Multi-parameter pure state estimation based on the right logarithmic derivative

Akio Fujiwara

METR 94-9 July 1994

Multi-parameter pure state estimation based on the right logarithmic derivative

Akio Fujiwara,
University of Tokyo*

Abstract

A statistical multi-parameter estimation theory for quantum pure state models is presented, based on right logarithmic derivatives. It gives a new information theoretical insight into the coherent states in quantum mechanics.

Keywords : quantum estimation theory, pure state, right logarithmic derivative, coherent 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 $\mathcal{H}$ with finite-dimensional real parameters $\theta = (\theta^i)_{i=1}^n$ which are to be estimated statistically. In order to avoid singularities, the conventional quantum estimation theory [1][2] has been often restricted to models that are composed of strictly positive density operators. It was Helstrom [3] who successfully introduced the symmetrized logarithmic derivatives for the one parameter estimation theory as a quantum counterpart of the logarithmic derivative in the classical estimation theory. The right logarithmic derivative is another successful counterpart introduced by Yuen and Lax [4] in the expectation parameter estimation theory for quantum gaussian models, which provided a theoretical background of optical communication theory. Quantum information theorists have also kept away from degenerated states, such as pure states, for mathematical convenience [5]. Indeed, the von Neumann entropy cannot distinguish the pure states, and the relative entropies diverge.

This is a companion to the paper [6] which try to construct an estimation theory for pure state models. In Sec. 2, we give a brief summary of the conventional quantum parameter estimation theory. In Sec. 3, we study a multi-parameter quantum estimation theory based on the right logarithmic derivative. The estimation theoretical significance of the coherent models is also clarified. In order to demonstrate the results, some examples are presented in Sec. 4.

2 Review of the conventional theory

Let $\mathcal{H}$ be a separable Hilbert space which corresponds to a physical system with inner product $\langle\phi|\psi\rangle$, $(\phi, \psi \in \mathcal{H})$, and $\mathcal{L}$ and $\mathcal{L}_{sa}$ be, respectively, the set of all the (bounded) linear operators and all the self-adjoint operators on $\mathcal{H}$. 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 $\mathrm{rank}\,\rho = 1$ or equivalently $\rho^2 = \rho$. In order to handle joint probability distributions of possibly mutually noncommuting observables, an extended framework of measurement theory is needed [1, p. 53] [2, p. 50]. A *generalized measurement* $\{M(B)\}_{B\in\mathcal{F}}$ on a measurable space $(\Omega, \mathcal{F})$ 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. $M(\bigcup_j B_j) = \sum_j M(B_j), \quad$ (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 $M(B_1)M(B_2) = 0$, $(^\forall B_1 \cap {}^\forall B_2 = \phi)$.

Given a statistical parametric model composed of strictly positive density operators:

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

Here, $\theta = (\theta^1, \cdots, \theta^n)$ 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 $\theta \in \Theta$. 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

$$v_\theta^{jk} = \int (\hat{\theta}^j - \theta^j)(\hat{\theta}^k - \theta^k) P_\theta^M(d\hat{\theta}). \tag{4}$$

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

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \frac{1}{2}[\rho_\theta L_{\theta,j} + L_{\theta,j}^* \rho_\theta]. \tag{5}$$

For instance,

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \frac{1}{2}[\rho_\theta L^S_{\theta,j} + L^S_{\theta,j}\rho_\theta], \quad L^S_{\theta,j} = L^{S*}_{\theta,j} \tag{6}$$

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

$$\frac{\partial \rho_\theta}{\partial \theta^j} = \rho_\theta L^R_{\theta,j} \tag{7}$$

defines the *right logarithmic derivative* (RLD) $L^R_{\theta,j}$ introduced by Yuen and Lax [4]. Thus, (5) defines a certain family 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

$$(A, B)_\rho = \mathrm{Tr}\, \rho\, BA^*. \tag{8}$$

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

$$\langle A, B\rangle_\rho = \frac{1}{2}\mathrm{Tr}\, \rho\, (BA^* + A^*B). \tag{9}$$

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

Proposition 1 *For any locally unbiased measurement M, the following inequality holds:*

$$V_\theta[M] \geq (\mathrm{Re}\, J_\theta)^{-1}, \tag{10}$$

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,*

$$V_\theta[M] \geq (J^R_\theta)^{-1} \tag{11}$$

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 (10) and (11) 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 as long as the SLD is used.

