Estimation of $SU(2)$ operation and dense coding: An information geometric approach

Akio Fujiwara
Department of Mathematics, Osaka University
Toyonaka, Osaka 560-0043, Japan

Abstract

This paper addresses quantum statistical estimation of operators $U \in SU(2)$ acting on $\mathbb{C}P^3$ as $\psi \mapsto (U \otimes I)\psi$ where $\psi \in \mathbb{C}^2 \otimes \mathbb{C}^2$. This is regarded as a continuous analogue of the dense coding. We first prove that the quantum Cramér-Rao lower bound takes the minimum, and is achievable, if and only if ψ is a maximally entangled state. We next show that an $SU(2)$ orbit on $\mathbb{C}P^3$ equipped with the standard Riemannian structure is isometric to $SU(2)/\{\pm I\} \cong SO(3)$ if and only if ψ is a maximally entangled state. These results provide an alternative view for the optimality of the use of a maximally entangled state.

PACS number: 03.67.-a, 03.65.Ta, 89.70.+c

1 Introduction

Let ψ be a maximally entangled vector on $\mathbb{C}^2 \otimes \mathbb{C}^2$, say,

$$\psi = \frac{1}{\sqrt{2}} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right), \tag{1}$$

and let $U_0 = I$, $U_1 = i\sigma_x$, $U_2 = i\sigma_y$, $U_3 = i\sigma_z$. Then $\langle (U_i \otimes I)\psi | (U_j \otimes I)\psi \rangle = \delta_{ij}$, so that one can distinguish reliably four vectors $\{(U_i \otimes I)\psi\}_{i=0}^3$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$, and hence four operators $\{U_i\}_{i=0}^3$ in $SU(2)$. This is the basic idea of the so called dense coding [1], and is a manifestation of improved distinguishability through entanglement.

The dense coding and its variants [2], as well as the celebrated quantum channel coding theorem [3], concern distinguishability among finitely many alternatives, and the proper quantum statistical framework for dealing with finite alternatives is the hypothesis testing [4] [5]. There is another, essentially different, framework in quantum statistics, called the parameter estimation [4] [6], in which one deals with continuously many alternatives. Among recent development in the latter framework is a quantum channel identification problem [7] [8], in which one seeks the best strategy of estimating an unknown quantum operation Γ acting on the set $\mathcal{S}(\mathcal{H})$ of quantum states on a Hilbert space $\mathcal{H}$. In quantum information theory, it is customary that a quantum channel is given *a priori*. In practice, however, one first identifies the quantum channel of interest, and then applies various information theoretic results to the channel. Identification of a quantum channel thus precedes every quantum information scheme, and its optimization is of fundamental importance in quantun information theory. As an illustrative example, we have explored in [7] the identification problem of a depolarization channel $\Gamma : \mathcal{S}(\mathbb{C}^2) \to \mathcal{S}(\mathbb{C}^2)$, and have observed improvement of distinguishability through quantum entanglement and a rather unexpected transitionlike behavior of the optimal estimation scheme.

In this paper, we explore the identification problem of a unitary channel

$$\Gamma_U : \mathcal{S}(\mathbb{C}^2) \longrightarrow \mathcal{S}(\mathbb{C}^2) : \rho \longmapsto U\rho U^*, \qquad (U \in SU(2)).$$

In particular, we focus on the estimation of the operator U through the extension

$$\Gamma_U \otimes I : \sigma \longmapsto (U \otimes I)\sigma(U \otimes I)^*$$

on $\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$, or its restriction

$$\Gamma_U \otimes I : |\psi\rangle\langle\psi| \longmapsto |(U \otimes I)\psi\rangle\langle(U \otimes I)\psi|$$

to the set $\partial_e\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ of pure states that is identified with the 3-dimensional complex projective space $\mathbb{C}P^3$. This is naturally regarded as a continuous analogue of the dense coding, and its analysis demonstrates the qualitative difference between distinguishability for finitely many alternatives and that for continuously many alternatives. The main results are summarized as follows.

