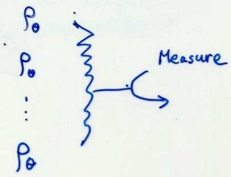
First ORder Asymptotic Theory of Quantum Statistical Estimation

& its application

K. Matsumoto

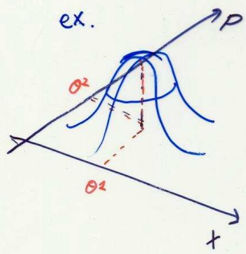
Quantum Computation & Information project, JST.

- When Collectiver measurement is effective?



- Consideration of non-commutativity from estimation theoretical poit of view.

ex.



Quantum State Estimation

$$\mathcal{M} = \{\rho_\theta \mid \theta \in \Theta \subseteq \mathbb{R}^m\},$$

ρ_θ in $\mathcal{H}$

POVM $\{M_\omega\}_{\omega\in\Omega}$, $\sum_{\omega\in\Omega} M_\omega = I$

M : measurement

$$\boxed{\theta}\triangleleft \rightsquigarrow \rho_\theta \hookrightarrow \omega \longmapsto \hat{\theta}(\omega)$$

Source. data estimate

① $\rho_\theta \in \mathcal{M}$ is assumed

② θ: unknown

EX.

$$\mathcal{M} = \{\rho \mid \rho > 0\}$$

$$\rho = \frac{1}{2}\begin{bmatrix} 1+\theta^1 & \theta^2+i\theta^3 \\ \theta^2-i\theta^3 & 1-\theta^1 \end{bmatrix} \qquad \theta = (\theta^1, \theta^2, \theta^3)$$

etc.

Asymptotic Setting (1)

measurement M_n in $\mathcal{H}^{\otimes n}$

$$\boxed{\theta}\triangleleft \rightsquigarrow n\left\{\boxed{\begin{matrix}\rho_\theta \\ \vdots \\ \rho_\theta\end{matrix}}\right. \hookrightarrow \omega_n \longrightarrow \hat{\theta}_n(\omega)$$

$= \rho_\theta^{\otimes n}$

estimate

ex. MLE (Maximum Likelyhood Estimate)

$n \gtrsim 10^2 \sim 10^3$

① Separable measurement • of practical importance

ρ_θ ρ_θ $\cdots$ ρ_θ

$M^1 \curvearrowright \omega^1$, $M^2 \curvearrowright \omega^2$, $\cdots$, $M^n \curvearrowright \omega^n$

$M_n = (M^1, \cdots, M^n)$

$\omega_n = (\omega^1, \cdots, \omega^n)$

② corrective measurement • of theoretical importance

$\boxed{\rho_\theta \ \rho_\theta \ \cdots \ \rho_\theta}$

$M_n \curvearrowright \omega_n$

Asymptotic Setting (2)

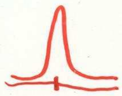
$E_\theta[d(\hat{\theta}_n,\theta)] = O(\frac{1}{n})$

measure off error

$\therefore \lim_{n\to\infty} n E_\theta[d(\hat{\theta}_n,\theta)] \to \min$

$= \sum g_{ij}(\hat{\theta}_n^i-\theta^i)(\hat{\theta}_n^j-\theta^j) + o(\|\hat{\theta}_n-\theta\|^2)$

negligeble



θ

$C_\theta(G)$

$C_\theta^Q(G)$

$G=[g_{ij}]$

$V_\theta[\hat{\theta}] = E_\theta[(\hat{\theta}_n^i-\theta^i)(\hat{\theta}_n^j-\theta^j)]$

mean square error matrix

in $\mathbb{R}^m$

$$\boxed{\lim_{n\to\infty} n E_\theta[\mathrm{Tr}\, G V_\theta[\hat{\theta}]] \to \min}$$

= Cramér Rao bound

condition: $\hat{\theta}_n$ must satisfy

$\hat{\theta}_n \to \theta \quad (n\to\infty)$ in probability

1dim - parameter model (Helstrom, Nagaoka)

