A new approach to the Cramér-Rao-type bound of the pure-state model

2002 J. Phys. A: Math. Gen. 35 3111

(http://iopscience.iop.org/0305-4470/35/13/307)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **35** (2002) 3111–3123 PII: S0305-4470(02)25703-0

A new approach to the Cramér–Rao-type bound of the pure-state model

K Matsumoto

Quantum Computation and Information Project, JST 5-28-3, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Received 8 June 2001, in final form 18 January 2002
Published 22 March 2002
Online at stacks.iop.org/JPhysA/35/3111

Abstract

This paper sheds light on non-commutativity in quantum theory as regards theoretical estimation. In it, we calculate the quantum Cramér–Rao-type bound for many cases, by use of a newly proposed powerful technique. We also discuss the use of collective measurement in statistical estimation.

PACS numbers: 03.67.−a, 03.65.Ta, 05.30.−d

1. Introduction

Quantum estimation theory, or the theory of statistical estimation of unknown density operators, was introduced by Helstrom [2, 3], motivated by the wish to optimize detection processes for optical communication systems. Recently, estimation of unknown states has been attracting the attention of many physicists [5–7].

In this paper, for simplicity, it is assumed that the unknown state is a member of a *model* $\mathcal{M} = \{\rho(\theta)|\theta \in \Theta \subset \mathbb{R}^m\}$, or a subset of the totality of the density operators, and that the finite-dimensional parameter θ is to be estimated statistically. Further, it is also assumed that a large number of i.i.d. samples of the unknown state are given, and we study the lower bound of the asymptotic error of the estimate.

Recently, there have been several studies based on the variation of the measuring precision with respect to the number of samples [7, 8, 29, 31, 32]. Nagaoka [9–11] studied, for the first time, asymptotic aspects of estimation of the density operator. He pointed out that the quasi-quantum Cramér–Rao-type (CR-type) bound, or the lower bound for the asymptotic mean square error of the estimate which does not use a collective measurement, can be 'single-letterized'. (Later, in 2000, Gill and Massar gave a similar theorem. The difference between these works is explained briefly in section 2.) His studies also include the investigation of the use of collective measurement in the statistical estimation of the one-dimensional parameter model. Massar and Popescu [8] and Hayashi [12] studied the asymptotic theory of the total-space model, or the totality of the pure states in a given Hilbert space, which we also study later. Hayashi and Matsumoto [13] showed that the quantum CR-type bound, which is also defined later, gives the lower bound for the asymptotic mean square error of the estimate which uses the collective measurement.

The determination of an explicit expression for the quasi-quantum CR-type bound has been an important aim since the foundation of quantum estimation theory. In fact, it was Helstrom [2, 3] who determined the bound of the one-dimensional model which consists of faithful states, or states whose density matrix is of full rank.

In the case of the multi-dimensional parameter model, this problem still remains challenging, because the underlying non-commutativity makes the problem complicated. In fact, when the author's research was starting, the bound was determined only for several specific models. For models which consist of faithful states, Yuen, Lax, and Holevo [14, 15] determined the bound of the Gaussian state model, and Nagaoka and Hayashi [16, 17] (and later on Gill and Massar [29] independently) solved the problem in the case of faithful spin-$\frac{1}{2}$ models. Fujiwara and Nagaoka [18–20] started the study of the quasi-quantum CR-type bound of the *pure-state model*, a model which consists of pure states, and determined the bound in the case of a one-dimensional parameter model and a two-dimensional coherent model.

In this paper, after the preliminaries in sections 2, 3, a new and powerful approach for determining the quasi-quantum CR-type bound of the pure-state model is proposed in section 4. Theorem theorem:commute and theorem theorem:simple, the key to this new approach, reduce the calculation of the quasi-quantum CR-type bound, which includes optimization of the positive-operator-valued measure (POM), to an easier problem, which just includes optimization of vectors of small dimension. In sections 6–8, by use of the method, the quasi-quantum CR-type bound is calculated for wide ranges of pure-state models.

One might suggest that the calculation of the quasi-CR-type bound is not that fundamental, for this is the bound achieved for some restricted class of measurement. Therefore, we also study the relation between the quasi-CR-type bound and the quantum CR-type bound, which is obtained by allowing use of any type of collective measurement. In section 2, the definitions and some basic facts about these bounds are stated, and in section 5, it is shown that these two bounds coincide with each other in any kind of pure-state model. Because of this identity, our new method of calculation of the quasi-quantum CR-type bound actually gives the quantum CR-type bound.

For the proof of this identity, we make use of the pure-state version of Holevo's bound, which is used in the study of the Gaussian model, and the fact that Holevo's bound is achievable in an arbitrary pure-state model. This does not mean that the determination of the bound is completed, for Holevo's bound still requires minimization about some variables.

The third topic in the paper is the relation between the non-commutativity and the eigenvalues of the linear transform $\boldsymbol{D}$, which is defined in section 3 in relation to complex structure. In section 8, we study the two-dimensional parameter model to show that the absolute value of the eigenvalues of $\boldsymbol{D}$ is a good measure of non-commutativity. In sections 6 and 7, we treat the model with an arbitrary dimension parameter, but only in special cases. Section 6 deals with the case where $\boldsymbol{D} = 0$, or the commutative case, while in section 7, the strongest non-commutativity case, or the case where the absolute values of the eigenvalues of $\boldsymbol{D}$ are maximal, is discussed.

2. The POM, asymptotic bounds, and locally unbiased measures

Let $\sigma(\mathbb{R}^m)$ be a σ-field in the space $\mathbb{R}^m$. Whatever measuring apparatus is used, or whatever calculation is being made, the probability that the estimate $\hat{\theta}$ lies in a measurable set B in $\mathbb{R}^m$ will be given by $\mathrm{P}_\theta^M(B) = \mathrm{tr}\, \rho(\theta) M(B)$, where M is a POM, or a mapping of a measurable set

$B \in \sigma(\mathbb{R}^m)$ to non-negative Hermitian operators in the separable Hilbert space $\mathcal{H}$, such that

$$M(\phi) = O, \qquad M(\mathbb{R}^m) = I,$$
$$M\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} M(B_i) \qquad (B_i \cap B_j = \phi, i \neq j),$$

(see [3, p 53], and [15, p 50]). Conversely, there is a measuring apparatus corresponding to any POM M [21, 22].

Denote the mean square matrix of M by $\mathrm{V}_\theta[M] = [\int(\hat{\theta}^i - \theta^i)(\hat{\theta}^j - \theta^j)\,\mathrm{tr}\,\rho(\theta)\,M(\mathrm{d}\hat{\theta})]$, and, as the measure of accuracy, let us take $\mathrm{Tr}\,G\mathrm{V}_\theta[M]$, where G is a strictly positive symmetric real matrix. If $G = \mathrm{diag}(g_1, \ldots, g_m)$, $\mathrm{Tr}\,G\mathrm{V}_\theta[M]$ is the weighted sum of mean square errors of the estimates $\hat{\theta}^i$ of each component θ^i of the parameter.

Let us suppose that we are given N i.i.d. pairs $\rho(\theta)^{\otimes N}$ of the unknown state $\rho(\theta)$. Then, if the estimate is appropriate, the mean square error of the estimate will be, at most, $\mathrm{O}(1/N)$. Therefore, the infimum of N times the mean square error of the estimate is of interest.

The sequence $\{M_N\}$ of POM, where M_N is a POM in $\mathcal{H}^{\otimes N}$, is said to be *asymptotically unbiased* if $\lim_{N\to\infty} \mathrm{E}_\theta[M_N] = \theta$ and $\lim_{N\to\infty} \int \hat{\theta}^j\,\mathrm{tr}\,\partial_k\rho(\theta)\,M_N(\mathrm{d}\hat{\theta}) = \delta^j_k$ hold true for any $\theta \in \Theta$.

