An estimation theoretical characterization of coherent states

Akio Fujiwara
Department of Mathematics, Osaka University
1-16 Machikane-yama, Toyonaka, Osaka 560-0043, Japan
email: fujiwara@math.wani.osaka-u.ac.jp

Hiroshi Nagaoka
Graduate School of Information Systems
The University of Electro-Communications
1-5-1 Chofugaoka, Chofu, Tokyo 182-8585, Japan
email: nagaoka@is.uec.ac.jp

Abstract

We introduce a class of quantum pure state models called the coherent models. A coherent model is an even dimensional manifold of pure states whose tangent space is characterized by a symplectic structure. In a rigorous framework of noncommutative statistics, it is shown that a coherent model inherits and expands the original spirit of the minimum uncertainty property of coherent states.

PACS numbers: 03.65.Bz, 42.25.Kb, 89.70.+c

I Introduction

Quantum estimation theory, originated in optical communications, offers a rigorous approach toward the optimization of detection processes in quantum communication systems [1] [2]. It aims to find, for a given smooth parametric family of density operators (a model) $\mathcal{P} = \{\rho_\theta\,;\,\theta = (\theta^1, \ldots, \theta^n) \in \Theta \subset \mathbf{R}^n\}$, the optimum measurement (positive operator-valued measure) $M = \{M(B)\,;\, B \text{ is a Borel set in } \mathbf{R}^n\}$ for the parameter θ under the unbiasedness condition: For all $\theta \in \Theta$,

$$\int \hat{\theta}^j \,\mathrm{Tr}\, \rho_\theta M(d\hat{\theta}) = \theta^j, \qquad j = 1, \ldots, n.$$

Here Tr denotes the operator trace. Normally a more tractable (weaker) condition is adopted, called the local unbiasedness condition: A measurement M is called locally unbiased at a given point θ if M satisfies at θ the above equality and its formal differentiation

$$\frac{\partial}{\partial \theta^i} \int \hat{\theta}^j \,\mathrm{Tr}\, \rho_\theta M(d\hat{\theta}) = \delta_i^j, \qquad i, j = 1, \ldots, n.$$

It is well-known that when $n = 1$, the quantum Cramér-Rao inequality with respect to the symmetric logarithmic derivative (SLD) offers the achievable lower bound (i.e., the bound attained by a certain measurement) of the variance of estimation. This is also regarded as a rigorous modification of the uncertainty relation. When $n \geq 2$, on the other hand, a matrix version of the SLD Cramér-Rao inequality itself does not always have an absolute significance because the lower bound cannot be attained in general unless the model has commutative SLDs. We therefore often deal with the minimization problem of the scalar quantity $\mathrm{tr}\, G V_\theta[M]$ with respect to M, where tr denotes the matrix trace on the parameter space Θ, G a real symmetric positive matrix representing the weight, and $V_\theta[M]$ the covariance matrix at θ with respect to a locally unbiased measurement M whose (i, j) entry is

$$(V_\theta[M])^{ij} = \int (\hat{\theta}^i - \theta^i)(\hat{\theta}^j - \theta^j) \mathrm{Tr}\, \rho_\theta M(d\hat{\theta}).$$

If there is a number C such that $\mathrm{tr}\, G V_\theta[M] \geq C$ holds for all M, C is called a Cramér-Rao type bound or simply a CR bound. The CR bound C may depend on both G and θ. The problem of finding the achievable CR bound is in general a hard one and has been solved only in a few special models such as the quantum gaussian model [3] [2] and the 2-dimensional spin 1/2 model [4] [5].

Holevo showed that if a model having the right logarithmic derivative (RLD) exhibits a certain "nice" property of a tangent space, the CR bound based on the RLD is expressed only in terms of the SLDs [2, p.280]. Moreover it was shown that this gives the achievable CR bound for the gaussian model of quantum oscillators. Motivated by these facts and that the SLD Fisher information is well-defined also for pure state models [6], we will introduce a class of pure state models called the coherent models [7] each having a "nice" tangent space, and will explore their parameter estimation theory.

The construction of the paper is as follows. In Section II, we explore some basic characteristics inherent in pure state models which are closely related with Holevo's commutation operator. In Section III, a special class of pure state models, called the coherent models, is introduced of which the SLD tangent space forms an invariant subspace with respect to the commutation operator. In Section IV, we derive a CR bound, called the generalized RLD bound, for a model that has an invariant SLD tangent space with respect to the commutation operator. Here the model is not assumed to be pure. In Section V, we show that for a coherent model, there exists a random measurement which attains the generalized RLD bound. In Section VI, the above results are demonstrated in two simple coherent models: a canonical squeezed state model and a spin coherent state model. The final Section VII gives conclusions.

II Commutation operator

In the study of noncommutative statistics, Holevo introduced useful mathematical tools called the square summable operators and the commutation operators associated with quantum states. We here give a brief summary: for details, consult [2]. Let $\mathcal{H}$ be a separable complex Hilbert space which corresponds to a physical system of interest, and let ρ be a fixed density operator. We define a real Hilbert space $\mathcal{L}_h^2(\rho)$ associated with ρ by the completion of $\mathcal{B}_h(\mathcal{H})$, the set of bounded self-adjoint operators, with respect to the pre-inner product $\langle X, Y\rangle_\rho = \operatorname{Re}\operatorname{Tr}\rho XY$. Letting $\rho = \sum_j s_j |\psi_j\rangle\langle\psi_j|$ be the spectral representation, an element $X \in \mathcal{L}_h^2(\rho)$ can be regarded as an equivalence class of such self-adjoint operators (called square summable operators) satisfying $\sum_j s_j \|X\psi_j\|^2 < \infty$ (so that $\psi_j \in \operatorname{Dom}(X)$ if $s_j \neq 0$) under the identification $X_1 \sim X_2$ if $X_1\psi_j = X_2\psi_j$ for $s_j \neq 0$. The space $\mathcal{L}_h^2(\rho)$ thus provides a convenient tool to cope with unbounded observables. Let $\mathcal{L}^2(\rho)$ be the complexification of $\mathcal{L}_h^2(\rho)$. Note that $\mathcal{L}^2(\rho)$ is also regarded as the completion of $\mathcal{B}(\mathcal{H})$, the set of bounded operators, with respect to the pre-inner product

$$\langle X, Y\rangle_\rho = \frac{1}{2}\operatorname{Tr}\rho\,(YX^* + X^*Y).$$

Thus $\mathcal{L}^2(\rho)$ is regarded as a complex Hilbert space with the inner product $\langle\cdot,\cdot\rangle_\rho$. We further introduce two sesquilinear forms on $\mathcal{B}(\mathcal{H})$ by

$$(X, Y)_\rho = \operatorname{Tr}\rho YX^*, \qquad [X, Y]_\rho = \frac{1}{2i}\operatorname{Tr}\rho\,(YX^* - X^*Y),$$

and extend them to $\mathcal{L}^2(\rho)$ by continuity.

