Bayesian Nagaoka-Hayashi Bound for Multiparameter Quantum-State Estimation Problem

鈴木淳 (Jun Suzuki)

Graduate School of Informatics and Engineering
The University of Electro-Communications

2024.2.22, quantum TUT workshop

UEC

科研費 KAKENHI

Introduction

- Bayesian lower bounds for parameter estimation about quantum states
- Quantum versions (Personick PhD thesis 1969)

  Quantum versions of van Tree, SLD type, Bhattacharyya, Barankin bounds

  SLD type bound: Optimal bound for single (scalar) parameter estimation
- Various bounds were proposed so far

- <u>Contribution:</u>

  A new class of Bayesian bounds

  - Bayesian Holevo bound

  - Bayesian Nagaoka-Hayashi bound

  JS, IEICE Transactions: special section (2024). [arXiv: 2302.14223]

Bayesian estimation in classical statistics

Classical Setting 1

- $M = \{p_\theta \,|\, \theta \in \Theta\}$ (n-parameter model on $\mathcal{X}$, $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$)
- $\pi(\theta)$ (Prior distribution on Θ)
- $p_X(x) := \int_\Theta d\theta \, \pi(\theta) p_\theta(x)$ (Marginal, evidence)
- $q(\theta|x) := \dfrac{\pi(\theta)p_\theta(x)}{p_X(x)}$ (Posterior)
- $\hat{\theta} = (\hat{\theta}_i)$: $\mathcal{X} \to \Theta$ (Estimator)
- Bayesian Mean Square Error Matrix:
  $$R_B[\hat{\theta}] := \left[\int_\Theta d\theta \int_{\mathcal{X}} dx \, \pi(\theta) p_\theta(x) (\hat{\theta}_j(x) - \theta_j)(\hat{\theta}_k(x) - \theta_k) \right]$$
- Optimal estimator: $\hat{\theta}_{\mathrm{opt},i}(x) := \int_\Theta d\theta \, \theta_i \, q(\theta|x) = E_q[\theta_i]$

Classical Setting 2

Bayesian lower bound for R_B

- van Tree inequality: $R_B[\hat{\theta}] \geq (J_B)^{-1}$

$$J_B := J_\pi + \int_\Theta d\theta\, \pi(\theta) J(\theta)$$

$$J_\pi := \left[\int_\Theta d\theta\, \pi(\theta) \frac{\partial \log \pi(\theta)}{\partial \theta_i} \frac{\partial \log \pi(\theta)}{\partial \theta_j}\right]$$

$$J(\theta) := \left[\int_\mathcal{X} dx\, p_\theta(x) \frac{\partial \log p_\theta(x)}{\partial \theta_i} \frac{\partial \log p_\theta(x)}{\partial \theta_j}\right]$$

- Key point: Inner product and the covariance inequality.

  Cf. Gill-Levit bound [Gill and Levit, *Bernoulli* **1**, 59 (1995)]

Tsang, *Physics-inspired forms of the Bayesian Cramér-Rao bound*, Physical Review A, **102**, 062217 (2020)

Classical Setting 3

Optimal Bayes Risk (Alternative expression):

$$R_B[\hat{\theta}_{\mathrm{opt}}] = M - \mathcal{K}$$

$$M := \left[\int_\Theta d\theta\, \pi(\theta)\, \theta_j \theta_k\right] \quad \text{(2nd moment matrix for } \theta\text{)}$$

$$\mathcal{K} := \left[\int_\mathcal{X} dx\, p_X(x) \hat{\theta}_{\mathrm{opt},j}(x) \hat{\theta}_{\mathrm{opt},k}(x)\right] \quad \text{(2nd modment for } \hat{\theta}_{\mathrm{opt}}\text{)}$$

- Bayesian logarithmic derivative function: $\hat{\theta}_{\mathrm{opt},j}(x)\, p_X(x) = \int_\Theta d\theta\, \theta_j \pi(\theta) p_\theta(x)$

$$\Leftrightarrow \int_\Theta d\theta\, (\hat{\theta}_{\mathrm{opt},j}(x) - \theta_j) \pi(\theta) p_\theta(x) = 0$$