Let us define $\mathcal{CM}_0$ to be the totality of adaptive measurements, where POM $M_{(1)}$ in $\mathcal{H}$ is applied to the first sample, and the measurement which corresponds to POM $M_{(2)}$ in $\mathcal{H}$ is applied to the second sample, and so on. (The choice of $M_{(k)}$ is generally dependent on the outcome of measurements $M_{(1)}, \ldots, M_{(k-1)}$.) Also, we denote by $\mathcal{CM}_1$ the totality of separable measurements, or measurements such that, for any measurable set B, $M_N(B)$ is a convex combination of operators which is written as $\bigotimes_{j=1}^N A_j$ in terms of non-negative operators A_j $(j = 1, \ldots, N)$ in $\mathcal{H}$.

Although it was pointed out by Bennett *et al* [30] that $\mathcal{CM}_0$ is strictly smaller than $\mathcal{CM}_1$, their minimum mean square errors are the same up to $\mathrm{O}(1/N)$. More concretely, it is known that

$$\left\{\lim_{N\to\infty} N\,\mathrm{Tr}\,G\mathrm{V}_\theta[M_N] | \{M_N\} \in \mathcal{CM}_0, M_N\text{: asymptotically unbiased}\right\} \tag{1}$$
$$= \inf\left\{\lim_{N\to\infty} N\,\mathrm{Tr}\,G\mathrm{V}_\theta[M_N] | \{M_N\} \in \mathcal{CM}_1, M_N\text{: asymptotically unbiased}\right\} \tag{2}$$
$$= \inf\left\{\mathrm{Tr}\,G\mathrm{V}_\theta[M] | M \text{ is locally unbiased at } \theta\right\}, \tag{3}$$

where a POM M is said to be *locally unbiased at* $\theta \in \Theta$, if

$$\mathrm{E}_\theta[M] = \int \hat{\theta}\,\mathrm{tr}\,\rho(\theta)\,M(\mathrm{d}\hat{\theta}) = \theta,$$
$$\int \hat{\theta}^j\,\mathrm{tr}\,\partial_k\rho(\theta)\,M(\mathrm{d}\hat{\theta}) = \delta^j_k, \qquad j, k = 1, \ldots, m.$$

Equality (1) = (3) was obtained by Nagaoka [9, 13], and (2) = (3) (or an equivalent relation) was obtained by Gill and Massar [29]. Notice that (3) is much easier to treat than (1) and (2), for (3) is an optimization over measurements in $\mathcal{H}$, while (1) and (2) are optimizations over measurements in $\mathcal{H}^{\otimes N}$. We denote (1) = (2) = (3) by $C_\theta(G)$, and call this the quasi-quantum CR bound. The name 'quasi-' comes from the fact that the bound is achieved by some restricted class of measurements. (Correctly speaking, in Gill and Massar's paper [29], they express (3) using Fisher information, which is very useful in the following calculations in their paper. However, this point is not essential in our context.)

We also define,

$$C^Q_\theta(G) \equiv \inf\left\{\lim_{N\to\infty} N\,\mathrm{Tr}\,G\mathrm{V}_\theta[M_N] | M_N\text{: asymptotically unbiased}\right\}, \tag{4}$$

and call this the quantum CR bound. Notice that in this optimization, we also consider collective measurement.

It is proved in [13] that

$$C_\theta^Q(G) = \lim_{N\to\infty} N C_\theta^N(G), \tag{5}$$

where $C_\theta^N(G)$ is the quasi-CR bound of the model $\{\rho(\theta)^{\otimes N}|\theta \in \Theta\}$. Because [13] is hard to access, a sketch of the argument is given.

From mostly the same consideration as in [9, 29], we have $\lim_{N\to\infty} N \operatorname{Tr} G\mathrm{V}_\theta[M_N] \geqslant C_\theta^Q(G)$. Achievability is proved as follows.

Regard N i.i.d. pairs $\rho(\theta)^{\otimes N}$ of the unknown state $\rho(\theta)$ as N_1 i.i.d pairs $(\rho(\theta)^{\otimes N_2})^{\otimes N_1}$ of $\rho(\theta)^{\otimes N_2}$ $(N = N_1N_2)$, and perform the optimal separable measurement for the model $\{\rho(\theta)^{\otimes N_1}|\theta \in \Theta\}$. Then, $N_1 C_\theta^{N_1}(G)$ is achieved. For any ϵ, if N_1 is large enough, we have $C_\theta^Q(G) - \epsilon \leqslant N_1 C_\theta^{N_1}(G)$ from the definition (5) of $C_\theta^Q(G)$. This concludes the proof that we can construct the sequence of measurement which achieves $C_\theta^Q(G) - \epsilon$ for any ϵ.

In the pure-state model, as is shown in section 5, $C_\theta^Q(G) = C_\theta(G)$ holds, and the achievability is rather trivial.

So far, we have considered mean square error only. However, if a measure $g(\hat\theta, \theta)$ of error satisfies the natural regularity condition such that

$$\left| g(\theta + \mathrm{d}\theta, \theta) - \sum_{i,j} G_{i,j}\, \mathrm{d}\theta^i\, \mathrm{d}\theta^j \right| \leqslant A(\theta, \epsilon)\|\mathrm{d}\theta\|^3, \qquad \forall\epsilon > 0,\ \forall\|\mathrm{d}\theta\| < \epsilon,$$

and if we assume that the estimate $\hat\theta$ will be asymptotically normal, then, denoting $\int f(\hat\theta) \operatorname{tr} \rho(\theta)\, M(\mathrm{d}\hat\theta)$ by $\mathrm{E}_\theta[f(\hat\theta), M]$, we have $\mathrm{E}_\theta[g(\hat\theta_{,N}, \theta), \tilde M_N] = \operatorname{Tr} G\mathrm{V}_\theta[\tilde M_N] + \mathrm{o}(1/N)$, which implies that the infimum of $\lim N\mathrm{E}_\theta[g(\hat\theta^N, \theta), M_N]$ is also given by the quantum CR-type bound $C_\theta^Q(G) = C_\theta(G)$.

3. Tangent space

In this section, we introduce notation and concepts concerning the tangent space $\mathcal{T}_{\rho(\theta)}(\mathcal{M})$, for these are used to characterize non-commutativity later on. From here on, the argument θ is often dropped.

Let $\mathcal{P}_1$ denote the totality of density operators of pure states in $\mathcal{H}$. A map π from $\mathcal{H}$ to $\mathcal{P}_1$ is defined by $\pi(|\phi\rangle) = |\phi\rangle\langle\phi|$, and its differential map is denoted by π_*. We identify a tangent vector with a differential operator in the usual manner, and for $X \in \mathcal{T}_\rho(\mathcal{P}_1)$ and for $Y \in \mathcal{T}_{|\phi\rangle}(\mathcal{H})$, $X\rho$ and $Y|\phi\rangle$ are seen as representations of X and Y, respectively.

The *horizontal lift* $|l_X\rangle$ of a tangent vector $X \in \mathcal{T}_\rho(\mathcal{P}_1)$ to $|\phi\rangle \in \pi^{-1}(\rho(\theta))$ is an element of $\mathcal{H}$ which satisfies $\pi_*(|l_X\rangle) = X\rho$ and $\langle l_X|\phi\rangle = 0$.