On the other hand, for the dimension $n \geq 2$, the matrix equalities in (10) and (11) 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

$$\mathrm{tr}\ GV_\theta[M] = \sum_{jk} g_{jk} v_\theta^{jk}. \tag{12}$$

If there is a constant C such that $\mathrm{tr}\ GV_\theta[M] \geq C$ 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 [7].

$$C^S = \mathrm{tr}\, G (J_\theta^S)^{-1}, \tag{13}$$

$$C^R = \mathrm{tr}\, G\, \mathrm{Re}\, (J_\theta^R)^{-1} + \mathrm{tr}\,\mathrm{abs}\, G\, \mathrm{Im}\, (J_\theta^R)^{-1}. \tag{14}$$

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 [4] (see also Holevo [2, p. 281] 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 [7] 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 Multi–parameter pure state model estimation theory

As was mentioned in the previous section, there is no prototype for general theory of quantum multi-parameter estimation theory. So, let us restrict ourselves here to seeking the estimation theory based on the RLD.

We first note that the sesquilinear forms $(\cdot,\cdot)_\rho$ and $\langle\cdot,\cdot\rangle_\rho$ defined by (8) (9) become pre-inner products on $\mathcal{L}$ when ρ is degenerated. Denote by $\mathcal{K}(\rho)$ the set of linear operators $K \in \mathcal{L}$ satisfying $(K,K)_\rho = 0$, which are called the kernel of the pre-inner product $(\cdot,\cdot)_\rho$. Also 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 following lemmas are fundamental [6].

Lemma 1 *Suppose ρ is pure. Then the following 3 conditions for linear operators $K \in \mathcal{L}$ are equivalent.*

(i) $(K,K)_\rho = 0$,

(ii) $\rho K = 0$,

(iii) $\mathrm{Tr}\,\rho K = 0$ *and* $\rho K + K^*\rho = 0$.

Lemma 2 *Suppose ρ is pure. Then the following 3 conditions for self-adjoint operators $K \in \mathcal{L}_{sa}$ are equivalent.*

(i) $\langle K,K\rangle_\rho = 0$,

(ii) $\rho K = 0$,

(iii) $\rho K + K\rho = 0$.

These lemmas are usefully employed in the pure state estimation theory. For instance, the SLD, also defined by (6), is determined up to the uncertainty of $K \in \mathcal{K}_{sa}$. Let us denote the totality of such SLD's by $\mathcal{T}^S(\rho)$. Then it is proved that for a pure state model ρ_θ, the SLD–Fisher information matrix $J_\theta^S = \left[\langle L_{\theta,j}^S, L_{\theta,k}^S\rangle_{\rho_\theta}\right]$ is uniquely determined on the quotient space $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$, and its (j,k) entry is given by

$$(J_\theta^S)_{jk} = 2\mathrm{Tr}\,(\partial_j\rho_\theta)(\partial_k\rho_\theta), \tag{15}$$

where $\partial_j = \partial/\partial\theta^j$, see [6]. Therefore, we may call the quotient space $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ the *SLD–tangent space*.

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

$$i(A\rho - \rho A) = \frac{1}{2}\left((\mathfrak{D}A)\rho + \rho(\mathfrak{D}A)\right), \qquad A,\ \mathfrak{D}A \in \mathcal{L}_{sa}. \tag{16}$$

This is an anti-symmetric super-operator such that $\langle A, \mathfrak{D}B\rangle_\rho = -\langle \mathfrak{D}A, B\rangle_\rho$ holds for all $A, B \in \mathcal{L}_{sa}$.

