Linear random measurements of two non-commuting observables

Akio Fujiwara

METR 94-10 July 1994

Linear random measurements of two non-commuting observables

Akio Fujiwara,
University of Tokyo[*]

Abstract

A Cramér–Rao type bound for two–parameter quantum statistical model is studied, based on linear random measurements. It is shown that infinitely many (continuous potency) measurements which attain a lower bound exist. The RLD–bound for pure coherent models is shown to be most informative.

Keywords : quantum estimation theory, random measurement, squeezed state

[*]Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan.

1 Introduction

A quantum statistical model is a family of density operators ρ_θ defined on a certain separable Hilbert space $\mathfrak{H}$ with finite-dimensional real parameters $\theta = (\theta^i)_{i=1}^n$ which are to be estimated statistically [1][2]. For one parameter case, it is rather easy to find the most informative lower bound of the variance of measurement, since the Cramér–Rao bound with respect to the SLD is always locally attainable (global feature is another problem) [3][4]. On the other hand, for multi-parameter case, it has been an awkward problem to find attainable lower bound for the variance of measurement because of non–commutativity of the SLD's. To the author's knowledge, there has been only two remarkable results that derived explicit attainable bounds in essentially non–commutative situations. The first one is the RLD bound for quantum gaussian models derived by Yuen and Lax [5][2], which provided a theoretical background of optical communication theory. Another one is related to the spin 2×2 matrix representation derived by Nagaoka [6][7][8].

In this paper, we study attainable lower bounds in the class of linear random measurements and obtain an explicit form. This result is applied to the multi-parameter estimation theory for pure state models.

2 Review of conventional theory

We first give a brief summary of the conventional quantum parameter estimation theory.

2.1 Quantum measurement theory

Let $\mathfrak{H}$ be a separable Hilbert space which corresponds to a physical system with inner product $\langle \phi | \psi \rangle$, $(\phi, \psi \in \mathfrak{H})$, and $\mathcal{L}$ and $\mathcal{L}_{sa}$ be, respectively, the set of all the (bounded) linear operators and all the self-adjoint operators on $\mathfrak{H}$. A quantum state is represented by a *density operator* $\rho \in \mathcal{L}_{sa}$ which satisfies $\rho \geq 0$ and $\mathrm{Tr}\,\rho = 1$. A state ρ is called *pure* if $\mathrm{rank}\,\rho = 1$. In order to handle simultaneous probability distributions of possibly mutually non-commuting observables, an extended framework of measurement theory is needed [1, p. 53] [2, p. 50]. A *generalized measurement* $\{M(B)\}_{B\in\mathfrak{F}}$ on a measurable space $(\Omega, \mathfrak{F})$ is an operator-valued set function which satisfy the following axioms:

1. $M(\phi) = 0, \quad M(\Omega) = I,$

2. $M(B) = M(B)^* \geq 0, \quad (^\forall B \in \mathcal{B})$,

3. $M(\bigcup_j B_j) = \sum_j M(B_j)$, (for all at most countable disjoint sequence $\{B_j\} \subset \mathcal{F}$).

By fixing a state ρ and a measurement M, the outcomes of the measurement form random variables, whose simultaneous distribution is given by $P_\rho^M(B) = \mathrm{Tr}\,\rho M(B)$. In particular, a measurement M is called *simple* if it satisfies, in addition to the above three axioms, $M(B)^2 = M(B)$, $(^\forall B \in \mathcal{B})$, or equivalently $M(B_1)M(B_2) = 0$, $(^\forall B_1 \cap {}^\forall B_2 = \phi)$. In the following, simple measurement is denoted by E.

Here we give an example of generalized measurement, which is called the random measurement [1, p. 70]. Given N measuring apparatuses, which yield outcomes from the same finite set whose elements being labeled with the integers j. Suppose the nth apparatus is described by a simple measurement $\{E_j^{(n)}\}_j$. Select an apparatus n $(n = 1, 2, \cdots, N)$ with probability ξ_n and apply to the physical system. Then the probability for obtaining the jth outcome is

$$P_\rho(j) = \sum_{n=1}^{N} \xi_n \mathrm{Tr}\,\rho E_j^{(n)} = \mathrm{Tr}\,\rho M_j,$$

where

$$M_j = \sum_{n=1}^{N} \xi_n E_j^{(n)}$$

form a generally non-orthogonal measurement.

2.2 Statistical estimation theory

Let

$$\mathcal{S} = \{\rho_\theta\ ;\ \rho_\theta = \rho_\theta^* > 0,\ \mathrm{Tr}\,\rho_\theta = 1,\ \theta \in \Theta \subset \mathbf{R}^n\} \tag{1}$$

be the statistical parametric model composed of strictly positive density operators. Here, θ is the parameter to be estimated statistically. An *estimator* for θ is identified to a generalized measurement which takes values on Θ. The expectation vector with respect to the measurement M at the state ρ_θ is defined as

$$E_\theta[M] = \int \hat{\theta} P_\theta^M(d\hat{\theta}).$$

The measurement M is called *unbiased* if $E_\theta[M] = \theta$ holds for all $\theta \in \Theta$, i.e.,

$$\int \hat{\theta}^j P_\theta^M(d\hat{\theta}) = \theta^j, \quad (j = 1, \cdots, n). \tag{2}$$

Differentiation yields

$$\int \hat{\theta}^j \frac{\partial}{\partial \theta^k} P_\theta^M(d\hat{\theta}) = \delta_k^j, \quad (j, k = 1, \cdots, n). \tag{3}$$