$\hat{\theta}_{\mathrm{opt},j}(x) \times$(marginal)$=$(marginal expectation of θ)

Quantum Estimation Theory

(Finite sample theory)

Helstrom bound (1967)
Yuen-Lax bound (1973)
Holevo bound (1982)
Nagaoka bound (1989)
Nagaoka-Hayashi bound (2021)

Conlon, JS, Lam, Assad, npj QI (2021)

Hayashi-Ouyang (‘**tight**’ bound, 2023)

Quantum Setting

- Model: $M = \{S_\theta \mid \theta \in \Theta\}$; $S_\theta \in \mathcal{L}(\mathcal{H}) = \mathbb{C}^{d\times d}$, $S_\theta \geq 0$, $\mathrm{Tr}\,(S_\theta) = 1$
- POVM: $\Pi = \{\Pi_x\}_{x\in\mathcal{X}}$; $\Pi_x \geq 0$, $\displaystyle\int_{\mathcal{X}} dx\,\Pi_x = I_d$ or $\sum_x \Pi_x = I_d$ (Identity matrix)
- Probability model: $p_\theta(x) := \mathrm{Tr}\,(S_\theta \Pi_x)$ (Probabilistic rule)
- $\hat{\theta}$: $\mathcal{X} \to \Theta$ (Estimator)
- MSE: $V_\theta[\Pi, \hat{\theta}] := \left[\displaystyle\int_{\mathcal{X}} dx\, p_\theta(x)(\hat{\theta}_j(x) - \theta_j)(\hat{\theta}_k(x) - \theta_k)\right]$
- A-optimal design problem for $W \geq 0$ (Cost, utility, weight matrix):

  $C_{\mathrm{opt}} := \inf_{\Pi,\hat{\theta}}$ Tr $\left\{W V_\theta[\Pi, \hat{\theta}]\right\}$ subject to $(\Pi, \hat{\theta})$: locally unbiased
- No closed formula for C_{opt} in general

Cf. Conic programming (Hayashi-Ouyang), Convex optimization (Zhang-JS in preparation)

Helstrom (SLD) bound (1967)

Symmetric logarithmic derivative (quantum score function) $L_{\theta,j}$:

$$\frac{\partial S_\theta}{\partial \theta_j} = \frac{1}{2}\left(S_\theta L_{\theta,j} + L_{\theta,j} S_\theta\right) \text{ (Lyapunov equation)}$$

$$C_{\text{opt}} \geq C_{\text{Hel}}$$

$$C_{\text{Hel}} := \mathsf{Tr}\left\{W J_{\text{SLD}}(\theta)^{-1}\right\}$$

$$J_{\text{SLD}}(\theta) := \left[\frac{1}{2}\mathrm{Tr}\left(S_\theta(L_{\theta,j}L_{\theta,k} + L_{\theta,k}L_{\theta,j})\right)\right]$$

- SLD quantum Fisher information (Real symmetric, positive semi-definite)
- Tight for scalar estimation ($n = 1$) and c-opt
- Key point: Inner product $\langle X, Y \rangle_S := \frac{1}{2}\mathrm{Tr}\left(S(X^\dagger Y + Y X^\dagger)\right)$

Yuen-Lax (RLD) bound (1973)

Right logarithmic derivative (quantum score function) $\tilde{L}_{\theta,j}$:

$$\frac{\partial S_\theta}{\partial \theta_j} = S_\theta \tilde{L}_{\theta,j}$$

$$C_{\text{opt}} \geq C_{\text{YL}}$$

$$C_{\text{YL}} := \mathsf{Tr}\left\{W \mathrm{Re}\, J_{\text{RLD}}(\theta)^{-1}\right\} + \mathsf{Tr}\left\{|W^{1/2} \mathrm{Im}\, J_{\text{RLD}}(\theta)^{-1} W^{1/2}|\right\}$$

$$J_{\text{RLD}}(\theta) := \left[\mathrm{Tr}\left(S_\theta \tilde{L}_{\theta,k} \tilde{L}^\dagger_{\theta,j}\right)\right]$$

