number of uses k.

Theorem 2 *There exists an explicit sequential quantum circuit such that unitary inversion can be implemented with probability*

$$\langle F \rangle_{seq}^{inv} \geq 1 - \left(1 - \frac{1}{d^2}\right)^{\lfloor \frac{k+1}{d} \rfloor}. \tag{4}$$

There exists an explicit sequential quantum circuit such that unitary transposition can be implemented with probability

$$\langle F \rangle_{seq}^{trans} \geq 1 - \left(1 - \frac{1}{d^2}\right)^{\lceil \frac{k}{d} \rceil}. \tag{5}$$

4 Computational results and the advantage of indefinite causality

We also present a computer assisted proof methods based on semidefinite programming which can be used to analyse parallel, sequential, and the general scenario. With this method, we can find explicit upper and lower bounds on the optimal averaged fidelity and show that indefinite order processes [5, 6] outperform sequential strategies in some scenarios.

References

[1] M. T. Quintino and D. Ebler, Deterministic transformations between unitary operations: Exponential advantage with adaptive quantum circuits and the power of indefinite causality, Quantum **6**, 679 (2022), arXiv:2109.08202 [quant-ph].

[2] M. T. Quintino, Q. Dong, A. Shimbo, A. Soeda, and M. Murao, Probabilistic exact universal quantum circuits for transforming unitary operations, Phys. Rev. A **100**, 062339 (2019), arXiv:1909.01366 [quant-ph].

[3] 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), arXiv:1810.06944 [quant-ph].

[4] A. Bisio, G. M. D'Ariano, P. Perinotti, and M. Sedlák, Optimal processing of reversible quantum channels, Physics Letters A **378**, 1797–1808 (2014), arXiv:1308.3254 [quant-ph].

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

[6] O. Oreshkov, F. Costa, and Č. Brukner, Quantum correlations with no causal order, Nature Communications **3**, 1092 (2012), arXiv:1105.4464 [quant-ph].

The National Quantum-Safe Network in Singapore

Jing Yan Haw[1] * Hao Qin[1] † Xiao Duan[3] Yu Cai[5] Ramana Murthy[1]
Nelly Ng[5] Biplab Sikdar[4] Christian Kurtsiefer[1 2] Michael Kasper[3]
Alexander Ling[1 2 ‡]

[1] *Centre for Quantum Technologies, National University of Singapore*
[2] *Department of Physics, National University of Singapore*
[3] *Fraunhofer Singapore Research Centre@NTU, Nanyang Technological University*
[4] *Department of Electrical and Computer Engineering, National University of Singapore*
[5] *School of Physical and Mathematical Sciences, Nanyang Technological University*

Abstract. We introduce the National Quantum-Safe Network, which is a nationwide collaborative field-deployed testbed aimed at demonstrating quantum-safe cryptography solutions. With several key aspects including testbed, security evaluation, standardisation, and ecosystem building, the network aims to achieve a vendor-neutral, multi-protocol platform that complies with international standards.

1 Introduction

As the quantum technology landscape evolves, it is difficult to predict when powerful quantum computers capable of breaking current cryptography will be available. Thus, applications and communication infrastructure handling high-value assets or requiring long-term protection needs to be equipped with quantum-safe security enhancements as soon as possible. Quantum Key Distribution (QKD), a tamper-evident secure communication technique based on quantum physics, whose security is independent of computation power, could potentially fulfil such requirement. In recent years, various QKD networks have been deployed worldwide [1, 2, 3]. The National Quantum-Safe Network (NQSN) in Singapore is a nationwide collaborative platform and a field-deployed testbed aimed at demonstrating quantum-safe cryptography solutions. The NQSN testbed, which links up academic, public and private members, targets trials for QKD network, augmented with the Post-Quantum Cryptography (PQC) technologies. PQC refers to mathematics-based cryptographic algorithms (software) which are believed to be secure against known attacks from quantum computers.

As shown in Figure 1, NQSN is a star type QKD network with the central node connected to the remote nodes across the island from east to west, and consists of four logical layers [4]: quantum layer, key management (KM) layer, network management layer and application layer (Figure 2). Beyond the terrestrial metropolitan area network setting, a satellite-based QKD is planned to be launched to serve as a moving node in the future phase. NQSN can be further linked up with other global QKD networks via the satellite node.

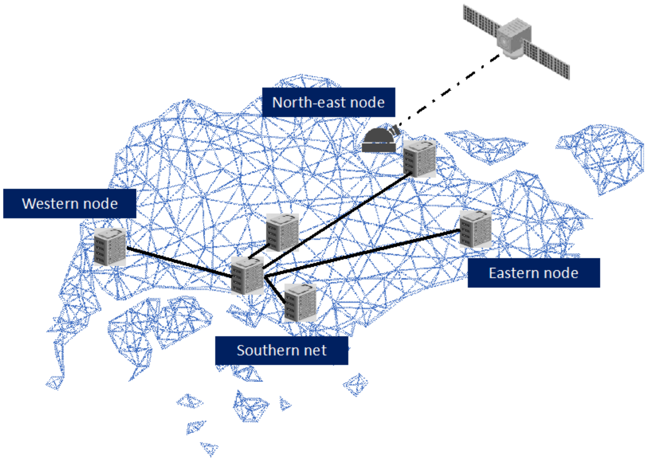


Figure 1: NQSN star-type QKD network.

2 Quantum Layer

In the quantum layer, commercial ready and production grade QKD devices will be deployed to connect the central node with each remote nodes, via the existing production grade fibre infrastructure. Different QKD protocols implementations from various QKD vendors will be considered with regard to the distances and losses featured by each point-to-point fibre link. The candidate protocols include BB84, coherent one way (COW), continuous-variable (CV) and entanglement-based (EB) QKD protocols. Under such operation, each pair of the QKD devices continuously outputs QKD keys to the KM layer, which are established over the quantum channel (fibre) and the classical channel (fibre or Ethernet). In this layer, it aims to achieve one of the main goals of the NQSN testbed - to serve as a vendor-neutral platform that demonstrates multi-QKD protocols, supporting different quantum channel conditions.

*jy.haw@nus.edu.sg
†hao.qin@nus.edu.sg
‡alexander.ling@nus.edu.sg

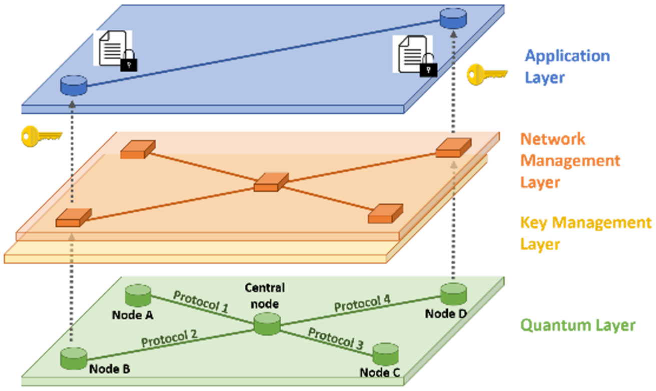


Figure 2: Different layers of NQSN. From Bottom to Top: Quantum Layer, Key Management Layer, Network Management Layer, Application Layer

3 Key Management Layer

In the KM layer, a high performance, customized, centralized key manager system will be installed in the central node and enable interoperability, connectivity, scalability of the QKD network. Remote key managers will also be paired with QKD devices in the remote nodes, connecting with the central node via KM links. With the star type configuration, the centralized key manager features a multi-input and multi-output interfaces. The input interface is to receive QKD keys provided by different QKD devices in the quantum layer [5]. The output interface is to supply keys to applications upon request by the application layer and network management layer (e.g. ETSI GS QKD 014 [6]). These key managers also process and store the received QKD keys into the formats that are required by specific applications. The key relaying and routing functions in the central key manager further enables symmetric key establishments between any of the two nodes in the network via KM links. Integration of PQC solutions such as hybrid key combination of QKD key and PQC exchanged key is among the considered architecture in the KM layer.

4 Network Management Layer

The network management layer is responsible for controlling and managing network resources across different nodes of the NQSN testbed network. A centralized network management server is in charge of the controlling and managing functions, which will be installed in the central node and hold a global view of the entire QKD network operation. The server gives controlling instructions to the quantum layer and KM layer to create the key delivery path across network nodes and to configure components such as switches, servers and QKD devices. The managing functions consist of monitoring and collecting performance parameters, detecting and reporting any fault events, collecting event logs for networking analysis from the quantum layer and the KM layer.

5 Application Layer

The application layer in the NQSN QKD network acts as an open platform that allows for the integration deployment of different applications at various layers of the Open Systems Interconnection (OSI) model. Some examples include physical encryption in the physical layer, link encryptor in the data link layer and IPSec in the network layer. These applications consume keys provided by the KM layer, via supported interfaces such as ETSI GS QKD 014 [6]. Different reference use cases and trials are explored for field trials, interoperability, and performance evaluation of quantum-safe technologies.

6 Quantum Security Lab

Along with the NQSN testbed, a testing lab dedicated to testing, evaluation and certification of QKD devices and their supporting units is established. The main objective of the lab is to verify the functionality and the security of the QKD network, and formalize certification framework towards industry applications of QKD technologies. Main activities include (i) research on quantum hacking and countermeasures, (ii) development of functional, performance and security evaluation methodologies, (iii) testing tools developments and building up certification capabilities with industry and academic labs. Lab facilities are also co-located within some of the nodes, which open possibilities for novel remote testing and evaluation methods on the quantum layer and other layers under a network configuration.

7 Standardisation

A quantum communication networks task force under local regulatory authority is also formed, with members from governmental agencies and industry partners. The main objective of the task force is to develop local standards to facilitate the deployment, operation, and adoption of the QKD technologies in different domains. The standardisation developments will also support certifications of QKD devices and other entities in the QKD network.

8 Reference Use Cases

One example reference use case is a demonstration of direct data centre interconnect (DCI) secured by QKD-keys, as shown in Fig. 3. In this demonstration, which was performed over two physically separated commercial data centres of ST Telemedia Global Data Centres (STT-GDC), we confirmed the feasibility of operating QKD systems (ID Quantique Cerberis XGR) over a production-grade fibre network (Netlink Trust). During the field test of the QKD devices, the secret key rate and the quantum bit error rate (QBER) are relatively stable and continuous over a fibre link of around 20 km. As shown in Figure 3, an average secret key rate of 2.39 kbps and QBER of 1.90% is achieved. A total of more than 2 Gbits of AES-256 keys are accumulated, with the rates of around 690 keys per minutes. A subset of the

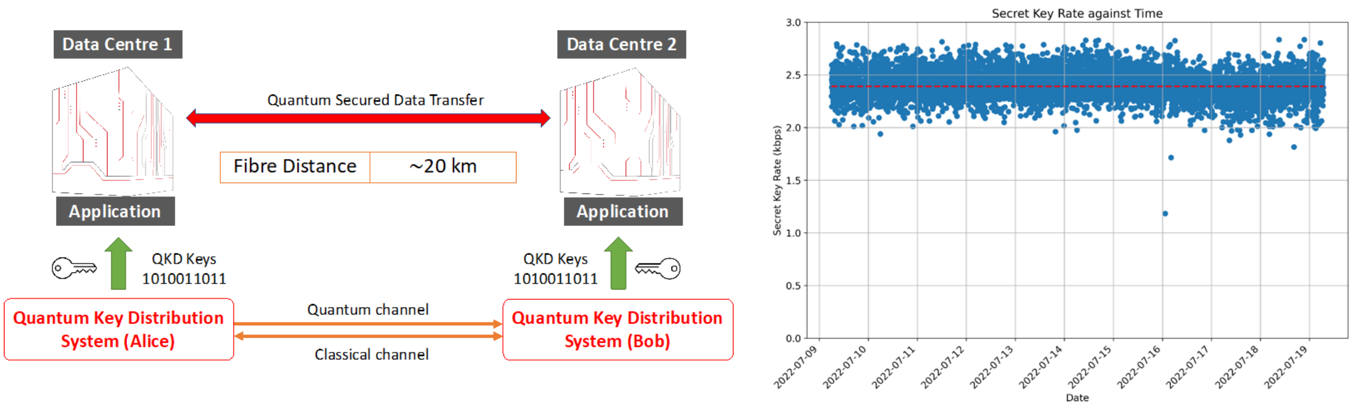


Figure 3: (Left) A quantum secured data transfer utilizing pairs of symmetric keys generated from QKD devices. (Right) The secret key rate over the trial period.

keys is used by a software-based Quantum Virtual Private Network (Q-VPN), which consumed the QKD symmetric keys to establish a VPN tunnel using QKD keys for quantum-secured file transfer.

9 Conclusion

The Singapore's NQSN focuses on developing a network that addresses several scopes, including multiple vendors, standards compliance and infrastructure requirements, and various implementations of QKD protocols. In particular, the NQSN seeks to significantly improve quantum-safe technologies by enhancing testing and assessment capabilities, promoting the broader availability of quantum-safe technology, and increasing awareness about it. End-users and stakeholders can take advantage of integrating quantum-safe security applications and solutions, catalysing future innovations and quantum-related products and services on a regional and global level.

10 Acknowledgements

This research is supported by the National Research Foundation, Singapore and A*STAR under its Quantum Engineering Programme (NRF2021-QEP2-04-P01). We acknowledge NQSN partners (www.nqsn.sg) for their support. We acknowledge ST Telemedia Global Data Centre for providing secured physical locations to host the QKD trial, Netlink Trust for the provisioning of the fibre network and ID Quantique SA for the loan of the QKD system. We thank KaiWei Qiu, Albert Huang & Charles Ci-Wen Lim for their participation in the reference use case.

References

[1] C. Elliott, et al., "Current status of the DARPA quantum network" In *Quantum Information and computation III, vol.5815, International Society for Optics and Photonics*, pages 138–149, 2005.

[2] M. Peev et al., "The SECOQC quantum key distribution network in Vienna" *New Journal of Physics*, vol. 11, no. 7, p. 075001, 2009.

[3] M. Sasaki et al., "Field test of quantum key distribution in the Tokyo QKD network" *Optics Express*, vol. 19, no. 11, pp. 10 387 – 10 409, 2011.

[4] ITU-T Recommendation Y.3800 (2019), plus Corrigendum 1 (2020), Overview on networks supporting quantum key distribution, 2020.

[5] ITU-T Recommendation Y.3803, Quantum key distribution networks – Key management.

[6] ETSI GS QKD 014, Quantum key distribution; Protocol and data format of key delivery API to Applications, 2019.

Quantum computation on a 19-qubit wide 2d nearest neighbour qubit array.

Alexis Shaw[1 2 *] Michael J. Bremner[1 2] Alexandru Paler[3] Daniel Herr[4]
Simon J. Devitt[2]

[1] *Centre for Quantum Computation and Communication Technology (CQC2T)*
[2] *Centre for Quantum Software and Information, School of Computer Science,*
Faculty of Engineering & Information Technology, University of Technology Sydney, NSW 2007, Australia
[3] *Aalto University, 02150 Espoo, Finland*
[4] *d-fine GmbH, An der Hauptwache 7, 60213, Frankfurt, Germany.*

Abstract. In this paper, we explore the relationship between the width of a qubit lattice constrained in one dimension and physical thresholds for scalable, fault-tolerant quantum computation. To circumvent the traditionally low thresholds of small fixed-width arrays, we deliberately engineer an error bias at the lowest level of encoding using the surface code. We then address this engineered bias at a higher level of encoding using a lattice-surgery surface code bus that exploits this bias, or a repetition code to make logical qubits with unbiased errors out of biased surface code qubits. Arbitrarily low error rates can then be reached by further concatenating with other codes, such as Steane $[\![7,1,3]\!]$ code and the $[\![15,7,3]\!]$ CSS code. This enables a scalable fixed-width quantum computing architecture on a square qubit lattice that is only 19 qubits wide, given physical qubits with an error rate of 8.0×10^{-4}. This potentially eases engineering issues in systems with fine qubit pitches, such as quantum dots in silicon or gallium arsenide.

Keywords: Quantum Error Correction, Fault Tolerance, Surface Code, Lattice Surgery

Quantum processor architectures have evolved significantly since their first conception, just over a quarter of a century ago. They started with early discussions on how to build basic gates with two-level systems and have evolved into recent plans for machines with millions of physical qubits. Architecture design must work around the peculiarities of their constituent qubits. Physical limitations on qubit size, gate speed, decoherence rates, temperature, control wiring, and infrastructure all have a major effect on the potential scalability of a given architecture.

Quantum dot qubits in silicon and gallium arsenide (GaAs) present interesting constraints when investigating quantum architecture. They have demonstrated relatively low decoherence rates [14, 16], high operation temperature [15], and high gate speeds [5]. Significantly, they offer the potential of high qubit density and ease of manufacture for large systems on a single wafer [12, 16, 1]. Yet, the size and small qubit pitch that could allow for high qubit density come with significant drawbacks.

Running control wires into gates that have qubit spacings on the order of nanometers in a two-dimensional geometry is extremely challenging. This is especially important because architectures using the highest threshold quantum error correcting codes, such as the surface or honeycomb codes, tend to assume two-dimensional lattices of unrestricted size. In response, quantum architectures that utilize three-dimensional fabrication have been proposed [13], even though the expected fabrication complexity is formidable. One approach to solving this is to restrict the width of the array, which limits the interconnect density because the system grows only in one dimension.

The idea of minimizing the width of a qubit array is certainly not new. Since the early days of fault tolerance, people have considered fixed or minimal-width arrays of qubits. However, previous approaches have always required undesirable tradeoffs, either increasing qubit requirements or requiring long-distance qubit interactions. For example, CSS codes were leveraged in the paper by Veldhorst et. al. [11], and the subsequent threshold was extremely small (less than 10^{-5}) due to the costs involved in interacting non-adjacent qubits in a nearest-neighbour array. Other examples include specific investigations into linear nearest neighbours in silicon [7] with a threshold of about 10^{-4} and the use of resonators to fold a square lattice into a bi-linear array [8]. Unfortunately, the latter approach required long resonators for interaction between qubits, and these resonators had unknown manufacturability and performance in spin systems given their length and complexity.

In this paper, we present a method of reducing the array width without requiring a lower error threshold or long-distance qubit interactions. Instead, we propose a coding structure that provides good reductions of width for realistic error rates, with an array width as low as 19 qubits for a physical error rate of 8×10^{-4}. It locally has the same nearest neighbour interactivity and the same threshold error rate as the surface code. Of course, this comes at the cost of significantly more qubits; however, this may be an acceptable trade-off for certain technologies, such as silicon quantum dots, where qubits are hoped to be relatively cheap.

[*] alexis@alexisshaw.com

Technical Contributions

The primary technical contribution of this work is to circumvent the traditionally low thresholds of small fixed-width arrays by engineering, not exploiting, an error bias at the lowest level of encoding using the surface code. We then address this engineered bias at a higher level of encoding using codes explicitly tailored for highly-biased noise. This allows us to reduce the error enough that another code can be concatenated over the top of this error correction scheme.

We consider two broadly different approaches, illustrated in Figure 1. In the first, we describe and analyze a fault-tolerant lattice surgery-based data bus inspired by the scheme of Herr et. al. [6] to implement higher-level codes above the surface code. key to implementing this bus with high efficiency is the engineered noise bias. In the second, we construct our higher-level codes on encoded data qubits with an engineered bias. We develop the required data and theory to describe these schemes and evaluate the relative trade-offs between array width and qubit density given a fixed target logical error rate and varying physical error rates.

The data-bus lattice surgery scheme measures a logical parity in the surface code using narrow surface code patches. We show the correctness of this approach, and evaluate its performance, including the minimum width and time required to measure that parity and the interaction between this and the error rates of square surface code patches of varying code distances.

We evaluate the performance of the the data bus scheme when concatenated with the Steane $[\![7,1,3]\!]$ and the $[\![15,7,3]\!]$ CSS codes when using the flag qubit compilation approach of Reichardt [3, 2, 10, 4, 9]. To do this, we design a table-based fault-tolerant decoder for different flagged qubit extraction circuits, simulate the behaviour of these codes with varying physical error rates, and then extract an equation of fit for each code. We also extract a pseudo threshold of $p \approx 4.52 \times 10^{-5}$ for the Steane $[\![7,1,3]\!]$ code and $p \approx 1.25 \times 10^{-6}$ for the $[\![15,7,3]\!]$ CSS code.

We then describe a complete compilation approach to concatenate multiple levels of both the Steane $[\![7,1,3]\!]$ and $[\![15,7,3]\!]$ CSS codes on top of the surface code when using the surface code bus to implement two-qubit interactions. We evaluate several approaches to compile these multiply concatenated circuits and choose the optimal approach for each number of levels of concatenation. We then calculate the relationship between the performance of the concatenated codes and the number of levels of concatenation for a target logical error rate of 10^{-15}, which was chosen as the approximate rate required to factor 2048bit numbers using Shor's algorithm. An excerpt of this performance evaluation is summarised in Table 1.

For the second scheme, we examine the possibility of using the rectangular patches to engineer a logical error bias in the data qubit patches. These are concatenated with the repetition code to create a memory with unbiassed errors and high enough logical fidelity that a nearest-neighbour implementation of a CSS code can be directly concatenated above it. A scheme is designed to measure the syndromes of two interlaced repetition codes that sit above the rectangular patch qubits assuming two rows of patches. This is used, along with our simulations on the performance of rectangular qubit patches, to evaluate the performance of this concatenated surface code, giving error rates and qubit counts for differing widths and lengths of rectangular surface code patches. We then determine the minimum widths required to exceed the pseudo-threshold for the Steane $[\![7,1,3]\!]$ code implemented on a logical nearest neighbour architecture that uses swap networks for two-qubit interactions and show that an array 27 qubits wide is sufficient for a physical error rate of 1.2×10^{-3}, an array 19 qubits wide is sufficient for an error rate of 8×10^{-4}, and an array 11 qubits wide is sufficient for an error rate of 2.5×10^{-4}.

Technical Version

The complete technical version of this paper can be found at https://arxiv.org/abs/2212.01550

References

[1] Ansaloni, F., Chatterjee, A., Bohuslavskyi, H., Bertrand, B., Hutin, L., Vinet, M., and Kuemmeth, F. Single-electron operations in a foundry-fabricated array of quantum dots. *Nature Communications 11*, 1 (Dec 2020), 6399.

[2] Chao, R., and Reichardt, B. W. Fault-tolerant quantum computation with few qubits. *npj Quantum Information 4*, 1 (Sep 2018), 42.

[3] Chao, R., and Reichardt, B. W. Quantum error correction with only two extra qubits. *Phys. Rev. Lett. 121* (Aug 2018), 050502.

[4] Chao, R., and Reichardt, B. W. Flag fault-tolerant error correction for any stabilizer code. *PRX Quantum 1* (Sep 2020), 010302.

[5] He, Y., Gorman, S. K., Keith, D., Kranz, L., Keizer, J. G., and Simmons, M. Y. A two-qubit gate between phosphorus donor electrons in silicon. *Nature 571*, 7765 (Jul 2019), 371–375.

[6] Herr, D., Paler, A., Devitt, S. J., and Nori, F. Time versus hardware: Reducing qubit counts with a (surface code) data bus, 2019. arxiv:1902.08117.

[7] Jones, C., Fogarty, M. A., Morello, A., Gyure, M. F., Dzurak, A. S., and Ladd, T. D. Logical qubit in a linear array of semiconductor quantum dots. *Phys. Rev. X 8* (Jun 2018), 021058.

[8] Mohiyaddin, F., Li, R., Brebels, S., Simion, G., Dumoulin Stuyck, N. I., Godfrin, C., Shehata, M., Elsayed, A., Gys, B., Kubicek, S., Jussot, J., Canvel, Y., Massar, S., Weckx, P., Matagne, P., Mongillo, M., Govoreanu,

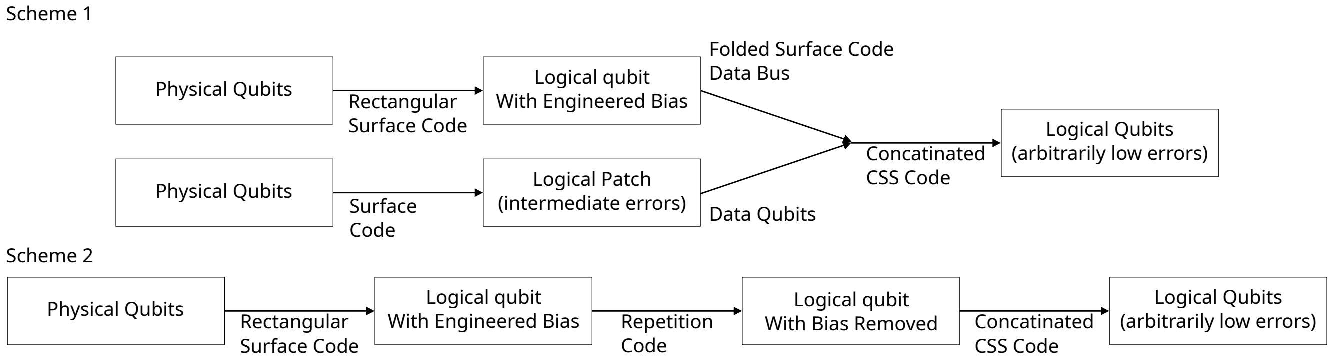

Scheme 1
Physical Qubits
Rectangular Surface Code
Logical qubit With Engineered Bias
Folded Surface Code Data Bus
Physical Qubits
Surface Code
Logical Patch (intermediate errors)
Data Qubits
Concatinated CSS Code
Logical Qubits (arbitrarily low errors)
Scheme 2
Physical Qubits
Rectangular Surface Code
Logical qubit With Engineered Bias
Repetition Code
Logical qubit With Bias Removed
Concatinated CSS Code
Logical Qubits (arbitrarily low errors)

Figure 1: Diagram describing the two schemes examine for reducing bus width

	d_{sc}	d_b	h	p_l	block size	qubit density
Surface Code - No Bus	27	NA	107	5.49×10^{-16}	5778	5778
Surface Code - Bus	31	5	71	1.66×10^{-16}	5112	5112
1L $[\![7,1,3]\!]$ on SC + Bus	21	7	55	3.96×10^{-17}	3.39×10^{4}	3.39×10^{4}
7L $[\![7,1,3]\!]$ on SC + Bus	11	7	35	8.27×10^{-21}	7.37×10^{9}	7.37×10^{9}
1L $[\![15,7,3]\!]$ on SC + Bus	23	7	59	2.74×10^{-17}	7.08×10^{4}	1.01×10^{4}
4L $[\![15,7,3]\!]$ on SC + Bus	17	7	47	3.84×10^{-17}	2.63×10^{8}	1.10×10^{5}

(a) $p = 1.0 \times 10^{-3}$

	d_{sc}	d_b	h	p_l	block size	qubit density
Surface Code - No Bus	15	NA	59	2.59×10^{-17}	1770	1770
Surface Code - Bus	15	3	35	4.93×10^{-16}	1260	1260
1L $[\![7,1,3]\!]$ on SC + Bus	11	3	27	2.46×10^{-18}	8.32×10^{3}	8.32×10^{3}
9L $[\![7,1,3]\!]$ on SC + Bus	5	3	19	5.10×10^{-22}	1.14×10^{11}	1.14×10^{11}
1L $[\![15,7,3]\!]$ on SC + Bus	11	3	27	1.67×10^{-16}	1.51×10^{4}	2.16×10^{3}
3L $[\![15,7,3]\!]$ on SC + Bus	9	3	23	1.44×10^{-20}	3.58×10^{5}	1.04×10^{4}

(b) $p = 1.0 \times 10^{-4}$

Table 1: Error rates for different code concatenation configurations, parameters, and physical error rates

B., and Radu, I. P. Large-scale 2d spin-based quantum processor with a bi-linear architecture. In *2021 IEEE International Electron Devices Meeting (IEDM)* (2021), pp. 27.5.1–27.5.4.

[9] Prabhu, P., and Reichardt, B. W. Fault-Tolerant Syndrome Extraction and Cat State Preparation with Fewer Qubits. In *16th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2021)* (Dagstuhl, Germany, 2021), M.-H. Hsieh, Ed., vol. 197 of *Leibniz International Proceedings in Informatics (LIPIcs)*, Schloss Dagstuhl – Leibniz-Zentrum für Informatik, pp. 5:1–5:13.

[10] Reichardt, B. W. Fault-tolerant quantum error correction for steane's seven-qubit color code with few or no extra qubits. *Quantum Science and Technology 6*, 1 (dec 2020), 015007.

[11] Stephens, A. M., and Evans, Z. W. E. Accuracy threshold for concatenated error detection in one dimension. *Phys. Rev. A 80* (Aug 2009), 022313.

[12] Vandersypen, L. M. K., Bluhm, H., Clarke, J. S., Dzurak, A. S., Ishihara, R., Morello, A., Reilly, D. J., Schreiber, L. R., and Veldhorst, M. Interfacing spin qubits in quantum dots and donors—hot, dense, and coherent. *npj Quantum Information 3*, 1 (Sep 2017), 34.

[13] Veldhorst, M., Eenink, H. G. J., Yang, C. H., and Dzurak, A. S. Silicon cmos architecture for a spin-based quantum computer. *Nature Communications 8*, 1 (Dec 2017), 1766.

[14] Xue, X., Russ, M., Samkharadze, N., Undseth, B., Sammak, A., Scappucci, G., and Vandersypen, L. M. K. Quantum logic with spin qubits crossing the surface code threshold. *Nature 601*, 7893 (Jan 2022), 343–347.

[15] Yang, C. H., Leon, R. C. C., Hwang, J. C. C., Saraiva, A., Tanttu, T., Huang, W., Camirand Lemyre, J., Chan, K. W., Tan, K. Y., Hudson, F. E., Itoh, K. M., Morello, A., Pioro-Ladrière, M., Laucht, A., and Dzurak, A. S. Operation of a silicon quantum processor unit cell above one kelvin. *Nature 580*, 7803 (Apr 2020), 350–354.

[16] Zwerver, A. M. J., Krähenmann, T., Watson, T. F., Lampert, L., George, H. C., Pillarisetty, R., Bojarski, S. A., Amin, P., Amitonov, S. V., Boter, J. M., Caudillo, R., Correas-Serrano, D., Dehollain, J. P., Droulers, G., Henry, E. M., Kotlyar, R., Lodari, M., Lüthi, F., Michalak, D. J., Mueller, B. K., Neyens, S., Roberts, J., Samkharadze, N., Zheng, G., Zietz, O. K., Scappucci, G., Veldhorst, M., Vandersypen, L. M. K., and Clarke, J. S. Qubits made by advanced semiconductor manufacturing. *Nature Electronics 5*, 3 (Mar 2022), 184–190.

On circuit complexity of quantum access models for encoding classical data

Xiao-Ming Zhang[1] *

[1] *Center on Frontiers of Computing Studies, Department of Computer Science, Peking University, Beijing, China*

Abstract. In quantum computing, data encoding is usually treated as an oracle, but its implementation is crucial in practice. We open the black-boxes of some typical access models. For sparse matrix encoding, the circuit complexity lower bound is polynomial with respect to matrix dimension, and a family of nearly-optimal construction is proposed. The circuit complexity is reduced exponentially if the matrix is the linear combination polynomial terms of efficient unitaries. All of our protocols are highly flexible, enabling trading circuit depth to ancillary qubits. Our methods are based on our improved state preparation and other operations that are of independent interest.

Keywords: block-encoding, sparse access models, circuit complexity optimization, state preparation

1 Introduction

The power of quantum computing is commonly studied in the framework of oracle based computation [1]. More specifically, a function $f(x)$ representing the classical data of interest is encoded by unitary U_f, which serves as an oracle during computation. To have indications to quantum advantages, the number of queries to U_f of a quantum algorithm is then compared to the number of queries to its classical counterpart of classical algorithms.

Treating access models as black-boxes is convenient, but the actual circuit complexity of an algorithm depends on the cost of each query. We open the black-boxes of some commonly used quantum access models and study how they can be constructed with Clifford+T gates. For a general $2^n = N$ dimensional sparse matrices, we consider its sparse-access input model (SAIM) [2–8, 12] and block-encoding [14, 15]. We provide both upper bound and lower bound of the circuit complexity. In particular, we show that for both access models, the gate count lower bound increase almost linearly with the matrix dimension, even under the sparsity assumption. We develop a family of construction algorithms with tunable qubit number ranging from $\Omega(n)$ to $O(N)$. In the full range of qubit number, we achieve nearly optimal circuit complexity. We then show that when the matrix can be represented by the linear combination of polynomial number of unitaries that can be implemented with polynomial-size quantum circuit, the block-encoding can be realized efficiently.