If (2) and (3) hold at a certain θ, M is called *locally unbiased* at θ. Obviously, M is unbiased iff M is locally unbiased at every $\theta \in \Theta$. A locally unbiased measurement is accompanied by n self-adjoint operators:

$$X^j = \int (\hat{\theta}^j - \theta^j) M(d\hat{\theta}), \quad (j = 1, \cdots, n). \tag{4}$$

The locally unbiasedness conditions (2) (3) are then re-expressed as

$$\mathrm{Tr}\, \rho_\theta X^j = 0, \quad \mathrm{Tr}\, \frac{\partial \rho_\theta}{\partial \theta^k} X^j = \delta_k^j, \quad (j, k = 1, \cdots, n). \tag{5}$$

Since these self-adjoint operators $\{X^j\}$ are usefully employed in the following, we may call them *locally unbiased operators*. Letting M be a locally unbiased measurement at θ, we define the covariance matrix $V_\theta[M] = [v_\theta^{jk}] \in \mathbf{R}^{n \times n}$ with respect to M at the state ρ_θ by

$$v_\theta^{jk} = \int (\hat{\theta}^j - \theta^j)(\hat{\theta}^k - \theta^k) P_\theta^M(d\hat{\theta}). \tag{6}$$

In order to obtain lower bounds for $V_\theta[M]$, let us consider a quantum analogue of the logarithmic derivative denoted by L_θ:

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \frac{1}{2}[\rho_\theta L_{\theta,j} + L_{\theta,j}^* \rho_\theta]. \tag{7}$$

For instance,

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \frac{1}{2}[\rho_\theta L_{\theta,j}^S + L_{\theta,j}^S \rho_\theta], \quad L_{\theta,j}^S = L_{\theta,j}^{S*} \tag{8}$$

defines the *symmetrized logarithmic derivative* (SLD) $L_{\theta,j}^S$ introduced by Helstrom [3], and

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \rho_\theta L_{\theta,j}^R \tag{9}$$

defines the *right logarithmic derivative* (RLD) $L^R_{\theta,j}$ introduced by Yuen and Lax [5]. Thus, (7) defines a certain class of logarithmic derivatives. Correspondingly, we define the quantum analogue of Fisher information matrix $J_\theta = [(L_{\theta,j}, L_{\theta,k})_{\rho_\theta}]$, where the inner product $(\cdot,\cdot)_\rho$ on $\mathcal{L}$ is defined by

$$(A, B)_\rho = \mathrm{Tr}\, \rho\, BA^*. \tag{10}$$

We also define another inner product on $\mathcal{L}$ as

$$\langle A, B\rangle_\rho = \frac{1}{2} \mathrm{Tr}\, \rho\, (BA^* + A^*B). \tag{11}$$

Then, the following quantum version of Cramér–Rao theorem holds.

Theorem 2.1 *For any locally unbiased measurement M, the following inequality holds:*

$$V_\theta[M] \geq (\mathrm{Re}\, J_\theta)^{-1}, \tag{12}$$

where $\mathrm{Re}\, J_\theta = (J_\theta + \overline{J_\theta})/2$. In particular, for the SLD, $J^S_\theta = \mathrm{Re}\, J_\theta = \left[\langle L^S_{\theta,j}, L^S_{\theta,k}\rangle_{\rho_\theta}\right]$ is called the SLD–Fisher information matrix. Moreover, for the RLD,

$$V_\theta[M] \geq (J^R_\theta)^{-1} \tag{13}$$

holds, where $J^R_\theta = \left[(L^R_{\theta,j}, L^R_{\theta,k})_{\rho_\theta}\right]$ is called the RLD–Fisher information matrix.

When the model is one dimensional, the inequalities in the theorem become scalar. In this case, it is shown that the lower bound $(\mathrm{Re}\, J_\theta)^{-1} = (J_\theta)^{-1}$ becomes most informative, i.e., it takes the maximal value, iff the SLD is adopted, and the corresponding lower bound $(J^S_\theta)^{-1} = 1/\mathrm{Tr}\, \rho_\theta (L^S_\theta)^2$ can be attained by the estimator $T = \theta I + L^S_\theta / J^S_\theta$, where I is the identity. Thus, the one parameter quantum estimation theory is quite analogous to the classical one when the SLD is used.

On the other hand, for the dimension $n \geq 2$, the matrix equalities in (12) and (13) cannot be attained in general, because of the impossibility of the exact simultaneous measurement of non-commuting observables. We must, therefore, abandon the strategy of finding the measurement that minimizes the covariance matrix itself. Rather, we often adopt another strategy as follows: Given a positive definite real matrix $G = [g_{jk}] \in \mathbf{R}^{n\times n}$, find the measurement M that minimizes the quantity

$$\mathrm{tr}\ GV_\theta[M] = \sum_{jk} g_{jk} v^{jk}_\theta. \tag{14}$$

If there is a constant C such that tr $GV_\theta[M] \geq C$ holds for all M, C is called a Cramér–Rao type bound or simply a CR bound, which may depend on both G and θ. For instance, it is shown that the folowing two quantities are both CR bounds [6].

$$\begin{aligned} C^S &= \mathrm{tr}\, G (J_\theta^S)^{-1}, \\ C^R &= \mathrm{tr}\, G \,\mathrm{Re}\,(J_\theta^R)^{-1} + \mathrm{tr}\,\mathrm{abs}\, G \,\mathrm{Im}\,(J_\theta^R)^{-1}. \end{aligned}$$

Here, for a matrix X, $\mathrm{Im}\, X = (X - \overline{X})/2i$ and $\mathrm{tr}\,\mathrm{abs}\, X$ denotes the absolute sum of the eigenvalues of X. Let us call these CR bounds, respectively, the SLD–bound and the RLD–bound. The most informative CR bound is the maximum value of such C for given G and θ. Yuen and Lax [5] proved that the above C^R is most informative for the gaussian model, and they explicitly constructed the optimum measurement which attains C^R. Holevo [2, p. 285] derived another CR bound which, though an implicit form, is not less informative than C^S and C^R. Nagaoka [6] investigated in detail the relation between these CR bounds. He also derived a new CR bound for 2 dimensional models, which is not less informative than Holevo's one, and obtained explicitly the most informative CR bound specific to the spin 1/2 model. The construction of the general quantum parameter estimation theory for $n \geq 2$ is left to future study.