The *commutation operator* $\mathcal{D}_\rho : \mathcal{L}^2(\rho) \to \mathcal{L}^2(\rho)$ with respect to ρ is defined by $[X, Y]_\rho = \langle X, \mathcal{D}_\rho Y\rangle_\rho$, which is formally represented by the operator equation $\rho(\mathcal{D}_\rho X) + (\mathcal{D}_\rho X)\rho = \frac{1}{i}(\rho X - X\rho)$. (To be precise, this definition is different from Holevo's original definition by a factor of 2.) The operator $\mathcal{D}_\rho$ is a complex-linear bounded skew-adjoint operator. Moreover, since the forms $[\cdot,\cdot]_\rho$ and $\langle\cdot,\cdot\rangle_\rho$ are real on the real subspace $\mathcal{L}_h^2(\rho)$, this subspace is invariant under the operation of $\mathcal{D}_\rho$. Thus $\mathcal{D}_\rho$ can also be regarded as a real-linear bounded skew-adjoint operator when restricted to $\mathcal{L}_h^2(\rho)$ as $\mathcal{D}_\rho : \mathcal{L}_h^2(\rho) \to \mathcal{L}_h^2(\rho)$.

Our main concern lies in the case where ρ is pure. In this case the above setting is considerably simplified as follows: Let $\rho = |\psi\rangle\langle\psi|$. Then for $X, Y \in \mathcal{L}^2(\rho)$,

$$\begin{aligned}
\langle X, Y\rangle_\rho &= \frac{1}{2}\{\langle Y^*\psi|X^*\psi\rangle + \langle X\psi|Y\psi\rangle\}, \\
[X, Y]_\rho &= \frac{1}{2i}\{\langle Y^*\psi|X^*\psi\rangle - \langle X\psi|Y\psi\rangle\}, \\
(X, Y)_\rho &= \langle Y^*\psi|X^*\psi\rangle.
\end{aligned}$$

Here $X\psi$, for example, stands for the vector $X_1\psi$ where X_1 is an arbitrary representative of X. (It is independent of the choice of a representative.) In particular, if $X, Y \in \mathcal{L}_h^2(\rho)$ we have

$$\langle X, Y\rangle_\rho = \mathrm{Re}\,\langle Y\psi|X\psi\rangle = \mathrm{Re}\,\langle X\psi|Y\psi\rangle, \tag{1}$$
$$[X, Y]_\rho = \mathrm{Im}\,\langle Y\psi|X\psi\rangle = -\mathrm{Im}\,\langle X\psi|Y\psi\rangle, \tag{2}$$
$$(X, Y)_\rho = \langle Y\psi|X\psi\rangle = \overline{\langle X\psi|Y\psi\rangle}. \tag{3}$$

It should be noted that operators X and Y (whether bounded or not) are identified with each other in $\mathcal{L}^2(\rho)$ iff $X\psi = Y\psi$ and $X^*\psi = Y^*\psi$. In particular, self-adjoint operators X and Y are identified in $\mathcal{L}_h^2(\rho)$ iff $X\psi = Y\psi$.

Lemma 1. *Let $\rho = |\psi\rangle\langle\psi|$. Then for all $X \in \mathcal{L}_h^2(\rho)$,*

$$(\mathcal{D}_\rho X)\psi = i(X - \langle\psi|X\psi\rangle I)\psi,$$

where I denotes the identity in $\mathcal{L}_h^2(\rho)$.

Proof For $X \in \mathcal{L}_h^2(\rho)$, let Z be the element in $\mathcal{L}_h^2(\rho)$ having a representative $Z_1 = i(|X\psi\rangle\langle\psi| - |\psi\rangle\langle X\psi|)$. Then $Z\psi = i(X - \langle\psi|X\psi\rangle I)\psi$. On the other hand, for $Y \in \mathcal{L}_h^2(\rho)$, we have

$$\langle Y\psi|Z\psi\rangle = i\{\langle Y\psi|X\psi\rangle - \langle\psi|X\psi\rangle\langle\psi|Y\psi\rangle\},$$

and hence $\langle Y, Z\rangle_\rho = [Y, X]_\rho$ because of (1) and (2). Thus $Z = \mathcal{D}_\rho X$, which completes the proof. □

Note that Lemma 1 does not imply $\mathcal{D}_\rho X = i(X - \langle\psi|X\psi\rangle I)$, since the right hand side is not a self-adjoint element in $\mathcal{L}^2(\rho)$ unless it equals 0.

Let us introduce a linear subspace

$$\mathcal{T}_h(\rho) = \{X \in \mathcal{L}_h^2(\rho)\ ;\ \langle I, X\rangle_\rho = 0\}$$

of $\mathcal{L}_h^2(\rho)$. Here ρ is not necessarily pure. This subspace is itself a real Hilbert space with the inner product $\langle\cdot,\cdot\rangle_\rho$. Now consider again the special case that ρ is pure: $\rho = |\psi\rangle\langle\psi|$. Then from Lemma 1, we obtain the important relation:

$$(\mathcal{D}_\rho X)\psi = (iX)\psi, \quad X \in \mathcal{T}_h(\rho). \tag{4}$$

This equation, combined with (1), implies that $\mathcal{D}_\rho$ is a unitary transformation on $(\mathcal{T}_h(\rho), \langle\cdot,\cdot\rangle_\rho)$. In particular, $\mathcal{D}_\rho$ is nondegenerate on $\mathcal{T}_h(\rho)$, and so is the skew-symmetric bilinear form $[\cdot,\cdot]_\rho$. In other words, the real linear space $\mathcal{T}_h(\rho)$ is regarded as a symplectic space [8] with the symplectic form $[\cdot,\cdot]_\rho$. We also note that $\mathcal{D}_\rho^2 = -\mathcal{I}$ on $\mathcal{T}_h(\rho)$ ($\mathcal{I}$ denotes the identity operator acting on $\mathcal{T}_h(\rho)$), since $\mathcal{D}_\rho$ is unitary and skew-adjoint. Indeed, equation (4) immediately leads to $(\mathcal{D}_\rho^2 X)\psi = -X\psi$ and hence $\mathcal{D}_\rho^2 X = -X$ for all $X \in \mathcal{T}_h(\rho)$, whereas $\mathcal{D}_\rho X \neq iX$ as mentioned earlier. In other words, $\mathcal{D}_\rho$ is an almost complex structure on $\mathcal{T}_h(\rho)$.

III Coherent model

Let $\mathcal{P} = \{\rho_\theta\ ;\ \theta = (\theta^1, \ldots, \theta^n) \in \Theta\}$ be an n-dimensional model, where ρ_θ are not necessarily pure for the present, and Θ is an open subset of $\mathbf{R}^n$. We assume the following regularity conditions:

(a) The parametrization $\theta \mapsto \rho_\theta$ is assumed to be appropriately smooth and nondegenerate so that the derivatives $\{\partial\rho_\theta/\partial\theta^j\}_{j=1}^n$ exist in trace-class and form a linearly independent set at each point θ.

(b) There exists a constant c such that

$$\left|\frac{\partial}{\partial\theta^j}\mathrm{Tr}\,\rho_\theta X\right|^2 \leq c\langle X, X\rangle_{\rho_\theta}$$

for all $X \in \mathcal{B}(\mathcal{H})$ and j.

From the condition (b), the linear functionals $X \mapsto (\partial/\partial\theta^j)\mathrm{Tr}\,\rho_\theta X$ can be extended to continuous linear functionals on $\mathcal{L}^2(\rho_\theta)$.