Denote by $|l_i(\theta)\rangle$ a horizontal lift of $\partial_i \in \mathcal{T}_{\rho(\theta)}(\mathcal{M})$; then $\mathrm{span}_{\mathbb{R}}\{|l_i\rangle | i = 1, \ldots, m\}$ is a representation of $\mathcal{T}_{\rho(\theta)}(\mathcal{M})$ because of the unique existence of the horizontal lift, which is proved as follows. $|l_X\rangle = (1/2)X\rho|\phi\rangle$ is easily checked to be a horizontal lift of X. To prove the uniqueness, it suffices to show that $\langle l|\phi\rangle = 0$ and $0 = |l\rangle\langle\phi| + |\phi\rangle\langle l|$ imply $|l\rangle = 0$. Multiplication by $|\phi\rangle$ of both sides of $0 = |l\rangle\langle\phi| + |\phi\rangle\langle l|$ proves the statement.

Using the horizontal lift, we introduce the inner product $\langle\cdot,\cdot\rangle_\rho$, which is called the *Fubini–Study metric* [23], in $\mathcal{T}(\mathcal{P}_1)$ via $\langle X, Y\rangle_\rho = \mathrm{Re}\,\langle l_X|l_Y\rangle$. Let U be an arbitrary unitary transform in $\mathcal{H}$, and denote by U_* its differential map. Then, we have $\langle U_*X, U_*Y\rangle_{U\rho U^*} = \mathrm{Re}\,\langle Ul_X|Ul_Y\rangle = \langle X, Y\rangle_\rho$, or the invariancy of $\langle\cdot,\cdot\rangle_\rho$ by unitary transform. Notice $\langle\cdot,\cdot\rangle_\rho$ is the only invariant metric, for $\mathrm{Re}\,\langle l_X|l_Y\rangle$ is the only strictly positive bilinear map from $\mathcal{H} \times \mathcal{H}$ to $\mathbb{R}$, which is invariant by the action of unitary operator (here, we ignored the difference by a constant factor). Limit the inner product $\langle\cdot,\cdot\rangle_\rho$ to $\mathcal{T}(\mathcal{M})$, and the metric tensor of the inner product $\langle\cdot,\cdot\rangle_{\rho(\theta)}$ is called the *SLD Fisher information matrix* and denoted by $J^S(\theta)$. Throughout the paper, we assume that $J^S(\theta)$ is strictly positive.

If we write $\mathsf{X}^*\mathsf{Y}$ means $[\langle x_i|y_j\rangle]$ for the ordered pairs $\mathsf{X} = [|x_1\rangle, |x_2\rangle, \ldots, |x_m\rangle]$ and $\mathsf{Y} = [|y_1\rangle, |y_2\rangle, \ldots, |y_m\rangle]$, $J^S = \mathrm{Re}\,\mathsf{L}^*\mathsf{L}$, where L is the ordered pair $[|l_1\rangle, |l_2\rangle, \ldots, |l_m\rangle]$. We also define the matrix $\tilde{J}$ by $\mathrm{Im}\,\mathsf{L}^*\mathsf{L}$.

Some historical remarks. Helstrom [2,3] defined the symmetric logarithmic derivative (SLD) Fisher information matrix by $J^S = [\mathrm{Re}\ \mathrm{tr}\,\rho L_i^S L_j^S]$ where L_i^S is the SLD defined as a solution to the equation

$$\partial_i\rho = \tfrac{1}{2}(L_i^S\rho + \rho L_i^S), \qquad L_i^S = (L_i^S)^\dagger. \tag{6}$$

Note that L_i^S is dependent on θ, and whenever it is necessary, we explicitly write $L_i^S(\theta)$. While the SLD is defined uniquely by (6) in the case where ρ is faithful, in the pure-state case, the SLD has the arbitrariness which corresponds to the kernel of ρ, and Fujiwara and Nagaoka [19] showed that $J^S(\theta)$ is uniquely defined regardless of this arbitrariness. Notice that in our definition, uniqueness of the SLD Fisher information matrix is trivial.

Let us define the linear transform $\boldsymbol{D}_\theta$ in $\mathcal{T}_{\rho(\theta)}(\mathcal{M})$ such that $\boldsymbol{D}_\theta X \in \mathcal{T}_{\rho(\theta)}(\mathcal{M})$ is the image of the projection of $\pi_*(\mathrm{i}|l_X\rangle) \in \mathcal{T}_{\rho(\theta)}(\mathcal{P}_1)$ onto $\mathcal{T}_{\rho(\theta)}(\mathcal{M})$ with respect to the inner product $\langle\cdot,\cdot\rangle_\rho$. Hereafter, we drop the argument θ for simplicity. The matrix which corresponds to $\boldsymbol{D}$ is $J^{S-1}\tilde{J}$, and the eigenvalue of $\boldsymbol{D}$ is of the form $\pm\mathrm{i}\beta_j$ or 0, where $0 \leqslant \beta_j \leqslant 1$. The model is said to be *quasi-classical* at θ if $\boldsymbol{D} = 0$, or equivalently, $\tilde{J} = 0$. The model is said to be *coherent* at θ if all of the eigenvalues of $\boldsymbol{D}$ are $\pm\mathrm{i}$. The dimension m of the parameter is even if the model is coherent. The definition of the coherency directly yields lemma 1.

Lemma 1. *The model $\mathcal{M}$ is coherent at θ if and only if* $\mathrm{span}_{\mathbb{R}}\{\mathrm{i}\mathsf{L}\}$ *is identical to* $\mathrm{span}_{\mathbb{R}}\mathsf{L}$*, or equivalently, if and only if span*$_{\mathbb{R}}\{\mathsf{L}, \mathrm{i}\mathsf{L}\}$ *is identical to span*$_{\mathbb{R}}\mathsf{L}$*, or equivalently, if and only if the dimension of* $\mathrm{span}_{\mathbb{C}}\mathsf{L}$ *is $m/2$.*

Lemma 2. *The model is coherent at θ if and only if* $|\det J^S| = |\det \tilde{J}|$.

Proof. $0 \leqslant \beta_j \leqslant 1$ leads to $|\det J^{S-1}\tilde{J}| \leqslant 1$, where the equality is established if and only if $\beta_j = 1$ for every j. Hence, we have the lemma. □

4. Reduction of the problem

In this section, we present the commuting theorem and the reduction theorem, and by use of these theorems, we reduce the determination of the CR-type bound, which is a minimization of a functional of the POM, to the minimization of function of finite numbers of finite-dimensional vectors. The proof of the commuting theorem is in the appendix.

Theorem 1 (commuting theorem). *If there exists an unbiased POM M such that*

$$\begin{aligned} |x^i\rangle &= \int (\hat{\theta}^i - \theta^i)\, M(\mathrm{d}\hat{\theta})\, |\phi\rangle, \\ \mathrm{V}[M] &= \mathrm{Re}\,\mathsf{X}^*\mathsf{X}, \end{aligned} \tag{7}$$

then

$$\mathrm{Im}\,\mathsf{X}^*\mathsf{X} = 0 \tag{8}$$

holds true. On the other hand, if (7) holds true, then there exists a projection-valued and unbiased POM E such that (7) holds, and the number of elements of its support is $m+2$, and $E(\{\hat{\theta}_0\})$ is a projection onto the orthogonal complement subspace of $\mathrm{span}_{\mathbb{C}}\{\mathsf{X}\}$.