3 Linear random measurement

Suppose we are given a two parameter model $\{\rho_\theta \;;\; \theta = (\theta^1, \theta^2) \in \mathbf{R}^2\}$, the corresponding SLD being $\{L^S_{\theta,1}, L^S_{\theta,2}\}$. In the following, we often drop the subscript θ for notational convenience since we only consider local properties of the model. If L^S_1 and L^S_2 commute, then we can estimate the parameters (θ^1, θ^2) in the same way as in the classical theory. Therefore, suppose $[L^S_1, L^S_2] \neq 0$. If one of the two parameters, say θ^2, is fixed, then we obtain a one parameter sub-model $\{\rho_{\theta^1\theta^2} \;;\; \theta^2 = const.\}$. For this sub-model, we have an optimum locally unbiased estimator $T^1 = \theta^1 I + L^S_1/(J^S)_{11}$ for the parameter θ^1 as was mentioned in the previous section, where $(J^S)_{11}$ is the $(1,1)$ component of the Fisher information matrix J^S. Therefore it is natural to ask whether the optimum estimation for the original two parameter model ρ_θ can be realized by a random measurement of two "observables" L^S_1 and L^S_2. More generally, let us investigate the infimum of $\mathrm{tr}\, GV[M]$ with respect to the random measurements M of linearly independent two observables A_1, A_2 in the linear span

$$\mathcal{L}^S = \{a_1 L_1^S + a_2 L_2^S \ ; \ a_1, a_2 \in \mathbf{R}\}.$$

Note that if we obtain the optimum measurement M for $G = I$, the corresponding locally unbiased operators being X^j, the solution for general weight G is $Y^j = \sum_{k=1}^n f_k^j X^k$, where $\sqrt{G} = [f_{jk}]$. We then consider only $G = I$ case in the following without loss of generality.

Denote the dual basis by $L^i = J^{ij} L_j^S$ where J^{ij} is the (i,j) entry of inverse of the SLD–Fisher information matrix. A pair of self-adjoint operators X^1, X^2 whose expectation vanish are locally unbiased with respect to the parameters θ^1, θ^2 iff $\langle L_i^S, X^j \rangle = \delta_i^j$ because of (5), where $\langle \cdot, \cdot \rangle = \langle \cdot, \cdot \rangle_{\rho_\theta}$. This condition is rewritten as $\langle L^i, X^j \rangle = \langle L^i, L^j \rangle$, which indicates that L^j is the orthogonal projection of X^j with respect to $\langle \cdot, \cdot \rangle$. We therefore restrict ourselves to the case $X^j = L^j$, $(j = 1, 2)$. Further, let $A_1, A_2 \in \mathcal{L}^S$ be linearly independent observables which are to be measured at random, assuming $\langle A_j, A_j \rangle = 1$ without loss of generality. Their spectral decompositions are written as

$$A_j = \sum_\xi a_j(\xi) E_j(\xi). \tag{15}$$

Let us construct their random measurement as follows. Select one of A_1, A_2 according to the probability p_1, p_2, respectively, and make an exact measurement of it (in the sense of von Neumann). The corresponding resolution of identity is defined by

$$M_j(\xi) = p_j E_j(\xi). \tag{16}$$

When we selected an observable A_j and obtained an outcome $a_j(\xi)$, we identify this result to a pair of real quantities $b_j^1(\xi)$ and $b_j^2(\xi)$, which are connected to L^1, L^2 by

$$b_j^k(\xi) = \frac{1}{p_j} \langle L^k, A^j \rangle a_j(\xi), \quad (j, k = 1, 2), \tag{17}$$

where $\{A^1, A^2\}$ is the dual basis of $\{A_1, A_2\}$ in $\mathcal{L}^S$ with respect to the inner product $\langle \cdot, \cdot \rangle$. Indeed, the following relation holds:

$$L^k = \sum_{j=1}^2 \sum_\xi b_j^k(\xi) M_j(\xi). \tag{18}$$

Then, by comparing the resolution (18) to (4), we can evaluate the equi-weighted trace of covariance matrix as

$$\begin{aligned}\operatorname{tr} V[M] &= \sum_{j=1}^{2}\sum_{\xi}\left[\left(b_j^1(\xi)\right)^2+\left(b_j^2(\xi)\right)^2\right]\operatorname{Tr}\rho M_j(\xi) \\ &= \sum_j \frac{1}{p_j}\left[\langle L^1, A^j\rangle^2+\langle L^2, A^j\rangle^2\right]. \end{aligned} \tag{19}$$