Our access model construction is based on the optimized realizations of some subroutines that are of independent interest, including quantum state preparation, select oracle for Pauli strings, sparse Boolean functions. For all operations listed, we obtained improved or at least comparable circuit complexities compared to the best-known realizations, and the protocols allow tunable qubit numbers.

2 Sparse-access input model

Let $N = 2^n$, we consider a sparse matrix $H \in \mathbb{C}^{N\times N}$ with at most $s = O(1)$ nonzero elements at each row and column. Let $H_{x,y}$ be the value of the elements at the xth row and yth column. Each $H_{x,y}$ is a d-digit integer ($d = O(1)$). We consider two unitaries O_H, O_F, which satisfies

$$O_H|x,y\rangle_{\mathrm{idx}}|z\rangle_{\mathrm{wrd}} = |x,y\rangle_{\mathrm{idx}}|z \oplus H_{x,y}\rangle_{\mathrm{wrd}}, \tag{1a}$$

$$O_F|x,k\rangle_{\mathrm{idx}} = |x, F(x,k)\rangle_{\mathrm{idx}}, \tag{1b}$$

where $F(x,k)$ is the column index of the kth nonzero element in row x. Due to its simplicity and generality, Eq. (1) becomes one of the standard access models in quantum computing, which is usually assumed to be available in processing classical data.

2.1 circuit complexity lower bound

Before we discuss the access model construction, we first study the lower bound of the circuit complexity. We first analyze the *capacity* of a quantum circuit with bounded resource, i.e. the number of unique unitaries that can be constructed with fixed number of elementary gates (or circuit depth and qubit number). Secondly, we analyze the *size* of the access model, i.e. the number of unique unitaries required to approximate the access model with arbitrary parameters. The circuit complexity can then be estimated by comparing the capacity of a quantum circuit and the size of the access model.

The capacity of quantum circuits without ancilla has been studied in Section 4.5.4 of [17]. We generalize the discussion to the scenarios allowing ancillary qubits. We assume that a finite two-qubit elementary gate set $\mathcal{G}_{\mathrm{ele}}$ is given, with $g \equiv |\mathcal{G}_{\mathrm{ele}}| = O(1)$. We require that all ancillary qubits are uncomputed at the end of the circuits, so the unitaries should be in the form of $U_{\mathrm{dat}} \otimes I_{\mathrm{anc}}$, where dat represents the data subspace, and anc represents the subspace containing all ancillary qubits. We have the following result.

Lemma 1 *Let $\mathcal{G}_C$ be the set containing all unitaries in the form of $U_{dat} \otimes I_{anc}$ that can be constructed with C*

*xmzhang93@pku.edu.cn

elementary gates in $\mathcal{G}_{ele}$ and unlimited ancillary qubits. Then, we have $\log|\mathcal{G}_C| = O\left((C\log(C+n))\right)$.

Lemma 2 *Let $\mathcal{G}'_{n_{anc},D}$ be the set containing all unitaries in the form of $U_{dat}\otimes I_{anc}$ that can be constructed with n_{anc} ancillary qubits and D circuit depth. Then, we have* $\log\left|\mathcal{G}'_{n_{anc},D}\right| = O\left(D(n+n_{anc})\right)$.

Lemma. 1, 2 are general and useful for not only the access model studied in this work. Based on Lemma. 1, 2, we have

Theorem 3 *Given an arbitrary finite two-qubit elementary gate set $\mathcal{G}_{ele}$. Let n_{anc}, D and C be the number of ancillary qubits, circuit depth and total number of gates in $\mathcal{G}_{ele}$ required to approximate the SAIM for an arbitrary sparse H to accuracy $\varepsilon < 1$. Then, we have $(n+n_{anc})D = \Omega(2^n n)$ and $C = \Omega(2^n)$.*

Theorem. 3 implies that a general SAIM can not be constructed with subexponential number of quantum gates. It is possible to trade ancillary qubit number for the circuit depth. However, the space and time complexities can not achieve sub-exponential scaling simultaneously.

2.2 Construction of SAIM

From the lower bound results above, there is little hope to have exponential quantum advantage when using SAIM. However, studying the construction of SAIM still has its great value. First of all, SAIM is represents the most general for of sparse classical data. It is in fact very rare to have structured classical data that can be encoded exponentially faster. Second, polynomial quantum speedup with respect to the matrix dimension N is still expected.

The construction of O_H is straightforward and we refer the readers to the technical version.

The construction of O_F can be realized in three steps. We introduce an n-qubit ancillary register (denoted as anc). We first perform the transformation $|x,k\rangle_{\rm idx}|0\rangle_{\rm anc} \rightarrow |x,k\rangle_{\rm idx}|F(x,k)\rangle_{\rm anc}$. Then, we apply swap gates between the ancillary register and half of the index register which encodes k, i.e. $|x,k\rangle_{\rm idx}|F(x,k)\rangle_{\rm anc} \rightarrow |x,F(x,k)\rangle_{\rm idx}|k\rangle_{\rm anc}$. Finally, we perform the transformation $|x,F(x,k)\rangle_{\rm idx}|k\rangle_{\rm anc} \rightarrow |x,F(x,k)\rangle_{\rm idx}|0\rangle_{\rm anc}$. The total circuit complexity is as follows.

Theorem 4 *Given n_{anc} ancillary qubits where $\Omega(n) \leqslant n_{anc} \leqslant O(Nnds)$, O_H can be constructed with $O(Nnds)$ count and $O\left(Nnds\frac{\log n_{anc}}{n_{anc}}\right)$ depth of Clifford+T gates.*

Given n_{anc} ancillary qubits where $\Omega(n) \leqslant n_{anc} \leqslant O(Nns\log s)$, O_F can be constructed with $O(Nns\log s)$ count and $O\left(Nns\log s\frac{\log n_{anc}}{n_{anc}}\right)$ depth of Clifford+T gates.

Compared to the circuit complexity lower bound obtained in Theorem. 3, our protocol has nearly optimal circuit complexities with respect to the matrix dimension up to a factor of n.

3 Block encodings

Given an N-dimensional matrix $H \in \mathbb{C}^{N\times N}$, we call unitary U a block encoding of H if $\alpha\left(\langle 0^{n_{\rm anc}}|\otimes\mathbb{I}_N\right)U\left(|0^{n_{\rm anc}}\rangle\otimes\mathbb{I}_N\right) = H$ for some $\alpha > 0$, and $\mathbb{I}_N$ is the N-dimensional identity. In practice, we may consider approximated construction of the block encoding. More specifically, we call unitary $\tilde{U}$ an $(\alpha, n_{\rm anc}, \varepsilon)-$block-encoding of H if

$$\left\|H - \alpha\left(\langle 0^{n_{\rm anc}}|\otimes\mathbb{I}_N\right)\tilde{U}\left(|0^{n_{\rm anc}}\rangle\otimes\mathbb{I}_N\right)\right\| \leqslant \varepsilon \qquad (2)$$

for some normalization factor $\alpha > 0$. Throughout our manuscript, $\|\cdot\|$ represents the Frobenius norm, i.e. Schatten 2-norm of either matrices or vectors. For a general H, the construction of its block-encoding requires $\Omega(\mathrm{Poly}(N))$ gate count. This is true even for sparse H.

On the other hand, when H has some other structures, the resource may be significantly reduced. In particular, we consider H in the form of linear combination of unitaries (LCU)

$$H = \sum_{p=0}^{P-1}\alpha_p u_p, \qquad (3)$$

where u_p are n-qubit unitaries that can be implemented with polynomial-size quantum circuit, and $P = O(\mathrm{poly}(n))$. In particular, the linear combination of Pauli strings

$$H = \sum_{p=0}^{P-1}\alpha_p H_p \qquad (4)$$

will be studied in details. Here, $\alpha_p > 0$, $P \geqslant 1$, $H_p = \bigotimes_{l=1}^{n} H_{p,l}$, and $H_{p,l} \in \{I,X,Y,Z\}$ are single-qubit Pauli operators. Eq. (4) is important as it corresponds to the Hamiltonian of quantum systems including the spin and molecular systems.

3.1 Construction of LCU-based Block-encoding

We now discuss the construction of the block-encoding of Eq. (3) and Eq. (4) in Sec. 2. Without loss of generality, we assume that $\log_2 P$ is an integer, and $\sum_{p=0}^{P-1}\alpha_p = 1$.

We define $\boldsymbol{\alpha} = [\alpha_1,\cdots,\alpha_P]$ and

$$|\boldsymbol{\alpha}\rangle = \sum_{p=1}^{P}\sqrt{\alpha_p}|p\rangle. \qquad (5)$$

Let $G_{|\boldsymbol{\alpha}\rangle}$ be the state preparation unitary for $|\boldsymbol{\alpha}\rangle$, and we define $\mathbb{G} \equiv G_{|\boldsymbol{\alpha}\rangle}\otimes\mathbb{I}_{2^n}$. We then define a Select oracle corresponding to Eq. (3) as

$$\mathrm{Select}(u_p) = \sum_{p=0}^{P-1}|p\rangle\langle p|\otimes u_p. \qquad (6)$$

Because

$$(\langle 0^a| \otimes \mathbb{I}_{2^n})\, \mathbb{G}^\dagger \mathrm{Select}(u_p) \mathbb{G}\, (|0^a\rangle \otimes \mathbb{I}_{2^n}) = H, \quad (7)$$

$\mathbb{G}^\dagger\mathrm{Select}(u_p)\mathbb{G}$ is a block-encoding of H with normalization factor $\alpha = 1$ [14]. The constructions of LCU-based block-encoding is then reduced to the constructions of quantum state preparation and Select oracle, both of which can be constructed with polynomial-size quantum circuits. So we have.

Theorem 5 *The block-encoding of H defined in Eq. (3) can be approximated to arbitrary accuracy with polynomial-size quantum circuits of Clifford+T gates.*

The exact circuit complexity of block-encoding depends on the specific form of u_p. We take the LCU for Pauli strings (Eq. (4)) as an example. The result is as follows, where $(n_{\mathrm{anc}}, \varepsilon)$-block-encoding is the abbreviation of $(1, n_{\mathrm{anc}}, \varepsilon)$-block-encoding.

Theorem 6 *With n_{anc} ancillary qubits where $\Omega(n) \leqslant n_{anc} \leqslant O(NP)$, the (n_{anc}, ε)-block-encoding of H defined in Eq. (4) can be constructed with $O\left(P(n + \log(1/\varepsilon))\right)$ count and $\tilde{O}\left(Pn\log(1/\varepsilon)\frac{\log n_{anc}}{n_{anc}}\right)$ depth of elementary gates in $\mathcal{G}_{Clf+T}$, where $\tilde{O}$ suppresses the doubly logarithmic factors of n_{anc}.*

Theorem. 6 is based on our improved protocols for quantum state preparation (Theorem. 7) and $\mathrm{Select}(H_p)$ (Theorem. 8), which are introduced in the following section.

4 Some useful subroutines

The construction of access models introduced above are based some subroutines, such as quantum state preparation, $\mathrm{Select}(H_p)$, and sparse Boolean memory. Their constructions are of independent interest.

4.1 Quantum state preparation

Give an n-qubit quantum state $|\psi\rangle$, we say that a $(n + n_{\mathrm{anc}})$ qubit unitary $G_{|\psi\rangle}$ is a quantum state preparation unitary with accuracy ε for $|\psi\rangle$ if

$$G_{|\psi\rangle}(|0^n\rangle \otimes |0^{n_{\mathrm{anc}}}\rangle) = |\tilde{\psi}\rangle \otimes |0^{n_{\mathrm{anc}}}\rangle \quad (8)$$

for some $\left\||\tilde{\psi}\rangle - |\tilde{\psi}\rangle\right\| \leqslant \varepsilon$. The quantum state preparation problem has been studied extensively [9–11, 13, 16, 18–23]. However, all existing algorithms have either $O(N\mathrm{poly}(n))$ or $O(N\mathrm{polylog}(n))$ Clifford+T gate count scaling with respect to N. We provide a family of improved quantum state preparation protocols with tunable ancillary qubit number. The result is summarized as follows.

Theorem 7 *With n_{anc} ancillary qubits where $\Omega(n) \leqslant n_{anc} \leqslant O(2^n)$, an arbitrary n-qubit quantum state can be prepared to precision ε with $O(N\log(1/\varepsilon))$ count and $\tilde{O}\left(N\log(1/\varepsilon)\frac{\log(n_{anc})}{n_{anc}}\right)$ depth of Clifford+T gates, where $\tilde{O}$ suppresses the doubly logarithmic factors of n_{anc}.*

To our best knowledge, Theorem. 7 is the first result achieving linear scaling of Clifford+T count with respect to N, and this is applied for arbitrary space complexity. The circuit depth lower than and comparable to the best-known results for $n_{\mathrm{anc}} = O(n)$ [20] and $n_{\mathrm{anc}} = O(N)$ [11]. Moreover, compared to [20] which also study the space-time trade-off of state preparation, our method improves the circuit depth scaling for a factor of $\tilde{O}(n_{\mathrm{anc}}/\log n_{\mathrm{anc}})$.

4.2 Select oracle for Pauli strings

Let $H_p = \bigotimes_{l=1}^{L} h_{x,l}$, where $p \in \{0, 1, \cdots, 2^m - 1\}$ and $h_{x,l} \in \{I, X, Y, Z\}$. The select oracle for Pauli strings is

$$\mathrm{Select}(H_x) = \sum_{x=0}^{2^m-1} |p\rangle\langle p| \otimes H_p. \quad (9)$$

We have the following result.

Theorem 8 *With n_{anc} ancillary qubits where $(m+L) \leqslant n_{anc} \leqslant O(ML)$, Eq. (9) can be realized with $O(ML)$ count and $O\left(ML\frac{\log n_{anc}}{n_{anc}}\right)$ depth of Clifford+T gates.*

Compared to the result in Ref [8] with $n_{\mathrm{anc}} = m$, our protocol we reduces the circuit depth for a factor of $O(\frac{\log n_{\mathrm{anc}}}{n_{\mathrm{anc}}})$ while maintaining the gate count and qubit number scaling.

4.3 Sparse Boolean memory

We consider a sparse Boolean function $B : \{0, 1\}^n \to \{0, 1\}^{\tilde{n}}$, which has totally s input digits q satisfying $B(q) \neq 0 \cdots 0$. Given an n-qubit index register (denoted as idx) and a $\tilde{n}$-qubit register (denoted as wrd), we define the sparse Boolean memory $\mathrm{Select}(B)$ as a unitary satisfying

$$\mathrm{Select}(B)|q\rangle_{\mathrm{idx}}|z\rangle_{\mathrm{wrd}} = |x\rangle_{\mathrm{idx}}|z \oplus B(x)\rangle_{\mathrm{wrd}}. \quad (10)$$

We have the following result.

Theorem 9 *With n_{anc} ancillary qubits where $\Omega(n) \leqslant n_{anc} \leqslant O(ns\tilde{n})$, Select$(B(q))$ can be realized with $O(ns\tilde{n})$ count and $O\left(ns\tilde{n}\frac{\log n_{anc}}{n_{anc}}\right)$ depth of Clifford+T gates.*

5 Conclusions

We have studied the circuit complexities of encoding sparse matrices and the block-encoding of LCU. We show that the circuit complexity lower bound for encoding sparse matrix is polynomial with respect to the matrix dimension, and provide a nearly-optimal construction protocol. For LCU-based block-encoding, we develop a construction protocol based on the improved implementation of quantum state preparation and select oracle for Pauli strings. Our protocols are based on Clifford+T gates and allow tunable ancillary qubit number.

The results obtain here are useful for determining the concrete circuit complexities of many algorithms that have been studied in the framework of query access model.

References

[1] Scott Aaronson. Open problems related to quantum query complexity. *ACM Transactions on Quantum Computing*, 2(4):1–9, 2021.

[2] Ryan Babbush, Dominic W Berry, Robin Kothari, Rolando D Somma, and Nathan Wiebe. Exponential quantum speedup in simulating coupled classical oscillators. *arXiv:2303.13012*, 2023.

[3] Dominic W Berry, Graeme Ahokas, Richard Cleve, and Barry C Sanders. Efficient quantum algorithms for simulating sparse hamiltonians. *Communications in Mathematical Physics*, 270:359–371, 2007.

[4] 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*, pages 283–292, 2014.

[5] Shantanav Chakraborty, András Gilyén, and Stacey Jeffery. The power of block-encoded matrix powers: Improved regression techniques via faster hamiltonian simulation. In *Proceedings of the 46th International Colloquium on Automata, Languages and Programming (ICALP)*, 2019.

[6] Andrew M Childs. On the relationship between continuous-and discrete-time quantum walk. *Communications in Mathematical Physics*, 294:581–603, 2010.

[7] Andrew M Childs and Robin Kothari. Simulating sparse hamiltonians with star decompositions. In *Theory of Quantum Computation, Communication, and Cryptography: 5th Conference, TQC 2010, Leeds, UK, April 13-15, 2010, Revised Selected Papers 5*, pages 94–103. Springer, 2011.

[8] Andrew M Childs, Robin Kothari, and Rolando D. Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM J. Comput.*, 46(6):1920–1950, 2017.

[9] B David Clader, Alexander M Dalzell, Nikitas Stamatopoulos, Grant Salton, Mario Berta, and William J Zeng. Quantum resources required to block-encode a matrix of classical data. *arXiv:2206.03505*, 2022.

[10] Lov Grover and Terry Rudolph. Creating superpositions that correspond to efficiently integrable probability distributions. *arXiv:quant-ph/0208112*, 2002.

[11] Kaiwen Gui, Alexander M Dalzell, Alessandro Achille, Martin Suchara, and Frederic T Chong. Spacetime-efficient low-depth quantum state preparation with applications. *arXiv:2303.02131*, 2023.

[12] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103(15):150502, 2009.

[13] Gui-Lu Long and Yang Sun. Efficient scheme for initializing a quantum register with an arbitrary superposed state. *Phys. Rev. A*, 64:014303, Jun 2001.

[14] Guang Hao Low and Isaac L Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.

[15] John M Martyn, Zane M Rossi, Andrew K Tan, and Isaac L Chuang. Grand unification of quantum algorithms. *PRX Quantum*, 2(4):040203, 2021.

[16] Mikko Möttönen, Juha J Vartiainen, Ville Bergholm, and Martti M Salomaa. Transformation of quantum states using uniformly controlled rotations. *Quantum. Inf. Comput.*, 5(6):467–473, 2005.

[17] Michael A Nielsen and Isaac Chuang. *Quantum computation and quantum information.* Cambridge University Press, Cambridge, 2000.

[18] Martin Plesch and Časlav Brukner. Quantum-state preparation with universal gate decompositions. *Phy. Rev. A*, 83(3):032302, 2011.

[19] Gregory Rosenthal. Query and depth upper bounds for quantum unitaries via Grover search. *arXiv:2111.07992*, 2021.

[20] Xiaoming Sun, Guojing Tian, Shuai Yang, Pei Yuan, and Shengyu Zhang. Asymptotically optimal circuit depth for quantum state preparation and general unitary synthesis. *arXiv:2108.06150v2*, 2021.

[21] Xiao-Ming Zhang, Man-Hong Yung, and Xiao Yuan. Low-depth quantum state preparation. *Phys. Rev. Res.*, 3:043200, Dec 2021.

[22] Xiao-Ming Zhang, Tongyang Li, and Xiao Yuan. Quantum state preparation with optimal circuit depth: Implementations and applications. *Phys. Rev. Lett.*, 129(23):230504, 2022.

[23] Zhicheng Zhang, Qisheng Wang, and Mingsheng Ying. Parallel quantum algorithm for hamiltonian simulation. *arXiv:2105.11889*, 2021.

Programming of channels in generalized probabilistic theories

Ryo Takakura[1] * Takayuki Miyadera[2]

[1] *Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo-ku, Kyoto, Japan*
[2] *Department of Nuclear Engineering, Kyoto University, Nishikyo-ku, Kyoto, Japan*

Abstract. For a target system and apparatus described by quantum theory, the quantum no-programming theorem indicates that a family of states (called programs) in the apparatus implements distinct unitaries to the target system through a global unitary only if the programs are orthogonal to each other. In this study, generalizing the programming scheme to generalized probabilistic theories (GPTs), we derive a similar theorem to the quantum no-programming theorem. We furthermore reveal that programming of reversible dynamics is related closely to a curious structure named a quasi-classical structure on the state space. Programming of irreversible dynamics (channels) in GPTs is also investigated.

Keywords: quantum foundations, generalized probabilistic theories, no-programming theorem

1 Introduction

(*The full paper of this study is [1].*) In the field of quantum information, implementing unitary dynamics to a target system is one of the most important tasks. Nielsen and Chuang proposed implementing unitary dynamics by means of "programmable gate array" [2]. In their scenario, an apparatus was considered besides the target system, and the desired unitaries on the target system were implemented by controlling states of the apparatus called "programs" and operating a global unitary to the total system. It could be marvelous if there exist an apparatus and a global unitary that realize an arbitrary number of unitaries on the system, but such a protocol was proved to be impossible: N unitaries on the target system can be programmed only if a distinguishable set of N states in the apparatus are used as the programs (the *quantum no-programming theorem*).

This study aims at investigating how general such a relation is between the possibility of programming unitary (reversible) dynamics and programs. We extend the programming scheme from quantum theory to *generalized probabilistic theories (GPTs)* [3, 4, 5], which are the most general framework of physics. It is proved similarly to the quantum case that a pair of distinct reversible dynamics in the system can be implemented only if the corresponding programs in the apparatus are perfectly distinguishable. We also study whether an apparatus can program a fixed number of reversible dynamics on an arbitrary target system. We reveal that a curious structure (named a *quasi-classical structure*) in the apparatus that quantum theory does not have enables the programming on any target system. Another extension of the quantum setting is also discussed on the scenario of programming irreversible dynamics, i.e., channels.

2 Generalized probabilistic theories (GPTs)

- Preliminaries. GPTs are physical theories where probabilistic mixtures of states and effects are possible. Mathematically, a GPT is a pair of sets $(\Omega, \mathscr{E})$, where

- Ω is a compact convex set in a real and finite-dimensional Euclidean space V such that its linear hull $span(\Omega)$ is identical with V and its affine hull $aff(\Omega)$ does not contain the origin O of V;
- $\mathscr{E}$ is the set of all elements e in the dual space V^* of V ($V \simeq V^*$) such that $e(\omega) \in [0,1]$ for all $\omega \in \Omega$ (we often write $e(\omega)$ also as $\langle e, \omega \rangle$);
- in particular, there is an element $u \in \mathscr{E}$ such that $u(\omega) = 1$ for all $\omega \in \Omega$.

Here we made several assumptions such as the finite dimensionality of V and the compactness of Ω for mathematical simplicity. The set Ω is called the *state space* of the theory and its elements are called *states*. We denote by Ω^{ext} the set of all extreme points of Ω. Elements in Ω^{ext} and $\Omega \backslash \Omega^{\mathrm{ext}}$ are called *pure* and *mixed states* respectively. On the other hand, the set $\mathscr{E}$ is called the *effect space* and its elements are called *effects*. The effect $u \in \mathscr{E}$ is called the *unit effect* and we call a family of effects $\{e_x\}_{x \in X}$ an *observable* if $\sum_{x \in X} e_x = u$.

Example 1 (quantum theory as a GPT) *From the perspective of GPTs, a quantum theory with the Hilbert space $\mathbb{C}^d$ ($d < \infty$) is expressed as $(\mathscr{S}(\mathbb{C}^d), \mathscr{E}(\mathbb{C}^d))$, where $\mathscr{S}(\mathbb{C}^d)$ and $\mathscr{E}(\mathbb{C}^d)$ are the sets of all density operators and effect operators on $\mathbb{C}^d$ respectively.*

Let us describe transformations between states. For state spaces Ω_1 and Ω_2 and their respective linear hulls V_1 and V_2, we define $\mathscr{C}(\Omega_1, \Omega_2) = \{\Lambda \colon V_1 \to V_2 \mid \Lambda(\Omega_1) \subseteq \Omega_2, \text{linear}\}$ and call its elements *channels*. When $\Omega_1 = \Omega_2 = \Omega$, we write the set simply as $\mathscr{C}(\Omega)$. A channel is called a *reversible dynamics* if it is bijective and

*ryo.takakura@yukawa.kyoto-u.ac.jp

$\Lambda(\Omega)=\Omega$, and the set of all reversible dynamics on Ω is written as $GL(\Omega)$. We remark that not all elements of $GL(\Omega)$ are physically possible: in quantum theory, reversible dynamics via anti-unitaries are prohibited.

- Bipartite systems. Let $(\Omega_1,\mathscr{E}_1)$ and $(\Omega_2,\mathscr{E}_2)$ be two GPTs, and consider their composite system. A fundamental assumption is that the composite is also a GPT, which we write by $(\Omega_{12},\mathscr{E}_{12})$. Requiring several axioms such as the no-signaling principle, we have

- the embedding vector space $V_{12}=span(\Omega_{12})$ is given by $V_{12}=V_1\otimes V_2$ with $V_i=span(\Omega_i)$ $(i=1,2)$, and the effect space $\mathscr{E}_{12}$ is embedded into $V_1^*\otimes V_2^*$;
- the independent preparations of states $\omega\in\Omega_1$ and $\xi\in\Omega_2$ in each system is described by $\omega\otimes\xi$, and the independent measurements of effects $e\in\mathscr{E}_1$ and $f\in\mathscr{E}_2$ by $e\otimes f$;
- the unit effect $u_{12}\in\mathscr{E}_{12}$ for Ω_{12} is given by $u_{12}=u_1\otimes u_2$, where u_i is the unit effect for Ω_i $(i=1,2)$;
- $\Omega_1\otimes_{min}\Omega_2\subseteq\Omega_{12}\subseteq\Omega_1\otimes_{max}\Omega_2$, where
$$\Omega_1\otimes_{min}\Omega_2=\{\mu\in V_1\otimes V_2\mid \mu=\sum_{k=1}^{n}p_k\omega_k\otimes\xi_k,\ \omega_k\in\Omega_1,\xi_k\in\Omega_2,\ p_i\geq 0,\ \sum_{k=1}^{n}p_k=1,\ n:\text{finite}\}$$
and
$$\Omega_A\otimes_{max}\Omega_B=\{\mu\in V_1\otimes V_2\mid \langle u_1\otimes u_2,\mu\rangle=1,\ \langle e\otimes f,\mu\rangle\geq 0 \text{ for all } e\in\mathscr{E}_1, f\in\mathscr{E}_2\},$$
and similarly $\mathscr{E}_1\otimes_{min}\mathscr{E}_2\subseteq\mathscr{E}_{12}\subseteq\mathscr{E}_1\otimes_{max}\mathscr{E}_2$.

The sets $\Omega_1\otimes_{min}\Omega_2$ and $\Omega_1\otimes_{max}\Omega_2$ are called the *minimal* and *maximal tensor products* of Ω_1 and Ω_2 respectively. We often use the notation $\Omega_1\otimes\Omega_2$ to represent a bipartite state space composed of Ω_1 and Ω_2.

Example 2 (quantum composite) *The standard composite $\mathscr{S}(\mathbb{C}^{d_1}\otimes\mathbb{C}^{d_2})$ of two quantum state spaces $\mathscr{S}(\mathbb{C}^{d_1})$ and $\mathscr{S}(\mathbb{C}^{d_2})$ indeed satisfies*
$$\mathscr{S}(\mathbb{C}^{d_1})\otimes_{min}\mathscr{S}(\mathbb{C}^{d_2})\subseteq\mathscr{S}(\mathbb{C}^{d_1}\otimes\mathbb{C}^{d_2})\subseteq\mathscr{S}(\mathbb{C}^{d_1})\otimes_{max}\mathscr{S}(\mathbb{C}^{d_2}).$$
The two inclusions are strict.

- Distinguishability. Let $(\Omega,\mathscr{E})$ be a GPT. A family of states $\{\omega_i\}_{i\in\mathscr{I}}$ with an index set $\mathscr{I}$ is called *perfectly distinguishable* if there exists an observable $\{e_i\}_i$ such that $e_i(\omega_j)=\delta_{ij}$. Likewise, $\{\omega_i\}_{i\in\mathscr{I}}$ is called *pairwise distinguishable* if any pair $\{\omega_i,\omega_j\}$ of its distinct elements is perfectly distinguishable. Although the two notions coincide with each other in classical and quantum theory, they are in general different: a perfectly distinguishable set of states is pairwise distinguishable, but the converse does not always hold in GPTs [6]. In finite dimensional cases, both perfectly and pairwise distinguishable sets of states are finite [1].

3 Programming of channels in GPTs

- Quantum theory. Let a quantum system and apparatus be with finite-dimensional Hilbert spaces $\mathscr{H}$ and $\mathscr{K}$ respectively. We consider programming unitary (reversible) dynamics on the system by choosing states of the apparatus. Let W be a unitary operator on $\mathscr{H}\otimes\mathscr{K}$. We say that a state $|\xi\rangle\in\mathscr{K}$ (called a program) in the apparatus implements a unitary dynamics U_ξ on the system through W if for any $|\varphi\rangle\in\mathscr{H}$
$$W(|\varphi\rangle\otimes|\xi\rangle)=(U_\xi|\varphi\rangle)\otimes|\xi'\rangle$$
holds, where $|\xi'\rangle\in\mathscr{K}$. It was proved that programs ξ and η can implement distinct unitaries U_ξ and U_η respectively only if $\langle\xi|\eta\rangle=0$ holds [2]. Thus the number of programs is at most $\dim\mathscr{K}$ and an arbitrary number of dynamics are not possible, while $\dim\mathscr{K}$ unitaries $\{U_n\}_{n=1}^{\dim\mathscr{K}}$ can be programmed with the global unitary $W=\sum_n U_n\otimes|n\rangle\langle n|$, where $\{|n\rangle\}_n$ is a CONS of $\mathscr{K}$.

- GPTs. Let a system and an apparatus described by GPTs $(\Omega_{sys},\mathscr{E}_{sys})$ and $(\Omega_{app},\mathscr{E}_{app})$ respectively, and $\Omega_{tot}:=\Omega_{sys}\otimes\Omega_{app}$ be their composite state space. We introduce a subset $GL_0(\Omega_{sys})$ of $GL(\Omega_{sys})$ such that any $\alpha\in GL_0(\Omega_{sys})$ satisfies $\alpha\otimes\mathrm{id}_{\Omega_{app}}\in GL(\Omega_{tot})$, where $\mathrm{id}_{\Omega_{app}}\in\mathscr{C}(\Omega_{app})$ is the identity channel on Ω_{app}. While any $\alpha\in GL(\Omega_{sys})$ satisfies this condition in the minimal and maximal tensor products, it does not hold in general composites. We also assume that $GL_0(\Omega_{sys})$ is a group with respect to the concatenation.

Let $\Lambda\in GL(\Omega_{tot})$ be a reversible dynamics on the total system. We say that a state (called a program) $\xi\in\Omega_{app}$ implements a reversible dynamics $\alpha_\xi\in GL_0(\Omega_{sys})$ on the system through Λ if for any $\omega\in\Omega_{sys}$ and $e\in\mathscr{E}_{sys}$
$$\langle e\otimes u_{app},\Lambda(\omega\otimes\xi)\rangle=\langle e,\alpha_\xi\omega\rangle \tag{1}$$
holds, where u_{app} is the unit effect for Ω_{app}. The condition (1) implies that the dynamics restricted on the system is α_ξ. In fact, if ω is pure, then (1) reduces to
$$\Lambda(\omega\otimes\xi)=\alpha_\xi\omega\otimes\xi'$$
with some $\xi'\in\Omega_{app}$. We remark that without loss of generality programs are assumed to be pure. Based on the setting above, we obtain the following theorem.

Theorem 1 *If states ξ and η of the apparatus implement distinct reversible dynamics of the system, then they are perfectly distinguishable.*

Theorem 1 is a generalization of the quantum result to GPTs: programs in the apparatus should be pairwise distinguishable, and only a finite number of reversible dynamics can be programmed on the system.