Theorem 2 (reduction theorem). *Let $\mathcal{M}$ be an m-dimensional manifold in $\mathcal{P}_1$, and* B *be a system* $\{|\phi'\rangle||l_i'\rangle, i = 1, \ldots, m\}$ *of vectors in* $\mathbb{C}^{2m+1}$ *such that* $\langle\phi'|l_j'\rangle = \langle\phi|l_j\rangle = 0$, $\langle l_i'|l_j'\rangle = \langle l_i|l_j\rangle$, *for any i, j. Then, for any locally unbiased POM M, there is a projection-valued POM E in* $\mathbb{C}^{2m+1}$ *such that*

$$|x^i\rangle = \int (\hat{\theta}^i - \theta^i)\, E(\mathrm{d}\hat{\theta})|\,\phi'\rangle \in \mathbb{C}^{2m+1}, \tag{9}$$

$$\langle x^i|\phi'\rangle = 0, \tag{10}$$

$$\mathrm{Re}\,\mathsf{X}^*\mathsf{L}' = I_m, \tag{11}$$

$$\mathrm{V}[M] = \mathrm{Re}\,\mathsf{X}^*\mathsf{X}, \tag{12}$$

where $\mathsf{X} = [|x^1\rangle, \ldots |x^m\rangle]$, $\mathsf{L}' = [|l_1'\rangle, \ldots |l_m'\rangle]$, *and I_m is an identity matrix in* $\mathbb{C}^m$.

Proof. For any locally unbiased POM M, there exists a Hilbert space $\mathcal{H}_M$ and a projection-valued POM E_M in $\mathcal{H}_M$ which satisfies $M(B) = PE(B)P$, where P is the projection onto $\mathcal{H}$, by virtue of Naimark's theorem (see [15, pp 64–8]). Note that E_M is also locally unbiased. Let us define $|y^i\rangle = \int(\hat{\theta}^i - \theta^i)\, E_M(\mathrm{d}\hat{\theta})|\phi'\rangle$. Mapping $\mathrm{span}_{\mathbb{C}}\{|\phi\rangle, |l_i\rangle, |y^i\rangle | i = 1, \ldots, m\}$ isometrically onto $\mathbb{C}^{2m+1}$ so that $\{|\phi\rangle, |l_i\rangle | i = 1, \ldots, m\}$ are mapped to $\{|\phi'\rangle, |l_i'\rangle | i = 1, \ldots, m\}$, we denote the images of $\{|y^i\rangle | i = 1, \ldots, m\}$ by $\{|x^i\rangle | i = 1, \ldots, m\}$. Then, by virtue of the commuting theorem, we can construct a projection-valued POM E in $\mathbb{C}^{2m+1}$ satisfying the equations (9)–(12). □

By virtue of the reduction theorem, the quantum CR-type bound is the minimization of $\mathrm{Re}\,\mathrm{Tr}\, G\mathsf{X}^*\mathsf{X}$, where $|x^i\rangle$ runs all the vectors in $\mathbb{C}^{2m+1}/\{\mathbb{C}|\phi'\rangle\}$ which satisfy (8) and (11). Now, the problem is simplified to a large extent, because we only need to treat vectors in $\mathbb{C}^{2m+1}$ instead of the POM.

5. Holevo's bound and uselessness of collective measurement

Holevo's bound is defined by

$$C^H(G) = \inf\{\mathrm{Tr}\, G\, \mathrm{Re}\, \mathsf{Y}^*\mathsf{Y} + \mathrm{Tr}\,\mathrm{abs}\, G\, \mathrm{Im}\, \mathsf{Y}^*\mathsf{Y} \,|\, |y_i\rangle \in \mathrm{span}_{\mathbb{C}}\{\mathsf{L}'\}\ (i = 1, \ldots, m), \\ \mathrm{Re}\,\mathsf{Y}^*\mathsf{L}' = I_m\},$$

in analogy with the faithful-model case (see [15, p 279]). Here, $\mathrm{Tr}\,\mathrm{abs}\, A$ means the sum of the absolute values of the eigenvalues of the matrix A. The proof of the following theorem is to be found in the appendix.

Theorem 3 ($C^H(G) = C(G)$). *Denote by $C^{H,N}$ the Holevo's bound of the model* $\{\rho(\theta)^{\otimes N} | \rho(\theta) \in \mathcal{M}\}$. *Then, we have*

$$NC^{H,N}(G) = C^H(G), \tag{13}$$

whose proof is in the appendix. The equation (13) leads to the following theorem, which means that the use of collective measurement does not improve the first-order asymptotic term of the error:

Theorem 4 ($C(G) = C^Q(G)$). *Because of this theorem, our new method for calculating the quasi-quantum CR-type bound actually gives the quantum CR-type bound.*

6. SLD CR-type inequality

When $G = \mathrm{diag}(g_1, \ldots, g_m)$, $C(G)$ is the infimum of the weighted sum of the mean square errors of the $\hat{\theta}^i$. Exchanging the infimum and the sum, we have

$$C^Q(G) = C(G) \geqslant \sum_{i=1}^{m} g_i \left(\inf_{M:\,\text{locally unbiased}} [\mathrm{V}[M]]_{ii} \right). \tag{14}$$

The equality in (14) does not always holds. For, because of non-commutativity, often there is no POM which estimates θ^i and θ^j simultaneously precisely. Hence, the difference between the two sides of the inequality (14) is considered to be an effect of non-commutativity. In this section, it is shown that this difference vanishes if and only if the model is quasi-classical.

The equality $C^Q(G) = C(G) = C^{\mathrm{H}}(G)$ leads to

$$C^Q(G) \geqslant \inf\{\mathrm{Re\ Tr}\, G\mathsf{Y}^*\mathsf{Y} \,|\, |y_i\rangle \in \mathrm{span}_{\mathbb{C}}\{\mathsf{L}'\}\ (i = 1, \ldots, m), \mathrm{Re}\, \mathsf{L}^*\mathsf{Y} = I_m\}.$$

Noting that $\mathsf{Y} = \mathsf{L}' J^{S-1}$ is the only element of $\mathrm{span}_{\mathbb{C}}\{\mathsf{L}'\}$ which satisfies (11), we have the *SLD CR-type inequality* [2, 3, 18, 19],

$$C^Q(G) \geqslant \mathrm{Tr}\, G(J^S)^{-1},$$

where for the equality to be established, the following is necessary:

$$\mathsf{X} = \mathsf{L}' J^{S-1} = \left[\sum_k [J^{S-1}]^{j,k} |l_k\rangle,\ j = 1, \ldots, m \right]. \tag{15}$$

As is proved in the appendix, when the matrix G is diagonal, we have

$$\mathrm{Tr}\, G J^{S-1} = \text{the right-hand side of (14)}. \tag{16}$$

Therefore, to prove our assertion, it suffices to check the necessary and sufficient condition for the equality in the SLD CR-type inequality to be achieved.

Theorem 5. *The equality in the SLD CR-type inequality is established if the model is quasi-classical. Conversely, if the model is quasi-classical, the lower bound is achieved by a projection-valued POM.*

Proof. If the equality is established, then by virtue of (A.1) and (15), it is proved that the model is quasi-classical. Conversely, if $\tilde{J} = 0$, then by virtue of the commuting theorem, there exists a projection-valued POM E such that $\sum_k [J^{S-1}]^{j,k} |l_k\rangle = \int (\hat{\theta}^j - \theta^j)\, E(\mathrm{d}\hat{\theta}) |\phi\rangle$. Elementary calculations show that the mean square matrix of this POM equals J^{S-1}. □

Notice that this theorem is true even if G is not diagonal. When the parameter is one dimensional, as is proved by Fujiwara and Nagaoka [19], the equality in the SLD CR-type inequality holds. For the horizontal lift $|l_i\rangle$, written as $(1/2)L_i^S|\phi\rangle$, theorem 5 and the commuting theorem lead to the following Fujiwara condition [24]: the equality in the SLD CR-type inequality is established if and only if SLDs $\{L_i^S | i = 1, \ldots, m\}$ can be chosen such that $[L_i^S, L_j^S] = 0,\ (i = 1, \ldots, m,\ j = 1, \ldots, m)$.