Given a model that satisfies (a) and (b), the *symmetric logarithmic derivative* (SLD) $L^S_{\theta,j}$ in the jth direction is defined by the requirement that

$$\frac{\partial}{\partial\theta^j}\mathrm{Tr}\,\rho_\theta X = \langle L^S_{\theta,j}, X\rangle_{\rho_\theta}, \qquad L^S_{\theta,j} \in \mathcal{L}^2(\rho_\theta)$$

for all $X \in \mathcal{L}^2(\rho_\theta)$. It is easily verified that $L^S_{\theta,j} \in \mathcal{L}^2_h(\rho_\theta)$; so the definition is formally written as $\partial\rho_\theta/\partial\theta^j = \frac{1}{2}(L^S_{\theta,j}\rho_\theta + \rho_\theta L^S_{\theta,j})$. The SLDs belong to $\mathcal{T}_h(\rho_\theta)$ since $\langle I, L^S_{\theta,j}\rangle_{\rho_\theta} = (\partial/\partial\theta^j)\mathrm{Tr}\,\rho_\theta = 0$, and the SLD Fisher information matrix defined by $J^S_\theta = \left[\langle L^S_{\theta,j}, L^S_{\theta,k}\rangle_{\rho_\theta}\right]$ gives a Cramér-Rao inequality $V_\theta[M] \geq (J^S_\theta)^{-1}$, where M is an arbitrary locally unbiased measurement for the parameter θ, see [2, p. 276].

In the rest of this section, we restrict ourselves to pure state models. Some remarks are in order. First, by differentiating the identity $\rho_\theta^2 = \rho_\theta$, we see that the element in $\mathcal{L}^2_h(\rho_\theta)$ having a representative $2\partial\rho_\theta/\partial\theta^j$ gives the SLD $L^S_{\theta,j}$. Thus for a pure state model, the condition (a) implies (b). Second, associated with a pure state model $\{\rho_\theta;\ \theta \in \Theta\}$ is, at least locally, a smooth family $\{\psi_\theta;\ \theta \in \Theta\}$ of normalized vectors in $\mathcal{H}$ such that $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. In what follows, we shall frequently use this representation.

A convenient way of finding SLDs for a pure state model ρ_θ is as follows: Let $L^A_{\theta,j}$ be the anti-symmetric logarithmic derivative (ALD) satisfying

$$\frac{\partial}{\partial\theta^j}\operatorname{Tr}\rho_\theta X = [L^A_{\theta,j}, X]_{\rho_\theta}, \qquad L^A_{\theta,j} \in \mathcal{T}_h(\rho_\theta)$$

for all $X \in \mathcal{L}^2(\rho_\theta)$, or formally $\partial\rho_\theta/\partial\theta^j = (L^A_{\theta,j}\rho_\theta - \rho_\theta L^A_{\theta,j})/2i$. (This definition is different from [6] by a factor of i.) Then the SLD is given by $L^S_{\theta,j} = -\mathcal{D}_\theta L^A_{\theta,j}$ where $\mathcal{D}_\theta = \mathcal{D}_{\rho_\theta}$, since $\langle L^S_{\theta,j}, X\rangle_{\rho_\theta} = [L^A_{\theta,j}, X]_{\rho_\theta}$. Note that since $\mathcal{D}^2_\theta = -\mathcal{I}$ on $\mathcal{T}_h(\rho_\theta)$, then $L^A_{\theta,j} = \mathcal{D}_\theta L^S_{\theta,j}$, which assures the existence and the uniqueness of the ALD for a pure state model. The advantage of the use of the ALD is this: Every pure state model can be expressed in the form $\rho_\theta = U_\theta\rho_0 U^*_\theta$ where $\{U_\theta\}_\theta$ is a smooth family of unitary operators (which do not necessarily form a group representation), so that the ALD is explicitly given by

$$L^A_{\theta,j} = 2i(A_{\theta,j} - \langle I, A_{\theta,j}\rangle_{\rho_\theta}),$$

where $A_{\theta,j}$ is the skew-adjoint element in $\mathcal{L}^2(\rho_\theta)$ having a representative $(\partial U_\theta/\partial\theta^j)U^*_\theta$, the local generator of U_θ. For a group covariant pure state model, the generator of the group is usually obvious.

Let $\mathcal{T}^S_\theta(\mathcal{P})$ be the real-linear subspace of $\mathcal{T}_h(\rho_\theta)$ spanned by the SLDs $\{L^S_{\theta,j}\}_j$. Since the tangent vectors of the manifold $\mathcal{P}$ at the point ρ_θ are faithfully represented by the elements of $\mathcal{T}^S_\theta(\mathcal{P})$ via the correspondence $(\partial/\partial\theta^j)_\theta \mapsto L^S_{\theta,j}$, we call $\mathcal{T}^S_\theta(\mathcal{P})$ the *SLD tangent space* of the model $\mathcal{P}$ at θ. A pure state model $\mathcal{P} = \{\rho_\theta\,;\,\theta\in\Theta\}$ is called *locally coherent* at θ if $\mathcal{T}^S_\theta(\mathcal{P})$ is $\mathcal{D}_\theta$-invariant. The model is called *coherent* if it is locally coherent for all $\theta\in\Theta$.

When the Hilbert space $\mathcal{H}$ is finite-dimensional, the totality of pure states forms a complex projective space and is an example of coherent model. The Riemannian metric on the model induced by the SLD Fisher information matrix J^S_θ is identical to the Fubini-Study metric up to a constant factor [6] and hence is a Kähler metric. The associated fundamental 2-form [9] in this case is nothing but the symplectic structure $[\cdot,\cdot]_\rho$.

Theorem 2. *Consider a pure state model of the form $\rho_\theta = U_{g(\theta)}\rho_0 U^*_{g(\theta)}$ where $\{U_g\,;\,g\in\mathcal{G}\}$ is a projective unitary representation of a Lie group $\mathcal{G}$ and $g(\cdot):\theta\mapsto g(\theta)$ is the parametrization of the elements of $\mathcal{G}$ by a local coordinate system satisfying $g(0)=e$ (: the unit element). This model is coherent iff it is locally coherent at ρ_0.*

Proof We only need to prove the if part. Let $\Lambda_\theta:\mathcal{G}\to\mathcal{G}$ be the left translation by $g(\theta)^{-1}$ which maps $h\mapsto g(\theta)^{-1}h$. Then its differential $(d\Lambda_\theta)_{g(\theta)}:T_{g(\theta)}(\mathcal{G})\to T_e(\mathcal{G})$ is represented by a nonsingular real matrix $a^k_j(\theta)$ such that $(d\Lambda_\theta)_{g(\theta)}\left([\partial g(\theta)/\partial\theta^j]_\theta\right) = \sum_k a^k_j(\theta)\,[\partial g(\theta)/\partial\theta^k]_{\theta=0}$. Now since $\rho_{\theta+\Delta\theta} = U_{g(\theta)}\,\rho_{\Delta\theta'}\,U^*_{g(\theta)}$, where $\Lambda_\theta(g(\theta+\Delta\theta)) = g(\Delta\theta')$, we find that $\partial\rho_\theta/\partial\theta^j = \sum_k a^k_j(\theta)\,U_\theta[\partial\rho_\theta/\partial\theta^k]_{\theta=0}U^*_\theta$. This implies that the SLDs at θ are given by $L^S_{\theta,j} = \sum_i a^k_j(\theta)\,U_\theta L^S_{0,k}U^*_\theta$. As a consequence

$$L^S_{\theta,j}\psi_\theta = \sum_k a^k_j(\theta)\,U_\theta L^S_{0,k}\psi_0. \tag{5}$$

Here we have set as $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$ with $\psi_\theta = U_\theta\psi_0$. Now suppose $\mathcal{P}$ is locally coherent at ρ_0. Then the vector $(\mathcal{D}_0 L^S_{0,k})\psi_0 = iL^S_{0,k}\psi_0$ (see (4)) belongs to the real linear span of