Remark 1 It is not proved in the theorem that N pairwise distinguishable states as programs are sufficient to realize N reversible dynamics. Rather, there exist a pair of state spaces and their composite where it does not hold.

- Universal programming. We study another scenario on programming reversible dynamics.

Definition 1 Let N be a positive integer. A state space Ω_{app} is said to have an *N-universal programming property* if for any state space Ω_{sys} and reversible dynamics $\{\alpha_n\}_{n=1}^{N} \subset GL(\Omega_{sys})$ there exist a composite state space $\Omega_{sys} \otimes \Omega_{app}$ and a reversible dynamics $\Lambda \in GL(\Omega_{sys} \otimes \Omega_{app})$ such that there are states $\{\xi_n\}_{n=1}^{N} \subset \Omega_{app}^{ext}$ implementing $\{\alpha_n\}_{n=1}^{N}$ through Λ.

The following notion is important to investigate universal programming properties.

Definition 2 A disjoint decomposition $\Omega^{ext} = \bigcup_{k\in K} \Omega^{ext}_{[k]}$ for the set Ω^{ext} of pure states of a state space Ω is called a *quasi-classical decomposition of degree* $|K|$ if there exists an observable $\{e_k\}_{k\in K}$ satisfying $\langle e_k, \omega_{k'}\rangle = \delta_{kk'}$ for $\omega_{k'} \in \Omega^{ext}_{[k']}$. A state space that admits a quasi-classical decomposition is said to have a *quasi-classical structure*.

There is a characterization of quasi-classical structures.

Proposition 2 *A decomposition $\Omega^{ext} = \bigcup_{k\in K} \Omega^{ext}_{[k]}$ is quasi-classical iff $\{\Omega^{ext}_{[k]}\}_k$ is "(linearly) independent" (i.e."perfectly distinguishable"). That is, when $\omega \in \Omega$ is expressed as $\omega = \sum_{k\in K} p_k\omega_k = \sum_{k\in K} q_k\omega'_k$, where $\omega_k, \omega'_k \in \Omega_{[k]}$ with $\Omega_{[k]}$ the convex hull of $\Omega^{ext}_{[k]}$ and $\{p_k\}_{k\in K}$ and $\{q_k\}_{k\in K}$ probability distributions on K, then $p_k = q_k$ holds for all $k \in K$.*

The state space of a finite-level classical system clearly has a quasi-classical structure, while the quantum state space $\mathscr{S}(\mathbb{C}^d)$ does not. There are other examples.

Example 3 *A state space $\Omega = \bigoplus_{n=1}^{N} \mathscr{S}(\mathbb{C}^{d_n})$ describing a quantum system with a superselection rule has a quasi-classical structure: $\Omega^{ext} = \bigcup_{n=1}^{N} \mathscr{S}^{ext}_{[n]}$ with*

$$\begin{aligned}\mathscr{S}^{ext}_{[1]} &= \mathscr{S}^{ext}(\mathbb{C}^{d_1}) \oplus 0 \oplus 0 \oplus \cdots,\\ \mathscr{S}^{ext}_{[2]} &= 0 \oplus \mathscr{S}^{ext}(\mathbb{C}^{d_2}) \oplus 0 \oplus 0 \oplus \cdots,\\ &\vdots\end{aligned}$$

is a quasi-classical decomposition of degree n.

Example 4 *Consider a state space Ω described in Figure 1. For the set of its pure states $\Omega^{ext} = \{\omega_1, \ldots, \omega_6\}$, there are quasi-classical decompositions*

$$\begin{aligned}\Omega^{ext} &= \{\omega_1,\omega_2,\omega_3\} \cup \{\omega_4,\omega_5,\omega_6\}\\ &= \{\omega_1,\omega_4\} \cup \{\omega_2,\omega_5\} \cup \{\omega_3,\omega_6\}.\end{aligned}$$

We can prove that quasi-classical structures enable universal programmings.

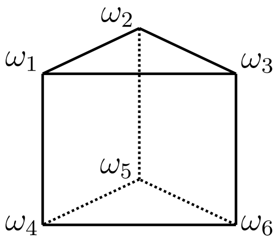


Figure 1: A state space shaped by a triangular prism.

Theorem 3 *A state space Ω_{app} has an N-universal programming property iff Ω_{app} has a quasi-classical structure such that $\Omega^{ext}_{app} = \bigcup_{k\in K} \Omega^{ext}_{app\,[k]}$ with $|K| \geq N$.*

- Irreversible universal programming. We investigate what observations can be obtained when programming more general state changes, i.e., channels. Similarly to Definition 1, we introduce the following notion.

Definition 3 Let N be a positive integer. A state space Ω_{app} is said to have an *irreversible N-universal programming property* if for any state space Ω_{sys} and channels $\{\tau_n\}_{n=1}^{N} \subset \mathscr{C}(\Omega_{sys})$, there exist a composite state space $\Omega_{sys} \otimes \Omega_{app}$ and a channel $\Theta \in \mathscr{C}(\Omega_{sys} \otimes \Omega_{app})$ such that there are states $\{\xi_n\}_{n=1}^{N} \subset \Omega_{app}^{ext}$ implementing N distinct channels $\{\tau_n\}_{n=1}^{N}$ of Ω_{sys} through Θ.

In this case, we obtain the following theorem.

Theorem 4 *A state space Ω_{app} has an irreversible N-universal programming property iff there exists a family of perfectly distinguishable states $\{\xi_n\}_{n=1}^{N}$ in Ω_{app}.*

This result can be compared with Theorem 3, where quasi-classical structures appear as a consequence of considering reversible dynamics instead of channels.

4 Conclusion

We generalized the quantum programming scheme to GPTs. We found that a family of reversible dynamics on a target system is programmable only if a pairwise distinguishable set of states in an apparatus is used as programs. This result seems to be a straightforward generalization of the quantum one, but it was derived for any composite between the minimal and maximal tensor products. We also considered changing the programming scenario itself: universal programmings of reversible dynamics and channels, and investigated when they are possible. It was shown that the former scheme is realizable iff the apparatus has a quasi-classical structure and the latter is possible iff the programs are perfectly distinguishable. We believe that the former result is important in that it is peculiar to GPTs beyond quantum theory: as Example 4 shows, state spaces with quasi-classical structures seem to have properties that the triangle (classical) and square theories have. Future study will be needed to characterize quasi-classical structures.

References

[1] T. Miyadera and R. Takakura. Programming of channels in generalized probabilistic theories. *J. Math. Phys.* 64: 042201, 2023.

[2] M. A. Nielsen and I. L. Chuang. Programmable Quantum Gate Arrays. *Phys. Rev. Lett.* 79: 321, 1997.

[3] L. Lami. Non-classical correlations in quantum mechanics and beyond. arXiv:1803.02902 [quant-ph], 2018.

[4] M. Plávala. General probabilistic theories: An introduction. arXiv:2103.07469 [quant-ph], 2021.

[5] R. Takakura. Convexity and uncertainty in operational quantum foundations. arXiv:2202.13834 [quant-ph], 2022.

[6] N. Brunner *et al*. Dimension of physical systems, information processing, and thermodynamics. *New J. Phys.* 16:123050, 2014.

Calculation of capacity with discrete-valued inputs using efficiently obtained eigenvalues

Shion Kitamura[1] * Tiancheng Wang[1 2] † Souichi Takahira[3] ‡
Tsuyoshi Sasaki Usuda[1] §

[1] *Graduate School of Information Science and Technology, Aichi Prefectural University, Aichi, Japan.*
[2] *Faculty of Informatics, Kanagawa University, Kanagawa, Japan.*
[3] *Faculty of Information Engineering, Meijo University, Aichi, Japan.*

Abstract. In quantum communication theory, the weighted Gram matrix with *a priori* probabilities is extremely important because it is the matrix representation of the density operator of a quantum information source. Recently, we have found that the eigenvalue problem of this matrix can be simplified for asymmetric quantum signals. In the present paper, we apply our results to the calculation of the capacity for these signals. We also discuss how much of the capacity of a continuous system, the ultimate limit of classical-quantum communication, can be achieved by discrete-valued inputs.

Keywords: Quantum communication, Quantum channel capacity, Gram matrix

1 Introduction

The classical capacity of a quantum (lossy) channel is attained by continuous inputs of coherent-states [1]. In 2000, Sohma and Hirota showed that the capacity is asymptotically achieved by binary coherent-state signals when the average number of photons is very small [2]. Subsequently, Ishida *et al.* applied up to 16-ary signals and showed that the average number of photons to achieve the capacity is widened [3]. Note that as multi-ary signals, asymmetric signals are required and we cannot use the formula for symmetric signals [4]. They called the maximum classical capacity attained by continuous inputs the "full capacity" and expressed "almost achieve the full capacity" if the ratio of the capacity with discrete-valued input to the full capacity is greater than 0.99 [3]. In the following, we use N_{th} to denote the maximum value of the average number of photons that almost reaches full capacity.

However, because of the difficulty of calculating the capacity for a large number of signals M, known N_{th}'s are still very small (e.g., $N_{\mathrm{th}} \sim 0.01$ when $M = 2$ and $N_{\mathrm{th}} \sim 1$ when $M = 16$). It is highly desirable to be able to calculate the capacity efficiently even when M is large in order to find signals that can almost achieve the full capacity over a wider range of the average number of photons. The capacity with discrete-valued inputs can be obtained by computing the eigenvalues of the weighted Gram matrix with *a priori* probabilities. Recently, by generalizing the results of [5], we found that the eigenvalue problem can be simplified [6, 7]. In the present paper, we apply our results to calculate the capacity with discrete-valued inputs and show examples of larger N_{th} (~ 10).

2 Preliminaries

2.1 Quantum signals and the density operator

For a Hilbert space $\mathcal{H}$ of a quantum system, a set of M-ary quantum-state signals is defined by

$$\mathcal{S} = \{|\psi_i\rangle \in \mathcal{H} \mid i = 1, 2, \ldots, M, \langle\psi_i|\psi_i\rangle = 1\}. \quad (1)$$

Let ξ_i be the *a priori* probability of $|\psi_i\rangle$ and $\xi = \{\xi_i \mid i = 1, 2, \ldots, M\}$. Then $(\mathcal{S}, \xi)$ is often referred to as the quantum information source, and

$$\hat{\rho} = \sum_{i=1}^{M} \xi_i |\psi_i\rangle\langle\psi_i| \quad (2)$$

is the so-called density operator of $(\mathcal{S}, \xi)$. Here, we introduce the weighted Gram matrix with *a priori* probabilities, G, that is a M-by-M matrix whose (i, j) component is $\langle\tilde{\psi}_i|\tilde{\psi}_j\rangle$, where $|\tilde{\psi}_i\rangle = \sqrt{\xi_i}|\psi_i\rangle$.

The importance of the weighted Gram matrix is that it is isomorphic to the density operator $\hat{\rho}$, i.e.,

$$G \cong \hat{\rho}, \quad (3)$$

*im233003@cis.aichi-pu.ac.jp
†wang@kanagawa-u.ac.jp
‡takahira@meijo-u.ac.jp
§usuda@ist.aichi-pu.ac.jp

which means that G is a matrix representation of $\hat{\rho}$. Therefore, G can be treated as $\hat{\rho}$ itself when analyzing it and the eigenvalues of $\hat{\rho}$ can be obtained by calculating those of G.

2.2 The von Neumann entropy and capacity

For the density operator $\hat{\rho}$, the von Neumann entropy is defined as

$$\chi(\xi) = -\mathrm{Tr}\left(\hat{\rho}\log_2 \hat{\rho}\right), \tag{4}$$

and the classical capacity of a quantum channel is defined as the value of $\chi(\xi)$ maximized with respect to *a priori* probabilities:

$$C = \max_{\xi} \chi(\xi). \tag{5}$$

From the fact (3), $\chi(\xi)$ can be calculated as

$$\chi(\xi) = -\mathrm{Tr}\left(G\log_2 G\right) = -\sum_j \lambda_j \log_2 \lambda_j, \tag{6}$$

where λ_j are the eigenvalues of G. Thus, for efficient computation of the capacity, it is important to calculate the eigenvalues of G efficiently.

3 Calculation results

Using the simplification of the eigenvalue problem for the weighted Gram matrix [6, 7], we calculate the capacity with discrete-valued inputs for asymmetric quantum signals. The von Neumann entropy can be calculated by its eigenvalues, but further maximization with respect to *a priori* probabilities of the signals is required to compute the capacity. This maximization can be achieved when the *a priori* probabilities of equidistant signals are equal [8]. Here, we perform the maximization by numerical search with a reduced set of *a priori* probabilities based on the partial symmetry of the signals.

3.1 The von Neumann entropy for QAM signals

Fig.1 shows a comparison between the von Neumann entropy with equal *a priori* probabilities of the QAM signals and the full capacity. The horizontal axis shows the average number of photons in received signals ηN_{S}, which means energy constraint [3]. The number of QAM signals is set to $M = 4, 16, 36, 64$. From Fig.1, even at $M = 64$, there is a clear gap between the von Neumann entropy with equal *a priori* probabilities and the full capacity when the average number of photons is very small, e.g., when $\eta N_{\mathrm{S}} = 1$ or 2. We verify how much of this gap can be filled by maximization with respect to *a priori* probabilities.

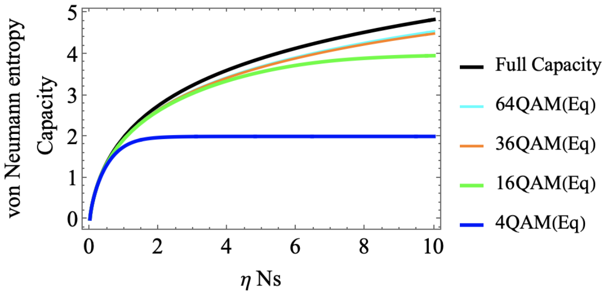


Figure 1: The von Neumann entropy with equal *a priori* probabilities for QAM signals and full capacity

3.2 The capacities for 16-ary signals

Here, we examine how much the full capacity can be achieved by the capacity with discrete-valued inputs for various 16-ary signals. It will be seen that even though the number of signals is the same, N_{th} differs depending on the signal constellations. We consider five types of signal constellations; 16PSK, three 16APSK, and 16QAM signals. Fig.2 shows the signal constellations of three types of 16APSK signals.

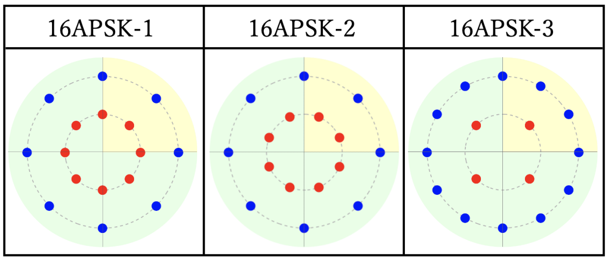


Figure 2: 16APSK signals on phase plane

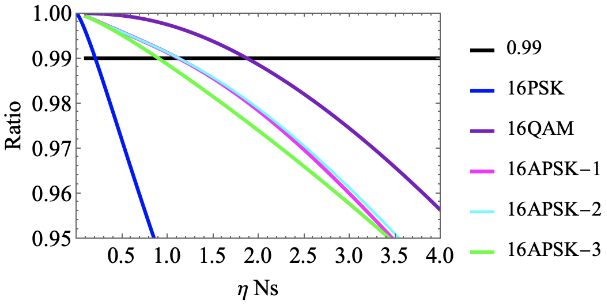


Figure 3: Ratio of capacities for 16-ary signals to full capacity

Fig.3 plots the ratio of the capacities with discrete-valued inputs of 16-ary signals to full capacity. From Fig.3, the average number of photons to achieve 99% full capacity are $N_{\mathrm{th}} \sim 0.2$ for 16PSK, $N_{\mathrm{th}} \sim 1$ for 16APSK, and $N_{\mathrm{th}} \sim 1.8$ for 16QAM. Regarding the signal constellations, 16PSK uses only one type of amplitude. In other words, only one type of concentric points is used in the phase plane. In contrast, 16APSK uses two types and 16QAM uses three types. This indicates that the full capacity can be achieved with a larger number of average photons for many types of amplitudes. In the class of 16APSK, 16APSK-1 and 16APSK-2 perform better than 16APSK-3. 16APSK-3 has a small number of inner signal points. In other words, the greater number of inner signal points, i.e., those with small amplitude, the larger capacity.

3.3 The capacity for 64QAM signals

Fig.4 shows the channel capacity with discrete-valued inputs for 64QAM in addition to the von Neumann entropy and the capacity shown in Fig.1.

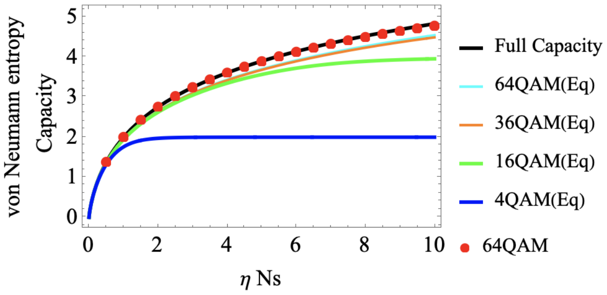


Figure 4: The capacity with discrete-valued inputs for 64QAM signals (red dots) on Fig.1 (solid lines)

All solid lines are the same as in Fig.1, and the capacity with discrete-valued inputs for 64QAM is indicated by red dots. Fig.1 plots the average number of photons below 10, and the capacity for 64QAM seems to almost achieve the full capacity over the entire range.

3.4 Discussion

In this section, the results of the capacity calculations are summarized, and the differences in achievement levels are discussed. Fig.5 shows the ratio of the capacity with discrete-valued inputs for each signal to the full capacity. Table 1 shows the maximum value of the average number of photons N_{th} to almost achieve the full capacity for each signal.

From Table 1, it can be seen that 64QAM almost achieves the full capacity up to approximately 9.33. By maximizing the von Neumann entropy with respect to *a priori* probabilities, N_{th} is approximately 4.7 times larger for 16QAM, 10.8 times larger for 36QAM, and 19.4 times larger for 64QAM. This result clearly demonstrates the significance of optimizing *a priori* probabilities, especially for large M.

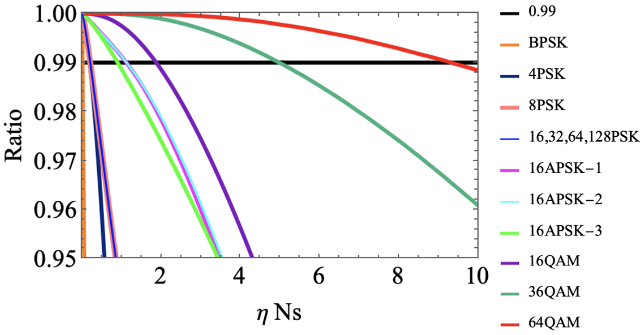


Figure 5: Ratio of capacities for each signal to the full capacity

Type of Signals	N_{th}
BPSK	0.01
4PSK	0.18
8,16,32,64,128PSK	0.20
16APSK-1	1.11
16APSK-2	1.12
16APSK-3	0.89
16QAM(Eq)	0.39
16QAM	1.87
36QAM(Eq)	0.46
36QAM	4.98
64QAM(Eq)	0.48
64QAM	9.33

Table 1: The maximum value of the average number of photons N_{th} to almost achieve the full capacity.

4 Conclusion

In the present paper, we calculated the channel capacity with discrete-valued inputs using the results in [6, 7] and considered the maximum value of the average number of photons N_{th} to almost achieve the full capacity as in [3]. First, we confirmed that N_{th} is less than 1 in the case of QAM with equal *a priori* probabilities, even for 64-ary signals, and then considered various signal constellations for 16-ary signals. As a result, it was found that increasing the types of amplitudes tends to yield larger capacity. Furthermore, we calculated the capacity with discrete-valued inputs by increasing the number of QAM signals and showed that $N_{\text{th}} > 9.3$, which is approximately five times wider than the previous result [3].

Acknowledgments: This work has been supported in part by JSPS KAKENHI Grant Number 20H00581, 20K20397, 21K04064, 22K20437, and The Nitto Foundation.

References

[1] V. Giovannetti, S. Guha, S. Lloyd, L. Maccone, J. H. Shapiro, and H. P. Yuen: "Classical capacity of the lossy bosonic channel: the exact solution," Phys. Rev. Lett. **90**, 027902, (2004).

[2] M. Sohma and O. Hirota: "Binary discretization for quantum continuous channels," Phys. Rev. **A62**, 052312, (2000).

[3] Y. Ishida, K. Kato, and T. S. Usuda: "Capacity of attenuated channel with discrete-valued input," Proc. of the 8th International Conference on Quantum Communication, Measurement and Computing (QCMC), (2006).

[4] T. S. Usuda and K. Shiromoto: "Analytical expression of s-th power of Gram matrix for group covariant signals and its application," AIP Conference Proceedings **1363**, T. Ralph and P. K. Lam (Eds.), American Institute of Physics, New York, pp.97-100, (2011).

[5] T. Wang, R. Miyazaki, S. Takahira, and T. S. Usuda: "Simplification of the Gram matrix eigenvalue problem for quantum signals formed by rotating signal points in a circular sector region," IEEJ Trans. on Electronics, Information and Systems **142**, pp.74-87, (2022). (in Japanese)

[6] S. Kitamura, T. Wang, S. Takahira, and T.S. Usuda: "Simplification of the generalized Gram matrix eigenvalue problem for asymmetric M-ary coherent-state signals," Proc. of 2022 Tokai-Section Joint Conference on Electrical, Electronics, Information, and Related Engineering, E1-1, (2022). (in Japanese)

[7] S. Kitamura, T. Wang, S. Takahira, and T.S. Usuda: "Simplification of the eigenvalue problem for density operators representing asymmetric quantum signals and its application to capacity calculations," Proc. of the 45th Symposium on Information Theory and its Applications (SITA), pp.169-174, (2022). (in Japanese)

[8] K. Kato, M. Osaki, and O. Hirota: "Derivation of classical capacity of quantum channel for discrete information source," Phys. Lett. **A251**, pp.157-163, (1999).

[9] C. E. Shannon: "A mathematical theory of communication," The Bell System Technical Journal **27**, pp.379-423, (1948).

[10] C. W. Helstrom: *Quantum detection and estimation theory*, Academic Press, New York, (1976).

[11] O. Hirota: *The foundation of quantum information science*, Morikita Publishing, Japan, (2002). (in Japanese)

[12] R. A. Horn and C. R. Jonson: *Matrix analysis*, Cambridge University Press, (1985).

Efficient stabilizer entropies for quantum computers

Tobias Haug[1] Soovin Lee *[1] M. S. Kim[1]

[1] *QOLS, Blackett Laboratory, Imperial College London SW7 2AZ, UK*

Abstract. Stabilizer entropies (SEs) are measures of nonstabilizerness or 'magic' that quantify the degree to which a state is described by stabilizers. SEs are especially interesting due to their connections to scrambling, localization, and property testing. However, applications have been limited so far as previously known measurement protocols for SEs scale exponentially with the number of qubits. Here, we introduce Tsallis-n SEs as efficient measures of nonstabilizerness for quantum computers. The number of measurements is independent of the number of qubits and the protocol is easy to implement via Bell measurements. The Tsallis SE is an efficient bound of various nonstabilizerness monotones which are intractable to compute beyond a few qubits. Using the IonQ quantum computer, we experimentally measure the Tsallis SE of random Clifford circuits doped with non-Clifford gates and give bounds for the stabilizer fidelity, stabilizer extent, and robustness of magic. As applications, we provide efficient algorithms to measure $4n$-point out-of-time-order correlators and multifractal flatness. Our results open up the exploration of nonstabilizerness with quantum computers.

Keywords: Nonstabilizerness, OTOC, magic, resource theory, NISQ

Stabilizer states and Clifford operations are essential to quantum information and quantum computing [1–3]. They are the cornerstone to run quantum algorithms on most fault-tolerant quantum computers, where Clifford operations are intertwined with non-Clifford gates [4, 5]. To characterize the amount of non-Clifford resources needed to realize quantum states and operations the resource theory of nonstabilizerness has been put forward [6–14]. Stabilizer entropies (SEs) [15] are measures of nonstabilizerness with efficient algorithms for matrix product states [16–18] which have enabled the study of nonstabilizerness in many-body systems [16–23].

Recently, SEs have also been related to various important properties of quantum systems. SEs can probe phase transitions in error-corrected circuits [24] and the entanglement spectrum [25] as well as determine the testing efficiency of purity [26] and fidelity [27]. SEs are also connected to the participation entropy [28], which are helpful to understand Anderson [29] and many-body localization [30]. Further, recent works established a fruitful connection between out-of-time-order correlators (OTOCs) and nonstabilizerness [27, 31, 32]. OTOCs describe scrambling in quantum systems [33, 34]. However, OTOCs are challenging to measure directly and often require an inverse of the time evolution [35]. Higher-order OTOCs and nonstabilizerness have been related to quantum chaos [31] and state certification [27].

The aforementioned properties make SEs highly interesting for experimental studies of quantum computers and simulators. However, the progress has so far been limited as all previously known measurement protocols for SEs scale exponentially with the number of qubits [14, 36].

Here, we introduce the Tsallis-n SE with integer $n > 1$ as an efficient measure of nonstabilizerness for quantum computers and simulators. Our algorithms are practical to implement via Bell measurements over two copies of the state, where for even n we additionally require access to the complex conjugate of the state. We leverage the relationship between SEs and OTOCs to devise an efficient protocol for $4n$-point OTOCs where for odd n we do not require an inverse time evolution. Further, the Tsallis SE enables the efficient measurement of the multifractal flatness. The Tsallis SE also provides efficiently computable bounds to other nonstabilizerness monotones, which are otherwise intractable beyond a few qubits. Finally, we experimentally measure the Tsallis SE on the IonQ quantum computer and demonstrate SEs as efficient bounds for the robustness of magic, stabilizer extent and stabilizer fidelity. Our work introduces methods to experimentally uncover the key features that characterize the power of quantum computers and simulators.

Importance for quantum information and computation

The Tsallis-n SE T_n as an efficient measure of nonstabilizerness with a measurement cost independent of qubit number N, which is an exponential

*soovinlee310@gmail.com

improvement over previous protocols [14, 36]. For integer $n > 1$, our protocol is asymptotically optimal with the number of copies scaling as $O(n\epsilon^{-2})$ and the classical post-processing time as $O(nN\epsilon^{-2})$. The protocol is easy to implement using Bell measurements which have been demonstrated for quantum computers and simulators [37–39].

Future work could find efficient protocols for even n without the need of complex conjugation, prove or refute the strong monotonicity condition for $n \geq 2$, and tighten the lower bound of SE for the stabilizer fidelity.

SEs hold promise for characterizing different properties of quantum states in experiments. We experimentally demonstrate SEs as a bound of nonstabilizerness monotones which otherwise are hard to compute beyond a few qubits. These monotones serve as lower bounds on state preparation complexity and characterize the runtime of classical simulation algorithms [8, 40].

Further, we show how to efficiently measure a class of $4n$-point OTOCs. Our protocol has the advantage that it does not require implementing time-reversal for odd $n > 1$, which has been a challenge in other experiments [33]. Our protocol can measure higher order OTOCs which promise to reveal more features compared to the usually considered 4-point OTOCs [41, 42]. We also give an efficient protocol to evaluate multifractal flatness which characterizes the distribution of basis states of wavefunctions.

Finally, our work enables the experimental study of phase transitions in SE which have been found for purity testing [26] and quantum error correction [24], as well as characterization of recent experiments on fault-tolerant encodings of magic states [43, 44].

References

1. Gottesman, D. *Stabilizer codes and quantum error correction. Caltech Ph. D* PhD thesis (Thesis, eprint: quant-ph/9705052, 1997).
2. Shor, P. W. *Fault-tolerant quantum computation* in *Proceedings of 37th conference on foundations of computer science* (1996), 56–65.
3. Kitaev, A. Y. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303,** 2–30. http://www.sciencedirect.com/science/article/pii/S0003491602000180 (2003).
4. Bravyi, S. & Kitaev, A. Universal quantum computation with ideal Clifford gates and noisy ancillas. *Phys. Rev. A* **71,** 022316. https://link.aps.org/doi/10.1103/PhysRevA.71.022316 (2 Feb. 2005).
5. Campbell, E. T., Terhal, B. M. & Vuillot, C. Roads towards fault-tolerant universal quantum computation. *Nature* **549,** 172–179. https://www.nature.com/articles/nature23460 (2017).
6. Campbell, E. T. Catalysis and activation of magic states in fault-tolerant architectures. *Phys. Rev. A* **83,** 032317. https://link.aps.org/doi/10.1103/PhysRevA.83.032317 (3 Mar. 2011).
7. Veitch, V., Mousavian, S. H., Gottesman, D. & Emerson, J. The resource theory of stabilizer quantum computation. *New J. Phys.* **16,** 013009. https://doi.org/10.1088/1367-2630/16/1/013009 (2014).
8. Howard, M. & Campbell, E. Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing. *Phys. Rev. Lett.* **118,** 090501. https://link.aps.org/doi/10.1103/PhysRevLett.118.090501 (9 Mar. 2017).
9. Wang, X., Wilde, M. M. & Su, Y. Quantifying the magic of quantum channels. *New J. Phys.* **21,** 103002. https://doi.org/10.1088/1367-2630/ab451d (2019).
10. Beverland, M., Campbell, E., Howard, M. & Kliuchnikov, V. Lower bounds on the non-Clifford resources for quantum computations. *Quantum Science Tech.* **5,** 035009. https://doi.org/10.1088/2058-9565/ab8963 (2020).
11. Jiang, J. & Wang, X. Lower Bound for the T Count Via Unitary Stabilizer Nullity. *Physical Review Applied* **19,** 034052 (2023).
12. Liu, Z.-W. & Winter, A. Many-Body Quantum Magic. *PRX Quantum* **3,** 020333. https://link.aps.org/doi/10.1103/PRXQuantum.3.020333 (2 May 2022).
13. Bu, K., Garcia, R. J., Jaffe, A., Koh, D. E. & Li, L. Complexity of quantum circuits via sensitivity, magic, and coherence. *arXiv:2204.12051.* https://arxiv.org/abs/2204.12051 (2022).

14. Haug, T. & Kim, M. Scalable Measures of Magic Resource for Quantum Computers. *PRX Quantum* **4,** 010301. https://link.aps.org/doi/10.1103/PRXQuantum.4.010301 (1 Jan. 2023).

15. Leone, L., Oliviero, S. F. E. & Hamma, A. Stabilizer Rényi Entropy. *Phys. Rev. Lett.* **128,** 050402. https://link.aps.org/doi/10.1103/PhysRevLett.128.050402 (5 Feb. 2022).

16. Haug, T. & Piroli, L. Quantifying nonstabilizerness of matrix product states. *Phys. Rev. B* **107,** 035148. https://link.aps.org/doi/10.1103/PhysRevB.107.035148 (2023).

17. Haug, T. & Piroli, L. Stabilizer entropies and nonstabilizerness monotones. *arXiv:2303.10152* (2023).

18. Lami, G. & Collura, M. Quantum Magic via Perfect Sampling of Matrix Product States. *arXiv:2303.05536.* https://arxiv.org/abs/2303.05536 (2023).

19. Oliviero, S. F. E., Leone, L. & Hamma, A. Magic-state resource theory for the ground state of the transverse-field Ising model. *Phys. Rev. A* **106,** 042426. https://link.aps.org/doi/10.1103/PhysRevA.106.042426 (4 Oct. 2022).

20. Odavić, J. *et al.* Complexity of frustration: a new source of non-local non-stabilizerness. *arXiv:2209.10541.* https://arxiv.org/abs/2209.10541 (2022).

21. Chen, L., Garcia, R. J., Bu, K. & Jaffe, A. Magic of Random Matrix Product States. *arXiv:2211.10350.* https://arxiv.org/abs/2211.10350 (2022).

22. Goto, K., Nosaka, T. & Nozaki, M. Probing chaos by magic monotones. *Physical Review D* **106,** 126009 (2022).

23. Piemontese, S., Roscilde, T. & Hamma, A. Entanglement complexity of the Rokhsar-Kivelson-sign wavefunctions. *Physical Review B* **107,** 134202 (2023).