- RLD quantum Fisher information (Complex Hermitian, positive semi-definite)
- Tight for Gaussian shift estimation
- Key point: Inner product $\langle X, Y \rangle^+_S := \mathrm{Tr}\left(SYX^\dagger\right)$

Cf. A family of bounds based on Petz monotone metric (quantum Fisher information)

Holevo bound (1982)

$$\boxed{C_{\rm opt} \geq C_{\rm Hol} \geq \max\{C_{\rm Hel}, C_{\rm YL}\}}$$

$$C_{\rm Hol} := \min_{X^n \in \mathcal{X}_\theta} h_\theta[X^n]$$

$$X^n = (X_1, \dots, X_n);\ X_j \in \mathcal{L}_h(\mathcal{H}) \quad \text{(Hermitian)}$$

$$\mathcal{X}_\theta := \left\{ X^n \,\middle|\, \mathrm{Tr}\,(S_\theta X_j) = 0,\ \mathrm{Tr}\left(\frac{\partial S_\theta}{\partial \theta_j} X_k\right) = \delta_{jk} \right\} \quad : \text{locally unbiasedness}$$

$$h_\theta[X^n] := \mathrm{Tr}\,\{W \mathrm{Re}\, Z_\theta[X^n]\} + \mathrm{Tr}\left\{ |\sqrt{W} \mathrm{Im}\, Z_\theta[X^n] \sqrt{W}| \right\}$$

$$Z_\theta[X^n] := [\mathrm{Tr}\,\{X_j S_\theta X_k\}] \quad n \times n \text{ positive semi-definite}$$

- Non-trivial optimization (semi-definite programming)
- Optimal rate for the first order asymptotics

  (Hayashi-Matsumoto, Guta-Khan, Yamagata *et al*, Yang *et al*...)

Nagaoka bound (1989)

$$\boxed{C_{\text{opt}} \geq C_{\text{Nag}} \geq C_{\text{Hol}} \geq \max\{C_{\text{Hel}}, C_{\text{YL}}\}}$$

- Tighter bound for two-parameter estimation

$$C_{\text{Nag}} := \min_{X^2 \in \mathcal{X}_\theta} n_\theta[X^2]$$

$$X^2 = (X_1, X_2) \in \mathcal{L}_h(\mathcal{H})^2 \quad \text{(Hermitian)}$$

$$\mathcal{X}_\theta := \left\{ X^2 \,\middle|\, \mathrm{Tr}\,(S_\theta X_j) = 0,\ \mathrm{Tr}\left(\frac{\partial S_\theta}{\partial \theta_j} X_k\right) = \delta_{jk} \right\} \quad : \text{ locally unbiasedness}$$

$$n_\theta[X^n] := \mathsf{Tr}\left\{ W \mathrm{Re}\, Z_\theta[X^2] \right\} + \sqrt{\mathrm{Det}\,\{W\}} \mathrm{Tr}\left(\left| S_\theta^{1/2} (X_1 X_2 - X_2 X_1) S_\theta^{1/2} \right| \right)$$

$$Z_\theta[X^n] := [\mathrm{Tr}\,\{X_j S_\theta X_k\}] \quad n \times n \text{ positive semi-definite}$$

Nagaoka-Hayashi bound (2021)

- Generalization of the Nagaoka bound for n-parameter estimation.
- Key idea: $\mathcal{H} \to \mathbb{H} := \mathbb{C}^n \otimes \mathcal{H}$ (extended Hilbert space)

$$\boxed{C_{\rm opt} \geq C_{\rm NH} \geq C_{\rm Hol} \geq \max\{C_{\rm Hel}, C_{\rm YL}\}}$$

$$C_{\rm NH} := \min_{X^n, \mathbb{L}} \left\{ \mathbb{T}\mathrm{r}\left\{\mathbb{S}_\theta \mathbb{L}\right\} \,\middle|\, \mathbb{L} \geq X X^\intercal \right\}$$

$$\mathbb{S}_\theta := W \otimes S_\theta \quad (\text{“State” on } \mathbb{H})$$

$$X^n = (X_1, \ldots, X_n);\ X_j \in \mathcal{L}_h(\mathcal{H}) \quad (\text{Hermitian})$$

$$\mathbb{L}: \text{ Hermitian on } \mathbb{H}, \text{and self-transpose} (\overset{\text{def}}{\Leftrightarrow} \forall jk,\ \mathbb{L}_{jk} = \mathbb{L}_{kj} \text{ for block rep.})$$

$$\mathcal{X}_\theta := \left\{ X^n \,\middle|\, \mathrm{Tr}\left(S_\theta X_j\right) = 0,\ \mathrm{Tr}\left(\frac{\partial S_\theta}{\partial \theta_j} X_k\right) = \delta_{jk} \right\} \quad \text{locally unbiasedness}$$

- Efficiently computable with SDP.