$\{L^S_{0,k'}\psi_0\}_{k'=1}^n$; hence the vector $(\mathcal{D}_\theta L^S_{\theta,j})\psi_\theta = iL^S_{\theta,j}\psi_\theta$ belongs to the real linear span of $\{L^S_{\theta,j'}\psi_\theta\}_{j'=1}^n$ because of (5) and the nonsingularity of the matrix $a_j^k(\theta)$. This implies that $\mathcal{P}$ is locally coherent at every point θ. $\square$

It is clear from the definition that if $\mathcal{P}$ is locally coherent at θ, then $\mathcal{T}^S_\theta(\mathcal{P})$ forms a symplectic space with the symplectic form being the restriction of $[\cdot,\cdot]_{\rho_\theta}$. In particular, the dimensionality of $\mathcal{T}^S_\theta(\mathcal{P})$ is necessarily even (say $n = 2m$), and it has a symplectic basis $\{\tilde{L}^S_{\theta,j}\}_{j=1}^{2m}$ satisfying

$$[\tilde{L}^S_{\theta,j}, \tilde{L}^S_{\theta,k}]_{\rho_\theta} = \begin{cases} -1, & \text{if } j \text{ is odd and } k = j+1 \\ 1, & \text{if } j \text{ is even and } k = j-1 \\ 0, & \text{otherwise.} \end{cases}$$

Furthermore, since $\mathcal{D}_\theta$ is unitary on $\mathcal{T}^S_\theta(\mathcal{P})$ with respect to the inner product $\langle\cdot,\cdot\rangle_{\rho_\theta}$, we can take $\{\tilde{L}^S_{\theta,j}\}$ to be orthonormal. Such a basis, which we shall call a *normalized ρ_θ-symplectic basis*, satisfies

$$\mathcal{D}_\theta \begin{bmatrix} \tilde{L}^S_{\theta,1} \\ \tilde{L}^S_{\theta,2} \\ \tilde{L}^S_{\theta,3} \\ \tilde{L}^S_{\theta,4} \\ \vdots \\ \tilde{L}^S_{\theta,2m-1} \\ \tilde{L}^S_{\theta,2m} \end{bmatrix} = \begin{bmatrix} 0 & 1 & & & & & \\ -1 & 0 & & & & & \\ & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{bmatrix} \begin{bmatrix} \tilde{L}^S_{\theta,1} \\ \tilde{L}^S_{\theta,2} \\ \tilde{L}^S_{\theta,3} \\ \tilde{L}^S_{\theta,4} \\ \vdots \\ \tilde{L}^S_{\theta,2m-1} \\ \tilde{L}^S_{\theta,2m} \end{bmatrix}. \tag{6}$$

Thus the SLD tangent space of a coherent model is decomposed into 2-dimensional $\mathcal{D}_\theta$-invariant subspaces. This suggests the importance of studying 2-dimensional coherent models.

Now, let us characterize a 2-dimensional coherent model.

Theorem 3. *For a 2-dimensional pure state model $\mathcal{P} = \{\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta| \ ; \ \theta \in \Theta\}$, the following three conditions are equivalent.*

(α) *$\mathcal{P}$ is locally coherent at θ.*

(β) *$L^S_{\theta,1}\psi_\theta$ and $L^S_{\theta,2}\psi_\theta$ are linearly dependent.*

(γ) *$L^A_{\theta,1}\psi_\theta$ and $L^A_{\theta,2}\psi_\theta$ are linearly dependent.*

Before going to the proof, we should remark that the condition (β) does not conflict with the fact that $L^S_{\theta,1}$ and $L^S_{\theta,2}$ are linearly independent due to the nondegeneracy of the parametrization $\theta \mapsto \rho_\theta$. Indeed, the linear independence of $\{L^S_{\theta,1}, L^S_{\theta,2}\}$ is concerned with the real linear structure of $\mathcal{L}^2_h(\rho_\theta)$ and is equivalent to the real linear independence of $\{L^S_{\theta,1}\psi_\theta, L^S_{\theta,2}\psi_\theta\}$. On the other hand, the condition (β) asserts the complex linear dependence of the same vectors.

Proof The proof relies essentially on (4). We only need to show that (α)$\Leftrightarrow$(β), since (β)$\Leftrightarrow$(γ) is obvious from the identity $L^S_{\theta,j}\psi_\theta = -(\mathcal{D}_\theta L^A_{\theta,j})\psi_\theta = -iL^A_{\theta,j}\psi_\theta$. Let $\varphi_j := L^S_{\theta,j}\psi_\theta$, and assume ($\alpha$) first. Then there exist real numbers x, y such that $\mathcal{D}_\theta L^S_{\theta,1} = xL^S_{\theta,1} + yL^S_{\theta,2}$. This is equivalent to $i\varphi_1 = x\varphi_1 + y\varphi_2$ and leads to (β). Assume (β) in turn. Recalling the real linear independence of $\{\varphi_1, \varphi_2\}$, we see that there exist real numbers x, y satisfying $\varphi_2 = (x + iy)\varphi_1$ with $y \neq 0$. It then follows that $i\varphi_1 = -(x/y)\varphi_1 + (1/y)\varphi_2$ and $\mathcal{D}_\theta L^S_{\theta,1} = -(x/y)L^S_{\theta,1} + (1/y)L^S_{\theta,2}$. Similarly $\mathcal{D}_\theta L^S_{\theta,2}$ is shown to be a real linear combination of $\{L^S_{\theta,1}, L^S_{\theta,2}\}$ and thus (α) is verified. □

The following corollary, whose proof is now straightforward from Theorem 3 and (4), offers a mostly useful method to treat group covariant coherent models as exemplifed in Section VI. Moreover the equation (7) in the corollary reveals a close connection with the conventional definition of coherent states [10]. Indeed, this fact gave a motive for the nomenclature of the coherent model.

Corollary 4. *Let $\mathcal{P} = \{\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|;\, \theta \in \Theta\}$ be a 2-dimensional pure state model and let $\mathcal{T}^A_\theta(\mathcal{P})$ be the real linear span of ALDs $\{L^A_{\theta,1}, L^A_{\theta,2}\}$ at θ. Then $\mathcal{P}$ is locally coherent at θ iff there exist nonzero elements X_1, X_2 in $\mathcal{T}^A_\theta(\mathcal{P})$ satisfying*

$$(X_1 + iX_2)\psi_\theta = 0. \tag{7}$$

Moreover, (7) is also necessary and sufficient for $\{cX_j\}_{j=1,2}$ to form a normalized ρ_θ-symplectic basis of $\mathcal{T}^S_\theta(\mathcal{P})$ $(= \mathcal{T}^A_\theta(\mathcal{P}))$ with a common normalizing constant c. Under the condition (7), the linear relations

$$L^A_{\theta,1} = c_{11}X_1 + c_{12}X_2, \quad L^A_{\theta,2} = c_{21}X_1 + c_{22}X_2$$

imply

$$L^S_{\theta,1} = c_{12}X_1 - c_{11}X_2, \quad L^S_{\theta,2} = c_{22}X_1 - c_{21}X_2.$$

IV Generalized RLD bound

Throughout this section we consider an n-dimensional model $\mathcal{P} = \{\rho_\theta\}$ of general (i.e., not necessarily pure) states satisfying the regularity conditions (a) and (b) presented in Section III.