Lemma 3 *Suppose ρ is pure. Then $\mathfrak{D}$ is regarded as a super–operator on the quotient space $\mathcal{L}_{sa}/\mathcal{K}_{sa}(\rho)$, and is defined by*

$$(\mathfrak{D}X)\,\rho = 2i(X - \mathrm{Tr}\,\rho X)\rho, \qquad (X \in \mathcal{L}_{sa}/\mathcal{K}_{sa}(\rho)). \tag{17}$$

Proof Let us denote two distinct images of $A \in \mathcal{L}_{sa}$ by $(\mathfrak{D}A)$ and $(\mathfrak{D}A)'$, then $K = (\mathfrak{D}A) - (\mathfrak{D}A)' \in \mathcal{L}_{sa}$ satisfies $K\rho + \rho K = 0$ and, from Lemma 2, $K \in \mathcal{K}_{sa}(\rho)$. Further, observing $\langle \mathfrak{D}K, \mathfrak{D}K\rangle_\rho = \langle -\mathfrak{D}^2 K, K\rangle_\rho$, $K \in \mathcal{K}_{sa}(\rho)$ implies $\mathfrak{D}K \in \mathcal{K}_{sa}(\rho)$. Therefore, $\mathfrak{D}$ is regarded as a super–operator on $\mathcal{L}_{sa}/\mathcal{K}_{sa}(\rho)$. Further, re-expressing (16) as

$$\rho\left[\frac{\mathfrak{D}X}{2} + i(X - \mathrm{Tr}\,\rho X)\right] + \left[\frac{\mathfrak{D}X}{2} + i(X - \mathrm{Tr}\,\rho X)\right]^* \rho = 0,$$

and using Lemma 1 together with the identity $\mathrm{Tr}\,\rho(\mathfrak{D}X) = 0$, we have an equivalent equation (17). ■

Let us return to the subject, i.e., the estimation theory based on the RLD. It may sound strange since the RLD defined by (7) does not exist for degenerated states. However, it is essential to notice that what we need is not the RLD itself but the inverse of the RLD–Fisher information matrix, as is understood by (11).

We start with the following theorem, which is a modification of the Holevo's result originally obtained in the strictly positive case [2, p. 280]. Hereafter, the subscripts θ of the SLD's are omitted for simplicity.

Theorem 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 $L^R_\theta(\varepsilon)$. If the SLD–tangent space $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ is $\mathfrak{D}$-invariant, then*

$$\lim_{\varepsilon\downarrow 0}\left(J^R(\varepsilon)\right)^{-1} = \left(J^S\right)^{-1} + \frac{i}{2}\left(J^S\right)^{-1} D \left(J^S\right)^{-1} \tag{18}$$

holds, where $J^R(\varepsilon) = \left[(L^R_j(\varepsilon), L^R_k(\varepsilon))_{\rho_\theta(\varepsilon)}\right]$ and $D = \left[i\,\mathrm{Tr}\,\rho_\theta[L^S_j, L^S_k]\right]$.

Proof Observing the identities

$$\frac{\partial\rho_\theta}{\partial\theta^j} = \frac{1}{2}(\rho_\theta L^S_j + L^S_j\rho_\theta) = \left(L^R_j\right)^*\rho_\theta$$

and

$$(A, B)_\rho = \langle A, (I + \frac{i}{2}\mathfrak{D})B\rangle_\rho,$$

we have, for all $X \in \mathcal{L}$,

$$\langle L_j^S, X\rangle_{\rho_\theta} = (L_j^R, X)_{\rho_\theta} = \langle (I + \frac{i}{2}\mathfrak{D})L_j^S, X\rangle_{\rho_\theta}.$$