24. Niroula, P. *et al.* Phase transition in magic with random quantum circuits. *arXiv:2304.10481* (2023).

25. Tirrito, E. *et al.* Quantifying non-stabilizerness through entanglement spectrum flatness. *arXiv preprint arXiv:2304.01175* (2023).

26. Leone, L., Oliviero, S. F., Esposito, G. & Hamma, A. Phase transition in Stabilizer Entropy and efficient purity estimation. *arXiv:2302.07895* (2023).

27. Leone, L., Oliviero, S. F. E. & Hamma, A. Nonstabilizerness determining the hardness of direct fidelity estimation. *Phys. Rev. A* **107,** 022429. https://link.aps.org/doi/10.1103/PhysRevA.107.022429 (2 Feb. 2023).

28. Turkeshi, X., Schirò, M. & Sierant, P. Measuring Magic via Multifractal Flatness. *arXiv:2305.11797* (2023).

29. Castellani, C. & Peliti, L. Multifractal wavefunction at the localisation threshold. *Journal of physics A: mathematical and general* **19,** L429 (1986).

30. Stéphan, J.-M., Furukawa, S., Misguich, G. & Pasquier, V. Shannon and entanglement entropies of one- and two-dimensional critical wave functions. *Phys. Rev. B* **80,** 184421. https://link.aps.org/doi/10.1103/PhysRevB.80.184421 (18 Nov. 2009).

31. Leone, L., Oliviero, S. F., Zhou, Y. & Hamma, A. Quantum chaos is quantum. *Quantum* **5,** 453. https://arxiv.org/abs/2102.08406 (2021).

32. Garcia, R. J., Bu, K. & Jaffe, A. Resource theory of quantum scrambling. *Proceedings of the National Academy of Sciences* **120,** e2217031120 (2023).

33. Xu, S. & Swingle, B. Scrambling dynamics and out-of-time ordered correlators in quantum many-body systems: a tutorial. *arXiv:2202.07060* (2022).

34. Dowling, N., Kos, P. & Modi, K. Scrambling is Necessary but Not Sufficient for Chaos. *arXiv:2304.07319* (2023).

35. Li, J. *et al.* Measuring out-of-time-order correlators on a nuclear magnetic resonance quantum simulator. *Physical Review X* **7,** 031011 (2017).

36. Oliviero, S. F., Leone, L., Hamma, A. & Lloyd, S. Measuring magic on a quantum processor. *npj Quantum Information* **8,** 148 (2022).

37. Islam, R. *et al.* Measuring entanglement entropy in a quantum many-body system. *Nature* **528,** 77–83 (2015).

38. Huang, H.-Y. *et al.* Quantum advantage in learning from experiments. *Science* **376,** 1182–1186 (2022).

39. Bluvstein, D. *et al.* A quantum processor based on coherent transport of entangled atom arrays. *Nature* **604,** 451–456 (2022).

40. Bravyi, S. *et al.* Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum* **3,** 181. https://quantum-journal.org/papers/q-2019-09-02-181/ (2019).

41. Roberts, D. A. & Yoshida, B. Chaos and complexity by design. *Journal of High Energy Physics* **2017,** 1–64 (2017).

42. Garcia, R. J., Zhou, Y. & Jaffe, A. Quantum scrambling with classical shadows. *Physical Review Research* **3,** 033155 (2021).

43. Ye, Y. *et al.* Near-Perfect Logical Magic State Preparation on a Superconducting Quantum Processor. *arXiv:2305.15972* (2023).

44. Gupta, R. S. *et al.* Encoding a magic state with beyond break-even fidelity. *arXiv:2305.13581* (2023).

Quantum space-time marginal problem: global causal structure from local causal information

Zhian Jia[1] [2] [*] Minjeong Song[3] [†] Dagomir Kaszlikowski[1] [2] [‡]

[1] *Centre for Quantum Technologies, National University of Singapore, Singapore 117543, Singapore*
[2] *Department of Physics, National University of Singapore, Singapore 117543, Singapore*
[3] *Nanyang Quantum Hub, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371, Singapore*

Abstract. Spatial and temporal quantum correlations can be unified in the framework of the pseudo-density operators, and quantum causality between the involved events in an experiment is encoded in the corresponding pseudo-density operator. We study the relationship between local causal information and global causal structure. A space-time marginal problem is proposed to infer global causal structures from given marginal causal structures where causal structures are represented by the pseudo-density operators; we show that there almost always exists a solution in this case. By imposing the corresponding constraints on this solution set, we could obtain the required solutions for special classes of marginal problems, like a positive semidefinite marginal problem, separable marginal problem, etc. The notion of quantum pseudo-channel is also introduced and we demonstrate that the quantum pseudo-channel marginal problem can be solved by transforming it into a pseudo-density operator marginal problem via the channel-state duality.

Keywords: Pseudo-density operator, Marginal problem, Quantum causality

1 Introduction

The relativity theory treats space and time on equal footing. However, in quantum mechanics, space and time play extremely different roles. Searching for a representation of quantum mechanics that treats space and time in a more even-handed fashion is thus a crucial problem and may shed new light on the notion of quantum space-time and quantum causality. There have been a variety of proposals for space-time states, process matrix [16], consistent history [8], entangled histories [5], and quantum-classical game [9], superdensity operators [4], multi-time states [1], pseudo-density operator (PDO) [7], doubled density operator[12], etc.

Clarifying the relation between the whole and its parts is crucial in many areas of science. The question that considers in what situation the local information can be reproduced from a global structure is known as the marginal problem, which has a long history. This seemingly effortless problem is indeed highly nontrivial, there exist locally compatible distributions that do not have global solutions. And the problem has been shown to be NP-hard [17]. In quantum mechanics, states are represented by density operators, and thus the marginal problems are rephrased in terms of density operators. The question of whether a given set of marginals (reduced density operators) is compatible with a global density operator is called a quantum state marginal problem, see, e.g [19], and references therein. This seemingly easy problem turned out to be challenging to solve in general, and it lies at the heart of many problems in quantum physics. The quantum state marginal problem was initially proposed in quantum chemistry with the name N-representability problem and it's regarded as one of the most prominent research challenges in quantum chemistry [3, 18]. The existence of absolutely maximally entangled states can be transformed into the existence of the solution for a specific quantum marginal problem [21]. The symmetric extension of a bipartite state can also be recast as a marginal problem [6, 15]. The monogamy of the maximally entangled states is equivalent to the disappearance of the solution for the corresponding marginal problem [2]. The marginal problem also plays a crucial role in investigating the quantum phases of matter [22]. The marginal problem essentially characterizes the compatibility of quantum states, this can also be generalized to quantum channels and quantum measurements [10, 11]. In this work, we investigate the marginal problem for space-time states and higher-order dynamics.

2 Pseudo-density operator

The pseudo-density operator formalism concerns the following scenario [7]: we have a quantum system distribution over space and we choose to measure some (generalized) Pauli measurements over some qudit (x) at some particular instant in time (t). We introduce a tensor product structure among all space-time events $\mathcal{A} = \{E(x_i,t_i)\}_{i=1}^n$. Thus the total space is $\mathcal{H}_\mathcal{A} = \otimes_i \mathcal{H}[E(x_i,t_i)]$. In this way, we obtain a state of the system that is distributed over space-time [7] (the general qudit definition is given in Ref. [13])

$$R_\mathcal{A} = \frac{1}{d^n} \sum_{\mu_1,\cdots,\mu_n=0}^{d-1} T^{\mu_1\cdots\mu_n} \otimes_{j=1}^n \sigma_{\mu_j}, \qquad (1)$$

where $T^{\mu_1\cdots\mu_n} = \langle \{\sigma_{\mu_j}\}_{j=1}^n \rangle$ is the expectation value of a collection of Pauli measurements. This $R_\mathcal{A}$ is called a PDO. Notice that when all qudits are measured at the same instant of time, we obtain the normal Bloch representation of a multipartite state [20]. We will denote the set of all PDOs for an event set $\mathcal{A}$ as $\mathbf{PDO}(\mathcal{A})$.

[*] giannjia@foxmail.com
[†] song.at.qit@gmail.com
[‡] phykd@nus.edu.sg

Theorem 1. For any n-event set $\mathcal{A}$, any PDO $R_{\mathcal{A}}$ can be expressed as a quasi-probabilistic mixture of pure space-time product states

$$R_{\mathcal{A}} = \sum_{a_1,\cdots,a_n} p(a_1,\cdots,a_n)|a_1,\cdots,a_n\rangle\langle a_1,\cdots,a_n|, \quad (2)$$

where $|a_1,\cdots,a_n\rangle = |a_1\rangle\otimes\cdots\otimes|a_n\rangle$.

Definition 2. Consider an n-event space-time scenario $\mathcal{A} = \{E_1,\cdots,E_n\}$, we still assign a local Hilbert space $\mathcal{H}_{E_i}$ for each event E_i. The local state vectors are independent, viz., they are in product-form $|a_1,\cdots,a_n\rangle = |a_1\rangle\otimes\cdots\otimes|a_n\rangle$. The correlations are captured by the negativity of quasi-probability distribution $\vec{p} = (p_1,\cdots,p_n)$,

$$W_{\mathcal{A}} = \sum_{i=1}^{k} p(a_1,\cdots,a_n)|a_1,\cdots,a_n\rangle\langle a_1,\cdots,a_n|. \quad (3)$$

3 Quantum space-time causal marginal problem

In the conventional space-time causal marginal problem, we ask that given a family of sets of events $\mathfrak{M}_{\mathcal{A}} = \{\mathcal{A}_1,\cdots,\mathcal{A}_k\}$, called a marginal scenario of $\mathcal{A} = \cup_i\mathcal{A}_i$, if there exists a global causal structure $R(\mathcal{A})$ over all events contained in $\mathcal{A}$ which is compatible with causal structures $R(\mathcal{A}_i)$ for all $i = 1,\cdots n$. This is clearly trivial, we only need to check if all $R(\mathcal{A}_i)$ are compatible. If they are compatible, there always exists a solution.

Theorem 3. The deterministic classical causal marginal problem always has a solution.

Definition 4 (PDO marginal problem). Consider a marginal scenario consisting of a family of event sets $\mathcal{A}_1,\cdots,\mathcal{A}_n$ with their corresponding PDOs $R_{\mathcal{A}_1},\cdots,R_{\mathcal{A}_n}$, such that they are compatible. The PDO marginal problem asks if there exists a global PDO $R_{\mathcal{A}}$ with $\mathcal{A} = \cup_i\mathcal{A}_i$ such that $R_{\mathcal{A}_i} = \mathrm{Tr}_{\mathcal{A}\setminus\mathcal{A}_i} R_{\mathcal{A}}$ for all $i = 1,\cdots,n$.

From the previous discussion, we see that an n-event PDO is determined by a rank-n tensor $T^{\mu_1,\cdots,\mu_n}$. Taking the partial trace over some event subset, we obtain the new tensor for the reduced PDO by just setting the corresponding indices as zero. For example, for $T^{\mu_1\mu_2\mu_3}$, tracing over the third event, the tensor of the reduced PDO is just $T^{\mu_1\mu_2 0}$. This substantially simplifies the problem.

Theorem 5 (**Herm**$_1$ marginal problem). Consider the marginal problem $\{R_{\mathcal{A}_i}\}_{i=1}^n$ with $R_{\mathcal{A}_i} \in \mathbf{PDO}(\mathcal{A}_i)$ and $\mathcal{A} = \cup_{i=1}^n\mathcal{A}_i$. In $\mathbf{Herm}_1(\mathcal{A})$, there always exists a solution R which is the solution to the marginal problem. In other words, the marginal problem in $\mathbf{Herm}_1(\mathcal{A})$ is trivial.

3.1 Space-time separable marginal problem

Definition 6 (space-times separable marginal problem). For a marginal scenario $\mathfrak{M}_{\mathcal{A}}$ consisting of a given collection of event sets $\{\mathcal{A}_i\}$ with their corresponding separable space-time separable states $\{W_{\mathcal{A}_i}\}$, the space-times separable marginal problem asks if there exists a space-time separable state $W_{\mathcal{A}}$ for $\mathcal{A} = \cup_i\mathcal{A}_i$ such that all $W_{\mathcal{A}_i}$ can be reproduced by taking marginals.

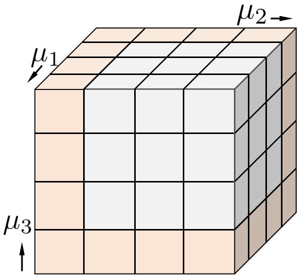


Figure 1: The illustration of the proof for **Herm**$_1$ marginal problem, the cube represents the tensor $T^{\mu_1\mu_2\mu_3}$ of marginal problem solution $R_{\mathcal{A}}$. The light gray boxes represent the free parameter, while the light red boxes represent the parameters fixed by reduced PDOs $R_{\mathcal{A}_1}, R_{\mathcal{A}_2}, R_{\mathcal{A}_3}$.

By the theorem 5, we know that there always exists a set of quasi-probabilistic separable solution

$$\begin{aligned}&\mathbf{Marg}(\mathfrak{M}_{\mathcal{A}})\\ =&\{W_{\mathcal{A}} = \sum_{a_1,\cdots,a_n} p(a_1,\cdots,a_n)\otimes_{i=1}^n |a_i\rangle\langle a_i|\},\end{aligned} \quad (4)$$

where all $p(a_1,\cdots,a_n)$ are quasi-probability distributions. To obtain the positive semidefinite solution set, we need first impose the positive semidefinite condition

$$\begin{aligned}&\mathbf{Marg}^{\mathrm{pos}}(\mathfrak{M}_{\mathcal{A}})\\ =&\{W_{\mathcal{A}} \in \mathbf{Marg}(\mathfrak{M}_{\mathcal{A}})|\,\mathrm{Tr}(W_{\mathcal{A}}Y) \geq 0, \forall Y \geq 0\}.\end{aligned} \quad (5)$$

The second step is to choose the separable ones from these positive semidefinite solutions. However, there is a more efficient approach to filter the solution from $\mathbf{Marg}(\mathfrak{M}_{\mathcal{A}})$ using the polytope approximation of $\mathbf{Sep}(\mathcal{A})$. Suppose that we have n space-time separable states $R_1,\cdots,R_n$, they can generate a convex polytope $\mathbb{P} = \mathbf{Sep}(R_1,\cdots,R_n) = \mathrm{Conv}(R_1,\cdots,R_n)$. By Minkowski-Weyl theorem, this polytope can be rewritten as a bounded intersection of half-spaces $\mathbb{P} = \cap_{i=1}^m\mathbb{H}_i$. Each half-space is determined by an Hermitian operator K_i, namely $\mathbb{H}_i = \{R \in \mathbf{Herm}|\langle R, K_i\rangle \geq 0\}$. The marginal problem solution contained in this polytope is thus

$$\begin{aligned}&\mathbf{Marg}^{\mathbb{P}}(\mathfrak{M}_{\mathcal{A}})\\ =&\{W_{\mathcal{A}} \in \mathbf{Marg}(\mathfrak{M}_{\mathcal{A}})|\,\mathrm{Tr}(W_{\mathcal{A}}K_i) \geq 0, \forall i\}.\end{aligned} \quad (6)$$

In this way, we obtain an operational method to solve the space-time separable marginal problem, which can be implemented numerically.

3.2 Space-time symmetric extension

Theorem 7. For any two-event space-time state (for which PDO is a special example) W_{AB}, the symmetric extension $W_{AB_1\cdots B_k}$ always exists in the space of all quasi-probabilistic mixture of space-time product state.

This technique can also be applied to extendibility for m-event $W_{A_1\cdots A_k B_1\cdots B_l}$ with respect to $B_1\cdots B_l$. Notice the above corollary means that any $W \in \mathbf{Herm}_1$ is extendible in $\mathbf{Herm}_1$. This can also be transformed into a marginal problem and be proved using theorem 5. Suppose we have a collection of space-time state $W_{AB} = W_{AB_1} = \cdots W_{AB_{n-2}}$, theorem 5 ensures that there exists a non-empty solution set $\mathbf{Marg}(W_{AB}, W_{AB_1}, \cdots, W_{AB_{n-2}})$. Then we can add more constraints to filter the solutions we need as we have done in the previous subsection.

3.3 Polygamy of space-time correlations

For spatial quantum correlations, it's well known that there are monogamy relations for entanglement, quantum steering, and Bell nonlocality. The monogamy relation can be reformulated using a quantum marginal problem, e.g., a singlet state cannot be shared by three parties simultaneously, if Alice and Bob share the singlet state, then the state between Alice and Carol must not be a singlet state. This means that the marginal scenario $\mathfrak{M} = \{\psi^-_{AB}, \psi^-_{AC}\}$ has no solution. However, for space-time correlations, the monogamy relation will be broken, an example has been given in Ref. [14]. Here, using the marginal problem framework, we see that polygamy is a general phenomenon for space-time states.

3.4 Classical quasi-probability marginal problem

Definition 8. Consider a set of quasi-random variables $\mathcal{A} = \{X_1, \cdots, X_n\}$, then a classical marginal scenario $\mathfrak{M}_{\mathcal{A}}$ on $\mathcal{A}$ is a non-empty collection $\{\mathcal{A}_1, \cdots \mathcal{A}_k\}$ of subsets of $\mathcal{A}$ together with a set of compatible quasi-probability distributions $\{p(X \in \mathcal{A}_i)\}_{i=1}^k$.

Definition 9 (classical space-time state marginal problem). For a set of classical space-time states $\{W_{\mathcal{A}_i}\}_{i=1}^k$ which are compatible with each other up to local unitary operations, find a classical space-time state $W_{\mathcal{A}}$ such that all $W_{\mathcal{A}_i}$ are local unitary equivalent the reduced states of $W_{\mathcal{A}}$.

Theorem 10. The quasi-probability classical marginal problem for a marginal behavior $\mathfrak{M}_{\mathcal{A}}$ is equivalent to the classical space-time state marginal problem $\{W_{\mathcal{A}_i}\}$.

4 Quantum pseudo-channel

4.1 Quantum pseudo-channel as higher-order maps

Definition 11 (QPC). Consider the space of all bounded operators over the Hilbert space $\mathcal{H}_{\mathcal{A}_X} = (\mathbb{C}^d)^{\otimes n_X}$ with $X = I, O$ ('in' and 'out'), a pseudo-density channel is a linear map $\Phi : \mathbf{B}(\mathcal{H}_{\mathcal{A}_I}) \to \mathbf{B}(\mathcal{H}_{\mathcal{A}_O})$ such that $\Phi(R_{\mathcal{A}_I}) \in \mathbf{PDO}(\mathcal{A}_O)$ for all $R_{\mathcal{A}_I} \in \mathbf{PDO}(\mathcal{A}_I)$, *viz.*, it maps PDO to PDO. We denote the corresponding set of QPC as $\mathbf{QPC}(\mathcal{A}_I, \mathcal{A}_O)$.

The above definition of QPC can naturally be generalized to space-time states, which we will call space-time channels. From the definition of a QPC Φ, we see that Φ must satisfy: (i) it's Hermiticity-preserving (HP); (ii) it's trace-preserving (TP).

4.2 Marginal quantum pseudo-channel

The notion of marginal quantum operation and quantum channel is introduced in [11]. This can be naturally generalized to the QPC. Suppose that $\mathcal{A}$ and $\mathcal{B}$ are input and out event sets of the QPC $\Phi_{\mathcal{B}|\mathcal{A}}$. The marginal is defined with respect to a bipartition of both the input and output event sets. Let $\mathcal{X} \subset \mathcal{A}$ and $\mathcal{Y} \subset \mathcal{B}$, the marginal QPC $\Phi_{\mathcal{Y}|\mathcal{X}}$ is defined as follows: for arbitary $R_{\mathcal{A}} \in \mathbf{PDO}(\mathcal{A})$ we have

$$\mathrm{Tr}_{\mathcal{Y}^c}\, \Phi_{\mathcal{B}|\mathcal{A}}(R_{\mathcal{A}}) = \Phi_{\mathcal{Y}|\mathcal{X}}(\mathrm{Tr}_{\mathcal{X}^c}(R_{\mathcal{A}})), \tag{7}$$

where $\mathcal{X}^c$ and $\mathcal{Y}^c$ are complements of $\mathcal{X}$ and $\mathcal{Y}$ in $\mathcal{A}$ and $\mathcal{B}$. We will denote this marginal QPC as $\mathrm{Tr}_{\mathcal{Y}^c|\mathcal{X}^c}\, \Phi_{\mathcal{B}|\mathcal{A}} = \Phi_{\mathcal{Y}|\mathcal{X}}$.

Hereinafter, for convenience of discussion, we will use a normalized Choi-Jamiołkowski representation of $\Phi_{\mathcal{B}|\mathcal{A}}$,

$$J(\Phi_{\mathcal{B}|\mathcal{A}}) = \frac{1}{d_{\mathcal{A}}} \Phi_{\mathcal{B}|\mathcal{A}}(E_{ij}) \otimes E_{ij}. \tag{8}$$

It's clear that $\Phi_{\mathcal{B}|\mathcal{A}}(R)/d_{\mathcal{A}} = \mathrm{Tr}_{\mathcal{A}}[J(\Phi_{\mathcal{B}|\mathcal{A}})(\mathbb{I} \otimes R^T)]$. We will call this correspondence channel-state duality. Using the channel state duality, we can translate this defining condition (7) into a state form (see, e.g., [11, Appendix A] and references therein)

$$\mathrm{Tr}_{\mathcal{Y}^c}\, J(\Phi_{\mathcal{B}|\mathcal{A}}) = J(\Phi_{\mathcal{Y}|\mathcal{X}}) \otimes \frac{\mathbb{I}_{\mathcal{X}^c}}{d_{\mathcal{X}^c}}. \tag{9}$$

Since we take a different convention for the Choi-Jamiołkowski map, there is no dimension factor here in our expression. This implies that the Choi map for the marginal channel is indeed the marginal state $J(\Phi_{\mathcal{Y}|\mathcal{X}}) = \mathrm{Tr}_{\mathcal{Y}^c|\mathcal{X}^c}\, J(\Phi_{\mathcal{B}|\mathcal{A}})$.

4.3 Quantum pseudo-channel marginal problem

Definition 12 (QPC marginal problem). Given a collection of QPC $\{\Phi_{\mathcal{B}_i|\mathcal{A}_i}\}$, suppose that they are compatible with each other, the QPC marginal problem asks if there exists a global QPC from event set $\mathcal{A} = \cup_i \mathcal{A}_i$ to $\mathcal{B} = \cup_i \mathcal{B}_i$ which can reproduce all QPCs by taking marginals.

From channel-state duality, $J(\Phi_{\mathcal{B}|\mathcal{A}})$ is Hermitian if and only if $\Phi_{\mathcal{B}|\mathcal{A}}$ is HP. $\Phi_{\mathcal{B}|\mathcal{A}}$ is TP implies that $\mathrm{Tr}_{\mathcal{B}}\, J(\Phi_{\mathcal{B}|\mathcal{A}}) = \mathbb{I}_{\mathcal{A}}/d_{\mathcal{A}}$, thus $\mathrm{Tr}\, J(\Phi_{\mathcal{B}|\mathcal{A}}) = 1$. When $\Phi_{\mathcal{B}|\mathcal{A}}$ is HPTP, $J(\Phi_{\mathcal{B}|\mathcal{A}}) \in \mathbf{Herm}_1$. As shown in subsection 4.2, the compatibility of two QPCs on their overlap is indeed the same as the compatibility of states corresponding to them.

Theorem 13 (**HPTP** marginal problem). For a collection of compatible QPC $\{\Phi_{\mathcal{B}_i|\mathcal{A}_i}\}$, there always exists a solution for the marginal problem in $\mathbf{HPTP}(\mathcal{A}, \mathcal{B})$.

Acknowledgements. — We acknowledge Fabio Costa, James Fullwood, Arthur Parzygnat and Xiangjing Liu for bringing our attention to some relevant works on quantum causal discovery and quantum causal inference. This work is supported by the National Research Foundation and the Ministry of Education in Singapore.

References

[1] Yakir Aharonov, Sandu Popescu, Jeff Tollaksen, and Lev Vaidman. Multiple-time states and multiple-time measurements in quantum mechanics. *Phys. Rev. A*, 79:052110, May 2009.

[2] Valerie Coffman, Joydip Kundu, and William K. Wootters. Distributed entanglement. *Phys. Rev. A*, 61:052306, Apr 2000.

[3] A. J. Coleman. Structure of fermion density matrices. *Rev. Mod. Phys.*, 35:668–686, Jul 1963.

[4] Jordan Cotler, Chao-Ming Jian, Xiao-Liang Qi, and Frank Wilczek. Superdensity operators for space-time quantum mechanics. *Journal of High Energy Physics*, 2018(9):1–57, 2018.

[5] Jordan Cotler and Frank Wilczek. Entangled histories. *Physica Scripta*, 2016(T168):014004, 2016.

[6] A. C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Distinguishing separable and entangled states. *Phys. Rev. Lett.*, 88:187904, Apr 2002.

[7] Joseph F Fitzsimons, Jonathan A Jones, and Vlatko Vedral. Quantum correlations which imply causation. *Scientific reports*, 5(1):1–7, 2015.

[8] Robert B Griffiths. Consistent histories and the interpretation of quantum mechanics. *Journal of Statistical Physics*, 36(1):219–272, 1984.

[9] Gus Gutoski and John 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.

[10] Erkka Haapasalo, Tristan Kraft, Nikolai Miklin, and Roope Uola. Quantum marginal problem and incompatibility. *Quantum*, 5:476, 2021.

[11] Chung-Yun Hsieh, Matteo Lostaglio, and Antonio Acín. Quantum channel marginal problem. *Phys. Rev. Res.*, 4:013249, Mar 2022.

[12] Zhian Jia and Dagomir Kaszlikowski. The spatiotemporal doubled density operator: a unified framework for analyzing spatial and temporal quantum processes. *to be published*, 2023.

[13] Zhian Jia, Minjeong Song, and Dagomir Kaszlikowski. Quantum space-time marginal problem: global causal structure from local causal information. *arXiv preprint arXiv:2303.12819*, 2023.

[14] Chiara Marletto, Vlatko Vedral, Salvatore Virzì, Enrico Rebufello, Alessio Avella, Fabrizio Piacentini, Marco Gramegna, Ivo Pietro Degiovanni, and Marco Genovese. Theoretical description and experimental simulation of quantum entanglement near open time-like curves via pseudo-density operators. *Nature Communications*, 10(1):182, 2019.

[15] Miguel Navascués, Flavio Baccari, and Antonio Acin. Entanglement marginal problems. *Quantum*, 5:589, 2021.

[16] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature communications*, 3(1):1–8, 2012.

[17] Itamar Pitowsky. Quantum probability quantum logic. 1989.

[18] Mary Beth Ruskai. n-representability problem: Conditions on geminals. *Phys. Rev.*, 183:129–141, Jul 1969.

[19] Christian Schilling. The quantum marginal problem. In *Mathematical Results in Quantum Mechanics: Proceedings of the QMath12 Conference*, pages 165–176. World Scientific, 2015.

[20] Lu Wei, Zhian Jia, Dagomir Kaszlikowski, and Sheng Tan. Antilinear superoperator and quantum geometric invariance for higher-dimensional quantum systems. *arXiv preprint arXiv:2202.10989*, 2022.

[21] Xiao-Dong Yu, Timo Simnacher, Nikolai Wyderka, H Chau Nguyen, and Otfried Gühne. A complete hierarchy for the pure state marginal problem in quantum mechanics. *Nature Communications*, 12(1):1012, 2021.

[22] Bei Zeng, Xie Chen, Duan-Lu Zhou, and Xiao-Gang Wen. *Quantum Information Meets Quantum Matter–From Quantum Entanglement to Topological Phase in Many-Body Systems.* Springer, 2019.

Comparison of capacity and coding rates in CPPM-type quantum cipher with outer codes

Yuki Wakahara[1] * Souichi Takahira[1] † Shogo Usami[1] ‡ Tsuyoshi Sasaki Usuda[2] §

[1] *Graduate School of Science and Engineering, Meijo University*
1-501, Shiogamaguchi, Nagoya-shi, Aichi, 468-0073, Japan.
[2] *School of Information Science and Technology, Aichi Prefecture University*
1522-3, Ibaragabasama, Nagakute-shi, Aichi, 480-1198, Japan.

Abstract. We have shown in AQIS2022 that the application of single parity-check codes as outer codes to CPPM-type KCQ enhances the quantum gain when the average number of photons is as small as 5 or less in the case of 2-CPPM. By comparing the applied codes with the channel capacities of a legitimate receiver and an attacker, this paper shows that the coding was effective in terms of the wire-tap channel.

Keywords: quantum cryptography, wire-tap channel, capacity, KCQ, CPPM

1 Introduction

In the principle of quantum cryptography KCQ (keyed communication in quantum noise) [1], the optimum quantum measurement of a legitimate receiver with a key and that of an attacker without a key are different. Therefore, secure communication is accomplished since there is obvious difference in capability between their measurements. The difference corresponds to the so-called quantum gain. There are two types of KCQ protocols; single-mode and multi-mode types. The former can only provide a maximum quantum gain of 6dB, whereas a quantum gain of the latter is unlimited. The CPPM (coherent pulse position modulation) is a typical multi-mode type KCQ. However, as stated in section VIB of [1], the CPPM has the weakness that its bandwidth increases exponentially with the number of slots. It was also mentioned in [1], that this problem can be solved by applying properly designed error correcting codes because of the fact that the capacities of the legitimate receiver and the attacker are different. It is well-known that the capacity argument provides only an asymptotic limit and the design of codes of finite lengths is different issue.

We have shown in AQIS2022 that applying single parity-check codes as outer codes to CPPM-type KCQ enhances the quantum gain when the average number of photons is as small as 5 or less in the case of 2-CPPM. In this paper, we first introduce the above results. Then, we calculate the capacity of 2-CPPM signals and verify that the (5,4) or (6,5) coding is effective in terms of the wire-tap channel [3] when the average number of photons N_s is $0.5 < N_s < 2.5$.

2 CPPM-Type KCQ Cryptography

2.1 Protocol Overview

We refer to the sender as Alice, the legitimate receiver as Bob, and the attacker as Eve. The details of the protocols are left to the literature; here are the main points. For simplicity, assume $M = 2$ in the M-ary CPPM. Alice encrypts the 2-PPM signal $\{|\phi_0\rangle = |\alpha\rangle|0\rangle, |\phi_1\rangle = |0\rangle|\alpha\rangle\}$ corresponding to the transmitted classical information $\{0,1\}$ into a CPPM signal by applying a unitary operator $U(k)$ depends on a key $k(=1,2,\dots)$, and transmits it via the quantum communication channel. The signal transmitted in this case is as follows.

$$|\psi_{i,k}\rangle = U(k)|\phi_i\rangle, \quad (i=0,1). \tag{1}$$

Then Bob obtains the 2-PPM signal by applying $U(k)^\dagger$ which is determined by the key. Assume that he uses a direct detection receiver or a quantum optimal receiver. On the other hand, Eve does not know the key and cannot obtain the 2-PPM signal, so she attempts to measure the 2-CPPM signal with a heterodyne receiver. Here, we consider Yuen's upper bound evaluation method, that is Eve is granted fictionally the key after her measurement.