Let $\mathcal{L}^2_+(\rho)$ denote the completion of $\mathcal{B}(\mathcal{H})$ with respect to the pre-inner product $(\cdot,\cdot)_\rho$. Since $(X,X)_\rho \leq 2\langle X,X\rangle_\rho$, then $\mathcal{L}^2(\rho) \subset \mathcal{L}^2_+(\rho)$. The *right logarithmic derivative* (RLD) $L^R_{\theta,j}$ in the jth direction of a model $\mathcal{P} = \{\rho_\theta\}$, when it exists, is defined by the requirement that

$$\frac{\partial}{\partial\theta^j}\operatorname{Tr}\rho_\theta X = (L^R_{\theta,j}, X)_{\rho_\theta}, \qquad L^R_{\theta,j} \in \mathcal{L}^2_+(\rho_\theta)$$

for all $X \in \mathcal{L}^2_+(\rho_\theta)$, or formally $\partial\rho_\theta/\partial\theta^j = (L^R_{\theta,j})^*\rho_\theta = \rho_\theta L^R_{\theta,j}$. The covariance matrix of an arbitrary locally unbiased estimator M is then bounded from below as

$$V_\theta[M] \geq (J^R_\theta)^{-1}, \tag{8}$$

where $J^R_\theta = \left[(L^R_{\theta,j}, L^R_{\theta,k})_{\rho_\theta}\right]$ is the RLD Fisher information matrix [2, p. 279]. When a real positive definite matrix G is specified as the weight for the estimation accuracy, the total deviation is bounded from below as

$$\operatorname{tr} G V_\theta[M] \geq C^R := \operatorname{tr} G \operatorname{Re} (J^R_\theta)^{-1} + \operatorname{tr} \operatorname{abs} G \operatorname{Im} (J^R_\theta)^{-1}, \tag{9}$$

where $\operatorname{tr}\operatorname{abs} A$ denotes the absolute sum of the eigenvalues of matrix A, see [2, p. 284]. The RLD thus gives a CR bound and plays a crucial role in optical communication theory [3] [2].

The RLD exists iff there is a constant c such that

$$\left| \frac{\partial}{\partial \theta^j} \operatorname{Tr} \rho_\theta X \right|^2 \leq c(X, X)_{\rho_\theta} \tag{10}$$

for all $X \in \mathcal{B}(\mathcal{H})$. Thus the RLD does not always exist for a model satisfying the weaker condition (b). In particular it never exists for a pure state model $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. To see this, let us fix a θ arbitrarily and take a vector $x \in \mathcal{H}$ such that $\langle \psi_\theta | x \rangle = 0$ and $\langle \partial\psi_\theta / \partial\theta^j | x \rangle \neq 0$. (This is indeed possible because ψ_θ and $\partial\psi_\theta/\partial\theta^j$ are linearly independent owing to $(\partial/\partial\theta^j)\langle\psi_\theta|\psi_\theta\rangle = 0$ and $(\partial/\partial\theta^j)|\psi_\theta\rangle\langle\psi_\theta| \neq 0$.) Then $X = |x\rangle\langle\psi_\theta|$ satisfies $(X, X)_{\rho_\theta} = 0$ and $(\partial/\partial\theta^j)\operatorname{Tr}\rho_\theta X \neq 0$. It is, however, important to notice that what is needed in estimation theory is not the RLD itself but the inverse of the RLD Fisher information matrix as indicated by (8) and (9).

In his book [2, p. 280], Holevo has shown that when a model satisfying the regularity conditions (a) and (10) has a $\mathcal{D}_\theta$-invariant SLD tangent space, the $(J^R_\theta)^{-1}$ is expressed only in terms of SLDs; so is the CR bound (9). We generalize this result to a wider class of models that satisfy only the weaker conditions (a) and (b).

Theorem 5. *Suppose we are given an n-dimensional model $\mathcal{P} = \{\rho_\theta\}$ having a $\mathcal{D}_\theta$-invariant SLD tangent space $\mathcal{T}^S_\theta(\mathcal{P})$. Then for all locally unbiased measurements M at θ,*

$$V_\theta[M] \geq \left(J^S_\theta\right)^{-1} + i\left(J^S_\theta\right)^{-1} D_\theta \left(J^S_\theta\right)^{-1},$$

where $D_\theta = \left[[L^S_{\theta,j}, L^S_{\theta,k}]_{\rho_\theta}\right]$.

Proof Let us introduce a family of inner products on $\mathcal{L}^2(\rho_\theta)$ having a parameter $\varepsilon \in (0, 1]$:

$$(X, Y)^{(\varepsilon)}_{\rho_\theta} = (1 - \varepsilon)(X, Y)_{\rho_\theta} + \varepsilon\langle X, Y\rangle_{\rho_\theta}.$$

Since

$$\varepsilon\langle X, X\rangle_{\rho_\theta} \leq (X, X)^{(\varepsilon)}_{\rho_\theta} \leq (2 - \varepsilon)\langle X, X\rangle_{\rho_\theta},$$

there exists, for each ε, a unique operator $L^{(\varepsilon)}_{\theta,j} \in \mathcal{L}^2(\rho_\theta)$ which satisfies

$$\frac{\partial}{\partial\theta^j}\operatorname{Tr}\rho_\theta X = (L^{(\varepsilon)}_{\theta,j}, X)^{(\varepsilon)}_{\rho_\theta}$$

for all $X \in \mathcal{L}^2(\rho_\theta)$. Then in a quite similar way to the derivation of the quantum Cramér-Rao inequality, we have

$$V_\theta[M] \geq \left(J_\theta^{(\varepsilon)}\right)^{-1}, \qquad J_\theta^{(\varepsilon)} = \left[(L_{\theta,j}^{(\varepsilon)}, L_{\theta,k}^{(\varepsilon)})_{\rho_\theta}^{(\varepsilon)}\right]. \tag{11}$$

Now observing the identity $(X,Y)_{\rho_\theta}^{(\varepsilon)} = \langle X, Y\rangle_{\rho_\theta} + i(1-\varepsilon)[X,Y]_{\rho_\theta}$, and using the definition of $\mathcal{D}_{\rho_\theta} = \mathcal{D}_\theta$, we see that for all $Y \in \mathcal{L}^2(\rho_\theta)$,

$$\frac{\partial}{\partial\theta^j}\operatorname{Tr}\rho_\theta Y = \langle L_{\theta,j}^S, Y\rangle_{\rho_\theta} = (L_{\theta,j}^{(\varepsilon)}, Y)_{\rho_\theta}^{(\varepsilon)} = \langle\{\mathcal{I} + i(1-\varepsilon)\mathcal{D}_\theta\}L_{\theta,j}^{(\varepsilon)}, Y\rangle_{\rho_\theta}.$$

Then $L_{\theta,j}^S = \{\mathcal{I} + i(1-\varepsilon)\mathcal{D}_\theta\}L_{\theta,j}^{(\varepsilon)}$, hence $(L_{\theta,j}^{(\varepsilon)}, L_{\theta,k}^{(\varepsilon)})_{\rho_\theta}^{(\varepsilon)} = \langle L_{\theta,j}^S, \{\mathcal{I} + i(1-\varepsilon)\mathcal{D}_\theta\}^{-1}L_{\theta,k}^S\rangle_{\rho_\theta}$. Let us introduce Dirac's notation $|L_{\theta,j}^S\rangle$ for the Hilbert space $(\mathcal{L}^2(\rho_\theta), \langle\cdot,\cdot\rangle_{\rho_\theta})$, and let $\Gamma_\theta := \left[|L_{\theta,1}^S\rangle, \cdots, |L_{\theta,n}^S\rangle\right]$. Then $\Gamma_\theta^*\Gamma_\theta = J_\theta^S$ and $\Gamma_\theta^*\mathcal{D}_\theta\Gamma_\theta = D_\theta$. And the matrix $J_\theta^{(\varepsilon)}$ can be written in the form $J_\theta^{(\varepsilon)} = \Gamma_\theta^*\{\mathcal{I} + i(1-\varepsilon)\mathcal{D}_\theta\}^{-1}\Gamma_\theta$. Thus from the assumption that $\mathcal{T}_\theta^S(\mathcal{P})$ is $\mathcal{D}_\theta$-invariant, the inverse of $J_\theta^{(\varepsilon)}$ is explicitly given by

$$\begin{aligned}\left(J_\theta^{(\varepsilon)}\right)^{-1} &= \left(J_\theta^S\right)^{-1}\Gamma_\theta^*\left\{\mathcal{I} + i(1-\varepsilon)\mathcal{D}_\theta\right\}\Gamma_\theta\left(J_\theta^S\right)^{-1} \\ &= \left(J_\theta^S\right)^{-1} + i(1-\varepsilon)\left(J_\theta^S\right)^{-1}D_\theta\left(J_\theta^S\right)^{-1}.\end{aligned} \tag{12}$$