Fujiwara's condition is much harder to check than ours, because to deny the attainability of the bound, $[L_i^S, L_j^S] \neq 0$ needs to be checked for all possible variations of SLDs. Still, it should be noted that Fujiwara's theorem implies that non-commutativity of the theory is the reason that the two sides of (14) are not necessarily equal. Hence, we can say metaphorically that the equality in the inverse of SLD Fisher information matrix is attainable if and only if any two components of the parameter 'commute' at θ.

Another 'ground' for this metaphor is the following. Often, a model is defined by an initial state and generators:

$$\rho(\theta) = \pi(|\phi(\theta)\rangle), \qquad \partial_i |\phi(\theta)\rangle = \mathrm{i} H_i(\theta) |\phi(\theta)\rangle, \qquad \theta \in \Theta \subset \mathbb{R}^m.$$

Then, the horizontal lift $|l_i\rangle$ can be written as $\mathrm{i}(H_i - \langle\phi|H|\phi\rangle)|\phi\rangle$. Therefore, the equality in the inequality (14) is established if there are generators such that $[H_i(\theta), H_j(\theta)] = 0$.

Example. $H_i(\theta)$ is the position shift operator P_i of the ith particle.

Example. We define *spin rotation model* $\mathcal{M}_{s,m}$ [26] by

$$\rho(\theta) = \pi\{\exp[\mathrm{i}\theta^1(\sin\theta^2 S_x - \cos\theta^2 S_y)]|s, m\rangle\}, \qquad 0 \leqslant \theta^1 < \pi,\ 0 \leqslant \theta^2 < 2\pi,$$

where S_x, S_y, S_z are spin operators, $|s, m\rangle$ is the simultaneous eigenstate of S_z and the total spin, $S_z|s, m\rangle = m|s, m\rangle$, and $(S_x^2 + S_y^2 + S_z^2)|s, m\rangle = s(s+1)|s, m\rangle$. s is a half-integer, and $m = -s, -s+1/2, \ldots, s-1/2, s$. Letting the eigenvalues of $\boldsymbol{D}$ be $\pm\mathrm{i}\beta$, we have $\beta = m/(s^2 + s - m)$. Therefore, when $s = 1, 2, \ldots, m = 0$, the model is quasi-classical.

7. The coherent model

In contrast with the previous section, the coherent case, where the absolute values of eigenvalues of $\boldsymbol{D}$ are maximal, is studied here. The coherent model is worthy of attention first because there are several physically important models which are coherent, and secondly because the coherent model can be viewed as the model with the strongest non-commutativity as is shown in section 8 in the case of the two-dimensional parameter model.

Fujiwara and Nagaoka [20] determined the quantum CR-type bound of the two-parameter coherent model. In the following, more generally, we treat the bound of the coherent model with an arbitrary-dimensional parameter.

Noting that the only system Y of vectors in $\mathrm{span}_{\mathbb{R}}\{\mathsf{L}'\}$ which satisfies $\mathrm{Re}\,\mathsf{Y}^*\mathsf{L}' = I_m$ is $\mathsf{Y} = \mathsf{L}' J^{S-1}$ in the coherent model case, lemma 1 leads to

$$C^Q(G) = C^H(G) = \mathrm{Tr}\, G J^{S-1} + \mathrm{Tr\,abs}\, G J^{S-1} \tilde{J} J^{S-1}.$$

In particular, if the measure $g(\cdot,\cdot)$ of the error is invariant under unitary transformation, as is mentioned in section 3, we can put $g(\theta + \mathrm{d}\theta, \theta) = \sum_{i,j} J^S_{i,j}\, \mathrm{d}\theta^i\, \mathrm{d}\theta^j + \mathrm{o}(||\mathrm{d}\theta||^2)$ without loss of generality, and we have

$$\inf \lim_{N\to\infty} N g(\hat{\theta}, \theta) = C^Q(J^S) = C(J^S) = 2m.$$

Example (squeezed-state model). The *squeezed-state model* [25] is defined by

$$\rho(z, \xi) = \pi(D(z)S(\xi)|0\rangle), \qquad z, \xi \in \mathbb{C},$$

where $S(\xi) = \exp[(1/2)(\xi a^{+2} - \overline{\xi} a^2)]$, $D(z) = \exp(z a^+ - \overline{z} a)$, and $|0\rangle$ is the vacuum state. Letting $z = 2^{-1/2}(\theta^1 + \mathrm{i}\theta^2)$, $Q = 2^{-1/2}(a + a^+)$, and $\xi = \theta^3 \mathrm{e}^{-2\mathrm{i}\theta^4}$, where $0 \leqslant \theta^3,\ 0 \leqslant \theta^4 < \pi$, we have

$$J^S = \frac{1}{2}\begin{bmatrix} \cosh 2\theta^3 - \sinh 2\theta^3 \cos 2\theta^4 & \sinh 2\theta^3 \sin 2\theta^4 & 0 & 0 \\ \sinh 2\theta^3 \sin 2\theta^4 & \cosh 2\theta^3 + \sinh 2\theta^3 \cos 2\theta^4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \sinh^2 2\theta^3 \end{bmatrix},$$

$$\tilde{J} = \frac{1}{2}\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sinh 2\theta^3 \\ 0 & 0 & \sinh 2\theta^3 & 0 \end{bmatrix}.$$

By lemma 2, coherency is checked: $|\det J^S| = |\det \tilde{J}| = \frac{1}{4}\sinh^2 2\theta^3$.

Example. The models $\{\rho(z)|\rho(z) = \rho(z, \xi_0), z \in \mathbb{C}\}$, and $\{\rho(\xi)|\rho(z) = \rho(z_0, \xi), \xi \in \mathbb{C}\}$, where $\rho(z, \xi)$ is defined in the previous example, are coherent. The coherency is checked by use of the lemma 2.

Example. In the spin rotation model $\mathcal{M}_{s,m}$, Fujiwara and Nagaoka [20] have shown that if $m = s$, the model is coherent.

Example (total-space model). *The total-space model* is the space of all the pure states $\mathcal{P}_1$ in finite-dimensional Hilbert space $\mathcal{H}$ [1]. If $\mathrm{span}_{\mathbb{R}}\{\mathsf{L}\}$ is invariant upon multiplication by the imaginary unit i, then by virtue of lemma 1, the model is coherent. Massar and Popescu [8] and Hayashi [12] studied this case, taking the measure of the error to be a unitary invariant, while in our case the measure of the error can be arbitrary.

8. The model with a two-dimensional parameter

So far we have studied two special cases. In this section, we study the intermediate case in the model with a two-dimensional parameter.

Let us define

$$\mathcal{V} = \{V | V = \mathrm{Tr}\, G\, \mathrm{V}[M],\ M \text{ is locally unbiased}\},$$
$$\inf \mathcal{V} = \{V | V = \operatorname*{argmin}_{V \in \mathcal{V}} \mathrm{Tr}\, GV,\ \exists G > 0\}.$$

As is proved in the appendix, we have

$$\det\left(\sqrt{J^S} V \sqrt{J^S} - I_2\right)^{1/2} + \left(\frac{1}{\beta^2} - 1\right)^{1/2} \mathrm{Tr}\left(\sqrt{J^S} V \sqrt{J^S} - I_2\right)^{1/2} = 1, \qquad \beta \neq 0,$$
$$V = J^{S-1}, \qquad \beta = 0. \tag{17}$$

In the following, we use notation like $\inf \mathcal{V}_\theta(\mathcal{M})$, $\beta_\theta(\mathcal{M})$, $C_\theta^Q(G,\ \mathcal{M})$, $J^S(\theta, \mathcal{M})$ to reveal the dependence of these values on θ and $\mathcal{M}$.