Then $L_j^S = (I + \frac{i}{2}\mathfrak{D})L_j^R$ and

$$J^R(\varepsilon) = \left[\langle L_j^S(\varepsilon), (I + \frac{i}{2}\mathfrak{D}(\varepsilon))^{-1}L_k^S(\varepsilon)\rangle_{\rho_\theta(\varepsilon)}\right].$$

Since $I + \frac{i}{2}\mathfrak{D}(\varepsilon)$ is symmetric with respect to $\langle\cdot,\cdot\rangle_{\rho_\theta}$, the following lemma immediately leads us to (18), by setting $V = \mathcal{L}_{sa}$, $x_j(\varepsilon) = L_j^S(\varepsilon)$, and $\mathfrak{A}(\varepsilon) = I + \frac{i}{2}\mathfrak{D}(\varepsilon)$. ■

Lemma 4 *Let V be a d-dimensional linear space (possibly $d = \infty$). Given a family of $n(< d)$ linearly independent vectors $\{x_j(\varepsilon)\}_{j=1}^n$ in V, a family of inner products $\langle\cdot,\cdot\rangle_\varepsilon$, and a family of symmetric operators $\mathfrak{A}(\varepsilon)$ on V with respect to the inner product, having a parameter $\varepsilon \geq 0$. Suppose $\mathfrak{A}(\varepsilon)$ is invertible for $\varepsilon > 0$ but $\mathfrak{A}(0)$ is not. Further, the linear span of $\{x_j(0)\}_{j=1}^n$ is $\mathfrak{A}(0)$-invariant in V. Denote $n \times n$ matrices*

$$J^R(\varepsilon) = \left[\langle x_j(\varepsilon), \mathfrak{A}^{-1}(\varepsilon)x_k(\varepsilon)\rangle_\varepsilon\right], \qquad J(\varepsilon) = \langle x_j(\varepsilon), x_k(\varepsilon)\rangle_\varepsilon.$$

Then

$$\lim_{\varepsilon\downarrow 0}\left(J^R(\varepsilon)\right)^{-1} = J^{-1}(0)\left[\langle x_j(0), \mathfrak{A}(0)x_k(0)\rangle_0\right]J^{-1}(0). \tag{19}$$

Proof Let us denote, by $W^\perp(\varepsilon)$, the orthogonal complement of $W(\varepsilon) = \operatorname{span}\{x_j(\varepsilon)\}_{j=1}^n$ with respect to the inner product $\langle\cdot,\cdot\rangle_\varepsilon$ in V. Further, let $\{y_j(\varepsilon)\}_{j=n+1}^d$ be a basis of $W^\perp(\varepsilon)$ and construct a basis $\{z_j(\varepsilon)\}_{j=1}^d$ of V by combining them as

$$z_j(\varepsilon) = \begin{cases} x_j(\varepsilon), & j = 1,\cdots,n, \\ y_j(\varepsilon), & j = n+1,\cdots,d. \end{cases}$$

Consider enlarged $d \times d$ matrices

$$\mathfrak{J}^R(\varepsilon) = \left[\langle z_j(\varepsilon), \mathfrak{A}^{-1}(\varepsilon)z_k(\varepsilon)\rangle_\varepsilon\right], \qquad \mathfrak{J}(\varepsilon) = \left[\langle z_j(\varepsilon), z_k(\varepsilon)\rangle_\varepsilon\right].$$

Since $V = W(\varepsilon) \oplus W^\perp(\varepsilon)$ is, of course, $\mathfrak{A}(\varepsilon)$-invariant, the inverse of $\mathfrak{J}^R(\varepsilon)$ is explicitly given as

$$\left(\mathfrak{J}^R(\varepsilon)\right)^{-1} = \mathfrak{J}^{-1}(\varepsilon)\left[\langle z_j(\varepsilon), \mathfrak{A}(\varepsilon)z_k(\varepsilon)\rangle_\varepsilon\right]\mathfrak{J}^{-1}(\varepsilon).$$