Cf. Properties (asymptotics, gap persistency, etc): arXiv: 2208.07386

Quantum Bayesian Estimation

Personick (van Tree, SLD type) (1969, 1971)
Helstrom-Liu-Gordon (SLD type) (1970)
Holevo (RLD type) (1970's)
Rubio-Dunningham (SLD type) (2020)
(Sidhu-Kok, Demkowicz-Dobrzanski *et al* (2020))
Tsang (Gill-Levit 1995) (2020)

Bayes Risk in Quantum Setting

- Model: $M = \{S_\theta \,|\, \theta \in \Theta\}$; $S_\theta \in \mathbb{C}^{d\times d}$, $S_\theta \geq 0$, $\mathrm{Tr}\,(S_\theta) = 1$
- $\pi(\theta)$ (Prior distribution)
- Measurement: $\Pi = \{\Pi_x\}_{x\in\mathcal{X}}$; $\Pi_x \geq 0$, $\int_{\mathcal{X}} dx\, \Pi_x = I_d$ (Identity matrix)
- $\hat{\theta}$: $\mathcal{X} \to \Theta$ (Estimator)
- A-optimal Bayes risk for $\{W(\theta)\}_{\theta\in\Theta}$ (Weight field):

$$R_B[\Pi, \hat{\theta}] := \int_\Theta d\theta \int_{\mathcal{X}} dx\, \pi(\theta) \mathsf{Tr} \left\{ W(\theta) V_\theta[\Pi, \hat{\theta}] \right\}$$

- For θ-independent case: $R_B[\Pi, \hat{\theta}] = \mathsf{Tr} \left\{ W\, V_B[\Pi, \hat{\theta}] \right\}$
- A-optimal Bayesian design problem

$$\mathcal{C}_{\mathrm{opt}} := \inf_{\Pi, \hat{\theta}} R_B[\Pi, \hat{\theta}] \qquad \text{(No unbiasedness condition)}$$

Personick Bound (SLD type)

Theorem
For any POVM and estimator, $V_B[\Pi, \hat{\theta}] \geq M - \mathcal{K}_{\mathrm{SLD}}$. *Hence,* $\mathcal{C}_{\mathrm{opt}} \geq \mathcal{C}_{\mathrm{Per}}$ *where* $\mathcal{C}_{\mathrm{Per}} := \mathrm{Tr}\left\{W(M - \mathcal{K}_{\mathrm{SLD}})\right\}$. *Further, this is tight for single-parameter estimation.*

$$M = \left[\int_\Theta d\theta\, \theta_j \theta_k \pi(\theta)\right]$$

$$\mathcal{K}_{\mathrm{Per}} = \left[\frac{1}{2}\mathrm{Tr}\left\{S_{\mathrm{B}}(\mathcal{L}_j\mathcal{L}_k + \mathcal{L}_k\mathcal{L}_j)\right\}\right] \text{ (real symmetric, positive)}$$

$$S_{\mathrm{B}} = \int_\Theta d\theta\, S_\theta \pi(\theta) \text{ (marginal state)}$$

$$D_{\mathrm{B},j} = \int_\Theta d\theta\, \theta_j S_\theta \pi(\theta) \text{ (marginal expectation of } \theta\text{)}$$

$$D_{\mathrm{B},j} = \frac{1}{2}\left(S_{\mathrm{B}}\mathcal{L}_j + \mathcal{L}_j S_{\mathrm{B}}\right) \quad \text{(Bayesian quantum logarithmic derivative)}$$

RLD Type Bound (Holevo)