Theorem 6. *In the two-dimensional model, if* $\beta_\theta(\mathcal{M}) > \beta_{\theta'}(\mathcal{M}') = \beta'$ *and* $J^S(\theta, \mathcal{M}) = J^S(\theta', \mathcal{M}') = J$ *holds, then for every* $V \in \inf \mathcal{V}_\theta(\mathcal{M})$*, there exists* $V' \in \inf \mathcal{V}_{\theta'}(\mathcal{M}')$ *such that* $V > V'$*, which implies* $C_\theta^Q(G\, \mathcal{M}) > C_{\theta'}(G\, \mathcal{M}')$ *for any* $G > 0$*.*

Proof. For $V \in \inf \mathcal{V}_\theta(\mathcal{M})$, $V' = a(V - J^{S-1}) + J^{S-1}$, where $0 < a < 1$ is a solution to

$$a^2 \det(\sqrt{J} V \sqrt{J} - I_2)^{1/2} + a\left(\frac{1}{\beta'^2} - 1\right)^{1/2} \mathrm{Tr}(\sqrt{J} V \sqrt{J} - I_2)^{1/2} = 1,$$

is a member of $\inf \mathcal{V}_{\theta'}(\mathcal{M}')$, and satisfies $V > V'$. □

These theorems imply that if β is larger, the difference between the two sides of (14) is larger, and, remembering the discussion in section 6, non-commutativity of the model is stronger. If we take the measure $g(\cdot, \cdot)$ of the error to be unitary invariant, we obtain

$$\inf \lim_{N \to \infty} N g(\hat{\theta}, \theta) = C^Q(J^S) = \frac{4}{1 + (1 - \beta^2)^{1/2}}, \tag{18}$$

which is increasing in β (for the proof of the equation, see the appendix). All of these arguments support the assertion that β is a good measure of non-commutativity. In particular, the quasi-classical model is 'commutative', and the coherent model is the model with 'strongest non-commutativity'.

It is pointed out that $\tilde{J}$ is a curvature form of the celebrated Berry phase, and that in the two-dimensional parameter model, β is Berry's phase per unit area, where the unit of area is measured with respect to the Fubini–Study metric [27].

Example (shifted-number-state model [26]). The *shifted-number-state model* $\mathcal{M}_n$ is defined by $\rho(\theta) = \pi[D(\theta^1 + \mathrm{i}\theta^2)|n\rangle]$, where $|n\rangle$ is the nth number state. Then, we have $J^S(\theta, \mathcal{M}_n) = (n + (1/2))I_2$, and $\beta_\theta(\mathcal{M}_n) = 1/(2n + 1)$. As n tends to infinity, $\beta_\theta(\mathcal{M}_n)$ goes to 0 and $\mathcal{M}_n$ approaches a quasi-classical model.

Example. In the spin rotation model, we have $\beta_\theta(\mathcal{M}_{s,m}) = m/(s^2+s-m^2)$. If $m = \alpha s$, where $\alpha < 1$ is a constant, $\beta_\theta(\mathcal{M}_{s,m})$ tends to 0 as $s \to \infty$, and the model tends to be quasi-classical. However, if $m = s$, the model $\mathcal{M}_{s,m}$ is coherent for any s.

Abe [26] calculated the Gaussian curvature of $\mathcal{M}_n$ and $\mathcal{M}_{s,m}$, and showed they tend to 0 in the classical limits, $n \to \infty$ and $s \to \infty$, respectively. When $m = s$, however, $\mathcal{M}_{s,m}$ remains coherent and never becomes commutative in the classical limit, although Gaussian curvature vanishes.

Among the shifted-number-state models, $\mathcal{M}_0$ has strongest non-commutativity. In addition, its quantum CR-type bound is largest, i.e., $C^Q_\theta(\mathcal{M}_0) \leqslant C^Q_\theta(\mathcal{M}_n)$ for any n, because of theorem 6 and the following theorem:

Theorem 7. *In the two-dimensional model, if* $\beta_\theta(\mathcal{M}) = \beta_{\theta'}(\mathcal{M}')$ *and* $J = J^S(\theta, \mathcal{M}) < J^S(\theta', \mathcal{M}') = J'$ *holds, then for every* $V \in \inf \mathcal{V}_\theta(\mathcal{M})$*, there exists* $V' \in \inf \mathcal{V}_{\theta'}(\mathcal{M}')$ *such that* $V > V'$*, which implies* $C^Q_\theta(G\,\mathcal{M}) > C^Q_{\theta'}(G\,\mathcal{M}')$ *for any* $G > 0$.

Proof. Letting V be a member of $\inf \mathcal{V}_\theta(\mathcal{M})$ and O be an orthogonal matrix, $V' = J'^{-(1/2)} O J^{1/2} V J^{1/2} O^{\mathrm{T}} J'^{-(1/2)}$ is a member of $\inf \mathcal{V}_{\theta'}(\mathcal{M}')$. Choose O such that $J^{1/2} V' J^{1/2}$ and $O J^{1/2} V J^{1/2} O^{\mathrm{T}}$ commute. Then, noting that $\det JV > \det JV'$ and $\mathrm{Tr}\, JV > \mathrm{Tr}\, JV'$ are true, we have $J^{1/2} V J^{1/2} > J^{1/2} V' J^1/2$, which means $V > V'$. □

This seems contradictory, because often the vacuum state is often referred to as 'the minimum-uncertainty state'. However, notice that the optimal POM for estimating the mean values of position and momentum is, in general, not the joint measurement of position and momentum in Holevo's sense (see [15, p 120]). Therefore, the uncertainty relation does not set a limit on the quantum CR-type bound in a straightforward manner. This point will be discussed elsewhere [28].

9. Conclusions

A new technique for determining the quasi-quantum CR-type bound of the pure-state model is proposed. As the quasi-quantum CR-type bound is equal to the quantum CR-type bound in the pure-state case, this method gives the quantum CR-type bound at the same time. By use of the reduction theorem in section 4, we can reduce the optimization of the POM to the optimization of finite-dimensional vectors.

The method is successfully applied to the proof of the achievability of Holevo's bound, and to the calculation of the explicit form of the quantum CR-type bound for various models.

The fact that the quantum CR-type bound is equal to the quasi-quantum CR-type bound means that the use of collective measurement is not effective in pure-state models, so far as the first-order term of the error is concerned.

The investigation of the quantum CR-type bound demonstrated that eigenvalues of the eigenvalue of the linear transform $\boldsymbol{D}$ nicely characterize the non-commutativity of the model.

Acknowledgments

The author is grateful to Professors A Fujiwara, H Nagaoka, and Dr M Hayashi for inspiring discussions. The author is also grateful to Professor Imai for providing a pleasant environment for the research.

Appendix. Proofs

Proof of the commuting theorem. The following inequalities are proved in almost the same manner as the faithful-state case (see [15, pp 88–9]):

$$\mathrm{V}[M] \geqslant \mathrm{Re}\,\mathsf{X}^*\mathsf{X}, \qquad \mathrm{V}[M] \geqslant \mathsf{X}^*\mathsf{X}. \tag{A.1}$$

Therefore, if (7) holds, we have $\mathrm{Re}\,\mathsf{X}^*\mathsf{X} \geqslant \mathsf{X}^*\mathsf{X}$, which implies (8).