This matrix is well-defined even for $\varepsilon = 0$. Decompose $\left(\mathfrak{J}^R(\varepsilon)\right)^{-1}$ into blocks:

$$\begin{bmatrix} P & O \\ O & Q \end{bmatrix}\begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} P & O \\ O & Q \end{bmatrix} = \begin{bmatrix} PAP & PBQ \\ QCP & QDQ \end{bmatrix},$$

where the three matrices in the left-hand side correspond to $\mathfrak{J}^{-1}(\varepsilon)$, $[\langle z_j(\varepsilon), \mathfrak{A}(\varepsilon) z_k(\varepsilon)\rangle_\varepsilon]$, and $\mathfrak{J}^{-1}(\varepsilon)$, respectively, P, A are $n\times n$ matrices, and Q, D are $(d-n)\times(d-n)$ matrices. Further, it is easy to see that $B = C = O$ for $\varepsilon = 0$, since $W(0)$ is $\mathfrak{A}(0)$-invariant and $\mathfrak{A}(0)$ is symmetric. Then $\lim_{\varepsilon\downarrow 0}\left(\mathfrak{J}^R(\varepsilon)\right)^{-1}$ becomes a block diagonal matrix, and the limit of the first $n\times n$ block PAP approaches (19). ■

Note that $\mathrm{Tr}\,\rho_\theta[L^S_j, L^S_k]$ in Theorem 1 also independent of the uncertainty of the SLD. Therefore, Theorem 1 asserts that the inverse of the RLD–Fisher information matrix can be obtained directly from the 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 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*

$$\mathfrak{D}\begin{bmatrix} \tilde{L}^S_1 \\ \tilde{L}^S_2 \\ \tilde{L}^S_3 \\ \tilde{L}^S_4 \\ \vdots \\ \tilde{L}^S_{2m-1} \\ \tilde{L}^S_{2m} \end{bmatrix} = \begin{bmatrix} 0 & 2 & & & & & \\ -2 & 0 & & & & & \\ & & 0 & 2 & & & \\ & & -2 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & 2 \\ & & & & & -2 & 0 \end{bmatrix}\begin{bmatrix} \tilde{L}^S_1 \\ \tilde{L}^S_2 \\ \tilde{L}^S_3 \\ \tilde{L}^S_4 \\ \vdots \\ \tilde{L}^S_{2m-1} \\ \tilde{L}^S_{2m} \end{bmatrix}. \tag{20}$$

Proof Since $\mathrm{Tr}\,\rho(\mathfrak{D}X) = 0$ holds for all $X \in \mathcal{L}_{sa}/\mathcal{K}_{sa}$, (17) is rewritten as

$$[-4(X - \mathrm{Tr}\,\rho X)]\rho = 2i[\mathfrak{D}X - \mathrm{Tr}\,\rho(\mathfrak{D}X)]\rho.$$

Comparing this equation to (17) with X replaced by $\mathfrak{D}X$, we have

$$\mathfrak{D}^2 X = -4(X - \mathrm{Tr}\,\rho X).$$

In particular, $\mathfrak{D}^2 X = -4X$ holds for every X which satisfies $\mathrm{Tr}\,\rho X = 0$. We may, therefore, write this relation as $\mathfrak{D}^2 = -4$ on the SLD–tangent space $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ for short.