Combining (11) and (12), and taking the limit $\varepsilon \downarrow 0$, we have the theorem. $\square$

Theorem 5 asserts that even for a model that does not have the RLDs, the $\lim_{\varepsilon\downarrow 0}(J_\theta^{(\varepsilon)})^{-1}$ indeed gives a generalization of $(J_\theta^R)^{-1}$ as long as the SLD tangent space is $\mathcal{D}_\theta$-invariant. Then by using Theorem 5 and an analogous argument to the derivation of (9), we obtain the CR bound

$$C^R = \operatorname{tr} G\left(J_\theta^S\right)^{-1} + \operatorname{tr}\operatorname{abs} G\left(J_\theta^S\right)^{-1}D_\theta\left(J_\theta^S\right)^{-1}, \tag{13}$$

for models each having a $\mathcal{D}_\theta$-invariant SLD tangent space $\mathcal{T}_\theta^S(\mathcal{P})$. This may be called a generalized RLD bound. We will show in the next section that this bound is achievable in a coherent model.

V Optimal estimation for 2-dimensional coherent models

We now proceed to a parameter estimation for a pure coherent model. In particular, taking into account the symplectic structure (6) of the SLD tangent space, we restrict ourselves to a 2-dimensional case. We note that as long as we are concerned with the achievable CR bound at each point on the model $\{\rho_\theta\}$, we can take the weight as $G = I$ without loss of generality. In fact, let M be a locally unbiased measurement for the parameter $\theta = (\theta^1, \theta^2)$ and let $p(\hat\theta^1, \hat\theta^2)d\hat\theta = \operatorname{Tr}\rho_\theta M(d\hat\theta)$ be the corresponding joint distribution. The coordinate transformation $\eta^i = \sum_j h^i{}_j\theta^j$, where $H = [h^i{}_j]$ is a real regular matrix, then induces another measurement $N(d\hat\eta)$ which corresponds to the joint distribution $q(\hat\eta^1, \hat\eta^2)d\hat\eta = p(\hat\theta^1, \hat\theta^2)d\hat\theta$ and is locally unbiased for the parameter $\eta = (\eta^1, \eta^2)$. In this case, $\mathrm{tr}\, V_\eta[N] = \mathrm{tr}\ ({}^tHH)V_\theta[M]$. Thus the parameter estimation for θ with the weight $G = {}^tHH$ is equivalent to that for η with the weight I.

Now suppose we are given a 2-dimensional coherent model $\mathcal{P} = \{\rho_\theta\ ;\ \theta = (\theta^1, \theta^2) \in \Theta\}$. Let $\{L^i\}$ be the dual basis of the SLDs: $L^i = \sum_j J^{ij} L^S_{\theta,j}$ with J^{ij} being the (i,j) entry of $(J^S_\theta)^{-1}$. Then

$$\left(J^S_\theta\right)^{-1} = \begin{bmatrix} \langle L^1, L^1\rangle_{\rho_\theta} & \langle L^1, L^2\rangle_{\rho_\theta} \\ \langle L^2, L^1\rangle_{\rho_\theta} & \langle L^2, L^2\rangle_{\rho_\theta} \end{bmatrix}$$

and

$$\left(J^S_\theta\right)^{-1} D_\theta \left(J^S_\theta\right)^{-1} = \begin{bmatrix} 0 & \left[L^1, L^2\right]_{\rho_\theta} \\ \left[L^2, L^1\right]_{\rho_\theta} & 0 \end{bmatrix}$$

Thus the generalized RLD bound (13) for $G = I$ can be rewritten in the form

$$C^R = \langle L^1, L^1\rangle_{\rho_\theta} + \langle L^2, L^2\rangle_{\rho_\theta} + 2\left|[L^1, L^2]_{\rho_\theta}\right|. \tag{14}$$

We will show that the bound C^R is achievable. In what follows, we fix a $\theta = (\theta^1, \theta^2)$ arbitrarily.

Let us consider a random measurement as follows. We first introduce a linear transformation $\phi : \mathcal{T}^S_\theta(\mathcal{P}) \longrightarrow \mathcal{T}^S_\theta(\mathcal{P})$ by

$$\phi(X) = \langle L^1, X\rangle_{\rho_\theta} L^1 + \langle L^2, X\rangle_{\rho_\theta} L^2.$$

Since ϕ is symmetric and positive definite, it has positive eigenvalues λ_1, λ_2, and mutually orthogonal unit eigenvectors A_1, A_2 satisfying $\phi(A_\nu) = \lambda_\nu A_\nu$, $\nu = 1, 2$. We next take positive numbers p_1, p_2 satisfying $p_1 + p_2 = 1$. Now letting

$$\int \xi E_\nu(d\xi), \qquad \nu = 1, 2$$

be the spectral decompositions of arbitrarily fixed representatives of A_ν, we define a generalized measurement

$$M(\nu, d\xi) = p_\nu E_\nu(d\xi).$$

This has the following physical interpretation: Select one of the two "observables" A_1, A_2 according to the probability p_1, p_2, respectively, and measure it in a usual sense.

Now suppose we have selected A_ν and have obtained an outcome ξ. We identify this result with a pair of real quantities

$$\hat\theta^i(\nu, \xi) = \theta^i + \frac{\xi}{p_\nu}\langle L^i, A_\nu\rangle_{\rho_\theta}, \qquad i = 1, 2.$$

The pair $\{\hat\theta^i(\nu, \xi)\}_{i=1,2}$ satisfies the local unbiasedness condition at θ:

$$\sum_{\nu=1}^{2} \int \hat\theta^i(\nu, \xi)\ \mathrm{Tr}\, \rho_\theta M(\nu, d\xi) = \theta^i, \qquad i = 1, 2 \tag{15}$$

$$\sum_{\nu=1}^{2} \int \hat\theta^i(\nu, \xi)\ \frac{\partial}{\partial\theta^j} \mathrm{Tr}\, \rho_\theta M(\nu, d\xi) = \delta^i_j, \qquad i, j = 1, 2. \tag{16}$$

To prove (15), we used the fact that $A_\nu \in \mathcal{T}_\theta^S(\mathcal{P})$, i.e., $\langle I, A_\nu \rangle_{\rho_\theta} = 0$. To prove (16), observe that

$$\int \xi \frac{\partial}{\partial \theta^j} \mathrm{Tr}\, \rho_\theta E_\nu(d\xi) = \langle L^S_{\theta,j}, A_\nu \rangle_{\rho_\theta},$$

so that the left hand side of (16) becomes

$$\sum_{\nu=1}^{2} \langle L^i, A_\nu \rangle_{\rho_\theta} \langle L^S_{\theta,j}, A_\nu \rangle_{\rho_\theta} = \langle L^i, L^S_{\theta,j} \rangle_{\rho_\theta} = \delta^i_j.$$