Conversely, let us assume that (8) holds true. By Schmidt's orthonormalization, from $\{|\phi\rangle, |x^1\rangle, \ldots, |x^m\rangle\}$ we obtain the orthonormal system $\{|b^i\rangle | i = 1, \ldots, m+1\}$ of vectors such that $|x^i\rangle = \sum_{j=1}^{m+1} \lambda^i_j |b^j\rangle, \lambda^j_i \in \mathbb{R}, i = 1, \ldots, m,\ j = 1, \ldots, m+1$. Let $O = [o^i_j]$ be an $(m+1) \times (m+1)$ real orthogonal matrix such that $\langle\phi| \sum_{j=1}^{m+1} o^i_j |b^j\rangle \neq 0$. Then, denoting $\sum_{j=1}^{m+1} o^i_j |b^j\rangle$ by $|b'^i\rangle$, and $\sum_{j=1}^{m+1} \lambda^i_j o^k_j / \langle b'^k|\phi\rangle$ by $\hat{\theta}^i_k$, we obtain an unbiased POM which meets the conditions as follows:

$$E(\{\hat{\theta}_k\}) = |b'^k\rangle\langle b'^k| \qquad (k = 1, \ldots, m+1), \qquad E(\{0\}) = I - \sum_{k=1}^{m+1} |b'^k\rangle\langle b'^k|.$$

Proof of theorem 3. Decompose $|x_i\rangle \in \mathbb{C}^{2m+1}/\{\mathbb{C}|\phi'\rangle\}$ into $|x_i\rangle = |y_i\rangle + |z_i\rangle$, where $|y_i\rangle \in \mathrm{span}_{\mathbb{C}}\{\mathsf{L}'\}$ and $|z_i\rangle \in \mathbb{C}^{2m+1}/\mathrm{span}_{\mathbb{C}}\{\mathsf{L}', |\phi'\rangle\}$, and denote $[|y_1\rangle, \ldots, |y_m\rangle]$ and $[|z_1\rangle, \ldots, |z_m\rangle]$ by Y and Z, respectively. Then, let us consider the minimization of $\mathrm{Re}\ \mathrm{Tr}\,\mathsf{X}^*\mathsf{X}G = \mathrm{Re}\ \mathrm{Tr}\,\mathsf{Y}^*\mathsf{Y}G + \mathrm{Re}\ \mathrm{Tr}\,\mathsf{Z}^*\mathsf{Z}G$ with Y fixed. Let us define

$$\mathrm{Lag}(\mathsf{X}) = \mathrm{Re}\ \mathrm{Tr}\,\mathsf{X}^*\mathsf{X}G + \mathrm{Tr}\,\mathrm{Im}\,\mathsf{X}^*\mathsf{X}\Lambda,$$

where Λ is an $m \times m$ antisymmetric real matrix. Differentiation with respect to Z yields $\mathsf{Z}(G - \mathrm{i}\Lambda) = 0$, which, multiplied by Z^* on both sides, leads to

$$\begin{aligned} \sqrt{G}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G} &= -\sqrt{G}\,\mathrm{Im}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}\sqrt{G^{-1}}\Lambda\sqrt{G^{-1}}, \\ \sqrt{G}\,\mathrm{Im}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G} &= \sqrt{G}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}\sqrt{G^{-1}}\Lambda\sqrt{G^{-1}}. \end{aligned} \tag{A.2}$$

Therefore, we have,

$$\sqrt{G}\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}[I_m + (\sqrt{G^{-1}}\Lambda\sqrt{G^{-1}})^2] = 0. \tag{A.3}$$

For $\sqrt{G}\,\mathrm{Im}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}$ and $\sqrt{G^{-1}}\Lambda\sqrt{G^{-1}}$ antisymmetric, and $\sqrt{G}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}$ symmetric and strictly positive, equation (A.2) implies that

$$\sqrt{G}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G} = O\,\mathrm{diag}(|a_1 b_1|, \ldots, |a_m b_m|)O^{\mathrm{T}}, \tag{A.4}$$

where O is an orthogonal matrix, and a_i, b_i $(i = 1, \ldots, m)$ are the eigenvalues of $\sqrt{G}\,\mathrm{Im}\,\mathsf{Z}^*\mathsf{Z}\sqrt{G}$ and $\sqrt{G^{-1}}\Lambda\sqrt{G^{-1}}$, respectively. Equations (A.3) and (A.4) lead to $\mathrm{Tr}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}G = \sum_{i=1}^m |a_i|$. Noting that $\mathrm{Im}\,\mathsf{X}^*\mathsf{X} = \mathrm{Im}\,\mathsf{Y}^*\mathsf{Y} + \mathrm{Im}\,\mathsf{Z}^*\mathsf{Z} = 0$, we have $\mathrm{Tr}\,\mathrm{Re}\,\mathsf{Z}^*\mathsf{Z}G = \mathrm{Tr}\,\mathrm{abs}\,\mathrm{Im}\,\mathsf{Y}^*\mathsf{Y}G$, which leads to

$$\inf_{\mathsf{Z}} \mathrm{Tr}\,\mathrm{Re}\,\mathsf{X}^*\mathsf{X}G = \mathrm{Tr}\,G\,\mathrm{Re}\,\mathsf{Y}^*\mathsf{Y} + \mathrm{Tr}\,\mathrm{abs}\,G\,\mathrm{Im}\,\mathsf{Y}^*\mathsf{Y}, \tag{A.5}$$

or $C^H(G) = C(G)$.

Proof of equation (13). Consider the model $\{\rho(\theta)^{\otimes N}|\rho(\theta) \in \mathcal{M}\}$, and denote by $|l_i^N\rangle$ the horizontal lift of the tangent vector $\partial_i[\rho(\theta)^{\otimes N}]$. Then, we have $\langle l_i^N|l_j^N\rangle = N\langle l_i|l_j\rangle$ $(i = 1,\ldots,m,\ j = 1,\ldots,m)$, which implies

$$C^{H,N}(G) = \inf_{\mathsf{Y}}\{\mathrm{Tr}\, G\,\mathrm{Re}\,\mathsf{Y}^*\mathsf{Y} + \mathrm{Tr\,abs}\, G\,\mathrm{Im}\,\mathsf{Y}^*\mathsf{Y}|\mathrm{Re}\,\mathsf{Y}^*\sqrt{N}\mathsf{L}' = I_m\}$$
$$= \frac{1}{N}\inf_{\sqrt{N}\mathsf{Y}}\{\mathrm{Tr}\, G\,\mathrm{Re}\,(\sqrt{N}\mathsf{Y})^*(\sqrt{N}\mathsf{Y})$$
$$+\ \mathrm{Tr\,abs}\, G\,\mathrm{Im}\,(\sqrt{N}\mathsf{Y})^*(\sqrt{N}\mathsf{Y})|\mathrm{Re}\,(\sqrt{N}\mathsf{Y})^*\mathsf{L}' = I_m\}$$
$$= \frac{1}{N}\inf_{\mathsf{Y}}\{\mathrm{Tr}\, G\,\mathrm{Re}\,\mathsf{Y}^*\mathsf{Y} + \mathrm{Tr\,abs}\, G\,\mathrm{Im}\,\mathsf{Y}^*\mathsf{Y}|\mathrm{Re}\,\mathsf{Y}^*\mathsf{L}' = I_m\}$$

and we have equation (13).

Proof of (16). Let $E^{(i)}$ be a projection-valued POM such that $\int_{\mathbb{R}} x\, E^{(i)}(\mathrm{d}x) = \sum_{j=1}^m [J^{S-1}]^{ij} L_j^S$, and $M_{\boldsymbol{p}}$ be a locally unbiased POM at θ such that

$$M_{\boldsymbol{p}}(B) = p_i E^{(i)}(\{\theta^i + p_i(x-\theta^i)|x \in B^i\}), \qquad B = \{\theta^1\}\times\cdots\times B^i\times\cdots\times\{\theta^m\},$$

where $\boldsymbol{p} = [p_i]$ $(i = 1,\ldots,m)$ satisfies $\sum_i p_i = 1$ and $p_i \geqslant 0$, and B^i $(i = 1,\ldots,m)$ are an arbitrary measurable set. Then, we have $[V[M_{\boldsymbol{p}}]]_{ii} = (1/p_i)[J^{S-1}]^{ii}$ which leads to

$$\inf\{[V[M]]_{ii}|\ M \text{ is locally unbiased at } \theta\} \leqslant \inf\left\{[V[M_{\boldsymbol{p}}]]_{ii}\,\middle|\,\sum_i p_i = 1,\ p_i \geqslant 0\right\} = [J^{S-1}]^{ii}.$$

This equation, combined with SLD CR-type inequality, leads to the theorem.