2.2 Error Probability for 2-CPPM Signals

Let P^{opt} be the error probability when Bob measures the 2-PPM signal using a quantum optimal

*223426025@ccmailg.meijo-u.ac.jp
†takahira@meijo-u.ac.jp
‡susami@meijo-u.ac.jp
§usuda@ist.aichi-pu.ac.jp

receiver. The error rate of a quantum optimal receiver for binary signals is well known [4], and for 2-PPM signals $\{|\phi_0\rangle, |\phi_1\rangle\}$ it is as follows.

$$\begin{aligned} P^{\mathrm{opt}} &= \frac{1}{2}\left(1-\sqrt{1-|\langle\phi_0|\phi_1\rangle|^2}\right) \\ &= \frac{1}{2}\left(1-\sqrt{1-e^{-2|\alpha|^2}}\right). \end{aligned} \tag{2}$$

Let P^{het} be the error probability when Eve measures the 2-CPPM signal using a heterodyne receiver. P^{het} is as follows [1, 5].

$$\begin{aligned} P^{\mathrm{het}} &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(y-\sqrt{2|\alpha|^2})^2}{2}\right] Q(y)dy, \\ Q(y) &= 1-[\Phi(y)], \\ \Phi(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} \exp\left[-v^2/2\right] dv. \end{aligned} \tag{3}$$

2.3 Channel Capacity for 2-CPPM Signals

Note that Bob can convert 2-CPPM back to 2-PPM with a key-based unitary transformation, and a quantum collective decoding can be performed on the encoded quantum signal. Let

$$\rho = \frac{1}{2}\left(|\phi_0\rangle\langle\phi_0| + |\phi_1\rangle\langle\phi_1|\right) \tag{4}$$

be the density operator corresponding to the 2-PPM signal. Then Bob's capacity is

$$\begin{aligned} C_{\mathrm{Bob}} &= -\operatorname{Tr}\rho\log\rho \\ &= -\lambda_+\log\lambda_+ - \lambda_-\log\lambda_-, \end{aligned} \tag{5}$$

where $\lambda_\pm = \frac{1}{2}(1 \pm \langle\phi_0|\phi_1\rangle)$ are eigenvalues of ρ. Assuming that Eve fictionally knows the key after her heterodyne measurement, the upper bound of Eve's capability can be evaluated by soft-decision decoding for coded signals affected by a Gaussian noise. The corresponding capacity is

$$C_{\mathrm{Eve}} = \max\{H(Y) - H(Y|X)\}. \tag{6}$$

Maximization of the right-hand side of Eq.(6) is performed with respect to the distribution of the input X and is achieved when the input distribution is uniform because the channel is symmetric. The $H(Y)$ and $H(Y|X)$ in Eq.(6) are the output entropy and conditional output entropy, respectively, and

$$H(Y) = -\int_{-\infty}^{\infty} p_y(y)\log p_y(y)dy, \tag{7}$$

$$H(Y|X) = -\int_{-\infty}^{\infty} p_{y|x}(y|x)\log p_{y|x}(y|x)dy, \tag{8}$$

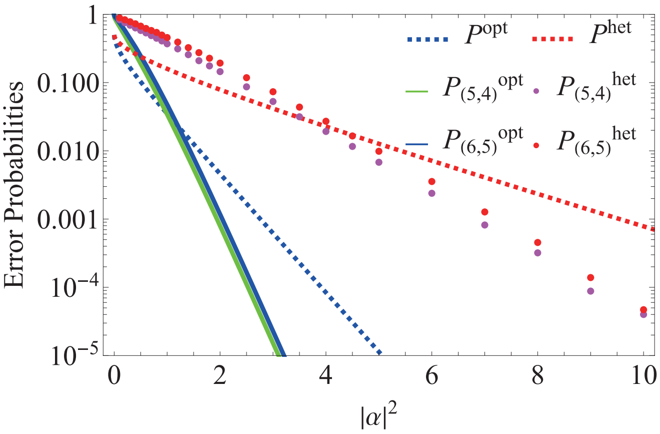


Figure 1: Error probability for uncoded and (5,4) and (6,5) single parity coded signals, measured by Bob using the optimal receiver and by Eve using heterodyne receiver

when the distribution of input X is uniform. Here,

$$p_{y|x}(y|x) = \frac{1}{\sqrt{\pi}}e^{-y^2} \tag{9}$$

$$p_y(y) = \frac{1}{2\sqrt{\pi}}\left(e^{-y^2} + e^{-(y-\sqrt{2|\alpha|^2})^2}\right). \tag{10}$$

In general, $C_{\mathrm{Eve}} < C_{\mathrm{Bob}}$, and the rate R of encoding should be set as

$$C_{\mathrm{Eve}} < R < C_{\mathrm{Bob}}, \tag{11}$$

for the security [3].

3 Codings and their Effects

This section describes the coding proposed in AQIS2022. Here, we explain using the (2,1) repetition code that is the simplest code. First, Alice generates a codeword state:

$$\begin{aligned} &|\phi_0\rangle|\phi_0\rangle = |\alpha\rangle|0\rangle|\alpha\rangle|0\rangle : 00, \\ \text{or } &|\phi_1\rangle|\phi_1\rangle = |0\rangle|\alpha\rangle|0\rangle|\alpha\rangle : 11. \end{aligned} \tag{12}$$

Alice then encrypts it. For example, if the generated codeword state is $|\phi_0\rangle|\phi_0\rangle$ and encryption is done using $U(k=1), U(k=2)$, then the encrypted state is $|\psi_{0,1}\rangle|\psi_{0,2}\rangle$. Bob can decrypt it to $|\phi_0\rangle|\phi_0\rangle$, but Eve cannot.

The results of applying the (5,4) single parity-check code and the (6,5) single parity-check code are shown in Fig.1. From Fig.1, when the average number of photons $N_s = |\alpha|^2$ is small (near 0 to 4), Eve's error probability is worsened by coding. Comparing the (5,4) single parity-check code and the (6,5) single parity-check code, it can be seen

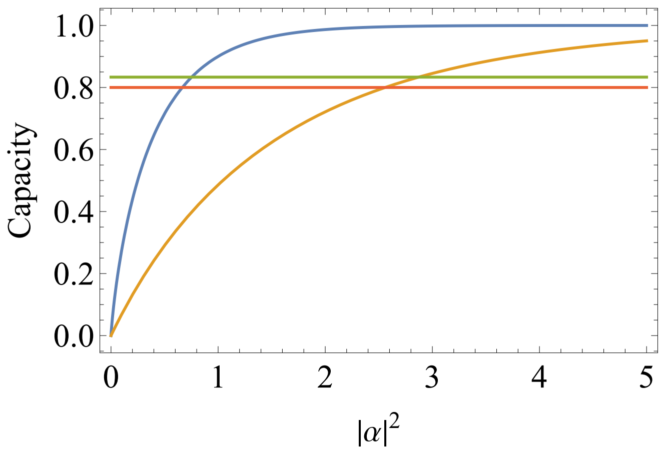


Figure 2: Channel capacities for 2-CPPM signals

that Eve's error probability is worse when the (6,5) single parity-check code is applied. Therefore, further increasing the codeword length is expected to further deteriorate Eve's error probability. On the other hand, Bob's error probability improves over uncoded at the average number of photons above 1.5. Therefore, by operating at the average number of photons near 2, the quantum gain can be widened by coding.

4 Capacity and discussion

Fig.2 shows the channel capacities for Bob and Eve (C_{Bob} and C_{Eve}). Needless to say, $C_{\mathrm{Eve}} < C_{\mathrm{Bob}}$ always holds. The two horizontal lines are the constants 4/5 and 5/6, which represent the rates of the (5,4) and (6,5) codes (denoted $R_{(5,4)}, R_{(6,5)}$).

From Fig.2, when the average number of photons satisfies

$$0.5 < |\alpha|^2 < 2.5, \tag{13}$$

then

$$C_{\mathrm{Eve}} < R_{(5,4)}, R_{(6,5)} < C_{\mathrm{Bob}}, \tag{14}$$

and the rate R satisfies the condition of Eq.(11), which is expected to be effective in coding to improve the security. In section 3, we mentioned that the coding is effective when the average number of photons is near 2, which is consistent with the results for the capacity.

5 Conclusion

We introduced that the quantum gain can be enhanced when the average number of photons is near 2 by applying the (5,4) or (6,5) single parity-check code as an outer code to 2-CPPM type KCQ. By comparing the capacity and coding rates of a legitimate receiver and an attacker for 2-CPPM, we show that coding is expected to be effective for the (5,4) and (6,5) codes with the average number of photons N_s of about $0.5 < N_s < 2.5$. Although the capacity is usually taken to tell an asymptotic behavior of the coding rate for very long codeword lengths, it is found to be useful in protocol design even for short codeword lengths such as 5 and 6. In the future, we will investigate whether choosing the code to be applied to the KCQ based on the capacities of a legitimate receiver and an attacker can enhance security at the required the average number of photons.

Acknowledgments: This work has been supported in part by JSPS KAKENHI Grant Number 20H00581, 20K20397, 21K04064, 22K20437, and The Nitto Foundation.

References

[1] H. P. Yuen, "KCQ: A new approach to quantum cryptography I. general principles and key generation," arXiv:quant-ph/0311061v6, (2004).

[2] Y. Wakahara, S. Takahira, S. Usami, and T. S. Usuda, "Applying outer codes to CPPM-type quantum cipher," 22nd Asian Quantum Information Science Conference (AQIS2022), (2022).

[3] A. D. Wyner, "The wire-tap channel," Bell System Technical Journal **54**, pp.1355-1387, (1975).

[4] C. W. Helstrom, *Quantum detection and estimation theory*, Academic Press, New York, (1976).

[5] M. Sohma and O. Hirota, "Coherent pulse position modulation quantum cipher supported by secret key," Tamagawa Univ. Quant. ICT Res. Inst. Bulletin **1**, 1, pp.15-19, (2011).

Observation and analysis of photoconductive charging effects in a semiconductor-based ion trap chip

Daun Chung[1 2 8 *] Woojun Lee[1 2 3 8 †] Honggi Jeon[2 4] Beomgeun Cho[1 2]
KwangYeul Choi[1 2 5] SeungWoo Yoo[1 2 5] Changhyun Jung[2 5 6] Junho Jeong[2 5 6]
Dong-Il “Dan” Cho[2 5 6] Taehyun Kim[1 2 3 5 7 ‡]

[1] *Dept. of Computer Science and Engineering, Seoul National University, Seoul 08826, South Korea*
[2] *Automation and Systems Research Institute, Seoul National University, Seoul 08826, South Korea*
[3] *Institute of Computer Technology, Seoul National University, Seoul 08826, South Korea*
[4] *Dept. of Physics and Astronomy, Seoul National University, Seoul 08826, South Korea*
[5] *Inter-university Semiconductor Research Center, Seoul National University, Seoul 08826, South Korea*
[6] *Dept. of Electrical and Computer Engineering, Seoul National University, Seoul 08826, South Korea*
[7] *Institute of Applied Physics, Seoul National University, Seoul 08826, South Korea*
[8] *These authors have contributed equally to this work.*

Abstract. Charging effects in ion traps are one of the main culprits of motional decoherence, complicating the physical control of ions and limiting the performance of qubit operations. The most commonly reported phenomenon is the long-term dielectric charging effect that occurs from stray charges produced on the surfaces of insulating materials in the vicinity of the ions. Here, we report on a photoconductive charging effect, a rather poorly studied subject to date, which mainly arises in semiconductor-based microfabricated ion trap chips. A Raman transition setup that utilizes the trapped ion as a quantum sensor is used to measure the stray field originating the exposed semiconductor surfaces in our ion trap chip. Observations are interpreted through a surface photovoltage (SPV) model which reproduces results that are in good agreement with experimental data.

Keywords: Ion trap, charging, semiconductor, surface photovoltage

1 Introduction

Ion traps have been demonstrated to be promising platforms for the physical realization of quantum computers [1, 2]. In particular, microfabricated ion trap chips are leveraging semiconductor fabrication technologies to make ever compact and elaborate ion traps, while scaling up the platform at the same time. Such chips, however, suffer from charging effects from semiconductors whose exact origin remains illusive despite numerous studies reporting its existence and experimental methods to mitigate such effects [3, 4, 5, 6, 7].

Here, we report on the photoconductive charging effect which is distinct from the more recognized dielectric charging [8, 9, 10] which arises from stray photo-induced charges residing on the surfaces of insulating materials. Whereas the latter is visible over large time scales ranging from several seconds to hours, the former is detectable in shorter time scales within the range of a few microseconds to tens of milliseconds in our ion trap chip. Since the time scale is comparable to that of the time evolution of the ion, significant motional decoherence occurs as the ion is displaced from its null position due to the stray field, thereby limiting the performance of the quantum operation.

In our study, we performed thorough measurements of the stray electric field generated from surfaces of the exposed semiconductor substrate, p-type silicon, by using a $^{171}\mathrm{Yb}^{+}$ trapped ion as a quantum sensor [16]. A beam configuration for Raman transitions was set up in a way such that the ion displacement in the direction of the stray field could be extracted. Large ion displacements were observed, even in the NIR wavelengths, raising a need for a correction to the common belief spread in the ion trap community that NIR lasers are typically harmless to ion trap performance.

We present a photoconductive charging model in the context of surface photovoltage (SPV) [11, 12, 13, 14] that can fully explain the general features observed in our ion trap system. The model is based on the semiconductor equations [15] subject to surface conditions that involve surface states that can exchange charges with the bulk through surface absorption (of light) and recombination. We find that surface conditions on the exposed silicon surfaces in our chip is very different from the typically oxidized state of the silicon surface. The formation of such surface states is presumed to have occurred through the microfabrication process.

2 Experimental setup and measurement method

The structure of our silicon-based microfabricated $^{171}\mathrm{Yb}^{+}$ trap chip and experimental setup for measurement of the stray electric field using the Raman transition is shown in Fig. 1 (for more detailed descriptions of the chip architecture, refer to Ref [17]). The direct current (dc) and radio frequency (rf) electrodes to create the trap potential are on the surface of the chip and are made

[*] acornzet@snu.ac.kr
[†] ujun0909@snu.ac.kr
[‡] taehyun@snu.ac.kr

of aluminium (Al), which is additionally coated by gold (Au) near the trap region to avoid oxidation. There is an extended loading slot along the trap axial direction with a width of 80 µm which penetrates the chip. The exposed silicon surface which causes the laser-induced photoconductive charging resides under the electrodes of the chip.

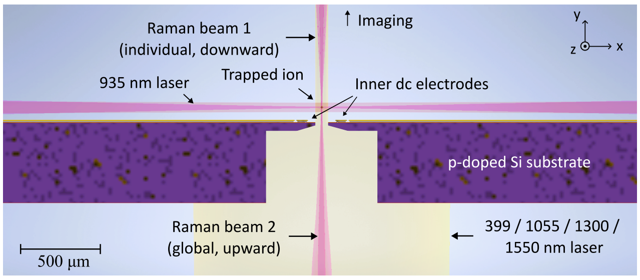


Figure 1: Schematic cross-sectional illustration of the microfabricated chip with parallel and perpendicularly incident laser beams

A quantitative estimation of the magnitude of the stray field is perfomed by measuring the change in the qubit transition strength induced by micromotion of the ion, which arises by the ion's displacement from the equilibrium of the rf potential [18]. The qubit transition is achieved by the Raman transition from $|0\rangle = |{}^2S_{1/2}, F = 0, m_F = 0\rangle$ state to $|1\rangle = |{}^2S_{1/2}, F = 1, m_F = 1\rangle$ state, driven by a pair of beatnote-locked 355-nm pulsed laser beams. The Raman beams are aligned perpendicular to the chip so that the ion displacement in the normal direction with respect to the surface of the chip could be detected as it is shifted by the stray field generated from the silicon substrate underneath. As the ion is pushed away from the null position by a stray field, it experiences phase modulation relative to the laser, induced by the micromotion oscillating at the frequency of the rf field, which reduces the excitation strength to the $|1\rangle$ qubit state. By tuning the voltages of the inner electrodes to compensate this stray field, its magnitude and sign can be determined.

3 Results

Image of the ion displacement by an NIR laser Prior to quantitative estimation of the laser-induced stray field, ion's displacement by the field which is caused by scattering of an NIR laser, or the repumping 935-nm laser in our setup, was witnessed in imaging of the ion. The 935-nm laser passes through the chip surface with a separation of 80 µm. Fig. 2 shows images of the ion captured by an electron-multiplying charge-coupled device (EMCCD) after collected by an imaging lens of 0.6 NA, becoming defocused as it is displaced from the null position by scattering of the 935-nm laser with an increasing input power from 100 µW to 1.8 mW (see Fig. 1 for the configuration of the laser). The defocusing is clearly noticible and the maximum displacement of the ion is estimated to be around 8 µm which can be determined by translating the EMCCD until the ion is re-focused.

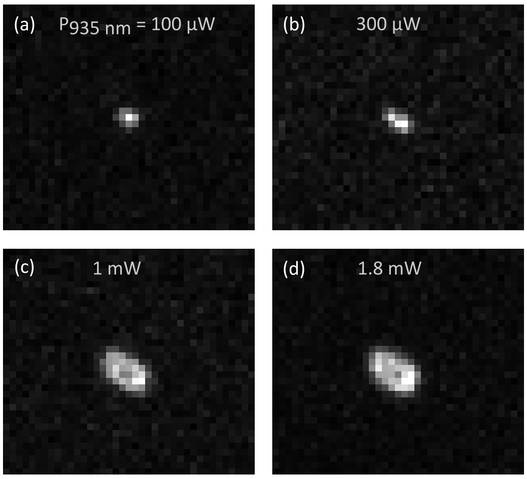


Figure 2: (a)-(d) Images of the ion displaced by the stray field from silicon substrate induced by scattered light as the power of the 935-nm laser was increased.

Spectral measurement of the SPV and theoretical results Ion displacements provide information about the stray field through which the SPV at the silicon surface can be estimated. Fig. 3 (a) shows the spectral signal of the SPV rescaled to the value of the probing compensation voltage applied on the electrodes in the ion trap chip. Lasers of three wavelengths, two in the NIR range, 1055 nm and 1300 nm, and one in the UV range, 399 nm, have been uniformly illuminated on the silicon surface exposed at the backside of the chip. The SPV was measured to be stronger for 1055 nm than 399 nm, in contrast to the common perception in the ion trap community that UV light is the dominant wavelength responsible for producing charging effects.

This is not unusual when we consider surface absorption from surface states whose theoretical formulations are provided in the references [19, 20, 21, 22]. Charge carriers initially populated in the surface states are optically excited to the conduction bands and diffuse into the bulk where they recombine with other free charge carriers. For certain surface states, the charge redistribution can result in an enhancement of the surface potential, hence a positive SPV for a p-type semiconductor as in our case. The magnitude of the SPV increases as more charge carriers are depopulated from the surface states, ultimately reaching saturation at zero occupation.

The theoretical absorption spectrum appropriate for our system (determined by the so called quantum defect number ν) is plotted in Fig. 3 (b). The numerical values of the SPV solved by our photoconductive charging model is plotted in Fig. 3 (a) in dashed lines. Reported values have been used for bulk parameters, whereas the surface state parameters have been fitted to agree with experimental data. The quantum defect number was estimated to be $\nu = 1$ from the spectral response of the SPV. This implies that the surface states responsible for the signals observed in our ion trap chip are hydrogenic in character [22].

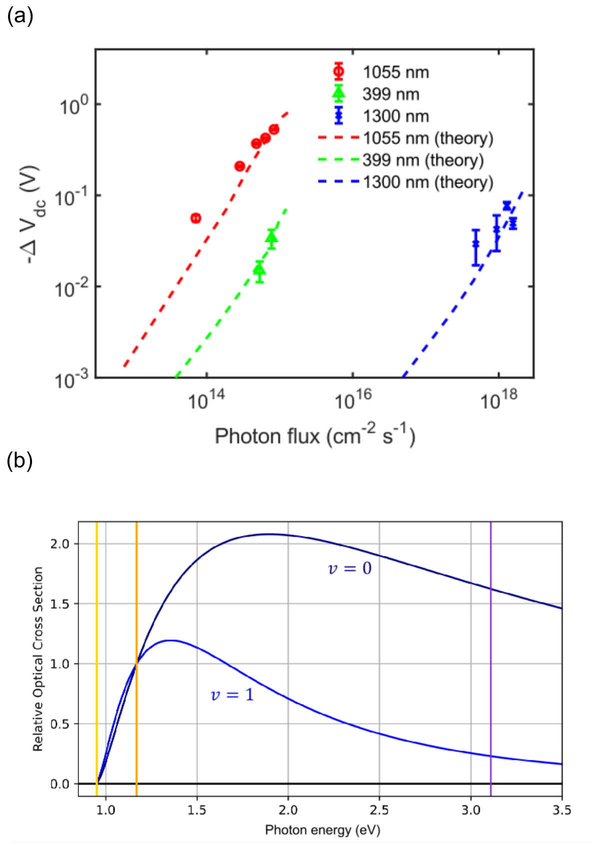


Figure 3: (a) Spectral response of laser-induced electric fields. The error bars in the plots indicate the 95 % confidence intervals of the fitting in each measurement. (b) Optical cross section by photon energy (normalized with respect to 1055 nm) for quantum defect numbers $\nu = 0, 1$ and the values for the selected wavelengths in our measurement (yellow-1300 nm, orange-1055 nm, purple-399 nm). Our observations are explained by the spectrum corresponding to $\nu = 1$.

4 Conclusion

A study on photoconductive charging effects in a semiconductor-based microfabricated ion trap chip was conducted. Significant displacements of trapped ions due to stray fields from exposed silicon surfaces were measured and understood through a photoconductive charging model based on the SPV effect. The framework presented for estimating the SPV is not limited to the surface conditions used in our work, but may be extended to predict behaviors for other types of surfaces as well.

References

[1] D. Stick, W. K. Hensinger, S. Olmschenk, M. J. Madsen, K. Schwab, and C. Monroe. Ion trap in a semiconductor chip. *Nat. Phys.*, 2:36–39, 2006.

[2] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, D. M. Lucas. High-fidelity preparation, gates, memory, and readout of a trapped-ion quantum bit. *Phys. Rev. L.*, 113(22):220501, 2014.

[3] K. K. Mehta,, A. M. Eltony, C. D. Bruzewicz, I. L. Chuang, R. J. Ram, J. M. Sage, J. Chiaverini. Ion traps fabricated in a CMOS foundry. *Appl. Phys. L.*, 105(4):044103, 2014.

[4] M. Niedermayr, K. Lakhmanskiy, M. Kumph, S. Partel, J. Edlinger, M. Brownnutt, R. Blatt. Cryogenic surface ion trap based on intrinsic silicon. *New J. Phys.*, 16(11):113068, 2014.

[5] D. Stick, K. M. Fortier, R. Haltli, C. Highstrete, D. L. Moehring, C. Tigges, M. G. Blain. Demonstration of a microfabricated surface electrode ion trap. Arxiv, 2018.

[6] M. G. .Blain, R. Haltli, P. Maunz, C. D. Nordquist, M. Revelle, D. Stick. Hybrid MEMS-CMOS ion traps for NISQ computing. *Quantum. Sci. Tech.*, 6:034011, 2021.

[7] S. Auchter, C. Axline, C. Decaroli, M. Valentini, L. Purwin, R. Oswald, R. Matt, E. Aschauer, Y. Colombe, P. Holz, T. Monz, R. Blatt, P. Schindler, C. Rossler, J. Home. Industrially microfabricated ion trap with 1 eV trap depth. *Quantum. Sci. Tech.*, 7:035015, 2022.

[8] M. Harlander, M. Brownnutt, W. Hänsel, R. Blatt. Trapped-ion probing of light-induced charging effects on dielectrics. *New J. Phys.*, 12(9):093035, 2010.

[9] S. X. Wang, G. Hao Low, N. S. Lachenmyer, and Y. Ge, and P. F. Herskind, and I. L. Chuang. Laser-induced charging of microfabricated ion traps. *J. Appl. Phys.*, 110(10):104901, 2011.

[10] A. Härter, A. Krükow, A. Brunner, J. Hecker Denschlag. Long-term drifts of stray electric fields in a Paul trap. *App. Phys. B.*, 114(1):275–281, 2014.

[11] W. H. Brattain, J. Bardeen. Surface properties of Germanium. *Bell Sys. Tech. J.*, 32(1):1–41, 1953.

[12] C. G. B. Garrett, W. H. Brattain. Physical Theory of Semiconductor Surfaces. *Phys. Rev.*, 99(2):376–387, 1955.

[13] E. .O .Johnson. Large-signal surface photovoltage studies with Germanium. *Phys. Rev.*, 111(1):153–166, 1958.

[14] L. Kronik, Y. Shapira. Surface photovoltage phenomena: theory, experiment, and applications. *Surf. Sci. Rep.*, 37(1-5):1–206, 1999.

[15] W. Van Roosbroeck. Theory of the flow of electrons and holes in germanium and other semiconductors. *Bell Sys. Tech. J.*, 29(4):560–607, 1950.

[16] R. Maiwald, D. Leibfried, J. Britton, J. C. Bergquist, G. Leuchs, and D. Wineland. Stylus ion trap for enhanced access and sensing. *Nat. Phys.*, 2:551–554, 2009.

[17] C. Jung, W. Lee, J. Jeong, M. Lee, Y. Park, T. Kim, and D. ”Dan” Cho. A microfabricated ion trap chip with a sloped loading slot to minimize exposing trapped ions to stray charges. *Quantum Sci. Technol.*, 6(4):044004, 2021.

[18] D. J. Berkeland, J. D. Miller, J. C. Bergquist, W. M. Itano, and D. J. Wineland Minimization of ion micromotion in a Paul trap. *J. Appl. Phys.*, 83:5025-–5033, 1998.

[19] W. Shockley, W. T. Read. Jr. Statistics of the recombinations of holes and electrons. *Phys. Rev. B.*, 87(5):835-842, 1952.

[20] R. N. Hall. Recombination processes in semiconductors. *IEEE*, 106(17S):923–931, 1959.

[21] Y. K. Hsieh, H. C. .Card. Limitation to Shockley–Read–Hall model due to direct photoionization of the defect states. *J. Appl. Phys.*, 65(6):2409–2415, 1989.

[22] S. Chaudhuri. Optical-transition cross sections involving impurities in semiconductors. *Phys. Rev. B.*, 26(12):6593–6602, 1982.

Interplay of nonlocality and incompatibility breaking qubit channels

Swati Kumari,[1,*] Javid Naikoo,[2] Sibasish Ghosh,[3] and A. K. Pan[4]

[1]Department of Physics and Center for Quantum Frontiers of Research & Technology (QFort), National Cheng Kung University, Tainan 701, Taiwan
[2]Centre for Quantum Optical Technologies, Centre of New Technologies, University of Warsaw, Banacha 2c, 02-097 Warsaw, Poland
[3]Optics and Quantum Information Group, The Institute of Mathematical Sciences, HNBI, CIT Campus, Taramani, Chennai 600113, India
[4]Department of physics, Indian Institute of Technology Hyderabad Kandi,Telengana, India

Incompatibility and nonlocality are not only of foundational interest but also act as important resources for quantum information theory. In CHSH (Clauser–Horne–Shimony–Holt) scenario, the incompatiblity of a pair of observables is known to be equivalent to Bell nonlocality. Here, we investigate these notions in the context of qubit channels. The Bell-CHSH inequality has a greater perspective – compared to any genuine tri-partite nonlocality scenario – while determining about the interplay between nonlocality breaking qubit channels and incompatibility breaking qubit channels. In Bell CHSH scenario, we prove that if the conjugate of a channel is incompatibility breaking, then the channel is itself nonlocality breaking and vice versa. However, this equivalence is not straightforwardly generalized to multi-partite systems, due to the absence of an equivalence relation between incompatiblity and nonlocality in the multi-partite scenario. We investigate this relation in tripartite scenario by considering some well known states like GHZ and W states and using the notion of Mermin and Svetlichny nonlocality. By subjecting the parties in question to unital qubit channels, we identify the range of state and channel parameters for which incompatiblity coexists with nonlocality. Further, we identify the set of unital qubit channels that is Mermin/Svetlichny nonlocality breaking *irrespective* of the input state.

I. INTRODUCTION

Nonlocality is one of the profound notions in quantum mechanics [1] and is often talked in conjunction with incompatibility of observables. Recent developments in quantum information theory have found nonlocality as a useful phenomenon underpinning many advantages afforded by various quantum information processing tasks [2]. Nonlocality can also be considered as a potential quantum resource for information processing, such as in developing quantum protocols to reduce the amount of communication needed in certain computational tasks [2] and providing secure quantum communications [3, 4]. Incompatibility, like nonlocality, is not merely of theoretical interest but of practical utility, for example, in order to explore the advantage of entanglement shared by two parties in a cryptography task, each party needs to carry out measurements that are incompatible, in the sense that these cannot be carried out simultaneously by a single measurement device. Incompatibility should not be confused with noncommutativity or the related concept of uncertainty principle. The notion of incompatibility is best understood in terms of joint measurability [5]. A collection of quantum measurements is jointly measurable, if it can be simulated by a single common quantum measurement device. If such a single common device cannot be constructed by a given set of quantum measurements, it then enables the set to be used as a quantum resource. This was first noted in [6] in the context of CHSH inequalities and later in the EPR steering, which is more explicit, when incompatibility appears as a quantum resource. Incompatibility is necessary and sufficient for the violation of the steering inequalities [7, 8]. The relation between incompatibility and contextuality has also been studied in references [9, 10]. Further, a set of observables that is pairwise incompatible, but not triplewise can violate the Liang-Spekkens-Wiseman noncontextuality inequality [11]. Recently, the connection between steerability and measurement incompatibility was studied in [12] in the context of the so called steerability equivalent observables. Thus, both nonlocality and incompatibility can be considered as quantum resources whose understanding is of utmost importance in view of emerging quantum technologies.

The interplay of nonlocality and incompatibility has been a subject matter of various studies. It is well known that any incompatible local measurements, performed by the constituent parties of a system, lead to the violation of Bell inequality provided they share a pure entangled state [1, 2]. Absence of either of them (i.e., entanglement and incompatibility) will not allow the system to exhibit nonlocality. It is important to mention here that a notion of quantum nonlocality without entanglement has been proposed in [13] which is different from Bell nonloclaity [1] and amounts to the inability of discriminating a set of product states by local operations and classical communication, while mutual orthogonality of the states assures their perfect global descrimination.

Further, for any pair of dicohotomic incompatible observables, there always exists an entangled state which enables the violation of a Bell inequality [6]. The relationship of incompatibility and nonlocality is sensitive to the dimension of system, for example, increasing the dimension beyond two, the incompatible observables do not necessarily lead to the violation of Bell type inequalities implying that the measurement incompatibility can not guarantee nonlocality in general [14, 15]. Here, we probe the interplay between incompatibility and nonlocality in tripartite case by using the well known Mermin and Svetlichny inequality [16]. Svetlichny inequal-

[*] swatipandey084@gmail.com

ity, unlike Mermin inequality, is a genuine measure of nonlocality that assumes nonlocal correlations between two parties which are locally related to a third party and is known to provide a suitable measure to detect tripartite nonlocality for W and GHZ class of states [17]. We refer the interested readers to [2, 18] for various facets of the multipartite nonlocality.

A nonlocality breaking channel (NBC) can be defined as a channel which when applied to a system (or part of it) leads to a state which is local [19], while as incompatibility breaking channel (IBC) is the one that turns incompatible observables into compatible ones [20, 21]. An IBC that renders any set of n incompatible observables compatible would be denoted by $n-\mathbf{IBC}$. The notion of NBC has been introduced in a similar spirit of well-studied entanglement breaking channel [22]. Every entanglement breaking channel is nonlocality breaking but the converse is not true. As an example, the qubit depolarising channel $\mathcal{E}(\rho) := p(\rho) + (1-p)\mathbf{I}/2$ is CHSH nonlocality breaking for all $\frac{1}{3} \leq p \leq \frac{1}{2}$, but not entanglement breaking [19]. Hence, based on this classification, nonlocality and entanglement emerge as different resources.

The equivalence of the steerability breaking channels and the incompatibility breaking channels was reported in [23] and CHSH nonlocality breaking channels were shown to be a strict subset of the steerability breaking channels [24]. The connection between Bell nonlocality and incompatibility of two observable is well understood, however, the question of the equivalence between NBC [19] and IBC [21] is rarely discussed. This motivates us to explore the relation between CHSH nonlocality breaking channels (CHSH-NBC) and $2-\mathbf{IBC}$, such that the action of one may be replaced by the other. The tripartite nonlocality has much more richer and complex structure and less is known about its synergy with incompatibility as compared to its bipartite counterpart. Mermin inequality assumes local-realistic correlations among all the three qubits; hence a violation would be a signature of the tripartite nonlocality shared among the qubits. However, biseparable states were shown to also violate the Mermin inequality [3, 25]. This motivated Svetlichny to introduce the notion of genuine tripartite nonlocality [16] and provided a set of inequalities sufficient to witness it. We make use of these notions of *absolute* and *genuine* nonlocality to figure out the ranges of state and channels parameters in which NBC and $2-\mathbf{IBC}$ coexist.