With this measurement M,

$$\begin{aligned}
\mathrm{tr}\, V_\theta[M] &= \sum_{\nu=1}^{2} \int \left[\left(\hat\theta^1(\nu,\xi) - \theta^1 \right)^2 + \left(\hat\theta^2(\nu,\xi) - \theta^2 \right)^2 \right] \mathrm{Tr}\, \rho_\theta M(\nu, d\xi) \\
&= \sum_{\nu=1}^{2} \frac{1}{p_\nu} \left[\langle L^1, A_\nu \rangle^2_{\rho_\theta} + \langle L^2, A_\nu \rangle^2_{\rho_\theta} \right].
\end{aligned}$$

In the second equality, we used the fact that

$$\int \xi^2\, \mathrm{Tr}\, \rho_\theta E_\nu(d\xi) = \langle A_\nu, A_\nu \rangle_{\rho_\theta} = 1.$$

Since, for given $\mu_1, \mu_2 > 0$, $\mu_1/p_1 + \mu_2/p_2$ takes the minimum $(\sqrt{\mu_1} + \sqrt{\mu_2})^2$ at $p_\nu = \sqrt{\mu_\nu}/(\sqrt{\mu_1} + \sqrt{\mu_2})$, we see

$$\begin{aligned}
\min_{\{p_\nu\}} \mathrm{tr}\, V_\theta[M] &= \left[\sqrt{\langle L^1, A_1 \rangle^2_{\rho_\theta} + \langle L^2, A_1 \rangle^2_{\rho_\theta}} + \sqrt{\langle L^1, A_2 \rangle^2_{\rho_\theta} + \langle L^2, A_2 \rangle^2_{\rho_\theta}} \right]^2 \\
&= \left[\sqrt{\langle A_1, \phi(A_1) \rangle_{\rho_\theta}} + \sqrt{\langle A_2, \phi(A_2) \rangle_{\rho_\theta}} \right]^2 \\
&= \left[\sqrt{\lambda_1} + \sqrt{\lambda_2} \right]^2 \\
&= \langle L^1, L^1 \rangle_{\rho_\theta} + \langle L^2, L^2 \rangle_{\rho_\theta} + 2\sqrt{\langle L^1, L^1 \rangle_{\rho_\theta} \langle L^2, L^2 \rangle_{\rho_\theta} - \langle L^1, L^2 \rangle^2_{\rho_\theta}}.
\end{aligned} \tag{17}$$

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 are independent of the choice of the basis which represents ϕ in a matrix form.

The random measurement presented above was first introduced in [5] by one of the present authors. In that paper, it was also shown that the problem of finding the achievable CR bound for an arbitrary 2-parameter faithful spin 1/2 model can be reduced to an easy minimization problem. Interestingly, the explicit solution of the minimization problem, i.e., the achievable CR bound, turns out to be identical to the quantity (17), although the model treated there is not pure nor has in general a $\mathcal{D}_\rho$-invariant tangent space.

Now we establish the relation between (14) and (17) for a coherent model.

Theorem 6. *For a 2-dimensional coherent model $\{\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|\}$, the lower bound (14) is identical to (17). In other words, the generalized RLD bound (14) is achievable.*

Proof By Theorem 3, $L^1\psi_\theta$ and $L^2\psi_\theta$ are linearly dependent. Therefore

$$\det \begin{bmatrix} \langle L^1\psi_\theta | L^1\psi_\theta\rangle & \langle L^1\psi_\theta | L^2\psi_\theta\rangle \\ \langle L^2\psi_\theta | L^1\psi_\theta\rangle & \langle L^2\psi_\theta | L^2\psi_\theta\rangle \end{bmatrix} = 0,$$

which leads to

$$(\mathrm{Im}\,\langle L^1\psi_\theta | L^2\psi_\theta\rangle)^2 = \langle L^1\psi_\theta | L^1\psi_\theta\rangle\langle L^2\psi_\theta | L^2\psi_\theta\rangle - (\mathrm{Re}\,\langle L^1\psi_\theta | L^2\psi_\theta\rangle)^2.$$

By (1) and (2), this can be read as

$$\left|[L^1, L^2]_{\rho_\theta}\right|^2 = \langle L^1, L^1\rangle_{\rho_\theta}\langle L^2, L^2\rangle_{\rho_\theta} - \langle L^1, L^2\rangle_{\rho_\theta}^2,$$

which proves the theorem. □

It should be noted that a more convincing result has been obtained by Matsumoto [11]. He proved that the CR bound (13) is achievable for a $2m$-dimensional coherent model with an arbitrary weight G.

It is also worth noting that the achievability of (14) is closely related to the Heisenberg uncertainty relation. By a coordinate transformation, we can assume that the SLD Fisher information matrix is diagonal at a fixed $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. Then there exist nonzero real numbers c_1, c_2 and normalized ρ_θ-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 by (7)

$$(c_1 L^1 + i c_2 L^2)\psi_\theta = 0.$$

This is nothing but the equality condition for the Heisenberg uncertainty relation. So we have

$$\langle L^1, L^1\rangle_{\rho_\theta}\langle L^2, L^2\rangle_{\rho_\theta} = \left|[L^1, L^2]_{\rho_\theta}\right|^2.$$

This equation, combined with the assumption that $\langle L^1, L^2\rangle_{\rho_\theta} = 0$, gives another proof of Theorem 6 for an orthogonal parametrization at ρ_θ.

VI Examples

In this section we calculate the achievable CR bounds for canonical and spin coherent models. Throughout this section, adjoint operators and complex conjugate numbers are denoted by † and $*$, respectively, according to the convention in physics. Also we use the same letter for both a square summable operator and the corresponding element in $\mathcal{L}_h^2(\rho)$.

VI.1 Canonical squeezed state model

The canonical squeezed state [12] [13] is defined by

$$\rho_{q,p} = D(q,p)|0\rangle_{\xi\xi}\langle 0|D^\dagger(q,p), \qquad (q,p \in \mathbf{R}),$$

where $D(q,p) = \exp(za^\dagger - z^* a)$ denotes the shift operator with $z = (q+ip)/\sqrt{2}$, and a and $a^\dagger$ are annihilation and creation operators, respectively, with $a = (Q+iP)/\sqrt{2}$. Further $|0\rangle_\xi = \exp[(\xi {a^\dagger}^2 - \xi^* a^2)/2]|0\rangle$ is the squeezed vacuum with $|0\rangle$ the Fock vacuum, and ξ a complex number which represents the squeezing ratio: let $\xi = se^{i\theta}$.