Our aim is to find the infimum of (19) with respect to $\{p_j\}$ and $\{A_j\}$. Observing the fact that $\mu_1/p_1+\mu_2/p_2$, $(p_1+p_2=1)$ takes its minimum $(\sqrt{\mu_1}+\sqrt{\mu_2})^2$ when and only when $p_j=\sqrt{\mu_j}/(\sqrt{\mu_1}+\sqrt{\mu_2})$, we have

$$\min_{\{p_j\}} \operatorname{tr} V[M] = \left[\sqrt{\langle L^1, A^1\rangle^2+\langle L^2, A^1\rangle^2}+\sqrt{\langle L^1, A^2\rangle^2+\langle L^2, A^2\rangle^2}\right]^2. \tag{20}$$

The problem is then reduced to the minimization of (20) with respect to A_1, A_2. The normalization conditions $\langle A_1, A_1\rangle = \langle A_2, A_2\rangle = 1$ impose the following constraints on A^1, A^2,

$$\langle A^1, A^1\rangle = \langle A^2, A^2\rangle = \frac{1}{1-\alpha^2}, \quad \langle A^1, A^2\rangle = -\frac{\alpha}{1-\alpha^2}, \tag{21}$$

where $\alpha = \langle A_1, A_2\rangle$.

Let us define a linear transformation $\phi : \mathcal{L}^S \longrightarrow \mathcal{L}^S$ by

$$\phi(W) = \langle L^1, W\rangle L^1 + \langle L^2, W\rangle L^2. \tag{22}$$

Since ϕ is symmetric and positive definite, it has positive eigenvalues λ_1, λ_2 and unit eigenvectors U_1, U_2, satisfying

$$\phi(U_j) = \lambda_j U_j, \qquad (j=1,2). \tag{23}$$

By expanding as $A^i = a^{ij}U_j$, the problem to be solved is written in the form

$$\begin{aligned} &\underset{A^1,A^2}{\text{minimize}}\left[\sqrt{\langle A^1, \phi(A^1)\rangle}+\sqrt{\langle A^2, \phi(A^2)\rangle}\right]^2 \\ =\ &\underset{A^1,A^2}{\text{minimize}}\left[\sqrt{\lambda_1(a^{11})^2+\lambda_2(a^{12})^2}+\sqrt{\lambda_1(a^{21})^2+\lambda_2(a^{22})^2}\right]^2. \end{aligned}$$

The constraints (21) become

$$(a^{11})^2+(a^{12})^2 = (a^{21})^2+(a^{22})^2 = \frac{1}{1-\alpha^2}, \quad (a^{11})(a^{21})+(a^{12})(a^{22}) = -\frac{\alpha}{1-\alpha^2}.$$

Introduce another parametrization by

$$
\begin{aligned}
a^{11} &= \frac{1}{\sqrt{1-\alpha^2}} \cos\theta, \qquad & a^{12} &= \frac{1}{\sqrt{1-\alpha^2}} \sin\theta, \\
a^{21} &= \frac{1}{\sqrt{1-\alpha^2}} \cos\varphi, \qquad & a^{22} &= \frac{1}{\sqrt{1-\alpha^2}} \sin\varphi.
\end{aligned}
$$

Then the remaining third constraint becomes $\cos(\theta-\varphi) = -\alpha$, and we have the constraint free minimization problem as

$$
\begin{aligned}
& \underset{\theta,\varphi,\alpha}{\text{minimize}}\ \frac{1}{1-\alpha^2}\left[\sqrt{\lambda_1\cos^2\theta+\lambda_2\sin^2\theta}+\sqrt{\lambda_1\cos^2\varphi+\lambda_2\sin^2\varphi}\right]^2 \\
= \ & \underset{\theta,\varphi}{\text{minimize}}\ \frac{1}{\sin^2(\theta-\varphi)}\left[\sqrt{a+b\cos 2\theta}+\sqrt{a+b\cos 2\varphi}\right]^2,
\end{aligned}
$$

where $a = (\lambda_1+\lambda_2)/2$, $b = (\lambda_1-\lambda_2)/2$. By using the inequality (see Appendix B)

$$
\frac{1}{\sin(\theta-\varphi)}\left[\sqrt{a+b\cos 2\theta}+\sqrt{a+b\cos 2\varphi}\right] \geq \sqrt{a+b}+\sqrt{a-b}, \tag{24}
$$

which hold for $0<\theta-\varphi<\pi$, we obtain the desired infimum

$$
\begin{aligned}
& \min_{\{A_j\}}\min_{\{p_j\}} \operatorname{tr} V[M] \\
& \quad = \left[\sqrt{\lambda_1}+\sqrt{\lambda_2}\right]^2 \\
& \quad = \langle L^1, L^1\rangle + \langle L^2, L^2\rangle + 2\sqrt{\langle L^1, L^1\rangle\langle L^2, L^2\rangle - \langle L^1, L^2\rangle^2}. \qquad (25)
\end{aligned}
$$

The last equality follows from the fact that the trace $\lambda_1+\lambda_2$ and the determinant $\lambda_1\lambda_2$ of the linear transformation ϕ is independent of the choice of the basis which represents ϕ in a matrix form. Since the equality in (24) is satisfied on a sinusoidal periodic curve, there exist continuous potency of optimum measurements that attain the minimum (25). The lower bound (25) is first appeared in [7][8], although its meaning is not clearly stated.

4 Applications

In this section, we give two applications to demonstrate the results obtained in the previous section.