Take an arbitrary element e_1 of the SLD–tangent space $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$, and let $e_2 = \mathfrak{D}e_1$. Then from the assumption, $e_2 \in \mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$, and $\mathfrak{D}e_2 = -4e_1$ since $\mathfrak{D}^2 = -4$. Therefore, $\mathcal{S}_1(\rho_\theta) = \mathrm{span}\,\{e_1, e_2\}$ is a $\mathfrak{D}$-invariant subspace of $\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)$ and

$$\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta) = \mathcal{S}_1(\rho_\theta) \oplus \mathcal{S}_1(\rho_\theta)^\perp,$$

where $\mathcal{S}_1(\rho_\theta)^\perp$ is the orthogonal complement of $\mathcal{S}_1(\rho_\theta)$ with respect to $\langle \cdot, \cdot \rangle_{\rho_\theta}$. Repeating the same procedure to $\mathcal{S}_1(\rho_\theta)^\perp$, we have

$$\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta) = \mathcal{S}_1(\rho_\theta) \oplus \mathcal{S}_2(\rho_\theta) \oplus \cdots \oplus \mathcal{S}_m(\rho_\theta).$$

In particular, $\dim[\mathcal{T}^S(\rho_\theta)/\mathcal{K}_{sa}(\rho_\theta)] = 2m$.

We next investigate the structure of 2-dimensional $\mathfrak{D}$-invariant subspace $\mathcal{S}_1(\rho_\theta) = \mathrm{span}\,\{e_1, e_2\}$. Expressing the operation of $\mathfrak{D}$ in a matrix form

$$\mathfrak{D} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} x & y \\ z & w \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}, \qquad (x, y, z, w \in \mathbf{R})$$

and using $\mathfrak{D}^2 = -4$, we have

$$x^2 + yz = -4, \quad y(x+w) = 0, \quad z(x+w) = 0, \quad w^2 + yz = -4.$$

These equations do not contradict the identities $\langle e_1, \mathfrak{D}e_1 \rangle_\rho = \langle e_2, \mathfrak{D}e_2 \rangle_\rho = 0$ iff $x + w = 0$, $y \neq 0$, $z \neq 0$. In this case

$$\mathfrak{D} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} x & y \\ -(x^2+4)/y & -x \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}.$$

Furthermore, the transformation of the basis

$$\begin{bmatrix} \tilde{L}_1^S \\ \tilde{L}_2^S \end{bmatrix} = \begin{bmatrix} 2/y & 0 \\ x/y & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

yields

$$\mathfrak{D} \begin{bmatrix} \tilde{L}_1^S \\ \tilde{L}_2^S \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -2 & 0 \end{bmatrix} \begin{bmatrix} \tilde{L}_1^S \\ \tilde{L}_2^S \end{bmatrix}.$$

Repeating the same procedure to other invariant subspaces, we have the theorem. ■

Definition 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 (20) is called* ρ_θ*–symplectic.*

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

$$\begin{array}{ll} \text{(i)} & \{\tilde{L}_j^S\}_{j=1,2} \textit{ is a } |\theta\rangle\langle\theta|\textit{–symplectic basis.} \\ \text{(ii)} & (\tilde{L}_1^S + i\tilde{L}_2^S)|\theta\rangle = 0. \end{array}$$

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

Proof We first assume (i). Letting $X = \tilde{L}_1^S$ in (17), we have

$$(\tilde{L}_1^S + \frac{i}{2}\mathfrak{D}\tilde{L}_1^S)\rho_\theta = 0.$$

Then, $\mathfrak{D}\tilde{L}_1^S = 2\tilde{L}_2^S$ yields (ii).

Next we assume (ii). Since $(\tilde{L}_1^S + i\tilde{L}_2^S)\rho_\theta = 0$,

$$2\tilde{L}_2^S\rho_\theta = 2i\tilde{L}_1^S\rho_\theta, \qquad -2\tilde{L}_1^S\rho_\theta = 2i\tilde{L}_2^S\rho_\theta.$$

Comparing these equations to (17), we have (i). ∎

Since the condition (ii) in Theorem 3 is similar to the definition of the coherent states in quantum theory [9], we shall make the following definition.

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

Thus the $\mathfrak{D}$-invariancy is equivalent to the coherency of the model. The next fact, a straightforward consequence of Theorem 3, characterizes a global structure.