Comparing the identity $b|z\rangle_\xi = \beta|z\rangle_\xi$ with Corollary 4, where $|z\rangle_\xi = D(q,p)|0\rangle_\xi$, $b = a\cosh s - a^\dagger e^{i\theta}\sinh s$, and $\beta = z\cosh s - z^* e^{i\theta}\sinh s$, we see that $\rho_{q,p}$ is a 2-dimensional coherent model, and a normalized $\rho_{q,p}$-symplectic basis is given by

$$\begin{aligned}
\tilde{L}_1^S &= \sqrt{2}[(Q-qI)(\cosh s - \cos\theta\sinh s) - (P-pI)\sin\theta\sinh s], \\
\tilde{L}_2^S &= \sqrt{2}[(P-pI)(\cosh s + \cos\theta\sinh s) - (Q-qI)\sin\theta\sinh s].
\end{aligned}$$

The SLDs at $\rho_{q,p}$ are calculated by operating $-\mathcal{D}_{q,p}$ to ALDs at $\rho_{q,p}$. By expanding ALDs $L_q^A = 2(P-pI)$, $L_p^A = -2(Q-qI)$ into linear combinations of $\tilde{L}_1^S$, $\tilde{L}_2^S$, and using the relations $\mathcal{D}_{q,p}\tilde{L}_1^S = \tilde{L}_2^S$, $\mathcal{D}_{q,p}\tilde{L}_2^S = -\tilde{L}_1^S$, we have

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

The corresponding SLD Fisher information matrix becomes

$$J_{q,p}^S = 2\begin{bmatrix} \cosh 2s - \cos\theta\sinh 2s & -\sin\theta\sinh 2s \\ -\sin\theta\sinh 2s & \cosh 2s + \cos\theta\sinh 2s \end{bmatrix}.$$

Then from (17), we have

$$\min_M \operatorname{tr} V_{q,p}[M] = \cosh 2s + 1.$$

VI.2 Spin coherent state model

The spin coherent state [14] [15] in the spin j representation is defined by

$$\rho_{\theta,\varphi} = R(\theta,\varphi)|j\rangle\langle j|R^\dagger(\theta,\varphi), \qquad (0 \le \theta \le \pi,\ 0 \le \varphi < 2\pi),$$

where (θ,φ) is the polar coordinate system (the north pole is $\theta = 0$ and x-axis corresponds to $\varphi = 0$), $R(\theta,\varphi) = \exp[i\theta(J_x\sin\varphi - J_y\cos\varphi)]$ the rotation through an angle $-\theta$ about an axis $(\sin\varphi, -\cos\varphi, 0)$, and $|j\rangle$ the highest weight state with respect to J_z that corresponds to the north pole.

Since $J_+|j\rangle = (J_x + iJ_y)|j\rangle = 0$, we find that $\rho_{\theta,\varphi}$ is a 2-dimensional coherent model, and a normalized $\rho_{0,0}$-symplectic basis is $\tilde{L}_1^S(0,0) = \sqrt{2/j}J_x$, $\tilde{L}_2^S(0,0) = \sqrt{2/j}J_y$. A normalized $\rho_{\theta,\varphi}$-symplectic basis is then calculated as

$$\tilde{L}_k^S(\theta,\varphi) = R(\theta,\varphi)\tilde{L}_k^S(0,0)R^\dagger(\theta,\varphi),$$

where $k = 1, 2$.

On the other hand, the generators of rotations about axes $(\sin\varphi, -\cos\varphi, 0)$ and $(\cos\varphi, \sin\varphi, 0)$ at $\theta = 0$ are $i(J_x \sin\varphi - J_y \cos\varphi)$ and $i(J_x \cos\varphi + J_y \sin\varphi)$, respectively. Therefore ALDs for the model at $\rho_{\theta,\varphi}$ are given by

$$\begin{aligned} L^A_\theta(\theta,\varphi) &= R(\theta,\varphi)\{-2(J_x \sin\varphi - J_y \cos\varphi)\}R^\dagger(\theta,\varphi) \\ &= -\sqrt{2j}\left\{\tilde{L}^S_1(\theta,\varphi)\sin\varphi - \tilde{L}^S_2(\theta,\varphi)\cos\varphi\right\}, \\ L^A_\varphi(\theta,\varphi) &= R(\theta,\varphi)\{-2(J_x \cos\varphi + J_y \sin\varphi)\sin\theta\}R^\dagger(\theta,\varphi) \\ &= -\sqrt{2j}\left\{\tilde{L}^S_1(\theta,\varphi)\sin\theta\cos\varphi + \tilde{L}^S_2(\theta,\varphi)\sin\theta\sin\varphi\right\}. \end{aligned}$$

The SLDs at $\rho_{\theta,\varphi}$ are calculated by operating $-\mathcal{D}_{\theta,\varphi}$ to ALDs, to obtain

$$\begin{aligned} L^S_\theta(\theta,\varphi) &= \sqrt{2j}\left\{\tilde{L}^S_1(\theta,\varphi)\cos\varphi + \tilde{L}^S_2(\theta,\varphi)\sin\varphi\right\}, \\ L^S_\varphi(\theta,\varphi) &= -\sqrt{2j}\left\{\tilde{L}^S_1(\theta,\varphi)\sin\theta\sin\varphi - \tilde{L}^S_2(\theta,\varphi)\sin\theta\cos\varphi\right\}. \end{aligned}$$

Since $\rho_{\theta,\varphi}$-symplectic basis $\{\tilde{L}^S_k(\theta,\varphi)\}_{k=1,2}$ is orthonormal, the SLD Fisher information matrix and the matrix D are easily calculated:

$$J^S_{\theta,\varphi} = \begin{bmatrix} 2j & 0 \\ 0 & 2j\sin^2\theta \end{bmatrix}, \qquad D_{\theta,\varphi} = \begin{bmatrix} 0 & -2j\sin\theta \\ 2j\sin\theta & 0 \end{bmatrix}.$$

We thus have

$$\min_M \operatorname{tr} V_{\theta,\varphi}[M] = \frac{1}{2j}\left(1 + \frac{1}{\sin\theta}\right)^2.$$

VII Conclusions

We introduced a class of quantum pure state models called the coherent models. They are characterized by a symplectic structure of the tangent space, and have a close connection with the conventional generalized coherent states in mathematical physics. A Cramér-Rao type bound for a coherent model was derived by an analogous argument to the derivation of the right logarithmic derivative bound. Moreover, by an argument of random measurement, this lower bound was found to be achievable.

Acknowledgment

We thank Keiji Matsumoto for helpful suggestions, with which the manuscript has been improved as compared with the early version [16] of this paper.

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).

[3] H. P. Yuen and M. Lax, IEEE Trans. **IT-19**, 740 (1973).

[4] H. Nagaoka, IEICE Technical Report **IT89-42**, 9 (1989).

[5] H. Nagaoka, Trans. Jap. Soc. Indust. Appl. Math. **1**, 43 (1991), (in Japanese).

[6] A. Fujiwara and H. Nagaoka, Phys. Lett. **A201**, 119 (1995).

[7] A. Fujiwara and H. Nagaoka, in *Quantum coherence and decoherence*, edited by K. Fujikawa and Y. A. Ono (Elsevier, Amsterdam, 1996), p. 303.

[8] V. Guillemin and S. Sternberg, *Symplectic techniques in physics* (Cambridge Univ. Press, Cambridge, 1984).

[9] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, II* (John Wiley, New York, 1969).

[10] J. R. Klauder and B. Skagerstam, *Coherent states* (World Scientific, Singapore, 1985).

[11] K. Matsumoto, *A geometrical approach to quantum estimation theory*, Ph.D. Thesis, University of Tokyo, 1998.

[12] H. P. Yuen, Phys. Rev. **A13**, 2226 (1976).

[13] C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmermann, Rev. Mod. Phys. **52**, 341 (1980).

[14] J. M. Radcliffe, J. Phys. A: Gen. Phys. **4**, 313 (1971).

[15] F. T. Arecchi, E. Courtens, R. Glimore, and H. Thomas, Phys. Rev. **6**, 2211 (1972).

[16] A. Fujiwara, Mathematical Engineering Technical Report **94-9**, **94-10**, University of Tokyo (1994).