Proof of (17) and (18). Define a function Lag(X) by

$$\mathrm{Lag}(\mathsf{X}) = \mathrm{Re\ Tr}\,\mathsf{X}^*\mathsf{X}G - 2\,\mathrm{Tr}[(\mathrm{Re}\,\mathsf{X}^*\mathsf{L}' - I_2)\Xi] + \mathrm{Tr\,Im}\,\mathsf{X}^*\mathsf{X}\Lambda, \tag{A.6}$$

where Ξ, Λ are real 2×2 matrices and Λ is antisymmetric. Then, differentiation of (A.6) with respect to X yields $\mathsf{X}(G - \mathrm{i}\Lambda) = \mathsf{L}'\Xi$. Multiplication by X^* and taking the real parts of both sides of this equation, together with (8) and (11), yield $\Xi = \mathrm{Re}\,\mathsf{X}^*\mathsf{X}G$. Therefore, putting $V = \mathrm{Re}\,\mathsf{X}^*\mathsf{X}$, we have

$$\mathsf{X}(G - \mathrm{i}\Lambda) = \mathsf{L}VG. \tag{A.7}$$

Equation (A.7), mixed with (8), leads to

$$(G - \mathrm{i}\Lambda)V(G - \mathrm{i}\Lambda) = GV\mathsf{L}^*\mathsf{L}VG. \tag{A.8}$$

Let V be a solution of (A.8). Then $\mathsf{X} = UV^{1/2}$, where U is a 5×2 complex matrix such that $U^*U = I_2$, satisfies (8) and (A.7). $\mathsf{X} = UV^{1/2}$ also satisfies (11), because $VG = \mathrm{Re}\,\mathsf{X}^*\mathsf{L}VG$ is obtained by multiplying by X^* and taking the real parts of both sides of (A.7). Hence, our task is to solve (A.8).

Putting $V' = J^{S(1/2)}VJ^{S(1/2)}$, $G' = J^{S-(1/2)}GJ^{S-(1/2)}$, $\Lambda' = J^{S-(1/2)}\Lambda J^{S-(1/2)}$, and $\tilde{J}' = J^{S-(1/2)}\tilde{J}J^{S-(1/2)}$, equation (A.8) can be written as

$$(G' - \mathrm{i}\Lambda')V'(G' - \mathrm{i}\Lambda') = G'V'(I_2 + \mathrm{i}\tilde{J}')V'G'. \tag{A.9}$$

Here, without loss of generality, we can put $\tilde{J}_{12} = -\tilde{J}_{21} = \beta$.

If we put $G = J^S$, or equivalently $G' = I_2$, equation (A.9) is easily solved, and we obtain (18).

Letting O be an arbitrary 2×2 real special orthogonal matrix, we have $O\tilde{J}O^{\mathrm{T}} = \tilde{J}$ and $O\Lambda O^{\mathrm{T}} = \Lambda$. Therefore, letting u, v be eigenvalues of a solution V' to (A.9), diag(u, v) is also a solution. Conversely, if diag(u, v) is a solution, Odiag$(u, v)O^{\mathrm{T}}$ is also a solution for every special orthogonal matrix O. After a few lines of calculations, the necessary and sufficient condition for Λ' and G' which ensures that (A.9) exists is

$$\beta[(u-1)(v-1)]^{1/2} \pm (1-\beta^2)^{1/2}[(u-1)^{1/2}+(v-1)^{1/2}] = \beta.$$

Drawing the graph (not shown), we see that the lower sign in the equation corresponds to the set of stationary points, and the upper sign gives inf $\mathcal{V}$. Finally, replacing $(u-1)^{1/2}(v-1)^{1/2}$ and $(u-1)^{1/2}+(v-1)^{1/2}$ by $\det(V'-I_2)^{1/2}$ and $\mathrm{Tr}(V'-I_2)^{1/2}$, respectively, we obtain (17).

References

[1] Matsumoto K 1996 *Mathematical Engineering Technical Report* no 96-109
(Matsumoto K 1997 *LANL Preprint* quant-ph/9711008)
Matsumoto K 1997 A geometrical approach to quantum estimation theory *Doctoral Thesis* University of Tokyo
[2] Helstrom C W 1967 *Phys. Lett.* A **25** 101
[3] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[4] Jones K R W 1991 *J. Phys. A: Math. Gen.* **24** 121
Jones K R W 1994 *Phys. Rev.* A **50** 3682
[5] D'Ariano G M 1997 *Quantum Communication, Computing, and Measurement* ed O Hirota *et al* (New York: Plenum) p 253
(D'Ariano G M 1997 *LANL Preprint* quant-ph/9701011)
[6] D'Ariano G M and Yuen H P 1996 *Phys. Rev. Lett.* **76** 2832
[7] Bužek V, Adam G and Drobný G 1996 *Phys. Rev.* A **54**
[8] Massar S and Popescu S 1995 *Phys. Rev. Lett.* **74** 1259
[9] Nagaoka H 1989 *Proc. SITA '89* p 577
[10] Nagaoka H 1992 *Proc. SITA '92* p 63
[11] Nagaoka H 1994 *Proc. IEEE Int. Symp. on Information Theory* p 118
[12] Hayashi M 1998 *J. Phys. A: Math. Gen.* **30** 4633–55
[13] Hayashi M and Matsumoto K 1998 *RIMS Kokyuroku* **1055** 96 (in Japanese)
[14] Yuen H P and Lax M 1973 *IEEE Trans. Inform. Theory* **19** 740
[15] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[16] Nagaoka H 1989 *IEICE Technical Report* IT89-42, p 9
Nagaoka H 1991 *Trans. Japan Soc. Indust. Appl. Math.* **1** 305 (in Japanese)
[17] Hayashi A 1997 *LANL Preprint* quant-ph/9704044
[18] Fujiwara A 1994 A geometrical study in quantum information systems *Doctoral Thesis* University of Tokyo
[19] Fujiwara A and Nagaoka H 1995 *Phys. Lett.* A **201** 119
[20] Fujiwara A and Nagaoka H 1996 *J. Math. Phys.* **40** 4227–39
[21] Steinspring W F 1955 *Proc. Am. Math. Soc.* **6** 211
[22] Ozawa M 1984 *J. Math. Phys.* **25** 79
[23] Kobayashi S and Nomizu K 1969 *Foundations of Differential Geometry* vols 1 and 2 (New York: Interscience)
[24] Fujiwara A 1993 private communication
[25] Yuen H P 1976 *Phys. Rev.* A **13** 226
[26] Abe S 1993 *Phys. Rev.* A **48** 4102
[27] Matsumoto K 1997 *Mathematical Engineering Technical Report* no 97-110
[28] Matsumoto K 1997 *Mathematical Engineering Technical Report* no 97-108
[29] Gill R and Massar S 2000 *Phys. Rev.* A **61** 042312
[30] Bennett C *et al* 1998 *LANL Preprint* quant-ph/9804053
[31] Slater P B 2000 *LANL Preprint* quant-ph/0006009
[32] Actin A *et al* 1999 *LANL Preprint* quant-ph/9904056