Theorem

For any POVM and estimator, $V_B[\Pi, \hat{\theta}] \geq M - \mathcal{K}_{\mathrm{RLD}}$. *Hence,* $\mathcal{C}_{\mathrm{opt}} \geq \mathcal{C}_{\mathrm{Hol}}$ *where*
$\mathcal{C}_{\mathrm{Hol}} := \mathsf{Tr}\{WM\} - \mathsf{Tr}\{W \mathrm{Re}\, \mathcal{K}_{\mathrm{RLD}}\} + \mathsf{Tr}\{|W^{1/2} \mathrm{Im}\, \mathcal{K}_{\mathrm{RLD}} W^{1/2}|\}$

$$\mathcal{K}_{\mathrm{RLD}} = \left[\mathrm{Tr}\left\{ S_{\mathrm{B}} \tilde{\mathcal{L}}_k \tilde{\mathcal{L}}_j^\dagger) \right\} \right] \text{ (complex Hermitian, positive)}$$

$$S_{\mathrm{B}} = \int_\Theta d\theta\, S_\theta \pi(\theta) \text{ (marginal state)}$$

$$D_{\mathrm{B},j} = \int_\Theta d\theta\, \theta_j S_\theta \pi(\theta) \text{ (marginal expectation of } \theta\text{)}$$

$$D_{\mathrm{B},j} = S_{\mathrm{B}} \tilde{\mathcal{L}}_j \quad \text{(Bayesian quantum logarithmic derivative)}$$

Remark: This bound is tight for Gaussian shift estimation with a Gaussian prior.

Result

Bayesian Holevo bound

Bayesian Ngaoka bound

Bayesian Nagaoka-Hayashi bound

$$R_B[\Pi, \hat{\theta}] \geq \mathcal{C}_{\text{B-NH}} \geq \mathcal{C}_{\text{B-Hol}} \geq \max\{\mathcal{C}_{\text{Per}}, \mathcal{C}_{\text{Hol}}\}$$

Bayesian Nagaoka-Hayashi bound

Theorem

For any POVM Π and estimator $\hat{\theta}$, the following inequality holds for the Bayes risk.

$$R_B[\Pi,\hat{\theta}] \geq \mathcal{C}_{\text{B-NH}}$$

$$\mathcal{C}_{\text{B-NH}} := \min_{\mathbb{L},X} \left\{ \mathbb{Tr}\left\{\bar{\mathbb{S}}\mathbb{L}\right\} - \mathbb{Tr}\left\{\bar{D}X^{\intercal}\right\} - \mathbb{Tr}\left\{X\bar{D}^{\intercal}\right\}\right\} + \overline{w}. \tag{1}$$

Here optimization is subject to:
$\forall jk, \mathbb{L}_{jk} = \mathbb{L}_{kj}$, $\mathbb{L}_{jk}$: *Hermitian,* X_j: *Hermitian, and* $\mathbb{L} \geq XX^{\intercal}$.

$$\mathbb{S}_{jk}(\theta) := W_{jk}(\theta) \otimes S_{\theta}, \qquad \bar{\mathbb{S}} := \int_{\Theta} d\theta\, \mathbb{S}(\theta)\pi(\theta) \tag{2}$$

$$D_j(\theta) := \sum_k W_{jk}(\theta)\theta_k S_{\theta}, \quad \bar{D} := \int_{\Theta} d\theta\, D(\theta)\pi(\theta) \tag{3}$$

$$w(\theta) := \sum_{j,k} \theta_j W_{jk}(\theta)\theta_k, \quad \overline{w} := \int_{\Theta} d\theta\, w(\theta)\pi(\theta) \tag{4}$$

Parameter Independent Weight Matrix

- To simplify further, we set $W(\theta) = W$ (No θ dependence)

$$\overline{\mathbb{S}} = W \otimes S_{\mathrm{B}},$$
$$\overline{D}_j = \sum_k W_{jk} D_{\mathrm{B},k},$$
$$\overline{w} = \mathsf{Tr}\left\{WM\right\},$$

where
$$S_{\mathrm{B}} = \int_{\Theta} d\theta\, \pi(\theta) S_\theta$$
$$D_{\mathrm{B},j} = \int_{\Theta} d\theta\, \pi(\theta) \theta_j S_\theta$$
$$M = \left[\int_{\Theta} d\theta\, \theta_j \theta_k \pi(\theta)\right]$$