II. RESULTS AND DISCUSSION

A. Equivalance of CHSH nonlocality breaking and incompatibility breaking channels

Our first result establishes an *equivalence* of the CHSH nonlocality breaking channel acting on one party, with its *dual* being an incompatibility breaking channel – in the context of $2-\mathbf{IBC}$s. The result can be summarized by the following two theorems:

Theorem 1. *If the conjugate of a qubit channel $\mathcal{E}$ is* $2-\mathbf{IBC}$ *, then the channel itself is CHSH-NBC.*

Proof. Consider the Bell-CHSH scenario, such that $[A_1, A_2] \neq 0$ and $[B_1, B_2] \neq 0$, i.e., the operators A_1, A_2 and B_1, B_2 are incompatible in *conjunction*. Let $\mathcal{E}^\dagger$ be the conjugate channel that is $2-\mathbf{IBC}$. Then the action of this channel on Alice's side makes A_1 and A_2 compatible, i.e., $[A_1, A_2] = 0$. Therefore, the Bell-CHSH inequality is not violated [6], and we have

$$\mathrm{Tr}\left[\rho\mathcal{B}\left(\mathcal{E}^\dagger[A_1], \mathcal{E}^\dagger[A_2], B_1, B_2\right)\right] \leq 2. \tag{1}$$

Or,

$$\mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_1]\otimes B_1\right] + \mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_2]\otimes B_1\right] + \mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_1]\otimes B_2\right] - \mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_2]\otimes B_2\right] \leq 2. \tag{2}$$

This can be viewed in the Schrödinger picture as

$$\mathrm{Tr}\left[(\mathcal{E}\otimes\mathbf{I})[\rho]A_1\otimes B_1\right] + \mathrm{Tr}\left[(\mathcal{E}\otimes\mathbf{I})[\rho]A_2\otimes B_1\right] + \mathrm{Tr}\left[(\mathcal{E}\otimes\mathbf{I})[\rho]A_1\otimes B_2\right] - \mathrm{Tr}\left[(\mathcal{E}\otimes\mathbf{I})[\rho]A_2\otimes B_2\right] \leq 2, \tag{3}$$

which tells us that the CHSH inequality is satisfied even when operators A_1, A_2 and B_1, B_2 are incompatible in *conjunction*. Therefore, the action of $\mathcal{E}$ (to be precise of $\mathcal{E}\otimes\mathbf{I}$) on state ρ is solely responsible for non-violation of the CHSH inequality. We conclude that $\mathcal{E}$ is CHSH-NBC. ■

Theorem 2. *If a qubit channel $\mathcal{E}$ is CHSH-NBC, then its conjugate is* $2-\mathbf{IBC}$.

Proof. Here we start with incompatible operators associated with the respective subsystems, $[A_1, A_2] \neq 0$ and $[B_1, B_2] \neq 0$ and assume that the channel $\mathcal{E}$ acting on Alice's side does not allow for the violation of CHSH inequality, that is

$$\mathrm{Tr}\left[(\mathcal{E}\otimes\mathbf{I})\,[\rho]\mathcal{B}\,(A_1, A_2, B_1, B_2)\right] \leq 2. \tag{4}$$

In other word, looking from the measurement point of view, in Heisenberg picture, we have

$$\mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_1]\otimes B_1\right] + \mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_1]\otimes B_2\right] + \mathrm{Tr}\left[\rho\mathcal{E}^\dagger[A_2]\otimes B_1\right] - \mathrm{Tr}[\rho\mathcal{E}^\dagger[A_2]\otimes B_2] \leq 2 \tag{5}$$

The above inequality holds for arbitrary state ρ which can even be an entangled state. Thus the non-violation of CHSH inequality is coming from the action of $\mathcal{E}^\dagger$ on the operators A_1 and A_2, making them compatible, $\left[\mathcal{E}^\dagger[A_1], \mathcal{E}^\dagger[A_2]\right] = 0$. We conclude that $(\mathcal{E}^\dagger\otimes\mathbf{I})$ is incompatibility breaking. ■

B. Nonlocality and incompatibility breaking channels in tripartite scenario

The conditions for Mermin and Svetlichny nonlocality breaking channel obtained by the application of a unital quantum channel to one party of a tripartite system, are depicted in Fig. 1. All the points below solid (black) and dashed (red)

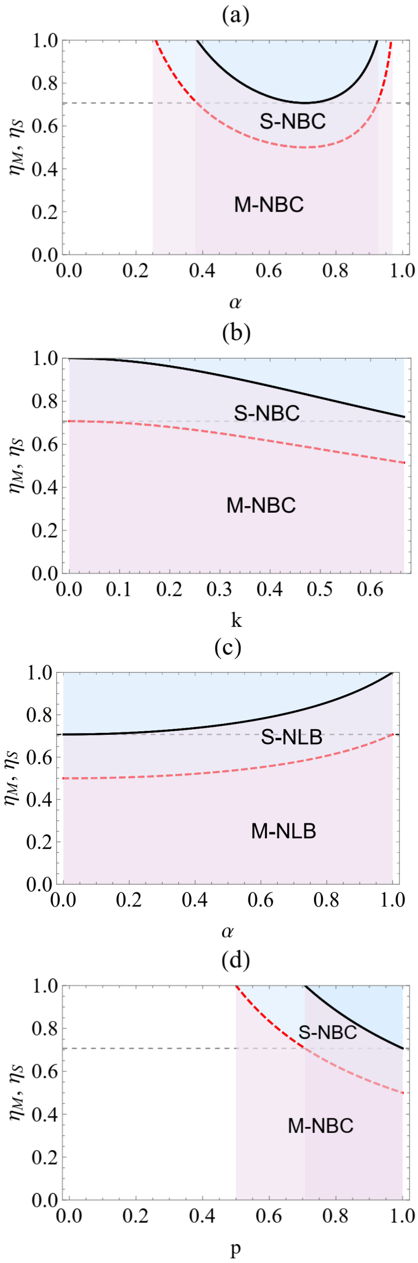


FIG. 1: (Color online): Region below the dashed (red) and solid (black) curve in (a), (b), (c), and (d) corresponds to Mermin and Svetlichny nonlocality breaking conditions (M-NBC and S-NBC) plotted against the (dimensionless) state coefficients. The pairwise incompatibility breaking condition pertains to all points below the horizontal dashed line.

curve correspond to nonlocality breaking channel, while as the points below the horizontal dashed line, $\eta = 1/\sqrt{2}$, pertain to pairwise incompatibility breaking. In all the four examples, Fig. 1 (a)-(d), the minimum value of η_S for which SI is violated is $1/\sqrt{2}$, suggesting that *genuine* nonlocal correlations can not be established if at least one pair of observable is compatible. The converse is not true, since there exist regions (above the horizontal dashed line and below the solid (black) curve) of Svetlichny nonlocality breaking even when the channel is not $2-$ **IBC**. Thus, these examples illustrate that *corresponding to* $2-$**IBC** *the conjugate channels are definitely S-NLB; however, the conjugate of S-NLB channels may not necessarily be a* $2-$**IBC**. However, in the context of Mermin nonlocality, even the first statement does not hold, that is, *existence of a* $2-$ **IBC** *does not necessarily gaurentee a conjugate channel that is M-NLB.* Also, the minimum η_M (that is maximum noise) for which a channel is M-NLB is always less by a factor of $1/\sqrt{2}$ than the minimum noise below which that channel is S-NLB. Note that instead of one party, if two or all the three parties are subjected to noise, the NBC conditions become

$$\eta_M \le \left(\frac{1}{\sqrt{2}\lambda_{max}}\right)^{1/n}, \qquad \eta_S \le \left(\frac{1}{\lambda_{max}}\right)^{1/n}, \tag{6}$$

where n corresponds to the number of qubits subjected to noise. Since $1/\lambda_{max} < (1/\lambda_{max})^{1/2} < (1/\lambda_{max})^{1/3}$ (with $\lambda_{max} > 1$), the solid (black) and dashed (red) curves in Fig. 1 (a)-(d), are shifted up, thereby decreasing the region of nonlocality with increase in n.

III. CONCLUSION

This work is devoted to a study of the interplay between nonlocality breaking and incompatibility breaking power of noisy quantum qubit channels. The action of quantum channels on projective measurements transforms them into noisy POVMs, characterized in particular by unsharpness parameters. As a consequence, noise tends to increase the compatibility of observers that are otherwise incompatible. In fact, pairwise incompatibility breaking is assured if the channel parameter is less than or equal to $1/\sqrt{2}$. To be specific, we consider bipartite and tripartite scenarios, with CHSH nonlocality in the former and Mermin and Svetlichny nonlocality in the later case. The degree of incompatibility breaking directly depends on the unsharpness parameters. Here, we showed that in Bell CHSH scenario, if the conjugate of a channel is incompatibility breaking then the channel is itself nonlocality breaking and the converse is also true. In tripartite scenario, however, such an equivalence between nonlocality breaking and incompatibility breaking does not exist. We then consider various examples of three qubit states and identify the state parameters for which the equivalence of nonlocality breaking corroborates with the pairwise incompatibility. In particular, it is illustrated that the conjugate of incompatiblity breaking channels are nonlocality breaking, however, the nonlocality breaking channels do not guarantee the existence of conjugate channels that are incompatibility breaking. This may be viewed as a useful feature of the Bell-CHSH inequality when it comes to the study incompatibility of observables. Further, from randomly generated three qubit states subjected to general unital channels, we conclude that no Mermin (Svetlichny) nonlocal correlations are supported for $\eta_M < 0.090$ ($\eta_S < 0.128$) .

[1] J. S. Bell, Physics Physique Fizika **1**, 195 (1964).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).

[3] V. Scarani and N. Gisin, Phys. Rev. Lett. **87**, 117901 (2001).

[4] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[5] T. Heinosaari, T. Miyadera, and M. Ziman, Journal of Physics A: Mathematical and Theoretical **49**, 123001 (2016).

[6] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).

[7] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).

[8] R. Uola, C. Budroni, O. Gühne, and J.-P. Pellonpää, Phys. Rev. Lett. **115**, 230402 (2015).

[9] S. KOCHEN and E. P. SPECKER, Journal of Mathematics and Mechanics **17**, 59 (1967).

[10] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-Å. Larsson, arXiv preprint arXiv:2102.13036 (2021).

[11] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, Physics Reports **506**, 1 (2011).

[12] H.-Y. Ku, C.-Y. Hsieh, S.-L. Chen, Y.-N. Chen, and C. Budroni, Nature Communications **13** (2022), 10.1038/s41467-022-32466-y.

[13] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **59**, 1070 (1999).

[14] E. Bene and T. Vértesi, New Journal of Physics **20**, 013021 (2018).

[15] F. Hirsch, M. T. Quintino, and N. Brunner, Phys. Rev. A **97**, 012129 (2018).

[16] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).

[17] A. Ajoy and P. Rungta, Phys. Rev. A **81**, 052334 (2010).

[18] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, Phys. Rev. A **88**, 014102 (2013).

[19] R. Pal and S. Ghosh, Journal of Physics A: Mathematical and Theoretical **48**, 155302 (2015).

[20] T. Heinosaari, J. Kiukas, D. Reitzner, and J. Schultz, Journal of Physics A: Mathematical and Theoretical **48**, 435301 (2015).

[21] T. Heinosaari and T. Miyadera, Journal of Physics A: Mathematical and Theoretical **50**, 135302 (2017).

[22] M. Horodecki, P. W. Shor, and M. B. Ruskai, Reviews in Mathematical Physics **15**, 629641 (2003).

[23] J. Kiukas, C. Budroni, R. Uola, and J.-P. Pellonpää, Phys. Rev. A **96**, 042331 (2017).

[24] H.-Y. Ku, J. Kadlec, A. Cernoch, M. T. Quintino, W. Zhou, K. Lemr, N. Lambert, A. Miranowicz, S.-L. Chen, F. Nori, and Y.-N. Chen, PRX Quantum **3**, 020338 (2022).

[25] D. Collins, N. Gisin, S. Popescu, D. Roberts, and V. Scarani, Phys. Rev. Lett. **88**, 170405 (2002).

[26] P. Busch, Phys. Rev. D **33**, 2253 (1986).

[27] M. A. Nielsen and I. Chuang, "Quantum computation and quantum information," (2002).

[28] S. Yu, N.-l. Liu, L. Li, and C. H. Oh, Phys. Rev. A **81**, 062116 (2010).

[29] T. Heinosaari, Journal of Physics: Conference Series **1638**, 012002 (2020).

[30] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).

[31] J. L. Cereceda, Phys. Rev. A **66**, 024102 (2002).

[32] P. Mitchell, S. Popescu, and D. Roberts, Phys. Rev. A **70**, 060101 (2004).

[33] S. Ghose, N. Sinclair, S. Debnath, P. Rungta, and R. Stock, Phys. Rev. Lett. **102**, 250404 (2009).

[34] Z. Su, L. Li, and J. Ling, Quantum Information Processing **17**, 1 (2018).

[35] M. A. Siddiqui and S. Sazim, Quantum Information Processing **18** (2019), 10.1007/s11128-019-2246-1.

[36] M. Li, S. Shen, N. Jing, S.-M. Fei, and X. Li-Jost, Phys. Rev. A **96**, 042323 (2017).

[37] M. Beth Ruskai, S. Szarek, and E. Werner, Linear Algebra and its Applications **347**, 159 (2002).

[38] J. Naikoo, S. Banerjee, A. K. Pan, and S. Ghosh, Phys. Rev. A **104**, 042608 (2021).

[39] G. Kar and S. Roy, Physics Letters A **199**, 12 (1995).

[40] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[41] D. M. Greenberger, M. A. Horne, and A. Zeilinger, "Going beyond bell's theorem," (2007), arXiv:0712.0921 [quant-ph].

[42] H. A. Carteret and A. Sudbery, Journal of Physics A: Mathematical and General **33**, 4981 (2000).

[43] J. Liu, Z.-w. Mo, and S.-q. Sun, International Journal of Theoretical Physics **55**, 2182 (2016).

[44] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, Phys. Rev. Lett. **115**, 030404 (2015).

[45] C. King and M. B. Ruskai, IEEE Transactions on information theory **47**, 192 (2001).

[46] A. Acín, A. Andrianov, E. Jané, and R. Tarrach, Journal of Physics A: Mathematical and General **34**, 6725 (2001).

[47] Y. Zhang, R. A. Bravo, V. O. Lorenz, and E. Chitambar, New Journal of Physics **22**, 043003 (2020).

A Scalable Fault-Tolerant Three-Dimensional Cluster State Construction Protocol using Linear Arrays of Emitters

Jintae Kim[1] * Isaac H. Kim[2] †

[1] *Department of Physics, Sungkyunkwan University, Suwon 16419, Korea*
[2] *Department of Computer Science, University of California, Davis, CA 95616, USA*

Abstract. Scalability in quantum computing poses challenges due to high overhead in building fault-tolerant quantum computers. A study by Wan et al. (PRX Quantum 2, 2021) proposed fault-tolerant protocols for constructing three-dimensional cluster states using emitters and waveguides. However, scalability was limited by delay line errors and waveguide length restrictions. In this paper, an improved protocol is presented, distributing cluster state construction across an array of emitter and waveguide pairs. Simulation results demonstrate that threshold values are tolerant with increased emitters and noisy gates, enabling scalability. The proposed protocol also reveals that desired logical error rates can be achieved with higher dephasing error rates than previously believed, emphasizing practical advantages.

Keywords: fault-tolerant quantum computation, measurement-based quantum computation, emitter system

1 Introduction

Significant progress has been made in experiments related to large-scale fault-tolerant quantum computation. An experiment conducted by Google Quantum AI [1] demonstrated that physical superconducting qubits exhibit sufficiently low error rates to realize a lower logical error rate through scaling up the surface code [2, 3, 4, 5] and employing error correction techniques. This achievement represents a crucial advancement towards fault-tolerant quantum computation, as described by the threshold theorem [3, 6, 7, 8, 9, 10].

However, the scaling up of the number of qubits poses a major challenge due to the substantial experimental overhead required. The manufacturing, calibration, and management of a large quantity of qubits present formidable obstacles. Therefore, the successful attainment of sufficient scalability in experimental setups is crucial for the realization of large-scale fault-tolerant quantum computation.

Wan et al. introduced a practical and scalable experimental setup along with protocols [11]. Compared to previous methods [12, 13, 14, 15, 16, 17, 18], these protocols utilize an emitter system, offering enhanced feasibility. Building upon the concept of time-domain multiplexing [19, 20, 21, 22, 23], the protocols achieve fault-tolerant construction of a three-dimensional cluster state capable of universal fault-tolerant measurement-based quantum computation using the surface code [12, 24, 25, 26, 27, 28]. The setup comprises an emitter and a waveguide that facilitates the flow of photons or phonons. By utilizing fewer components, the experimental overhead is expected to decrease, leading to lower logical error rates by enhancing individual components. The qubits become entangled through emitter-mediated interactions, forming the structure of the three-dimensional cluster state. Interactions occur successively as the qubits traverse the waveguide, exclusively between the emitter and the qubits. While the emitter interacts with all qubits, the resulting effective errors stemming from a single emitter or qubit error remain localized [11].

The Monte Carlo simulation results, employing a minimum-weight perfect matching (MWPM) decoder, demonstrate that the logical error rate of the $L \times L \times L$ three-dimensional cluster state construction under open boundary conditions (OBC) decreases as L increases. For Protocol A and B, the circuit error rates below the thresholds of 0.23% and 0.39%, respectively, yield reduced logical error rates [11]. However, the setup is also vulnerable to delay line errors, dephasing errors and loss errors, as the cumulative delay line errors with longer waveguides accompanying larger three-dimensional cluster states cannot be disregarded. The errors from cumulative delay lines are directly proportional to the cross-sectional area of the three-dimensional cluster state. The simulation results for Protocol B indicate that achieving logical error rates $10^{-5}, 10^{-10}, 10^{-15}$ for a static circuit error rate of 10^{-3} requires a dephasing error rate of 1.39×10^{-5}, 2.53×10^{-6}, 1.02×10^{-6} and a loss error rate of 1.4×10^{-4}, 2.4×10^{-5}, 9.5×10^{-6}, respectively, when dephasing errors and loss errors are considered separately. Notably, the required loss error rates are currently more achievable than what can be achieved with state-of-the-art waveguides [11].

In this paper, we propose enhanced three-dimensional cluster state construction protocols based on Wan et al.'s paper, which include multiple emitters and are more resistant to delay line errors. The protocols are fault-tolerant and composed of a number of emitter and waveguide pairs, still allowing for scalability with a small number of experimental components. Each emitter is associated with its own waveguide, allowing interactions between emitters and qubits within their respective waveguides, as well as interactions between neighboring ancilla qubits. The emitters mediate interactions between two data qubits within the same waveguide, facilitating the construction of the three-dimensional cluster state. Each waveguide stores a portion of the cluster state,

*jint1054@gmail.com
†ikekim@ucdavis.edu

and inter-waveguide qubit connections are established through emitter-mediated interactions. By confining emitter interactions to their two nearest neighbors, the emitter and waveguide pairs can be linearly aligned. The introduction of multiple emitters reduces the construction time of the three-dimensional cluster state, leading to shorter waveguides and a decrease in delay line errors.

2 The protocols

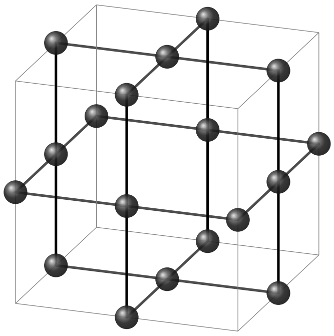


Figure 1: Unit cell of the bcc lattice.

We start by introducing the graph $G = (V, E)$, where the vertices are all numbered, and the edge between the vertices i and j is represented as $\{i, j\} = \{j, i\}$. Then, the cluster state $|\psi_G\rangle$ corresponding to the graph $G = (V, E)$ is defined as

$$|\psi_G\rangle = \prod_{(i,j)\in E} Z_{i,j} \otimes_{i'\in V} |+\rangle_{i'}, \tag{1}$$

where $Z_{i,j}$ is the controlled-Z gate on qubits i and j. Each vertex in the graph corresponds to a data qubit, while each edge represents a controlled-Z gate applied between two data qubits. The three-dimensional cluster state $|\psi_{G_{\mathrm{bcc}}}\rangle$, defined in Eq. (1), is represented by the graph $G_{\mathrm{bcc}} = (V_{\mathrm{bcc}}, E_{\mathrm{bcc}})$, which corresponds to the body-centered cubic (bcc) lattice [12]. The unit cell of the bcc lattice is illustrated in Fig. 1. This bcc structure can be interpreted as the foliation of the surface code cluster state with additional edges [28] and the three-dimensional cluster state can be utilized for the universal fault-tolerant measurement-based quantum computation [12, 24]. The construction of this state involves the initialization of data qubits and the implementation of quantum logic gates.

To construct the three-dimensional cluster state using multiple emitters, the experimental setup involves pairs of emitters and waveguides. In this setup, the ancilla qubit corresponds to the emitter, while the data qubits represent the qubits traveling along the waveguides. Interactions are allowed between the ancilla qubit and the data qubits within a waveguide, as well as between emitters. To facilitate the discussion, we initially consider the case of using two ancilla qubits, denoted as Q_1 and Q_2, to construct the three-dimensional cluster state $|G_{\mathrm{bcc}}\rangle$. We will later address the extension of this protocol to accommodate any number of ancilla qubits. The three-dimensional cluster state graph $G_{\mathrm{bcc}} = (V_{\mathrm{bcc}}, E_{\mathrm{bcc}})$ can be partitioned into two subgraphs, G_1 and G_2, with the inclusion of additional edges. By using $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, we may express G_{bcc} as

$$G_{\mathrm{bcc}} = (V_1 \cup V_2, E_1 \cup E_2 \cup E_{12}), \tag{2}$$

where the set of edges connecting vertices, one in G_1 and one in G_2, is represented by E_{12}. We additionally define V_{12} satisfying $V_{12} \subset V_1 \cup V_2$ and consisting of all the vertices appear in E_{12} for the protocol explanation. The qubits in the graphs G_1 and G_2 are identified by the subscripts 1 and 2, respectively. For the successful implementation of the protocol, it is crucial to divide G_1 and G_2 in a suitable manner. Specifically, each vertex in G_1 must be connected to at most one vertex in G_2, and vice versa. The identification of G_1 and G_2 that satisfy this condition can be easily accomplished, and an illustrative example is provided in Fig. 2.

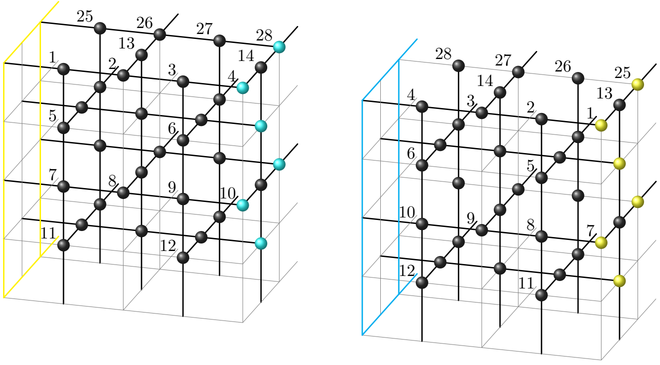


Figure 2: Figure of the part of $8\times8\times8$ three-dimensional cluster state under PBC is depicted. Subscripts of the data qubits are omitted for simplicity.

Figure 2 illustrates a $8\times8\times8$ three-dimensional cluster state that is partitioned into two subgraphs, namely G_1 (left graph) and G_2 (right graph), under periodic boundary conditions (PBC). In G_1, the blue data qubits correspond to the blue face in G_2 and connect to the data qubits in G_2. Similarly, in G_2, the yellow data qubits correspond to the yellow face in G_1 and connect to the data qubits in G_1. The black data qubits on the same yz plane as the blue or yellow data qubits do not connect to data qubits in the other subgraph.

The key strategy when employing multiple ancilla qubits is to construct subgraphs of G_1 and G_2 using ancilla qubits Q_1 and Q_2, respectively, for the vertices not included in V_{12}. This is accomplished by utilizing the additional protocol provided in the appendix and Protocol B in Wan et al.'s paper [11]. When encountering the link $i_1, j_2 \in E_{12}$, the connection between all edges involving vertices i_1 and j_2 is established using the interaction between the two ancilla qubits. Consequently, there are two distinct protocols, namely Protocol 1 and Protocol 2. Protocol 1 (Protocol 2) employs the protocol in the appendix (Protocol B) for the vertices not included in V_{12}.

Expanding the implementation to include more than two emitters necessitates the restriction that each vertex in one subgraph can connect to at most one vertex in other subgraphs. This avoids the requirement for interactions involving more than two ancilla qubits within a single procedure. For an optimized construction, each subgraph should be a three-dimensional cluster state of size $(L/n) \times L \times L$, with qubits numbered to match the numbers of the qubits they will connect to in the other subgraphs. If the numbers of qubits to be connected differ, the construction of one subgraph must occasionally be paused. To fulfill these requirements, the number of ancilla qubits, n, and L/n should be even numbers. Subgraphs with an odd number of emitters will always result in the pause of construction for some subgraphs.

3 Main results

Our protocols raise two main concerns: nonlocal effective errors and emitter-to-emitter errors. While it has been demonstrated that Protocol B (the protocol in the appendix) results in local effective errors for a single error in an emitter or qubit in all cases (most cases) [11], the situation becomes less clear when multiple emitters and emitter-to-emitter interactions are involved. Therefore, it is essential to investigate whether the threshold value remains constant as the number of emitters increases. Moreover, given the current technology, emitter-to-emitter interactions tend to introduce more errors compared to emitter-to-qubit interactions. Thus, it is crucial to assess the tolerance of the threshold value under significantly higher emitter-to-emitter interaction error rates.

To address these concerns and investigate the scalability of our protocols, we employ Monte Carlo simulations using a standard depolarizing model for the $L \times L \times L$ three-dimensional cluster state under PBC. We conduct a comprehensive evaluation of the logical error rates for our proposed protocols considering circuit errors and an even number of ancilla qubits. Additionally, we assess the logical error rates for two ancilla qubits, considering various emitter-to-emitter gate error rates expressed as multiples of the other gate rates.

The simulation results clearly show that the threshold values of the proposed protocols are tolerant to an increase in the number of emitters. Specifically, the threshold values for 2, 4, 6 emitters are 0.324%, 0.323%, 0.325% for Protocol 1 and 0.389%, 0.387%, 0.388% for Protocol 2, respectively. Additionally, the threshold values remain tolerant to a significantly higher error rate in emitter-to-emitter interactions [Fig. 3]. For instance, when the emitter-to-emitter gate error rate is ten times noisier than other gates, the threshold is decreased to 0.317% and 0.382%. Thus, while two nearest-neighbor emitter interaction may be noisier than other interactions, the proposed protocol can deal with this issue easily.

In our protocols, the cumulative delay line error is directly related to the cross-sectional area of the three-dimensional cluster state, but inversely proportional to the number of emitters. To demonstrate the scalability of our protocols, we estimate the delay line error rates required to achieve specific logical error rates using the maximum number of emitters available. We evaluate the logical error rates for an even number of ancilla qubits across different lattice sizes and dephasing error rates, assuming a static circuit error rate of $p = 10^{-3}$, which is below the threshold value. Furthermore, we estimate the dephasing error rates that can achieve specific logical error rates for the protocol employing the number of emitters proportional to the length of the three-dimensional cluster state.

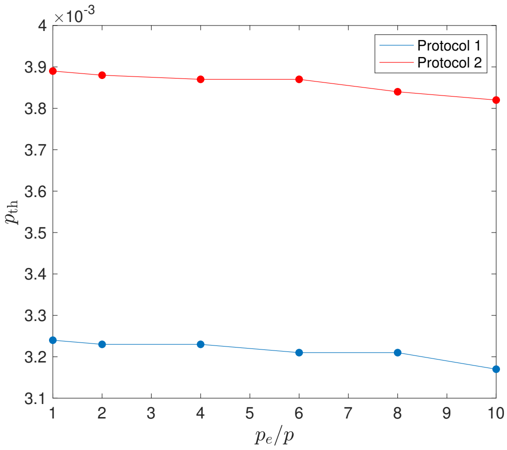


Figure 3: Threshold values for $p_e/p = 1, 2, 4, 6, 8, 10$ are depicted for Protocol 1 and Protocol 2. p_e and p denote the emitter-to-emitter gate error rate and the error rate of other gates, respectively.

We can expect the optimized three-dimensional cluster state constructing using the maximum number of emitters, $L/2$. In that circumstance, each waveguide has $2L^2$ qubits and the minimum logical error rate is proportional to . Simulation results show that achieving logical error rates of 10^{-5}, 10^{-10}, and 10^{-15} for a static circuit error rate of 10^{-3} requires a dephasing error rate of 2.05×10^{-4}, 9.58×10^{-5}, and 6.25×10^{-5}, respectively for Protocol 1. Compared to the values in Wan et al.'s paper, our protocol's values are better by at least 14 times. Notably, our protocol's performance is getting better as the target logical error rate becomes smaller.

In summary, our protocols exhibit scalability even with a limited number of emitter and waveguide pairs. The challenge posed by emitter-to-emitter interactions can be effectively managed, as the protocol is resilient to such noise. By restricting the interactions of emitters to their nearest neighbors, the entire system can be represented as a linear array of emitter and waveguide pairs. Furthermore, the protocols demonstrate that desired logical error rates can be achieved, even in the presence of higher dephasing error rates than those showed by previous simulation results, thereby indicating practical advantages.