Corollary 1 *Consider the pure state model of the form* $\rho_\theta = U_\theta\rho_0U_\theta^*$ *where* $\{U_\theta\}$ *forms a projective unitary group. This model is coherent iff* $\mathcal{T}^S(\rho_0)/\mathcal{K}_{sa}(\rho_0)$ *is* $\mathfrak{D}$*-invariant, i.e., the model has a* ρ_0*–symplectic basis. Indeed, if* $\{\tilde{L}_j^S\}_{j=1,2}$ *is a* ρ_0*–symplectic basis, then* $\{U_\theta\tilde{L}_j^SU_\theta^*\}_{j=1,2}$ *becomes a* ρ_θ*–symplectic basis.*

Now we investigate the RLD–bound for coherent models. From Theorem 2, it is sufficient to consider a two parameter coherent model. In this case, the RLD–bound can be explicitly obtained by substituting (18) into (14) as

$$C^R = C^S + \frac{\sqrt{\det G}}{\det J^S} \left| \mathrm{Tr}\, \rho_\theta [L_1^S, L_2^S] \right|. \tag{21}$$

It is also shown that the above CR–bound is most informative, i.e., we can construct explicitly the generalized measurement that attains this bound [8]. Thus, the coherent model has a nice property from an estimation theoretical viewpoint.

4 Examples

In this section, we give two examples of coherent model. The first one is the family of canonical coherent states $\rho_z = |z\rangle\langle z|$ in a one dimensional harmonic oscillator with frequency ω, where $z = (\omega q + ip)/2\hbar \in \mathbf{C}$, see [10][11][12]. This can be regarded as a 2-parameter pure state model which has real parameters q and p. It is shown that the representative elements of SLD are

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

and

$$\mathfrak{D} L_q^S = 2\omega L_p^S, \qquad \mathfrak{D} L_p^S = -\frac{2}{\omega} L_q^S.$$

Letting

$$\tilde{L}_q^S = \frac{\hbar}{2} L_q^S = \omega(Q-q), \qquad \tilde{L}_p^S = \frac{\hbar\omega}{2} L_p^S = P-p,$$

we have

$$\mathfrak{D}\tilde{L}_q^S = 2\tilde{L}_p^S, \qquad \mathfrak{D}\tilde{L}_p^S = -2\tilde{L}_q^S.$$

This indicates that $\{\tilde{L}_q^S,\ \tilde{L}_p^S\}$ forms a ρ_z–symplectic basis. Therefore, from Theorem 3,

$$(\tilde{L}_q^S + i\tilde{L}_p^S)|z\rangle = [\omega(Q-q) + i(P-p)]|z\rangle = 0,$$

which is nothing but the definition of canonical coherent states. Furthermore, from Theorem 1, we obtain

$$(J^R)^{-1} = \begin{bmatrix} \sigma_P^2 & i\hbar/2 \\ -i\hbar/2 & \sigma_Q^2 \end{bmatrix},$$

where $\sigma_P^2 = \hbar\omega/2$, $\sigma_Q^2 = \hbar/2\omega$, and the corresponding RLD–bound

$$g_P V_P[M] + g_Q V_Q[M] \geq g_P \sigma_P^2 + g_Q \sigma_Q^2 + \hbar\sqrt{g_P g_Q}$$

is identical to the pure state limit of the most informative CR bound obtained by Yuen and Lax [4] [2, p. 281].