Bayesian Holevo Bound

Corollary

For a parameter independent weight matrix, the Bayesian Holevo-type bound is

$$\mathcal{C}_{\text{B-Hol}} = \min_{X_j:\text{Hermitian}} \Big\{ \mathsf{Tr}\left\{W \,\mathrm{Re}\, \mathsf{Z}_{\mathrm{B}}[X]\right\} + \mathsf{Tr}\left\{ \left| W^{1/2}\, \mathrm{Im}\, \mathsf{Z}_{\mathrm{B}}[X] W^{1/2} \right| \right\} - \mathsf{Tr}\left\{W \mathsf{H}_{\mathrm{B}}[X]\right\} - \mathsf{Tr}\left\{W \mathsf{H}_{\mathrm{B}}[X]^{\mathsf{T}}\right\} \Big\} + \overline{w}, \quad (5)$$

where $\mathsf{Z}_{\mathrm{B}}[X]$ *and* $\mathsf{H}_{\mathrm{B}}[X]$ *are* $n \times n$ *matrices defined by*

$$\mathsf{Z}_{\mathrm{B},jk}[X] := \mathrm{Tr}\left\{S_{\mathrm{B}} X_k X_j\right\}$$
$$\mathsf{H}_{\mathrm{B},jk}[X] := \mathrm{Tr}\left\{D_{\mathrm{B},j} X_k\right\}$$

$$S_{\mathrm{B}} = \int_{\Theta} d\theta\, \pi(\theta) S_\theta,\; D_{\mathrm{B},j} = \int_{\Theta} d\theta\, \pi(\theta)\theta_j S_\theta$$

Bayesian Nagaoka Bound

Corollary

For two-parameter estimation with a parameter independent weight matrix, the Bayesian Nagaoka bound is expressed as

$$\mathcal{C}_{\text{B-Nag}} = \min_{X_j:\text{Hermitian}} \Big\{ \mathsf{Tr}\left\{ W \operatorname{Re} \mathsf{Z}_{\mathrm{B}}[X] \right\} + \sqrt{\operatorname{Det}\{W\}} \operatorname{TrAbs}\left\{ \left| S_{\mathrm{B}}^{1/2} (X_1 X_2 - X_2 X_1) S_{\mathrm{B}}^{1/2} \right| \right\} - \mathsf{Tr}\left\{ W \mathsf{H}_{\mathrm{B}}[X] \right\} - \mathsf{Tr}\left\{ W \mathsf{H}_{\mathrm{B}}[X]^{\mathsf{T}} \right\} \Big\} + \overline{w}. \quad (6)$$

In general, we have

$$\mathcal{C}_{\text{B-Nag}} \geq \mathcal{C}_{\text{B-Hol}}$$

Relation to Other Bounds

Theorem

(i) $\mathcal{C}_{\text{B-NH}} \geq \mathcal{C}_{\text{B-Hol}}$

(ii) $\mathcal{C}_{\text{B-Hol}} \geq \mathcal{C}_{\text{Per}}$, *and hence,* $\mathcal{C}_{\text{B-NH}} \geq \mathcal{C}_{\text{Per}}$

(iii) $\mathcal{C}_{\text{B-Hol}} \geq \mathcal{C}_{\text{Hol}}$, *and hence,* $\mathcal{C}_{\text{B-NH}} \geq \mathcal{C}_{\text{Hol}}$

In summary

$$\boxed{R_B[\Pi, \hat{\theta}] \geq \mathcal{C}_{\text{B-NH}} \geq \mathcal{C}_{\text{B-Hol}} \geq \max\{\mathcal{C}_{\text{Per}}, \mathcal{C}_{\text{Hol}}\}}$$

Summary and Outlook

- New tighter Bayesian bounds for the Bayes risk
- Bayesian Holevo bound and Bayesian Nagaoka-Hayashi bound
- Efficiently computable by semidefinite programming (SDP)
  (Python code will be available soon)

- Examples.
- Parameter dependent weight matrix
- Achievability