4.1 Gaussian models

A quantum Gaussian model is a model for the state of coherent lights with thermal noise [2, Chap. V]. The SLD's are

$$L_q^S = \frac{1}{\sigma_q^2}(Q-q), \qquad L_p^S = \frac{1}{\sigma_p^2}(P-p),$$

where $\sigma_q^2 = \langle Q-q, Q-q\rangle$ and $\sigma_p^2 = \langle P-p, P-p\rangle$ are their variances [2, p. 277]. Then,

$$J^S = \begin{bmatrix} 1/\sigma_q^2 & 0 \\ 0 & 1/\sigma_p^2 \end{bmatrix}, \qquad (J^S)^{-1} = \begin{bmatrix} \sigma_q^2 & 0 \\ 0 & \sigma_p^2 \end{bmatrix},$$

and

$$L^q = Q-q, \qquad L^p = P-p.$$

Therefore, the bound (25) becomes

$$\mathrm{tr}\, V[M] = \sigma_q^2 + \sigma_p^2 + 2\sigma_q\sigma_p.$$

This is greater than the most informative bound

$$\mathrm{tr}\, V[M] = \sigma_q^2 + \sigma_p^2 + \hbar$$

unless $\sigma_q\sigma_p = \hbar/2$, i.e., unless the system is in a minimum uncertainty state. Still more, even in the pure coherent state case, it is an open question whether the covariant measurement

$$M(p,q) = U[p,q]|0\rangle\langle 0|U^*[p,q],$$

which attain the above most informative bound, can be constructed directly from the continuous combination of random measurements.

4.2 Pure coherent models

In this section, we derive the RLD–bound for quantum coherent models [9] and show that it is attainable by random measurements. The prerequisites are summarized in Appendix A. A straightforward calculation leads to the following proposition.

Proposition 4.1 *The RLD–bound for a coherent model ρ_θ is*

$$C^R = C^S + \frac{\sqrt{\det G}}{\det J^S}\left|\mathrm{Tr}\,\rho_\theta[L_1^S, L_2^S]\right|. \qquad (26)$$

Since the identity

$$\left|\operatorname{Tr}\rho_\theta[L_1^S, L_2^S]\right| = \operatorname{Tr}\operatorname{Abs}\rho_\theta[L_1^S, L_2^S].$$

holds for any pure state models, the above CR–bound is in identical form to the Nagaoka's bound which is most informative when ρ_θ is represented in 2×2 matrix [6].

We derive here another results which holds not only in 2×2 matrix representation but also in arbitrary representation.

Definition 4.1 *The parameter θ of a model ρ_θ is called ρ_0–canonical if the SLD–Fisher information matrix with respect to θ is in diagonal form at ρ_0.*

This condition is not restrictive since, by a certain transformation of coordinate system, we can always diagonalize the SLD–Fisher information matrix.

Theorem 4.1 *Suppose the parameter θ of the coherent model ρ_θ is ρ_0–canonical. Then the corresponding RLD–bound is attainable at ρ_0 by a certain random measurement.*

Proof Some calculation lead us to another form of the RLD–bound (26) as

$$C^R = \langle L^1, L^1\rangle + \langle L^2, L^2\rangle + \left|\operatorname{Tr}\rho_\theta[L^1, L^2]\right|, \tag{27}$$

where we set $G = I$. On the other hand, from the assumption, there exists non–zero real numbers c_1, c_2 and normalized ρ_0–symplectic basis $\{\tilde{L}_1^S, \tilde{L}_2^S\}$ such that $L_j^S = c_j \tilde{L}_j^S$. Then $L^j = \tilde{L}_j^S / c_j$ and

$$(c_1 L^1 + i c_2 L^2)\rho_0 = 0,$$

which is nothing but the minimum uncertainty condition in the Heisemberg's uncertainty relation. Then

$$\langle L^1, L^1\rangle\langle L^2, L^2\rangle = \frac{1}{4}\left|\operatorname{Tr}\rho_0[L^1, L^2]\right|^2, \quad \langle L^1, L^2\rangle = 0. \tag{28}$$

Comparing (25) (27) and (28), we have the theorem. ■

This proof is deeply owed to the special choice of the coordinate system. It is not yet clear whether C^R can be attained in an arbitrary coordinate system.

Here we give an example of a two parameter coherent model which have non-diagonal SLD–Fisher information matrix, called the photon squeezed state model. Throughout this example, adjoint operators and complex conjugate numbers are denoted by $\dagger$ and $*$, respectively, according to the convention in physics. A photon squeezed state $|z\rangle_\xi$ is defined as

$$\begin{aligned} |z\rangle_\xi &= D(z)S(\xi)|0\rangle, \\ D(z) &= \exp(za^\dagger - z^* a), \\ S(\xi) &= \exp[\frac{1}{2}(\xi a^{\dagger 2} - \xi^* a^2)], \end{aligned}$$

where z, ξ are complex numbers, a and $a^\dagger$ are the annihilation and creation operators of a boson satisfying $[a, a^\dagger] = 1$, $[a, a] = [a^\dagger, a^\dagger] = 0$, and $|0\rangle$ is the Fock vacuum with respect to a. Letting $z = (q + ip)/\sqrt{2}$, we may regard the family of density operators

$$\rho_z = |z\rangle_{\xi\xi}\langle z| = D(z)\rho_0 D^*(z), \qquad \rho_0 = |0\rangle_{\xi\xi}\langle 0|$$

as a quantum parametric model which is parametrized by two real numbers q and p. The pre-inner products with respect to the position operator $Q = (a + a^\dagger)/\sqrt{2}$ and the momentum operator $P = (a - a^\dagger)/i\sqrt{2}$ are

$$\left[\begin{array}{cc} (Q,Q)_{\rho_z} & (Q,P)_{\rho_z} \\ (P,Q)_{\rho_z} & (P,P)_{\rho_z} \end{array}\right] = \frac{1}{2}\left[\begin{array}{cc} (\cosh 2s + \cos\theta \sinh 2s) & (\sin\theta \sinh 2s + i) \\ (\sin\theta\sinh 2s - i) & (\cosh 2s - \cos\theta\sinh 2s) \end{array}\right].$$