ρ_θ , θ : scalor

$\theta \in \Theta \subset \mathbb{R}$

Classical counterpart:

$p(x,\theta)$, $\theta\in\Theta\subset\mathbb{R}$

$V_\theta[\hat{\theta}] \ge \frac{1}{n}\frac{1}{J_\theta} + o(\frac{1}{n})$ (Cramér-Rao)

$J_\theta = E_\theta\left(\frac{d}{d\theta}\log p(x,\theta)\right)^2$ (Fisher Information)

logarhysmic derivative

Quantum case.

$J_\theta^S := \mathrm{Tr}\,\rho_\theta (L^S(\theta))^2$ (SLD Fisher Information)

SLD $L^S(\theta)$: The solution of the following eq.

$\frac{d\rho_\theta}{d\theta} = \frac{1}{2}(L^S(\theta)\rho_\theta + \rho_\theta L^S(\theta))$

If $[L^S(\theta),\rho_\theta]=0$, $L^S(\theta)=\frac{d}{d\theta}\log\rho_\theta$

Optimal Measurement $\Leftrightarrow$ projection onto eigenvectors of $L^S(\theta)$

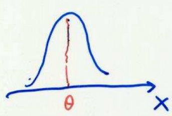
1-dim - parameter model (2)

(1) $C^Q = C$. (Collective measurement is NOT effective)

(2) $C_\theta^Q(G) = C_\theta(G) = G/J_\theta^S$

(3) Theory is quite parallel with classical theory

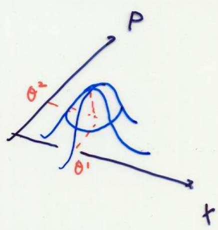
☺ Non-commutativity is absent.



Multi-parameter models

ρ_θ : θ is a vector.

$\theta \in \Theta \subset \mathbb{R}^m$



From the result of 1-dim parameter model theory,

$V_\theta[\hat{\theta}^i] \ge 1/[J_\theta^S]_{ii}$ ← Achievable, If and only if concentrate θ^i & forget other parameters.

$[J_\theta^S]_{ij} := \mathrm{Re}\,\mathrm{Tr}\,\rho_\theta L_i^S L_j^S$

SLD Fisher Information Matrix

$\therefore C_\theta^Q(G) \gtrsim \mathrm{Tr}\, G[J_\theta^S]^{-1}$

$C_\theta^*(G) \gtrsim$:

① Equality is achievable if $[L_i^S(\theta), L_j^S]=0$

② Gap between r.h.s & l.h.s. comes from non-commutativity

Pure State Models (Matsumoto, e-print 1998, JPA 2002)

$$\mathcal{M} = \{ \rho_\theta : \rho_\theta \text{ is a pure state } . \theta \in \Theta \}$$

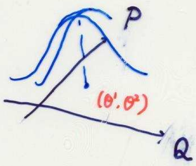
$\rho_\theta = |\phi_\theta\rangle\langle\phi_\theta|$

ex.1

$$|\phi_\theta\rangle = e^{-i(\theta^1 P - \theta^2 Q)} |\phi_0\rangle$$

P : Momentum Operator

Q : Position Operator



Estimate of (θ^1, θ^2) $\Leftrightarrow$ measurement of the average of the position & of the momentum

ex.2

$|\phi_\theta\rangle$: all the state vectors in $\mathbb{C}^d$

c.f. Massar & Popescu (1995) (← baysian)

Hayashi (1996) (← baysian & Minimax)

studied this ~~model~~ in non-asymptotic setting

Pure State Models : Obtained results

① $C^Q = C = C^H$

- Collective measurement is NOT effective, asymptotically

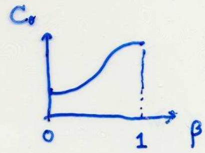
  In the case of ex2. Hayashi (1996) pointed out even feedback is not needed

- We can calculate C^Q, C !!