References

[1] Acharya, R., Aleiner, I., Allen, R., Andersen, T., Ansmann, M., Arute, F., Arya, K., Asfaw, A., Atalaya, J., Babbush, R., Bacon, D., Bardin, J., Basso,

J., Bengtsson, A., Boixo, S., Bortoli, G., Bourassa, A., Bovaird, J., Brill, L., Broughton, M., Buckley, B., Buell, D., Burger, T., Burkett, B., Bushnell, N., Chen, Y., Chen, Z., Chiaro, B., Cogan, J., Collins, R., Conner, P., Courtney, W., Crook, A., Curtin, B., Debroy, D., Del Toro Barba, A., Demura, S., Dunsworth, A., Eppens, D., Erickson, C., Faoro, L., Farhi, E., Fatemi, R., Flores Burgos, L., Forati, E., Fowler, A., Foxen, B., Giang, W., Gidney, C., Gilboa, D., Giustina, M., Grajales Dau, A., Gross, J., Habegger, S., Hamilton, M., Harrigan, M., Harrington, S., Higgott, O., Hilton, J., Hoffmann, M., Hong, S., Huang, T., Huff, A., Huggins, W., Ioffe, L., Isakov, S., Iveland, J., Jeffrey, E., Jiang, Z., Jones, C., Juhas, P., Kafri, D., Kechedzhi, K., Kelly, J., Khattar, T., Khezri, M., Kieferová, M., Kim, S., Kitaev, A., Klimov, P., Klots, A., Korotkov, A., Kostritsa, F., Kreikebaum, J., Landhuis, D., Laptev, P., Lau, K., Laws, L., Lee, J., Lee, K., Lester, B., Lill, A., Liu, W., Locharla, A., Lucero, E., Malone, F., Marshall, J., Martin, O., McClean, J., McCourt, T., McEwen, M., Megrant, A., Meurer Costa, B., Mi, X., Miao, K., Mohseni, M., Montazeri, S., Morvan, A., Mount, E., Mruczkiewicz, W., Naaman, O., Neeley, M., Neill, C., Nersisyan, A., Neven, H., Newman, M., Ng, J., Nguyen, A., Nguyen, M., Niu, M., O'Brien, T., Opremcak, A., Platt, J., Petukhov, A., Potter, R., Pryadko, L., Quintana, C., Roushan, P., Rubin, N., Saei, N., Sank, D., Sankaragomathi, K., Satzinger, K., Schurkus, H., Schuster, C., Shearn, M., Shorter, A., Shvarts, V., Skruzny, J., Smelyanskiy, V., Smith, W., Sterling, G., Strain, D., Szalay, M., Torres, A., Vidal, G., Villalonga, B., Vollgraff Heidweiller, C., White, T., Xing, C., Yao, Z., Yeh, P., Yoo, J., Young, G., Zalcman, A., Zhang, Y., Zhu, N. & AI, G. Suppressing quantum errors by scaling a surface code logical qubit. *Nature.* **614**, 676-681 (2023,2)

[2] Bravyi, S. & Kitaev, A. Quantum codes on a lattice with boundary. (arXiv,1998)

[3] Kitaev, A. Fault-tolerant quantum computation by anyons. *Annals Of Physics.* **303**, 2-30 (2003)

[4] Fowler, A., Mariantoni, M., Martinis, J. & Cleland, A. Surface codes: Towards practical large-scale quantum computation. *Phys. Rev. A.* **86**, 032324 (2012,9)

[5] Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *Journal Of Mathematical Physics.* **43**, 4452-4505 (2002)

[6] Gottesman, D. Theory of fault-tolerant quantum computation. *Phys. Rev. A.* **57**, 127-137 (1998,1)

[7] Gottesman, D. The Heisenberg Representation of Quantum Computers. (arXiv,1998)

[8] Gottesman, D. An Introduction to Quantum Error Correction and Fault-Tolerant Quantum Computation. (arXiv,2009)

[9] Aharonov, D. & Ben-Or, M. Fault-Tolerant Quantum Computation with Constant Error Rate. *SIAM Journal On Computing.* **38**, 1207-1282 (2008)

[10] Knill, E., Laflamme, R. & Wojciech H. Zurek Resilient Quantum Computation. *Science.* **279**, 342-345 (1998)

[11] Wan, K., Choi, S., Kim, I., Shutty, N. & Hayden, P. Fault-Tolerant Qubit from a Constant Number of Components. *PRX Quantum.* **2**, 040345 (2021,12)

[12] Raussendorf, R., Harrington, J. & Goyal, K. A fault-tolerant one-way quantum computer. *Annals Of Physics.* **321**, 2242-2270 (2006)

[13] Barrett, S. & Stace, T. Fault Tolerant Quantum Computation with Very High Threshold for Loss Errors. *Phys. Rev. Lett..* **105**, 200502 (2010,11)

[14] Larsen, M., Chamberland, C., Noh, K., Neergaard-Nielsen, J. & Andersen, U. Fault-Tolerant Continuous-Variable Measurement-based Quantum Computation Architecture. *PRX Quantum.* **2**, 030325 (2021,8)

[15] Bourassa, J., Alexander, R., Vasmer, M., Patil, A., Tzitrin, I., Matsuura, T., Su, D., Baragiola, B., Guha, S., Dauphinais, G., Sabapathy, K., Menicucci, N. & Dhand, I. Blueprint for a Scalable Photonic Fault-Tolerant Quantum Computer. *Quantum.* **5** pp. 392 (2021,2)

[16] Bartolucci, S., Birchall, P., Bombin, H., Cable, H., Dawson, C., Gimeno-Segovia, M., Johnston, E., Kieling, K., Nickerson, N., Pant, M., Pastawski, F., Rudolph, T. & Sparrow, C. Fusion-based quantum computation. (arXiv,2021)

[17] Tzitrin, I., Matsuura, T., Alexander, R., Dauphinais, G., Bourassa, J., Sabapathy, K., Menicucci, N. & Dhand, I. Fault-Tolerant Quantum Computation with Static Linear Optics. *PRX Quantum.* **2**, 040353 (2021,12)

[18] Fukui, K. & Takeda, S. Building a large-scale quantum computer with continuous-variable optical technologies. *Journal Of Physics B: Atomic, Molecular And Optical Physics.* **55**, 012001 (2022,2)

[19] Economou, S., Lindner, N. & Rudolph, T. Optically Generated 2-Dimensional Photonic Cluster State from Coupled Quantum Dots. *Phys. Rev. Lett..* **105**, 093601 (2010,8)

[20] Yokoyama, S., Ukai, R., Armstrong, S., Sornphiphatphong, C., Kaji, T., Suzuki, S., Yoshikawa, J., Yonezawa, H., Menicucci, N. & Furusawa, A. Ultra-large-scale continuous-variable cluster states multiplexed in the time domain. *Nature Photonics.* **7**, 982-986 (2013,12)

[21] Yoshikawa, J., Yokoyama, S., Kaji, T., Sornphiphatphong, C., Shiozawa, Y., Makino, K. & Furusawa, A. Invited Article: Generation of one-million-mode continuous-variable cluster state by unlimited time-domain multiplexing. *APL Photonics.* **1**, 060801 (2016)

[22] Pichler, H., Choi, S., Zoller, P. & Mikhail D. Lukin Universal photonic quantum computation via time-delayed feedback. *Proceedings Of The National Academy Of Sciences.* **114**, 11362-11367 (2017)

[23] Asavanant, W., Shiozawa, Y., Yokoyama, S., Charoensombutamon, B., Emura, H., Alexander, R., Takeda, S., Jun-Yoshikawa, Menicucci, N., Yonezawa, H. & Akira Furusawa Generation of time-domain-multiplexed two-dimensional cluster state. *Science.* **366**, 373-376 (2019)

[24] Raussendorf, R., Harrington, J. & Goyal, K. Topological fault-tolerance in cluster state quantum computation. *New Journal Of Physics.* **9**, 199-199 (2007,6)

[25] Raussendorf, R., Bravyi, S. & Harrington, J. Long-range quantum entanglement in noisy cluster states. *Phys. Rev. A.* **71**, 062313 (2005,6)

[26] Bombin, H., Dawson, C., Mishmash, R., Nickerson, N., Pastawski, F. & Roberts, S. Logical blocks for fault-tolerant topological quantum computation. (arXiv,2021)

[27] Walther, P., Resch, K., Rudolph, T., Schenck, E., Weinfurter, H., Vedral, V., Aspelmeyer, M. & Zeilinger, A. Experimental one-way quantum computing. *Nature.* **434**, 169-176 (2005,3)

[28] Bolt, A., Duclos-Cianci, G., Poulin, D. & Stace, T. Foliated Quantum Error-Correcting Codes. *Phys. Rev. Lett..* **117**, 070501 (2016,8)

Long-distance measurement-device-independent quantum key distribution using hybrid-entangled states

Soumyakanti Bose[1] * Jaskaran Singh[2] † Adán Cabello[2] ‡ Hyunseok Jeong[1] §

[1] *Department of Physics & Astronomy, Seoul National University, Gwanak-ro 1, Gwanak-gu, Seoul 08826, Korea.*
[2] *Departamento de Física Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain.*

Abstract. Here we introduce a method of generating high-rate high-quality entanglement between distant locations, a crucial requirement for quantum science and technology, by using hybrid-entanglement (HE) between continuous-variables (CV) and discrete-variables (DV). We show that by using the CV part for transmission and using the DV part for key generation, HE is especially useful in measurement-device-independent (MDI) quantum-key-distribution (QKD), as it can enhance the transmission distance up to 300 km with a secure keyrate almost an order of magnitude higher than the existing MDI-QKD protocols. Our results indicate that HE states are a feasible choice in optimaizing practical long-distance entanglement.

Keywords: AQIS, quantum key distribution, hybrid-entanglement, measurement-device-independent

1 Introduction

Generation of high-rate high-quality entanglement between distant locations is crucial for fundamental tests of quantum theory and for its applications. For example, it is needed for extending the current distances of loophole-free Bell tests [1, 2], quantum steering [3], and quantum teleportation [4]. It is also needed for increasing the transmission distance and the key rate in entanglement-based quantum key distribution (QKD) protocols [5, 6, 7]. Moreover, higher-rates in distant locations will allow us to achieve higher detection efficiencies by means of heralded qubit amplifiers [8] or photonic precertification schemes [9, 10, 11], whose practicality is currently limited by the rates achieved after transmission [12].

QKD, for which the security stems from the the laws of nature [14], promises a vivid example of quantum versus classical advantage [15, 16]. It has been extensively analyzed both theoretically and experimentally under variety of setups such as prepare-and-measure, entanglement-based, and measurement-device-independent (MDI) scenarios for physical systems that could be categorized in two different sets [17]. The first kind, formally referred to as discrete variable (DV) systems, are based on discrete degrees of freedom such as polarization of a photon or orbital angular momentum states of light. On the other hand, the second kind of physical systems, formally known as the continuous variable (CV) systems, exploit the continuous degrees of freedom such as quadrature distribution of a quantized light inside a cavity.

While both DV and CV based QKD offer their own set of advantages, they come with their disadvantages too [12]. For example, DV-QKD protocols offer composable security proofs with good key rate and are also compatible with other QKD systems that makes them an ideal choice for quantum networks. However, they require expensive single-photon-sources and precise single-photon measurements which are hard to perform. On the other hand, CV-QKD protocols generally require Gaussian states which are comparatively easier to prepare than DV systems. Moreover, they are robust against transmission losses and can potentially offer long transmission distances. However, their performance is limited by the requirement of almost ideal homodyne detectors at telecommunication wavelength. As a consequence, despite an extensive theoretical and experimental analysis on both types of systems, the quest for an optimal physical system for QKD remains open.

Nonetheless, there exists a different class of physical systems formally known as hybrid entangled (HE) states [23, 24, 25]. These systems, representing a strongly correlated [26, 27] cross-system entanglement between DV and CV states, play a crucial role in various quantum information processing tasks [28, 29, 30]. Although, such states have been efficiently generated in a wide range of experimental setups [31, 32, 33] their potential application in QKD still remains unexplored.

From the perspective of generating high-rate high quality entanglement between distant locations, MDI-QKD protocols are of great interest as they can effectively double the transmission distances. Here, we propose a novel MDI-QKD scheme, using HE states as an initial resource, that primarily hinges on generating a single-photon DV entangled state between two distant parties by exploiting CV entanglement swapping by a third party located midway. Our scheme offers three major advantages as compared to earlier DV and CV MDI-QKD protocols. These are: (i) Elimination of major limiting factors of DV-MDI-QKD, which include high precision Bell state or single-photon measurements as well as the photon-number-splitting attack by an eavesdropper, (ii) Elimination of the requirement of near-unit efficiency for the homodyne detectors used for key generation in CV-MDI-QKD and (iii) Long transmission distance in telecom wavelength stemming from the robustness of multiphoton coherent state against transmission losses.

By bringing forth the best of both DV and CV systems, we show that it is possible to achieve a total transmission

*soumyakanti.bose09@gmail.com
†jaskaran@us.es
‡adan@us.es
§h.jeong37@gmail.com

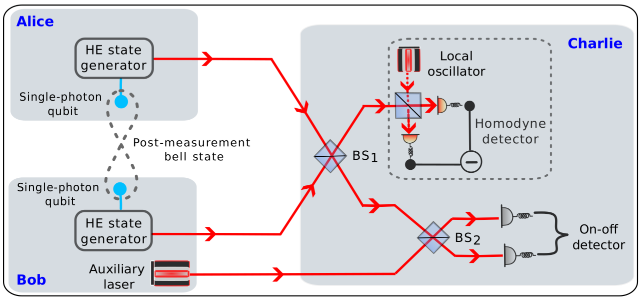


Figure 1: Schematic for generating DV entangled states between Alice and Bob using HE states. Alice and Bob first send the CV part of their individual HE states to Charlie, who then mixes the incoming signals at a balanced beam splitter (BS$_1$), and uses one of the output modes for homodyne measurement. The other outgoing signal of BS$_1$ is mixed with the additional coherent signal sent by Bob at another balanced beam splitter (BS$_2$) and is used for a post-selection measurement by on-off detectors. Upon declaration of the results by Charlie, Alice and Bob obtain a DV entangled pair which is used for secure key generation.

distance of ≈ 300 km by using practical homodyne detectors with efficiency $\eta_h = 0.55$ [18, 19, 20, 21, 22] and on-off detectors with efficiency $\eta_0 = 0.8$, thereby surpassing earlier results at telecommunication wavelengths on optical fibres. The use of realistic and widely available equipment makes our scheme readily implementable for an intercity QKD network using current state-of-the-art technologies.

2 Protocol

Let $|0\rangle$ and $|1\rangle$ correspond to photon number states in the Fock basis and $|\alpha\rangle$ correspond to a coherent state of a quantized light. We denote the photon number state and the coherent state as the DV system and the CV system respectively. A HE state is defined as entangled pair between the DV and CV degrees of freedom as [23]

$$|\psi\rangle_{a_1a_2} = \frac{1}{\sqrt{2}}\left(|0\rangle_{a_1}|\alpha\rangle_{a_2} + |1\rangle_{a_1}|-\alpha\rangle_{a_2}\right). \qquad (1)$$

We consider that two distant parties, Alice and Bob, each of them having access to bipartite HE states $|\psi\rangle_{a_1a_2}$ and $|\psi\rangle_{b_1b_2}$ given by Eq. (1). We provide a step-by-step description of the protocol, schematically represented in Fig. 1.

Step 1: Alice and Bob transmit the CV part of their systems (modes a_2 and b_2, respectively) to a third untrusted party, Charlie, who lies midway between them, through a lossy quantum channel with transmittance T $(0 \leq T \leq 1)$. Bob also transmits the state $\left|\sqrt{2}\alpha\right\rangle$, which we label by mode c, to Charlie separately through a similar quantum channel.

Step 2: Charlie mixes the two incoming modes a_2 and b_2 via a beam splitter (BS), labelled as BS$_1$, with two output modes which we can label as $a_2^{'}$ and $b_2^{'}$.

Step 3: In our protocol we are specifically interested in the vacuum state contributions from the mode $a_2^{'}$. To extract this contribution, Charlie mixes this mode though a second BS (BS$_2$) with mode c with output modes labelled as $a_2^{''}$ and $c^{'}$.

Step 4: Charlie now performs a projective measurement, $\mathcal{M} = \{\Pi_0, 1 - \Pi_0\}$, where $\Pi_0 = (1 - |0\rangle\langle 0|)_{a_2^{''}} \otimes (1 - |0\rangle\langle 0|)_{c^{'}}$. This measurement is accomplished by using on-off detectors on each of the modes $a_2^{''}$ and $c^{'}$. Charlie then publicly announces the outcome of the projective measurement which is considered to be successful only if the result Π_0 is obtained, i.e., both detectors click. In that case, the protocol continues. Otherwise, the measurement is deemed unsuccessful and the parties must repeat the aforementioned steps again. In order to model realistic detectors, we consider imperfect on-off detectors with efficiency η_0.

Step 5: After a successful projective measurement (as dictated in Step 4), Charlie performs a homodyne measurement on the mode $b_2^{'}$ and, again, announces the results publicly. We consider that homodyne measurements have efficiency η_h.

Step 6: After a public announcement of the results of a successful projective measurement and the homodyne measurement by Charlie, Alice and Bob end up with the final normalized single-photon-Bell-state in modes a_1 and b_1 as

$$\rho_{a_1b_1} = \frac{1}{2}\left(|01\rangle\langle 01| + |10\rangle\langle 10| + f\left[g|01\rangle\langle 10| + g^*|10\rangle\langle 01|\right]\right) \qquad (2)$$

with probability

$$p_0 = \frac{\left(1 - e^{-\eta_o T\alpha^2}\right)^2}{2}, \qquad (3)$$

where $f = e^{-4(1-T\eta_h)\alpha^2}$, $g = e^{4i\sqrt{T\eta_h}\alpha p}$, g^* is the conjugate of g, and p is the result of the homodyne measurement.

Step 7: For the case in which Alice and Bob share $\rho_{a_1b_1}$, they perform two-outcome Pauli measurements corresponding to σ_Z on their respective subsystems to generate a raw key. The length of the raw key that the parties can generate is quantified by the mutual information $I(A:B)$ between them for the observable σ_Z.

Step 8: Alice and Bob then estimate the amount of information that an adversary, Eve, can have on their raw key. This information is quantified by the Holevo bound $\chi(A:E)$ between Alice and Eve. As a consequence, we can write the final secure key rate as

$$r \geq p_0\left[I(A:B) - \chi(A:E)\right]. \qquad (4)$$

3 Results

We assume that Charlie is midway between Alice and Bob such that the total transmission distance is L. We assume that the transmittance of both the channels, Alice-Charlie and Bob-Charlie, is given by T such that $T = 10^{-l\frac{L/2}{10}}$, where $l = 0.2$ dB/km is the standard channel loss for telecom wavelength [34, 35].

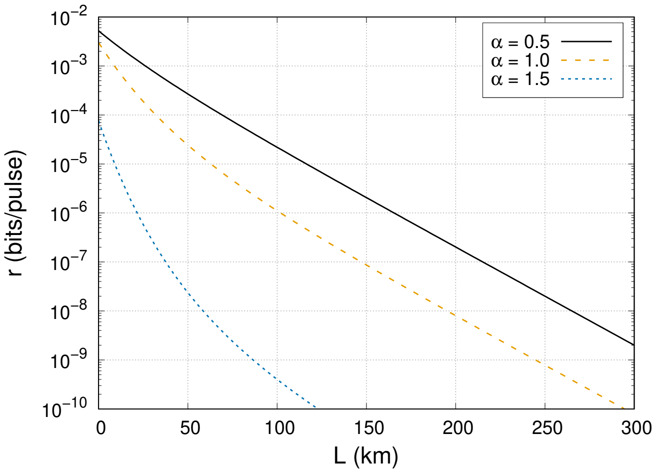


Figure 2: Secure key rate as a function of total transmission distance L for different values of α. We fix $\eta_h = 0.55$, $\eta_0 = 0.8$, and $p = \frac{\pi}{2}$.

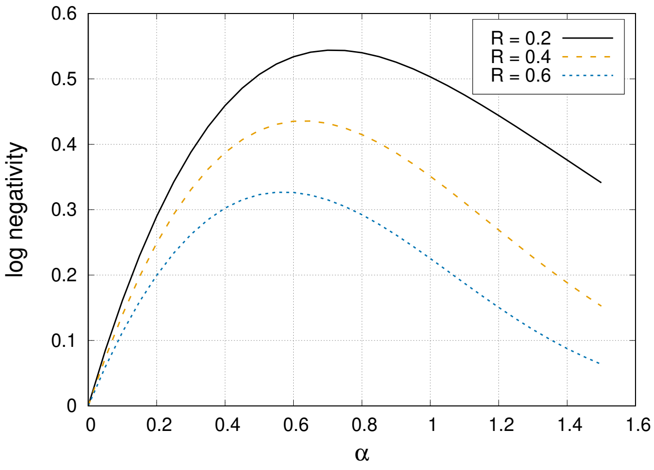


Figure 3: Logarithmic negativity of the HE state after its CV system is transmitted via a lossy quantum channel as a function of the coherent amplitude α.

In Fig. 2, we plot the secure key rate as a function the total transmission distance L. We choose the parameters $\eta_h = 0.55$, $\eta_0 = 0.8$, and $p = \frac{\pi}{2}$ to be as realistic as possible. We see that the total transmission distance can reach upto 300 km with a secure key rate of the order of 10^{-9} bits/pulses for $\alpha = 0.5$. However, as we increase α, the key rate is found to decrease drastically. We also find that decreasing the value of α below 0.5 also decreases the key rate as well as the maximum transmission distance leading to an optimal choice of $\alpha \sim 0.5$.

This optimal value for α is in agreement with the observation that it is also the value of α which is more robust to transmission losses. The effect of a quantum state passing through a noisy channel can be seen as the system undergoing photon loss. Let us denote the photon loss fraction by R such that $R = 0$ and $R = 1$ correspond to no photon loss and complete photon loss, respectively. In Fig. 3, we show that the entanglement content of an HE state undergoing photon loss over the CV part, as given by logarithmic negativity. For a significantly lossy channel, we find that the optimal value of α approaches $\alpha = 0.5$. This behaviour of HE states can be qualitatively understood in terms of the interplay between entanglement and the fragility of the initial HE state. Starting from the initial separable state at $\alpha = 0$, the HE state becomes more entangled as α increases. An increase in α also corresponds to an increase in the average number of photons, which can be understood as an increase in the mean energy of the system. However, with an increase in the mean energy, the state becomes more vulnerable to decoherence. This behaviour is similar to what is also shown in Ref. [36] for superposition of coherent states with increasing values of coherent amplitude. As a consequence, with increase in α beyond an optimal value, the HE state becomes extremely fragile under noise leading to a drop in entanglement when the multiphoton part passes through a noisy quantum channel.

4 Conclusion

We have shown that our scheme with HE-states is able to overcome the major drawbacks and disadvantages of both DV and CV MDI-QKD protocols. Most notably, it removes the requirement of using single photon detectors, Bell state measurements, and near perfect homodyne detectors, which are difficult to implement in an experimental realisation. Instead, our protocol provides the best of both worlds: longer transmission distances and higher key rate using devices that can be readily implemented in the laboratory. Our results indicate that it is possible to achieve a total transmission distances of $\sim$ 300 km with a secure keyrate of the order of 10^{-9} bits/pulse. However, there exists a critical choice of α (here it is ~ 0.5) that yields optimal result which could be attributed to the interplay between the fragility and entanglement for an HE state undergoing transmission losses.

It may further be noted that the probability of preparing the HE state is $\sim 50\%$ and with fidelity ≈ 0.75 for $\alpha = 0.5$ [31]. This could be a limiting factor in a practical realisation of our protocol. However, this problem can be mitigated by using other forms of HE states, most notably with the DV and CV modes corresponding to polarization and cat states respectively [33] which offer exceptionally good fidelity of preparation as well as probability of generation.

Nonetheless, the use of realistic instruments, widely available on the market, makes our protocol a viable alternative in implementing linear optics based secured cryptosystem at telecommunication wavelength within the reach of current state-of-art technology. Moreover, it could be reasonably extended to quantum networks [37, 38] as well as satellite-based secure communication [39].

References

[1] B. Hensen *et. al.* Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. Nature 526, 682 (2015).

[2] W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortegel, M. Rau, and H. Weinfurter. Event-ready Bell test using entangled atoms simultaneously closing detection and locality loopholes. Phys. Rev. Lett. 119, 010402 (2017).

[3] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. M. Wiseman, R. Ursin, and A. Zeilinger. Loophole-free Einstein–Podolsky–Rosen experiment via quantum steering. New J. Phys. 14, 053030 (2012).

[4] X.-S. Ma, T. Herbst, T. Scheidl, D. Wang, S. Kropatschek, W. Naylor, B. Wittmann, A. Mech, J. Kofler, E. Anisimova, V. Makarov, T. Jennewein, R. Ursin and A. Zeilinger. Quantum teleportation over 143 kilometres using active feed-forward. Nature 489, 269 (2012).

[5] D. P. Nadlinger *et. al.* Experimental quantum key distribution certified by Bell's theorem. Nature 607, 682 (2022).

[6] W. Zhang *et. al.* A device-independent quantum key distribution system for distant users. Nature 607, 687 (2022).

[7] W.-Z. Liu, Y.-Z. Zhang, Y.-Z. Zhen, M.-H. Li, Y. Liu, J. Fan, F. Xu, Q. Zhang, and J.-W. Pan. Toward a photonic demonstration of device-independent quantum key distribution. Phys. Rev. Lett. 129, 050502 (2022).

[8] N. Gisin, S. Pironio, and N. Sangouard. Proposal for implementing device-independent quantum key distribution based on a heralded qubit amplifier. Phys. Rev. Lett. 105, 070501 (2010).

[9] A. Cabello and F. Sciarrino. Loophole-free bell test based on local precertification of photon's presence. Phys. Rev. X 2, 021010 (2012).

[10] E. Meyer-Scott *et. al.* Certifying the presence of a photonic qubit by splitting it in two. Phys. Rev. Lett. 116, 070501 (2016).

[11] A. Z. Leger, S. Gambhir, J. L. egere, and D. R. Hamel. Amplification of cascaded downconversion by reusing photons with a switchable cavity. arXiv:2209.11668 (2022).

[12] E. Diamanti, H.-K. Lo, B. Qi, and Z. Yuan. Practical challenges in quantum key distribution. npj Quantum Inf. 2, 16025 (2016).

[13] S. Pirandola *et. al.*. Advances in quantum cryptography. Adv. Opt. Photon. 12, 1012 (2020).

[14] C. Portmann and R. Renner. Security in quantum cryptography. Rev. Mod. Phys. 94, 025008 (2022).

[15] 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).

[16] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. Rev. Mod. Phys. 74, 145 (2002).

[17] D. Bouwmeester, A. K. Ekert, and A. Zeilinger. The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation. 1st ed. (Springer Publishing Company, Incorporated, 2010)

[18] B. Qi, P. Lougovski, R. Pooser, W. Grice, and M. Bobrek. Generating the local oscillator "locally" in continuous-variable quantum key distribution based on coherent detection. Phys. Rev. X 5, 041009 (2015).

[19] P. Jouguet, S. Kunz-Jacques, A. Leverrier, P. Grangier, and E. Diamanti. Experimental demonstration of long-distance continuous-variable quantum key distribution. Nat. Photonics 7, 378 (2013).

[20] M. Zou, Y. Mao, and T.-Y. Chen. Rigorous calibration of homodyne detection efficiency for continuous-variable quantum key distribution. Opt. Express 30, 22788 (2022).

[21] G. Zhang *et. al.* An integrated silicon photonic chip platform for continuous-variable quantum key distribution. Nat. Photonics 13, 839 (2019).

[22] Y. Zhang, Z. Chen, S. Pirandola, X. Wang, C. Zhou, B. Chu, Y. Zhao, B. Xu, S. Yu, and H. Guo. Long-distance continuous-variable quantum key distribution over 202.81 km of fiber. Phys. Rev. Lett. 125, 010502 (2020).

[23] H. Jeong. Using weak nonlinearity under decoherence for macroscopic entanglement generation and quantum computation. Phys. Rev. A 72, 034305 (2005).

[24] B. He, Q. Lin, and C. Simon. Cross-Kerr nonlinearity between continuous-mode coherent states and single photons. Phys. Rev. A 83, 053826 (2011).

[25] M. Hosseini, S. Rebic, B. M. Sparkes, J. Twamley, B. C. Buchler, and P. K. Lam. Memory-enhanced noiseless cross-phase modulation. Light: Sci. Appl. 1, e40 (2012).

[26] K. Park, S.-W. Lee, and H. Jeong. Quantum teleportation between particlelike and fieldlike qubits using hybrid entanglement under decoherence effects. Phys. Rev. A 86, 062301 (2012).

[27] H. Kwon and H. Jeong. Violation of the Bell–Clauser-Horne-Shimony-Holt inequality using imperfect photodetectors with optical hybrid states. Phys. Rev. A 88, 052127 (2013).

[28] U. L. Andersen, J. S. Neergaard-Nielsen, P. van Loock, and A. Furusawa. Hybrid discrete and continuous-variable quantum information. Nat. Phys. 11, 713 (2015).

[29] S. Omkar, Y. S. Teo, and H. Jeong. Resource-efficient topological fault-tolerant quantum computation with hybrid entanglement of light. Phys. Rev. Lett. 125, 060501 (2020).

[30] S. Omkar, Y. S. Teo, S.-W. Lee, and H. Jeong. Highly photon-loss-tolerant quantum computing using hybrid qubits. Phys. Rev. A 103, 032602 (2021).

[31] H. Jeong, A. Zavatta, M. Kang, S.-W. Lee, L. S. Costanzo, S. Grandi, T. C. Ralph, and M. Bellini. Generation of hybrid entanglement of light. Nat. Photonics 8, 564 (2014).

[32] O. Morin, K. Huang, J. Liu, L. H. Jeannic, C. Fabre, and J. Laurat. Remote creation of hybrid entanglement between particle-like and wave-like optical qubits. Nat. Photonics 8 570 (2014).

[33] J. Wen, I. Novikova, C. Qian, C. Zhang, and S. Du. Hybrid entanglement between optical discrete polarizations and continuous quadrature variables. Photonics 8, 552 (2021).

[34] H.-K. Lo, M. Curty, and B. Qi. Measurement-device-independent quantum key distribution. Phys. Rev. Lett. 108, 130503 (2012).

[35] S. Pirandola *et. al.* High-rate measurement-device-independent quantum cryptography. Nat. Photonics 9, 397 (2015).

[36] C.-W. Lee and H. Jeong. Quantification of macroscopic quantum superpositions within phase space. Phys. Rev. Lett. 106, 220401 (2011).

[37] B. Frohlich, J. F. Dynes, M. Lucamarini, A. W. Sharpe, Z. Yuan, and A. J. Shields. A quantum access network. Nature 501, 69 (2013).

[38] Y.-L. Tang *et. al.* Measurement-device-independent quantum key distribution over un-trustful metropolitan network. Phys. Rev. X 6, 011024 (2016).

[39] W. Liang and R. Jiao. Satellite-based measurement-device-independent quantum key distribution. New J. Phys. 22 083074 (2020).

Efficient computation of nonstabilizerness with quantum computers and matrix product states

Tobias Haug[1] * Lorenzo Piroli[2] M. S. Kim[1]

[1] *QOLS, Blackett Laboratory, Imperial College London SW7 2AZ, UK*

[2] *Philippe Meyer Institute, Physics Department, École Normale Supérieure (ENS), Université PSL, 24 rue Lhomond, F-75231 Paris, France*

Abstract. Nonstabilizerness or ‘magic’ characterizes the amount of non-Clifford operations needed to prepare quantum states. It is a crucial resource for quantum computing and a necessary condition for quantum advantage. However, quantifying nonstabilizerness beyond a few qubits has been a major challenge. Here, we provide efficient methods to compute measures of nonstabilizerness for quantum computers [1] and matrix product states [2, 3]. We apply our methods to experimentally uncover the transition in nonstabilizerness with increasing non-Clifford resources on the IonQ quantum computer. Further, we study critical many-body systems and find nonstabilizerness is in general not extremal. Our methods open up the experimental and numerical study of nonstabilizerness of extensive quantum systems.

Keywords: Nonstabilizerness, magic, resource theory, critical many-body systems, NISQ

Nonstabilizerness, also known as magic, quantifies the minimal number of non-Clifford operations needed to prepare quantum states and is a necessary condition for quantum advantage [4, 5]. As typical measures of nonstabilizerness involve minimization procedures and a computational cost exponential in the number N of qubits [6], it has been notoriously hard to characterize the nonstabilizerness of many-body states [7]. In particular, the previously best known algorithms were limited to about 12 qubits [8].

In our submission, we efficiently compute measures of nonstabilizerness for quantum computers [1] and matrix product states (MPSs) [2, 3] respectively. Our results give us novel insights in the nonstabilizerness of quantum many-body states, as well provide practical applications for quantum computers which we demonstrate in experiment.

Measures of nonstabilizerness for quantum computers

In Ref. [1], we introduce “Bell magic” that can be efficiently measured on quantum computers with the numbers of samples scaling as $O(1)$ and classical post-processing time as $O(N)$. The measurement protocol is simple, requiring only two copies and Bell measurements. Our measurement protocol is robust against noise via a cost-free error mitigation strategy, which we demonstrate experimentally on the IonQ quantum computer. We apply our methods to experimentally study the transition nonstabilizerness of random Clifford circuits doped with T-gates. Further, we efficiently distinguish stabilizer and non-stabilizer states with low measurement cost, even in the presence of noise. In the context of noisy intermediate scale quantum computers [9], we provide a variational quantum algorithm to maximize Bell magic, which can remain trainable even for highly expressible circuits.

Measures of nonstabilizerness for matrix product states

In Ref. [2], we show that the recently introduced Stabilizer Rényi Entropies (SREs) [8] can be computed efficiently for MPSs. Specifically, given an MPS of bond-dimension χ, the replica trick allows us to compute the SRE with a computational cost of $O(N\chi^{12})$. We recently improved the cost to $O(N\chi^3)$ [3]. Our work opens up the numerical study of nonstabilizerness of many-body systems for extensive system sizes, which so far has been challenging [7]. Using our tools, we investigate the stabilizer entropy of the transverse-field Ising model, one of the most fundamental spin models. We analyze the SRE near criticality and investigate its dependence on the local computational basis. Surprisingly, the SRE is in general not extremal at the critical point, hinting at a more subtle connection between nonstabilizerness and criticality. Further, local unitary transformations can reduce the SRE even at criticality, implying that the nonstabilizerness of the Ising model ground state is

*tobias.haug@u.nus.edu

short-range.

Importance for quantum information and computation

- Quantum phase transitions are generally associated with long-range correlations. Thus, naively one would assume that quantum state preparation requires the most resources for states at the transition point. However, our work shows that the magic of many-body systems is not extremal at the critical point in general. This highly counter-intuitive results hints at a more subtle connection between the cost of quantum simulation and criticality.