Noting $\mathrm{Tr}\, \rho_z L_j^A = 0$ $(j = q, p)$, the ALD are readily obtained as

$$L_q^A = 2i(P - p), \qquad L_p^A = -2i(Q - q).$$

Further, we observe the relation

$$b|z\rangle_\xi = \beta|z\rangle_\xi,$$

where $\xi = se^{i\theta}$ and

$$\begin{aligned} b &= S(\xi) a S^{-1}(\xi) = a\cosh s - a^\dagger e^{i\theta}\sinh s, \\ \beta &= z\cosh s - z^* e^{i\theta}\sinh s. \end{aligned}$$

Then we immediately have a normalized ρ_z–symplectic basis as

$$\begin{aligned} \tilde{L}_q^S &= \sqrt{2}[(Q-q)(\cosh s - \cos\theta\sinh s) - (P-p)\sin\theta\sinh s], \\ \tilde{L}_p^S &= \sqrt{2}[(P-p)(\cosh s + \cos\theta\sinh s) - (Q-q)\sin\theta\sinh s]. \end{aligned}$$

On the other hand, from (29), the SLD is given as $L_j^S = \mathfrak{D}(-L_j^A/2i)$, $(j = q, p)$. Then, by expanding $\{L_j^A\}$ into linear combinations of $\{\tilde{L}_j^S\}$, and using the relations $\mathfrak{D}\tilde{L}_q^S = -2\tilde{L}_p^S$ and $\mathfrak{D}\tilde{L}_p^S = 2\tilde{L}_q^S$, we have

$$\begin{aligned} L_q^S &= 2[(Q-q)(\cosh 2s - \cos\theta \sinh 2s) - (P-p)\sin\theta \sinh 2s], \\ L_p^S &= 2[(P-p)(\cosh 2s + \cos\theta \sinh 2s) - (Q-q)\sin\theta \sinh 2s]. \end{aligned}$$

Then the SLD–Fisher information matrix becomes

$$\begin{aligned} J^S &= \begin{bmatrix} \langle L_q^S, L_q^S\rangle_{\rho_z} & \langle L_q^S, L_p^S\rangle_{\rho_z} \\ \langle L_p^S, L_q^S\rangle_{\rho_z} & \langle L_p^S, L_p^S\rangle_{\rho_z} \end{bmatrix} \\ &= 2\begin{bmatrix} \cosh 2s - \cos\theta \sinh 2s & -\sin\theta \sinh 2s \\ -\sin\theta \sinh 2s & \cosh 2s + \cos\theta \sinh 2s \end{bmatrix}. \end{aligned}$$

This is, up to a constant factor, identical to the Fubini–Study metric [10], and is also identical to the real part of the complex ALD–Fisher information matrix

$$\begin{aligned} J^A &= \begin{bmatrix} (L_q^A, L_q^A)_{\rho_z} & (L_q^A, L_p^A)_{\rho_z} \\ (L_p^A, L_q^A)_{\rho_z} & (L_p^A, L_p^A)_{\rho_z} \end{bmatrix} \\ &= 2\begin{bmatrix} \cosh 2s - \cos\theta \sinh 2s & -\sin\theta \sinh 2s + i \\ -\sin\theta \sinh 2s - i & \cosh 2s + \cos\theta \sinh 2s \end{bmatrix}, \end{aligned}$$

see also [11]. This result indicates that the coordinate system (q, p) is a non-orthogonal one. Let us change the coordinate system (q, p) into another one (q', p') such that the corresponding SLD $L_{q'}^S, L_{p'}^S$ are orthogonal at ρ_z with respect to the pre-inner product $\langle\cdot,\cdot\rangle_{\rho_z}$. For instance, the transformation of coordinate system

$$\begin{aligned} q' &= q(\cosh s - \cos\theta \sinh s) - p \sin\theta \sinh s \\ p' &= p(\cosh s + \cos\theta \sinh s) - q \sin\theta \sinh s \end{aligned}$$

lead to $L_{q'}^S = \sqrt{2}\tilde{L}_q^S, L_{p'}^S = \sqrt{2}\tilde{L}_p^S$. Then, according to Theorem 4.1, there exist random measurements that attain the RLD bound with respect to (q', p').

5 Conclusions

We first derived an explicit infimum of $\operatorname{tr} V[M]$ with respect to the linear random measurement M in the linear span of SLD's. The corresponding lower bound is attainable by infinitely many (continuous potency) measurements. The RLD–bound for pure coherent models is also studied, and found to be most informative.

Acknowledgment

I am mostly indebted to Professor H. Nagaoka.

Appendix

A Multi-parameter pure state model estimation theory

We present here a minimum of multi-parameter estimation theory based on the RLD. The sesquilinear form $\langle\cdot,\cdot\rangle_\rho$ defined by (11) become a pre-inner product on $\mathcal{L}$ when ρ is degenerated. Denote by $\mathcal{K}_{sa}(\rho)$ the set of self-adjoint operators $K \in \mathcal{L}_{sa}$ satisfying $\langle K, K\rangle_\rho = 0$, which are called the kernel of the pre-inner product $\langle\cdot,\cdot\rangle_\rho$. The SLD, also defined by (8), is determined up to an uncertainty of $K \in \mathcal{L}_{sa}$. Let us denote the totality of such SLD's by $\mathcal{T}^S(\rho)$. Then the following theorem holds.