② Let $d(\hat{\theta}, \theta) := 1 - F(\rho_{\hat{\theta}}, \rho_\theta)$ — $\hat{\theta}$: estimate, θ: true value, F: Fidelity

(2-1) if $m = 2$,

$$C_\theta = \frac{4}{1 + \sqrt{1-\beta^2}}$$



β : a parameter which characterize $\mathcal{M}$.

$$\beta := \frac{|\tilde{J}_{12}|}{\sqrt{\det J^S_\theta}}$$

J^S_θ : SLD Fisher Info. Matrix

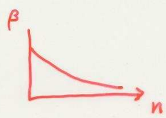
$$[\tilde{J}_\theta]_{ij} := \mathrm{Im}\, \mathrm{Tr}\, \rho_\theta [L^S_i, L^S_j]$$

ex1.

$$|\phi_\theta\rangle = e^{-i(\theta^1 P - \theta^2 X)} |n\rangle$$

$|n\rangle$: number state — $H = -\frac{\partial^2}{\partial x^2} + x^2$

$$\beta = \frac{1}{2n+1}$$



$n = 0$: non-commutativity is strongest.

i.e. simultanious estimation of position & momentum is very hard

$n \to$ large simultanious estimation of position & momentum becomes easier.

(2-2)

$$C_\theta = \sum_{i=1}^{[m/2]} \frac{2}{1 + \sqrt{1-\beta_i^2}}$$

β_i : eigenvalues of $J^{S-1} \cdot \tilde{J}$

Gaussian State Model (Hayashi 1998. QCM)

First example of $C^Q \neq C$, experimentally realizable

$$\rho_\theta = c_{\theta_3} U_{\theta^1,\theta^2} e^{-\theta^3 \hat{N}} U^*_{\theta^1,\theta^2}$$

$$U_{\theta^1,\theta^2} = e^{-i(\theta^1 P - \theta^2 X)}$$

$\int R_z\, |z\rangle\langle z|\, d^2 z$

$\hat{N}$: number operator

P, X : quadratures $[P, X] = i$

Optimal Separable Measurement.

$\boxed{H}$: measurement of complex amplitude $\{ |z\rangle\langle z| \}_{z \in \mathbb{C}}$

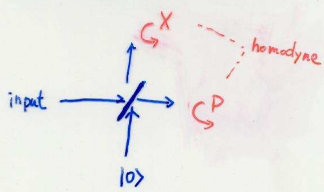
$|z\rangle$: coherent state.

$\rho_\theta \rightsquigarrow \boxed{H}$

$\rho_\theta \rightsquigarrow \boxed{H}$

$\vdots$

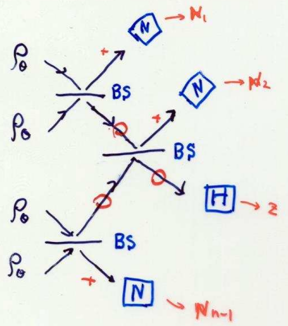
$\rho_\theta \rightsquigarrow \boxed{H}$



$$\hat{\theta}^1 + i\hat{\theta}^2 := \frac{1}{n} \sum_{i=1}^n z_i$$

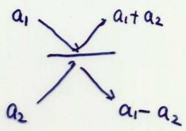
$$\hat{\theta}^3 := 1 \Big/ \left[\frac{1}{n-1} \{ z_i - (\hat{\theta}^1 + i\hat{\theta}^2) \}^2 - 1 \right]$$