- Nonstabilizerness of quantum states can exist on different length-scales. When nonstabilizerness is short-range, it can be removed by local transformations. In contrast, removing long-range nonstabilizerness requires deep circuits extending over many qubits [10]. This feature is crucial to understand the difficulty of quantum simulation [11]. We find that local transformations can substantially reduce the SRE of the ground state of the quantum Ising model, implying that its nonstabilizerness is short-range. Surprisingly, the short-range character of the SRE persists even at criticality.

- Is it always necessary to compute nonstabilizerness over the full quantum system, or are measurements over local subystems sufficient to predict its nonstabilizerness [12]? For translational invariant MPS, we prove that local measurements over subsystems of size ℓ can indeed be used to predict the nonstabilizerness of the full system. In particular, we show that the error scales as ℓ^{-1}.

- Variational quantum algorithms suffer from various limitations. One is that gradients can vanish exponentially with system size, which is the so-called barren plateau problem [13]. Barren plateaus are closely related to the expressibility of the ansatz, i.e. the ansatz explores the full Hilbertspace equally such that it forms a 2-design. In particular, when the ansatz circuit is expressible, barren plateau always appear for a general class of cost functions that consist of a polynomial number of Pauli strings [14]. We show that nonstabilizerness as a cost function can overcome this limitation in training variational quantum algorithms, i.e. it can remain trainable even if the ansatz is a 2-design. Measures of nonstabilizerness cannot be expressed as a polynomial number of Pauli strings, owing to the fact they require quantum memory to be measured, which is known to give advantages [15]. Beyond our work, exploiting this feature could yield other cost functions with similar desirable training properties.

- When randomly given either a stabilizer state and a state with high nonstabilizerness, can one efficiently distinguish those two states? This question has been recently answered with 'yes' for noise-free states [16]. Our work shows that efficient discrimination is robust even in the presence of noise [1].

References

1. Haug, T. & Kim, M. Scalable Measures of Magic Resource for Quantum Computers. *PRX Quantum* **4,** 010301. https://link.aps.org/doi/10.1103/PRXQuantum.4.010301 (1 Jan. 2023).
2. Haug, T. & Piroli, L. Quantifying nonstabilizerness of matrix product states. *Phys. Rev. B* **107,** 035148. https://link.aps.org/doi/10.1103/PhysRevB.107.035148 (2023).
3. Haug, T. & Piroli, L. Stabilizer entropies and nonstabilizerness monotones. *arXiv:2303.10152* (2023).
4. Howard, M., Wallman, J., Veitch, V. & Emerson, J. Contextuality supplies the 'magic'for quantum computation. *Nature* **510,** 351–355. https://www.nature.com/articles/nature13460 (2014).
5. Veitch, V., Mousavian, S. H., Gottesman, D. & Emerson, J. The resource theory of stabilizer quantum computation. *New J. Phys.* **16,** 013009. https://doi.org/10.1088/1367-2630/16/1/013009 (2014).
6. Howard, M. & Campbell, E. Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing. *Phys. Rev. Lett.* **118,** 090501. https://link.aps.org/doi/10.1103/PhysRevLett.118.090501 (9 Mar. 2017).

7. Liu, Z.-W. & Winter, A. Many-Body Quantum Magic. *PRX Quantum* **3,** 020333. https://link.aps.org/doi/10.1103/PRXQuantum.3.020333 (2 May 2022).

8. Leone, L., Oliviero, S. F. E. & Hamma, A. Stabilizer Rényi Entropy. *Phys. Rev. Lett.* **128,** 050402. https://link.aps.org/doi/10.1103/PhysRevLett.128.050402 (5 Feb. 2022).

9. Bharti, K. *et al.* Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.* **94,** 015004. https://link.aps.org/doi/10.1103/RevModPhys.94.015004 (1 Feb. 2022).

10. White, C. D., Cao, C. & Swingle, B. Conformal field theories are magical. *Phys. Rev. B* **103,** 075145. https://link.aps.org/doi/10.1103/PhysRevB.103.075145 (7 Feb. 2021).

11. Sewell, T. J. & White, C. D. Mana and thermalization: Probing the feasibility of near-Clifford Hamiltonian simulation. *Phys. Rev. B* **106,** 125130. https://journals.aps.org/prb/abstract/10.1103/PhysRevB.106.125130 (2022).

12. Oliviero, S. F. E., Leone, L. & Hamma, A. Magic-state resource theory for the ground state of the transverse-field Ising model. *Phys. Rev. A* **106,** 042426. https://link.aps.org/doi/10.1103/PhysRevA.106.042426 (4 Oct. 2022).

13. McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R. & Neven, H. Barren plateaus in quantum neural network training landscapes. *Nature communications* **9,** 4812 (2018).

14. Holmes, Z., Sharma, K., Cerezo, M. & Coles, P. J. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum* **3,** 010313 (2022).

15. Chen, S., Cotler, J., Huang, H.-Y. & Li, J. *Exponential separations between learning with and without quantum memory* in *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)* (2022), 574–585.

16. Gross, D., Nezami, S. & Walter, M. Schur–Weyl duality for the Clifford group with applications: Property testing, a robust Hudson theorem, and de Finetti representations. *Communications in Mathematical Physics* **385,** 1325–1393 (2021).

Computing ground state energies of molecules using variational quantum eigensolver on IBM superconducting quantum computers

Jun-Ho Lee[1] * HyukGeun Cha[1] † Seong-Hyok Sean Kim[1] ‡

[1] *Quantum AI Task, AI Laboratory, LG Electronics, Seoul 06772, Korea*

Abstract. Simulating quantum systems such as molecules and condensed matter will most likely be a killer application of quantum computers. We will introduce our recent efforts on quantum simulation using a noisy intermediate-scale quantum (NISQ) algorithm, variational quantum eigensolver (VQE), for computing the ground state energy of hydrogen molecules on the IBM superconducting quantum computers. For a single hydrogen molecule, ground state energy is converged within a few iterations and the converged value is close to the result that is obtained by classical computers. For two hydrogen molecules, however, obtaining converged results requires tens of iterations, and its converged value shows the difference from the results obtained by classical computers because of the increased circuit complexity. We also investigate the effect of error mitigation techniques in our VQE calculations.

Keywords: Quantum simulation, IBM Quantum, VQE, NISQ algorithm, error mitigation

Quantum computer has been regarded as a solution to perform certain tasks that are intractable by classical high-performance computers (HPC). Among diverse applications where HPC is used, simulating quantum systems such as molecules and condensed matter systems would most likely be benefited from quantum computing in the near term as understanding the quantum mechanical nature of the quantum systems by a quantum computer is the original idea proposed by R. Feynman [1] and the number of qubits and gates required to achieve quantum advantage is relatively small compared to other applications [2].

The advancements in quantum technologies have been significant in recent years and the milestone for achieving a fault-tolerance quantum computer has been accomplished step-by-step [3]. Many private companies and governments are also expanding their investigations to prepare for the age of quantum computing. IBM has been providing cloud service so that anyone who is interested in using superconducting quantum computers can access it from their laptop and its capability and services have been improved over time. But we are now in the noisy intermediate-scale quantum (NISQ) era where its capability is quite limited and does not provide any values over classical HPCs to the real-world industry yet. It is important to understand the performance limit of the current state-of-the-art NISQ devices and to find out how we could improve quantum algorithms and quantum devices to get better results.

Variational quantum eigensolver (VQE) is the prototype NISQ algorithm to compute electronic ground state energies using shallow quantum circuits with classical parameter optimization technique [4]. Recently, the development of ansatz for parameterized quantum circuits and novel optimization schemes have been proposed for improving variational algorithm [5, 6]. In this work, we focus our attention on computing ground state energies of hydrogen molecules using VQE on IBM superconducting quantum computer to benchmark how large a quantum system can reliably be addressed using the state-of-the-art IBM quantum computer. We also apply error mitigation techniques [7] in our quantum simulations to improve the accuracy of results. It turns out that the current IBM quantum computers are still limited in computing a large quantum system and significant hardware and algorithmic advancements should be achieved so that quantum computing markets are more attractive enough than the classical counterpart. This work will provide quantum community benchmark results by the IBM quantum computer on computing molecular ground state energy using the NISQ algorithm and direction to further developments.

In this work, we investigate the ground state energies of hydrogen molecules using the VQE algorithm on IBM superconducting quantum computers with error mitigation. We first explore a H_2 molecule that requires 2 qubits and 3 parameters to be optimized in quantum circuits ($\sim$ 10 gates). As the number of qubits and gates in quantum circuits is small, it gives us acceptable outputs that are close to the results obtained by classic computers. There are no substantial bottlenecks in finding optimal parameters used in quantum circuits and obtaining accurate results against circuit noise. However, as we increase system size as H_4, the situation changes. The number of qubits increases to 6 and the number of parameters to be optimized increases to 26 in the case of UCCSD ansatz, giving rise to difficulty in finding optimized parameters in variational processes and accumulated noise because of the deep quantum circuit. We apply diverse error mitigation techniques to improve calculation results but the error mitigation does not improve results significantly.

We used VQE algorithm [4] to compute ground state energies of hydrogen molecules. For a single hydrogen molecule, we set the distance between two hydrogen atoms of 0.735 Å. We used PySCF [8] as a driver and used STO-3G minimal basis to expand electronic

*junho5.lee@lge.com

†hyukgeun.cha@lge.com

‡seonghyok.kim@lge.com

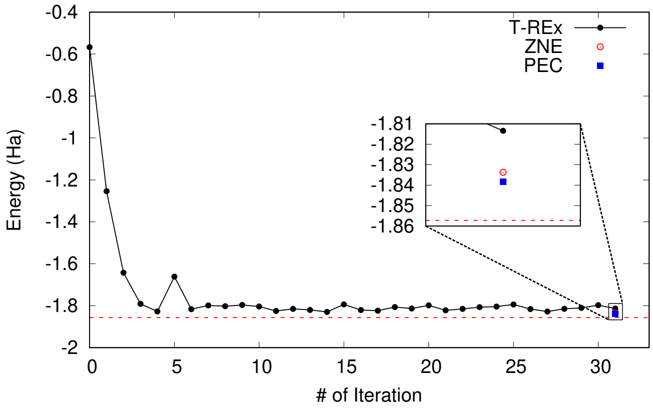


Figure 1: VQE calculation for a hydrogen molecule in IBM superconducting quantum computer. Error mitigation of T-REx has been applied during the variational process. After obtaining optimal parameters of quantum circuits from the calculation with T-REx (black line), two additional calculations using optimized parameters as initial parameters with zero-noise extrapolation (ZNE) (red-filled square) and probabilistic error cancellation (PEC) (blue-filled square) are done as denoted by filled squares. Red dashed line represents numpy result. For this simulation, we used `ibmq_kolkata` backend. Energy represents electronic energy only, not including nuclear energy.

orbitals and Parity mapper to transform fermionic operators to qubit operators. For parameterized quantum circuits, we used unitary coupled-cluster single and double (UCCSD) ansatz and hardware efficient ansatz (HAE) [9]. In HAE, we used a single representation of circuit and linear entanglement. We perform 1,000 shots for each quantum circuit. To optimize a set of parameters used in quantum circuits, we used the COBYLA optimizer. In order to obtain reference results, we used the NumPy minimum eigensolver algorithm. We apply error mitigation techniques such as Twirled readout error extinction (T-REx) [10], zero-noise extrapolation (ZNE), and probabilistic error cancellation (PEC) [11]. We used `ibmq_qasm_simulator` as a simulator and quantum devices of `ibmq_jakarta` and `ibmq_kolkata` backends.

We first perform VQE calculations for a single hydrogen molecule as shown in Fig. 1. It turns out that the quantum computer provides reasonable results (-1.81 Ha) compared to classical reference results (-1.86 Ha) in terms of accuracy, but not speed because of the variational approach that requires many iterations and computing overhead. In this case, we used UCCSD ansatz for the quantum circuit. The ground state energy converges to -1.81 Ha less than 10 iterations as there are only three parameters to be optimized in the quantum circuit. The noise is also not that significant in this calculation. We apply error mitigation such as T-REx (used as a default option), ZNE, and PEC. We perform single quantum computation (no iteration, but multiple shots for the given quantum circuit) applying ZNE and PEC using optimized parameters obtained by T-REx. We find that ZNE (-1.83 Ha) and PEC (-1.84 Ha) error mitigation slightly improve T-REx results. For the single hydrogen molecule case, noise in the result is not that significant, thereby observing the effect of error mitigation is not that clear. We enlarge the quantum system as four hydrogen atoms.

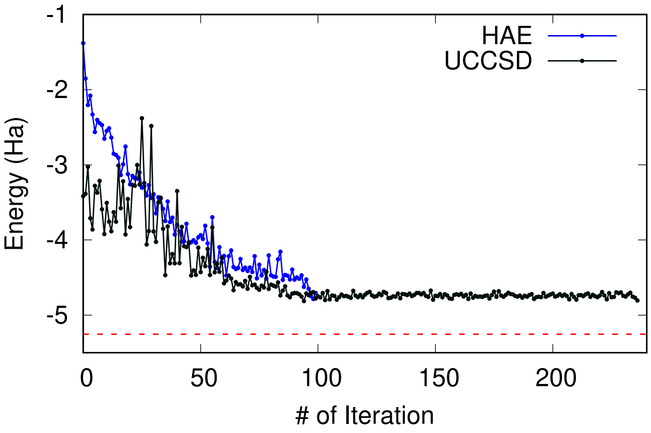


Figure 2: VQE calculation for H_4 using COBYLA optimizer. Hardware Efficient (blue) and UCCSD (black) ansatzes are compared.

We perform VQE calculations for two hydrogen molecules, H_4 in total so that we can systematically increase system size. We make the distance between two neighboring hydrogen atoms the same as 0.735 Å. In this case, the number of spin-orbitals is eight, and the number of qubits required to express a parameterized quantum circuit is six.

We perform VQE calculations on `ibmq_qasm_simulator` to see how they converge as shown in Fig. 2. For the COBYLA optimizer, HAE converges more monotonically than UCCSD ansatz which oscillates significantly at the early stage of iteration. More than ∼ 100 iterations, the UCCSD result converges to ∼ -4.81 Ha which is ∼ 8.6% deviation from the refernce value (∼ -5.26 Ha), thereby showing the limitation of the classic optimizer (or variational approach itself). We also try SPSA, ADAM, and L_BFGS_B optimizers, but COBYLA gives the best results given the computation setting mentioned above. Convergence of H_4 is quite slow compared to the H_2 case and the converged result of H_4 shows a large discrepancy compared to the reference result. This implies that even without the presence of noise of a real quantum computer, the difficulty of convergence of variational approach hinders practical applicability as the circuit complexity increases [12].

We then move to real quantum hardware. We find that without error mitigation there is substantial noise in the result although its iteration is converged to some number (∼ -3.50 Ha, pink line) compared to the simulator result (∼ -4.81 Ha, blue line) as shown in Fig. 3. We then perform VQE calculation with T-REx (∼ -3.13 Ha, red line) and get worse results than one without error mitigation. As we apply ZNE in our VQE calculation, the result be-

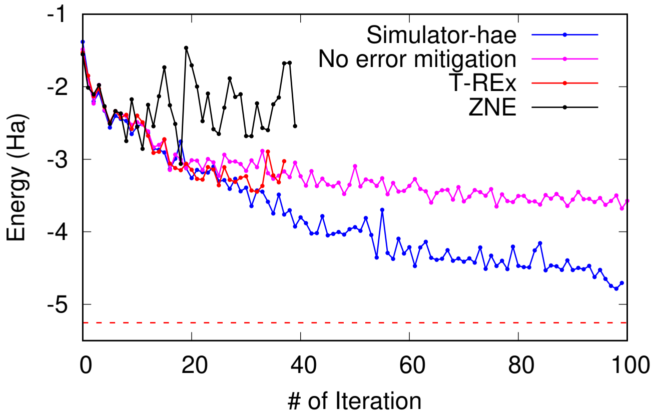


Figure 3: VQE calculations for H_4 with the same distance between nearest neighbor, forming a straight hydrogen chain. Hardware efficient ansatz and COBYLA optimizer are used. The blue line represents the simulator result. The pink line corresponds to results obtained by `ibmq_jakarta` backend without any error mitigation. Red and black lines represent T-REx and ZNE error mitigation results, respectively.

comes worse. It oscillates up to 40 iterations. This oscillatory behavior implies that iterative calculation with ZNE is not a proper way to improve results, contrary to previous results [13]. Instead, doing single quantum computation using optimized parameters obtained from T-REx or without error mitigation could be a proper way to get better results.

In summary, we perform VQE calculations for hydrogen molecules on IBM superconducting quantum computers. It turns out that as the system size increases, circuit complexity increases, giving rise to accumulated noise and difficulty in parameter optimization that hinders large-scale quantum simulation. Applying error mitigation also requires systematic investigation to get the right answer.

References

[1] R. Feynman. Simulating Physics with Computers. *Int. J. Theor. Phys.*, 21:467, 1982.

[2] T. Hoefler, T. Häner, and M. Troyer. Disentangling hype from practicality: On realistically achieving quantum advantage. *Commun. ACM*, 66:82, 2023.

[3] Google Quantum AI. Suppressing quantum errors by scaling a surface code logical qubit. *Nature*, 614:676, 2023.

[4] A. Peruzzo, J. McClean, P. Shadbolt, M. H. Yung, X. Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.*, 5:4213, 2014.

[5] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles Variational quantum algorithms. *Nat. Rev. Phys.*, 3:625, 2021.

[6] W. J. Huggins, B. A. O'Gorman, N. C. Rubin, D. R. Reichman, R. Babbush, and J. Lee. Unbiasing fermionic quantum Monte Carlo with a quantum computer. *Nature*, 603:416, 2022.

[7] Z. Cai, R. Babbush, S. C. Benjamin, S. Endo, W. J. Huggins, Y. Li, J. R. McClean, and T. E. O'Brien. Quantum error mitigation. arXiv.2210.00921

[8] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. McClain, S. Sharma, S. Wouters, and G. K.-L. Chan. PySCF: the Python-based simulations of chemistry framework. *WIREs Comput. Mol. Sci.* 8:e1340, 2018.

[9] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549:242, 2017.

[10] E. van den Berg, Z. K. Minev, and K. Temme. Model-free readout-error mitigation for quantum expectation values. *Phys. Rev. A*, 105:032620, 2022.

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

[12] L. Bittel and M. Kliesch. Training variational quantum algorithms is NP-hard. *Phys. Rev. Lett.* 127:120502, 2021.

[13] 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.

Distributed quantum computing with entanglement-assisted gate teleportation [1]

Jun-Yi Wu[1] * Kosuke Matsui[2] Tim Forrer[2] Akihito Soeda[3] [4]
Pablo Andres-Martinez[5] Daniel Mills[5] Luciana Henaut[5] Mio Murao[2]

[1] *Department of Physics, Tamkang University, 151 Yingzhuan Rd., Tamsui Dist., New Taipei City 25137, Taiwan, ROC*
[2] *The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan*
[3] *Principles of Informatics Research Division, National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
[4] *Department of Informatics, School of Multidisciplinary Sciences, SOKENDAI (The Graduate University for Advanced Studies), 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan*
[5] *Quantinuum, Terrington House, 13-15 Hills Road, Cambridge CB2 1NL, UK*

Abstract. The scalability of a quantum processor unit (QPU) in noisy intermediate-scale quantum computing is limited by its qubit number and connectivity. Utilizing entanglement-assisted local operations and classical communication (LOCC), distributed quantum computing (DQC) can break the topological limit of a QPU to enhance its scalability. To save entanglement cost in DQC for general quantum circuits, we propose an entanglement-efficient DQC protocol based on ”distributing enhanced by embedding”. The distributability and embeddability of a quantum circuit can be represented by graphs, based on which one can derive heuristic algorithms to reduce and determine a constructive upper bound on entanglement cost.

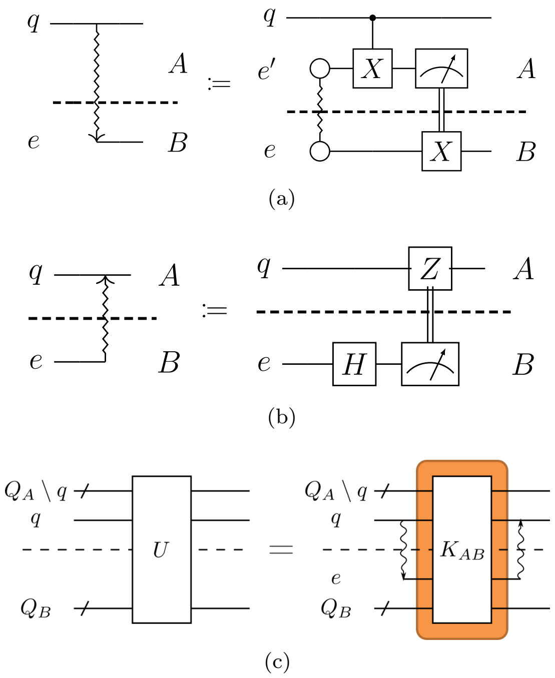
Keywords: AQIS, template

1 Background and motivation

The connectivity and the number of qubits on a QPU are believed to be key features to enhance quantum volume [2], which quantifies the scalability of a QPU. To scale up quantum computing, one has to overcome the intrinsic topological limits of a single QPU. To this end, one can employ entanglement shared between two QPUs to implement entanglement-assisted LOCC to extend the connectivity of qubits over multiple QPUs. Harnessing the topological structure of distributed entanglement over quantum internet [3], one can build a distributed quantum computing system with multiple QPUs to scale up quantum computing. As entanglement is costly resource in quantum networks, it is essential to develop a distributed quantum computing protocol that consumes as less entanglement as possible.

2 Distributing enhanced by embedding

For a circuit consisting of nonlocal control unitaries with the same control qubit, the DQC protocol introduced in [4] is optimal and utilizes only one pair of maximally entangled state, namely one ebit. Here, we refer to the protocol in [4] as EJPP protocol. For general circuits, we need to extend the EJPP protocol to a more general process called a packing process (Fig. 1 (c)), which starts with a pair of maximally entangled state (Fig. 1 (a)) and ends with measurement control operations (Fig. 1 (b)). The intermediate gates between the starting and ending are the kernel of a packing process. If the kernel is local, then the packing process is an entanglement-assisted LOCC, which we call a distributing process (Fig. 2 (a)).



(a)

(b)

(c)

Figure 1: The starting and ending processes: the symbol on the left side represents the operation given by the quantum circuit on the right side. A working qubit q and an auxiliary qubit e' belong to a local QPU A, while and an auxiliary qubit e belongs to another local QPU B. (a) The starting process. (b) The ending process. (c) A packing process.

*junyiwuphysics@gmail.com

The ultimate goal of distributed quantum computing is to transpile a quantum circuit into a circuit of distributing processes. Since each distributing process consumes an ebit, an entanglement-efficient DQC protocol should reduce the number of distributing processes. To this end, we introduce the so-called embedding process (Fig. 2 (b)), which embeds the original unitary into a packing process with additional local correcting gates.

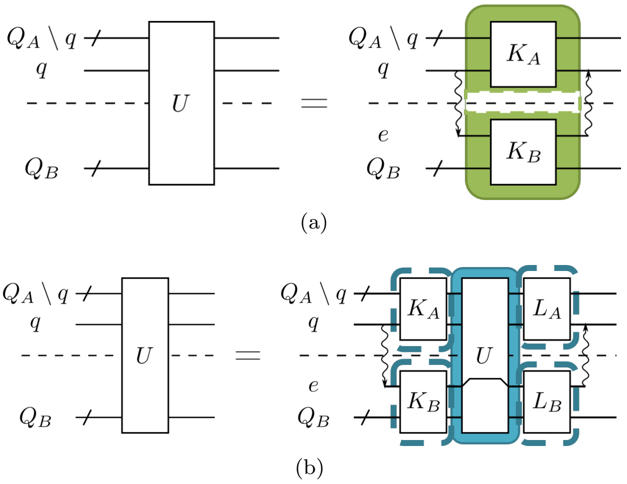


Figure 2: (a) A distributing process (b) An embedding process

Such an embedding process allows the merging of two non-sequential distributing processes. As it is shown in Fig. 3, the distributing processes of the control unitaries C_U and C_V are merged through the embedding of C_X. Such an embedding saves one ebit in DQC. We therefore introduce the so-call distributable packets, which are blue-highlighted on the left quantum circuit of Fig. 3. They fully describe the embedding-enhanced distributing structure of a quantum circuit.

3 Packing graph and conflict graph

From a set of embedding rules, one can then identify all the distributable packets of a quantum circuit. Each distributable packet is connected with another distributable packet through a nonlocal gate. From these connected the distributable packets one can induce a packing graph of a quantum circuit. As is it shown in Fig. 4 (a), the green-highlighted circuit nodes are the distributable packets, from which one obtains a packing graph in Fig. 4 (b). Each edge in the packing graph represents the nonlocal gates in the circuit.

To transpile a quantum circuit into distributing processes, we need to implement distributing processes on the vertices of the packing graph, such that all edges (nonlocal gates) are distributed. Technically speaking, we need to find the set of vertices that cover all the edges, namely the vertex cover of the packing graph. The minimum entanglement cost is then determined by the minimum vertex cover of the packing graph.

One important remark on our DQC protocol is that not all of the distributable packets can implement distributing processes at the same time. A distributing process may prevent the implementation of another distributing process. For example, in Fig. 4 (a), the distributing process on the packet 8 prevents the distributing process on the packet 0. Such incompatibility is due to the conflict between two embedding processes involved in the corresponding distributable packets, which are highlighted in violet in Fig. 4 (a). Such incompatibility is represented by a conflict edge between the two distributable packets. For a set of distributable packets selected from the minimum vertex cover of a packing graph, there may exist pairs of incompatible distributable packets represented by a set of conflict edges. Together with distributable packets, the set of conflict edges forms a conflict graph of the quantum circuit as it is shown in Fig. 4 (c).

One has to resolve the incompatibility in the selected distributable packets by removing the conflict edges in the conflict graph. To this end, one removes one of the embedding processes connected by a conflict edge. The removal of an embedding process leads to the splitting of the corresponding distributable packet and an increase of entanglement cost by one ebit. As it is shown in Fig. 5 (a), the conflict between the packets 0 and 8 is resolved by spliting the distributable packet 0 into $0a$ and $0b$. It results in an updated packing graph in Fig. 5 (b). The entanglement cost of the DQC of the circuit is then equal to the number of the minimum vertex cover of the updated packing graph, which is equal to 4.

4 The packing algorithm

The packing graph and conflict graph of a quantum circuit completely represent the distributability, embeddability and incompatibility of a quantum circuit. We have derived heuristic packing algorithms based on the packing graph and the conflict graph to find an entanglement-efficient packing strategy for a given quantum circuit distributed over two QPUs. The algorithms utilize the minimum vertex cover of bipartite graphs, which can be efficiently implemented on a classical computer. As an application, we employ our algorithm to find the DQC strategy of a 4-qubit UCC cluster circuit [5, 6]. It shows a reduction of entanglement cost from 64 ebits to 17 ebits, which saves 47 ebits.

5 Conclusion

As a conclusion, we have developed a protocol for DQC, which is based on two types of entanglement-assisted quantum gate teleportation, namely distributing processes and embedding processes. The protocol in this work can be summarized as “distributing enhanced by embedding”. We have developed heuristic packing algorithms to transpile a given quantum circuit into a distributed circuit assisted by a small number of entangled pairs. Such a protocol can facilitate large-scale quantum computing in a quantum network of QPUs.

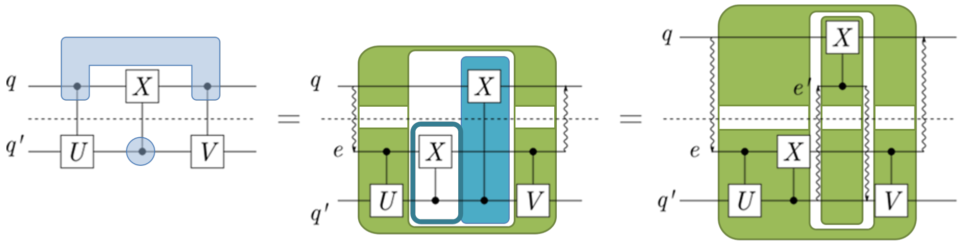
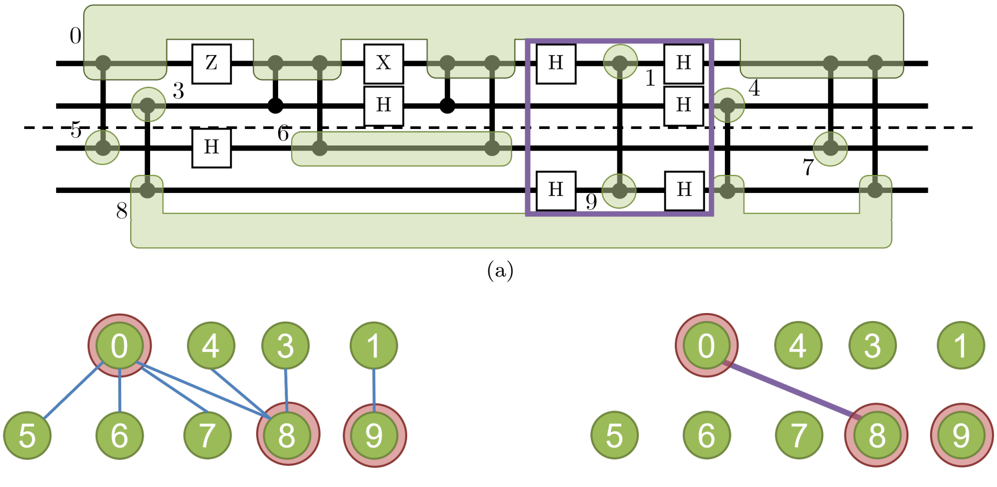


Figure 3: Merging of two distributing processes through an embedding process.

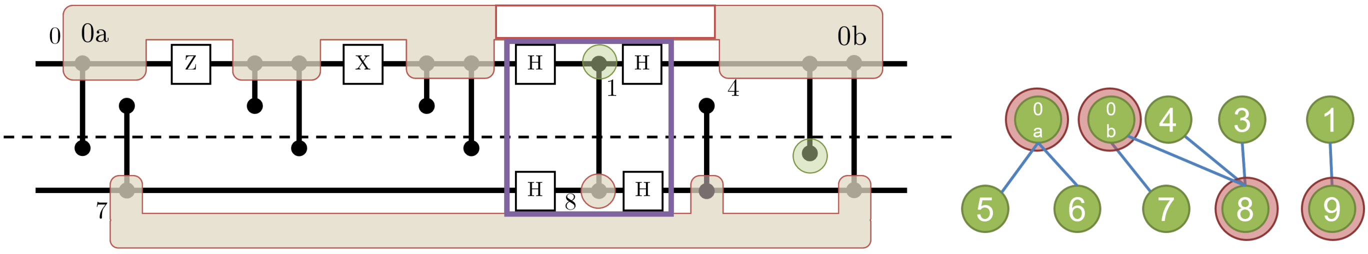


(a)

(b)

(c)

Figure 4: (a) Distributable packets. (b) Packing graph (c) Conflict graph



(a)

(b)

Figure 5: (a) Distributable packets after resolving the conflict. (b) Packing graph after resolving the conflict

References

[1] J.-Y. Wu, K. Matsui, T. Forrer, A. Soeda, P. Andrés-Martínez, D. Mills, L. Henaut, and M. Murao. Entanglement-efficient bipartite-distributed quantum computing with entanglement-assisted packing processes. *arXiv:2212.12688.*

[2] A. W. Cross, L. S. Bishop, S. Sheldon, P. D. Nation, and J. M. Gambetta. Validating quantum computers using randomized model circuits. *Phys. Rev. A*, 100:032328, 2019.

[3] H. J. Kimble. The quantum internet. *Nature*, 453:1023–1030, 2008.

[4] J. Eisert, K. Jacobs, P. Papadopoulos, and M. B. Plenio. Optimal local implementation of nonlocal quantum gates. *Phys. Rev. A*, 62:052317, 2000.

[5] A. G. Taube and R. J. Bartlett. New perspectives on unitary coupled-cluster theory. *International Journal of Quantum Chemistry*, 106(15):3393–3401, 2006.

[6] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1), 2014.