Theorem A.1 *Suppose ρ_θ is pure. Then the SLD–Fisher information matrix $J^S_\theta = \left[\langle L^S_{\theta,j}, L^S_{\theta,k}\rangle_{\rho_\theta}\right]$ is uniquely determined on the quotient space $\mathcal{T}^S(\rho_\theta)/K_{sa}(\rho_\theta)$, the SLD–tangent space, and its (j,k) entry becomes*

$$(J^S_\theta)_{jk} = 2\mathrm{Tr}\,(\partial_j\rho_\theta)(\partial_k\rho_\theta),$$

where $\partial_j = \partial/\partial\theta^j$. This metric is identical, up to a constant factor, to the Fubini–Study metric.

In the same way, the Holevo's commutation operator $\mathfrak{D}$ on $\mathcal{L}_{sa}$ [2] defined by

$$i(A\rho - \rho A) = \frac{1}{2}\left((\mathfrak{D}A)\rho + \rho(\mathfrak{D}A)\right), \qquad A, \mathfrak{D}A \in \mathcal{L}_{sa} \tag{29}$$

is regarded as a super-operator on $\mathcal{L}_{sa}/\mathcal{K}_{sa}(\rho)$.

It may sound strange to speak of the pure state estimation theory based on the RLD, since the RLD defined by (9) does not exist for degenerated states. However, it is important to notice that what we need is not the RLD itself but the inverse of the RLD–Fisher information matrix, as is understood by (13). The following lemma directly leads us to the inverse of the RLD–Fisher information matrix, which is a modification of the Holevo's result originally obtained in the strictly positive case [2, p. 280].

Lemma A.1 *Given a pure state model* ρ_θ. *Let* $\{\rho_\theta(\varepsilon)\ ;\ \varepsilon > 0\}$ *be a family of strictly positive density operators* $\rho_\theta(\varepsilon)$ *having a parameter* ε *which satisfy* $\lim_{\varepsilon\downarrow 0}\rho_\theta(\varepsilon) = \rho_\theta$, *and denote the corresponding RLD by* $L_\theta^R(\varepsilon)$. *If the SLD–tangent space* $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ *is* $\mathfrak{D}$*-invariant, then*

$$\lim_{\varepsilon\downarrow 0}\left(J^R(\varepsilon)\right)^{-1} = \left(J^S\right)^{-1} + \frac{i}{2}\left(J^S\right)^{-1} D \left(J^S\right)^{-1}$$

holds, where $J^R(\varepsilon) = \left[(L_j^R(\varepsilon), L_k^R(\varepsilon))_{\rho_\theta(\varepsilon)}\right]$ *and* $D = \left[i\,\mathrm{Tr}\,\rho_\theta[L_j^S, L_k^S]\right]$.

From this lemma, the inverse of the RLD–Fisher information matrix can be calculated directly from SLD, without using the diverging RLD–Fisher information matrix itself. Then, it may be important to investigate the condition for the SLD–tangent space to be $\mathfrak{D}$-invariant. The following theorem characterizes the structure of $\mathfrak{D}$-invariant SLD–tangent space.

Theorem A.2 *The* $\mathfrak{D}$*-invariant SLD–tangent space* $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ *has an even dimension and is decomposed into direct sum of 2 dimensional* $\mathfrak{D}$*-invariant subspaces. Moreover, by taking an appropriate basis of* $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$, *the operation of* $\mathfrak{D}$ *can be written in the form*

$$\mathfrak{D}\begin{bmatrix} \tilde{L}_1^S \\ \tilde{L}_2^S \\ \tilde{L}_3^S \\ \tilde{L}_4^S \\ \vdots \\ \tilde{L}_{2m-1}^S \\ \tilde{L}_{2m}^S \end{bmatrix} = \begin{bmatrix} 0 & 2 & & & & & \\ -2 & 0 & & & & & \\ & & 0 & 2 & & & \\ & & -2 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & 2 \\ & & & & & -2 & 0 \end{bmatrix}\begin{bmatrix} \tilde{L}_1^S \\ \tilde{L}_2^S \\ \tilde{L}_3^S \\ \tilde{L}_4^S \\ \vdots \\ \tilde{L}_{2m-1}^S \\ \tilde{L}_{2m}^S \end{bmatrix}. \tag{30}$$

Definition A.1 *The basis* $\{\tilde{L}_j^S\}_{j=1}^{2m}$ *of the SLD–tangent space* $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ *which is subject to the transformation law (30) is called* ρ_θ*–symplectic.*

From Theorem A.2, it is sufficient to consider a 2-dimensional $\mathfrak{D}$-invariant SLD tangent space. The following theorem gives the condition for the model to have a 2-dimensional $\mathfrak{D}$-invariant SLD–tangent space at ρ_θ.

Theorem A.3 *For the pure state model* $\{\rho_\theta = |\theta\rangle\langle\theta|\}$*, the following two conditions are equivalent.*

(i) $\{\tilde{L}_j^S\}_{j=1,2}$ *is a* $|\theta\rangle\langle\theta|$*–symplectic basis.*

(ii) $(\tilde{L}_1^S + i\tilde{L}_2^S)|\theta\rangle = 0.$

The linear span of such basis $\mathrm{span}\{\tilde{L}_1^S,\ \tilde{L}_2^S\}$ *is* $\mathfrak{D}$*-invariant.*

The condition (ii) is nothing but the definition of the coherent states in quantum theory. Thus the $\mathfrak{D}$-invariancy is equivalent to the coherency of the model.