Optimal Collective measure



$\boxed{N}$: photon counter

BS : beam spliter



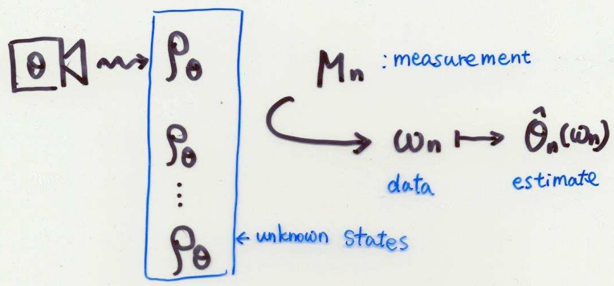
$\hat\theta^1 + i\hat\theta^2 = z$

$\hat\theta^3 = 1/\left[\frac{1}{n-1}\sum_{i=1}^{n-1} N_i\right]$

Problems in implimentation

1. photon counter → with these problems, still $C_\theta - C$ will be observable.
   - efficiency (~15% ?)
   - dark count (≪)
   - $|0\rangle$ $|1\rangle$ $|2\rangle$ $|3\rangle$ .. cannot 'count' ?
2. How to make a wave packet wait to interact with next comming wave packets without loss ? → Solved

Collective Setting



$$\lim_{n\to\infty} n E_\theta[\mathrm{Tr}\, G V_\theta[\hat\theta_n]] \to \min$$

- Over all Measurement in $\mathcal{H}^{\otimes n}$
- $\hat\theta_n \to \theta$ $(n\to\infty)$

$$C^H_\theta(G) := \min_{\substack{\{X^i\} \\ X^i = (X^i)^\dagger \\ \mathrm{Tr} \frac{\partial \rho_\theta}{\partial \theta^i} X^j = \delta^j_i}} \mathrm{Tr}\, G Z_\theta[\{X^i\}] + \mathrm{Tr\,abs}\,\mathrm{Im}\, G Z_\theta$$

where $Z_\theta[\{X^i\}] = [\mathrm{Tr}\,\rho_\theta X^i X^j]$

(Holevo - Belavkin's Bound)

$$\min \lim_{n\to\infty} n\, \mathrm{tr}\, G V_\theta[\hat\theta] \geq C^H_\theta(G)$$

in the following,
achievement of the equality is discussed.

Conjecture: in Corrective Setting (if we can use all corrective measurement)
Let $\mathcal{M} = \{\rho_\theta : \theta \in \Theta \subset \mathbb{R}^m\}$.
Arbitrary model.
then $C^H_\theta(G)$ is the asymptotic bound.

- This Conjecture is rigorously proved in the following cases.
  - Group representation of GL.
  - "Central Limit"

1. $\mathcal{H} = \mathbb{C}^2$ (Hayashi,)
2. pure state model. i.e. ρ_θ is pure $\forall \theta \in \Theta$ (Matsumoto, 1996)
3. total positive model. i.e. $\mathcal{M} = \{\rho \mid \rho > 0\}$ (Matsumoto, 1998)

- Naive prove of The assertion is done.
- (CR- bound) $\geq C^H_\theta(G)$ is known.

Full model

$\mathcal{H} \cong \mathbb{C}^d$

$$\rho_\theta = \rho_{\varphi,p} = U_\varphi \begin{pmatrix} p_1 & & 0 \\ & \ddots & \\ 0 & & p_d \end{pmatrix} U_\varphi^*$$

$p = (p_1, p_2, \cdots, p_{d-1})$

$\cancel{0 < p_1 < p_2 \cdots < p_{d-1} < p_d}$

$\sum_{i=1}^{d} p_i = 1 \qquad 0 \le p_d \le p_{d-1} \le \cdots \le p_2 \le p_1 \le 1$

$\{U_\varphi\}_\varphi = SU(d)$

Estimation

$\boxed{\theta} \rightsquigarrow \rho_\theta \xrightarrow{\text{measurement}} \omega$: data

$\hat\theta = \hat\theta(\omega)$: estimate.