Another example is the family of spin coherent states [13][14]. Let (θ, φ) be the polar coordinates where the north pole is $\theta = 0$ and x–axis corresponds to $\varphi = 0$. The spin coherent state $|\theta, \varphi\rangle$ is defined as

$$|\theta, \varphi\rangle = R[\theta, \varphi]|j\rangle = \exp\left[i\theta(J_x \sin\varphi - J_y \cos\varphi)\right]|j\rangle,$$

where $|j\rangle$ is the highest occupied state in the spin j system. It is shown that the SLD at the north pole in the direction of $\varphi = 0$ and $\varphi = \pi/2$ are, respectively, $2J_x$, $2J_y$ and the operation of $\mathfrak{D}$ becomes $\mathfrak{D}J_x = 2J_y$, $\mathfrak{D}J_y = -2J_x$. Therefore, $\tilde{L}_1^S = J_x$ and $\tilde{L}_2^S = J_y$ form a $|j\rangle\langle j|$–symplectic basis and from Theorem 3

$$\left(\tilde{L}_1^S + i\tilde{L}_2^S\right)|j\rangle = J_+|j\rangle = 0,$$

where $J_+ = J_x + iJ_y$ is the spin creation operator. This is nothing but the definition of the terminal state $|j\rangle$. From this fact, we can immediately conclude that the model which comprises the totality of the spin coherent states

$$\rho_{\theta,\varphi} = |\theta, \varphi\rangle\langle\theta, \varphi| = R[\theta, \varphi]|j\rangle\langle j|R[\theta, \varphi]^{-1}$$

has $\mathfrak{D}$-invariant SLD tangent space at every point on the sphere. Indeed, since $R[\theta, \varphi]$ form a compact Lie group, Corollary 1 asserts that

$$\left\{R[\theta, \varphi]\tilde{L}_1^S R[\theta, \varphi]^{-1},\ R[\theta, \varphi]\tilde{L}_2^S R[\theta, \varphi]^{-1}\right\}$$

form a $|\theta, \varphi\rangle\langle\theta, \varphi|$–symplectic basis. In particular, a 2-parameter spin 1/2 model has $\mathfrak{D}$-invariant SLD tangent space and

$$\left(J^R\right)^{-1} = \frac{1}{\sin^2\theta}\begin{bmatrix} \sin^2\theta & -i\sin\theta \\ i\sin\theta & 1 \end{bmatrix}.$$

The corresponding CR bound is

$$g_\theta V_\theta[M] + g_\varphi V_\varphi[M] \geq g_\theta + \frac{g_\varphi}{\sin^2\theta} + \frac{2}{\sin\theta}\sqrt{g_\theta g_\varphi}.$$

This bound is identical to the pure state limit of the most informative CR bound obtained by Nagaoka [7].

5 Conclusions

A statistical multi–parameter estimation theory for the pure state models was presented. We first considered the possibility of the estimation theory based on the right logarithmic derivatives. We next investigated $\mathfrak{D}$–invariancy of the SLD–tangent space, which lead us to the notion of coherent models, and derived explicitly the RLD–bound. Some examples were also given. The construction of the general quantum multi-parameter estimation theory is left to future study, as is the strictly positive model case.

Acknowledgment

I am mostly indebted to Professor H. Nagaoka.

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

[5] D. Petz, "Entropy in Quantum Probability I," Quantum Probability and Related Topics Vol. VII, 275-297 (World Scientific, 1992).

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

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

[8] A. Fujiwara, "Linear random measurements of two non-commuting observables," METR **94-10**, Univ. Tokyo (1994).

[9] J. R. Klauder and B. Skagerstam, *Coherent States – Applications in Physics and Mathematical Physics* (World Scientific, Singapore, 1985).

[10] R. J. Glauber, "The Quantum Theory of Optical Coherence," Phys. Rev., **130**, 2529-2539 (1963).

[11] R. J. Glauber, "Coherent and Incoherent States of the Radiation Field," Phys. Rev., **131**, 2766-2788 (1963).

[12] J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968).

[13] J. M. Radcliffe, "Some Properties of Coherent Spin States," J. Phys. A : Gen. Phys. **4**, 313-323 (1971).

[14] F. T. Arecchi, E. Courtens, R. Glimore, and H. Thomas, "Atomic Coherent States in Quantum Optics," Phys. Rev. **6**, 2211-2237 (1972).