Definition A.2 *The pure state model* $\{\rho_\theta = |\theta\rangle\langle\theta|\}$ *which satisfy the condition in Theorem A.3 is called coherent.*

B Proof of the inequality

Denote the right-hand side of (24) by $f(\theta,\varphi)$, and set $b = 1$ without loss of generality, i.e.,

$$f(\theta,\varphi) = \frac{1}{\sin(\theta-\varphi)}\left[\sqrt{a+\cos 2\theta} + \sqrt{a+\cos 2\varphi}\right], \quad (a > 1). \tag{31}$$

Extremizing conditions $\partial f/\partial\theta = 0$ and $\partial f/\partial\varphi = 0$ lead to

$$\left[a + \cos 2\theta + \sqrt{(a+b\cos 2\theta)(a+b\cos 2\varphi)}\right]\cot(\theta-\varphi) + \sin 2\theta = 0, \tag{32}$$

and

$$\left[a + \cos 2\varphi + \sqrt{(a+b\cos 2\theta)(a+b\cos 2\varphi)}\right]\cot(\theta-\varphi) - \sin 2\varphi = 0, \tag{33}$$

respectively. Subtracting these two equalities, we have

$$(\cos 2\theta - \cos 2\varphi)\cot(\theta-\varphi) + \sin 2\theta + \sin 2\varphi = 0,$$

but this is an identity. So the equalities (32) and (33) are the same. In other words, (32) is the only extremizing condition for (31). On the other hand, adding the equalities (32) and (33), we have,

$$\cos(\theta-\varphi) = -\frac{2}{f^2(\theta,\varphi)}\cos(\theta+\varphi), \tag{34}$$

which is also equivalent to either (32) or (33). We show that, under the condition $0 < \theta - \varphi < \pi$, (34) have a unique globally connected solution. Suppose (34) have possibly disconnected solutions (curves) labeled by C_n. Note that $f(\theta, \varphi)$ is constant on each C_n since $\partial f/\partial\theta = \partial f/\partial\varphi = 0$ along the curve. Let us denote the corresponding constant value of $2/f^2(\theta, \varphi)$ on C_n by A_n. Furthermore, let us change the coordinate system as $\theta - \varphi = \pi/2 - \delta$, $\theta + \varphi = \pi/2 - x$. Then

$$f(\theta, \varphi) = \hat{f}(x, \delta) = \frac{1}{\cos\delta}\left[\sqrt{a + \cos(x - \delta)} + \sqrt{a - \cos(x + \delta)}\right], \quad (a > 1). \tag{35}$$

and the extremizing condition (34) becomes

$$\sin\delta = -A_n \sin x, \quad (-\frac{\pi}{2} < \delta < \frac{\pi}{2}). \tag{36}$$

Therefore, we can regard C_n as the connected component of the solution of (36) which crosses the x-axis at $x = n\pi$ $(n = 0, \pm 1, \cdots)$. Since A_n is constant on each curve C_n, we can evaluate its value at the x-intercepts as

$$A_n = \frac{2}{\hat{f}^2(n\pi, 0)} = \frac{2}{(\sqrt{a+1} + \sqrt{a-1})^2}.$$

In particular, A_n does not depend on n and is less than unity. Therefore, the extremizing condition (36), which is rewritten as

$$\sin\delta = -\frac{2}{(\sqrt{a+1} + \sqrt{a-1})^2}\sin x,$$

has a unique globally connected solution

$$\delta = -\arcsin\left[\frac{2}{(\sqrt{a+1} + \sqrt{a-1})^2}\sin x\right]. \tag{37}$$

In other words, all the C_n's are identical with each other. It is evident that $\hat{f}(x, \delta)$ takes its minimum along the curve (37). Therefore,

$$f(\theta, \varphi) \geq \hat{f}(0, 0) = \sqrt{a+1} + \sqrt{a-1}.$$

References

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic Press, New York, 1976).

[2] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982) (in Russian, 1980).

[3] C. W. Helstrom, "Minimum Mean-Square Error Estimation in Quantum Statistics," Phys. Lett. **25A**, 101-102 (1967).

[4] H. Nagaoka, "On Fisher information of quantum statistical models," SITA '87, 241-246 (1987) (in Japanese).

[5] H. P. H. Yuen and M. Lax, "Multiple-Parameter Quantum Estimation and Measurement of Non-Selfadjoint Observables," IEEE Trans. **IT-19**, 740-750 (1973).

[6] H. Nagaoka, "A New Approach to Cramér-Rao Bounds for Quantum State Estimation," IEICE Technical Report **IT89-42**,9-14(1989).

[7] H. Nagaoka, "On a mathematical problem of matrices and quantum estimation theory," SITA '90, 681-686 (1991) (in Japanese).

[8] H. Nagaoka, "A generalization of the simultaneous diagonalization of Hermitian matrices and its relation to quantum estimation theory," Trans. Jap. Soc. Indust. Appl. Math. **1**, 43-56 (1991) (in Japanese).

[9] A. Fujiwara, "Multi-parameter pure state estimation based on the right logarithmic derivative," METR **94-9**, Univ. Tokyo (1994).

[10] S. Abe, "Quantized geometry associated with uncertainty and correlation," Phys. Rev. **A48**, 4102-4106 (1993).

[11] A. Fujiwara, "One-parameter pure state estimation based on the symmetric logarithmic derivative," METR **94-8**, Univ. Tokyo (1994).