$E_\theta[d(\hat\theta(\omega), \theta)] \to \min$

Estimation of p (1)

$$E_i = \begin{array}{c} \\ i \rangle \end{array} \begin{pmatrix} 0 & & \\ & 1 & \\ & & 0 \end{pmatrix}$$

$E_{i,\varphi} = U_\varphi E_i U_\varphi^*$ ← projection onto eigenvectors of $\rho_{p,\varphi}$

$\tilde M_{n,\varphi} = \{\tilde M_{n,\varphi,\omega}\}$ POVM

$\tilde M_{n,\varphi,\omega} = \bigotimes_{k=1}^{n} E_{\omega_k,\varphi}$

$\hat p_i = \frac{1}{n}\{\#\text{ of } \omega_k \text{ st. } \omega_k = i\}$

$n V_\theta[\hat p] = [(p_i \delta_{ij} - p_i p_j)]$

Cramér-Rao

$\lim n V_\theta[\hat p] \ge [p_i \delta_{ij} - p_i p_j]$,

if $\hat p \to p \ (n \to \infty)$

- Optimal
- Dependent on φ, which is unknown

→ This estimate is useless in this case

Estimation of p (2)

- indepen of φ
- Asymptotically as efficient as $\tilde M_{n,\varphi}$

group representation theoretic approach

$$\rho^{\otimes n} = \begin{bmatrix} * & & & & & 0 \\ & * & & & & \\ & & * & & & \\ & & & * & & \\ 0 & & & & * & \\ & & & & & * \end{bmatrix}$$

arbitrary

fixed

irreducible component

$\mathbb{C}^{d\otimes n} \cong \bigoplus_{\mathbf{n}} \bigoplus_i V_{\mathbf{n},i}$ ('eigen space of total spin')

$V_{\mathbf{n},i}$: irreducible component with Yung index $\mathbf{n}$ ('Total spin')

$\mathbf{n} = (n_1, n_2, \cdots, n_d)$, $n_1 > n_2 > \cdots > n_d$

$P^n_{\mathbf{n}}$: Projection onto $\bigoplus_i V_{\mathbf{n},i}$

Estimation of p (3)

Let $M_n = \{P^n_{\mathbf{n}}\}$, $\hat p_{i,n} = \frac{n_i}{n}$

then, ..

① $\lim_{n\to\infty} n E_\theta[(\hat p_{i,n} - p_i)(\hat p_{j,n} - p_j)] = \delta_{ij} p_i - p_i p_j$

Asymptotically optimal

② $\sum_{\mathbf{n}} P^n_{\mathbf{n}} \rho_\theta^{\otimes n} P^n_{\mathbf{n}} = \rho_\theta^{\otimes n}$

→ 'Do not distroy' the state if data is not observed

③ $\sum_{\mathbf{n} \in R_{\varepsilon,p}} P^n_{\mathbf{n}} \rho_\theta^{\otimes n} P^n_{\mathbf{n}} \doteqdot \rho_\theta^{\otimes n}$

Difference is $O(e^{-nD})$

$R_{\varepsilon,p} = \{\mathbf{n} : \|\frac{\mathbf{n}}{n} - p\| < \varepsilon\}$

→ universal source coding

Theoretical Applications.

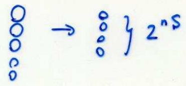
① Universal Data Compression

Josza . Holodekies . ... upper limit of entropy is known

Hayashi & Matsumoto ... arbitrary i.i.d. source.
(PRA, to be published)

- Optimality of the coding scheme is proven by estimation theory

Data compression $\Rightarrow$ estimate of entropy



② Universal Distortion-free Concentration

$$\begin{array}{ccc} \circ\!\sim|\phi\rangle\sim\!\circ & & \circ\!\sim|\Phi_+\rangle\sim\!\circ \\ \circ\!\sim|\phi\rangle\sim\!\circ & \Rightarrow & \vdots \\ \vdots & & \\ \circ\!\sim|\phi\rangle\sim\!\circ & & \circ\!\sim|\Phi_+\rangle\sim\!\circ \end{array}$$

unknown state

perfect Bell pair

Quantum Statistics has Just Started!

- More incorpolation with experiment.
- many unsolved problems
  - estimation of process
  - non-i.i.d case
  - ⋮