(i) The quantum Cramér-Rao lower bound for the family $\{(U \otimes I)\psi\}_{U \in SU(2)}$ of output states takes the minimum, and is achievable, if and only if ψ is a maximally entangled state (Theorems 3, 4).

(ii) The manifold of output states (i.e., an $SU(2)$ orbit) equipped with the quantum Fisher metric is isometric to $SU(2)/\{\pm I\} \cong SO(3)$ if and only if ψ is a maximally entangled state (Theorem 6).

These results provide an alternative view for the optimality of the use of a maximally entangled state.

The paper is organized as follows. In Section 2, we formulate a statistical estimation problem of $SU(2)$, and analyze it from a noncommutative statistical point of view. In Section 3, we introduce a general framework of information geometry on a projective Hilbert space $P(\mathcal{H})$. The standard Riemannian structure of $P(\mathcal{H})$ is derived as a special example. In Section 4, a Riemannian geometric study of $SU(2)$ orbits on $\mathbb{C}P^3$ is presented. These results are discussed in a unified manner in Section 5. Throughout the paper, the symbols '$\simeq$' and '$\cong$' stand for 'isomorphic' and 'isometric' respectively.

2 Statistical estimation of $SU(2)$

Suppose an unknown operation Γ acting on $\mathcal{S}(\mathbb{C}^2)$ is noiseless, in that there is a unitary operator $U \in SU(2)$ such that $\Gamma(=: \Gamma_U) : \rho \mapsto U\rho U^*$, and our problem is to estimate the unknown U. A general scheme of estimating an unknown quantum operation Λ acting on $\mathcal{S}(\mathcal{H})$ is this: input a well-prepared state $\sigma \in \mathcal{S}(\mathcal{H})$ to Λ and estimate the dynamical change $\sigma \mapsto \Lambda(\sigma)$ by performing a certain measurement on the output state $\Lambda(\sigma)$. When Λ belongs to a parametric family $\{\Lambda_\theta; \theta \in \Theta\}$ of operations, the problem amounts to finding an optimal input σ and an optimal estimator for the parameter θ of the family $\{\Lambda_\theta(\sigma); \theta \in \Theta\}$ of output states [7].

In our problem, the group $SU(2)$ is a 3-dimensional manifold and is parametrized, for example, as

$$U = U_{(\phi,\alpha,\beta)} := \begin{bmatrix} e^{i\alpha}\cos\phi & -e^{i\beta}\sin\phi \\ e^{-i\beta}\sin\phi & e^{-i\alpha}\cos\phi \end{bmatrix}, \qquad \left(0 \le \phi \le \frac{\pi}{2},\ 0 \le \alpha < 2\pi,\ 0 \le \beta < 2\pi\right).$$

Since, for any $\rho \in \mathcal{S}(\mathbb{C}^2)$, the family $\{\Gamma_U(\rho); U \in SU(2)\}$ of output states is at most 2-dimensional, we must extend Γ_U on an enlarged Hilbert space $(\mathbb{C}^2)^{\otimes n}$, $(n \geq 2)$, in order for the parametrization of output states to be nondegenerate. In this paper, we focus on the extension $\Lambda_U := \Gamma_U \otimes I$, i.e.,

$$\Lambda_{U_\theta} : \mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2) \longrightarrow \mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2) : \sigma \longmapsto (U_\theta \otimes I)\sigma(U_\theta \otimes I)^* \tag{2}$$

where $\theta := (\theta^1, \theta^2, \theta^3) := (\phi, \alpha, \beta)$. Since $\Lambda_{-U} = \Lambda_U$, we might as well express that our problem is to estimate the parameter θ of the quotient group $SU(2)/\{\pm I\} \simeq SO(3)$. Consequently, the estimation of $SU(2)$ operation must be a local one: the domain Θ of the parameter θ to be estimated forms a local chart of $SU(2)$ on which the parametrization $\theta \mapsto \Lambda_{U_\theta}(\sigma)$ is one-to-one.

Let us proceed to the parameter estimation for the family (2). Our task was to find an optimal input σ and an optimal estimator for the parametric family $\{\Lambda_{U_\theta}(\sigma); \theta \in \Theta\}$ of output states. One of the most fundamental result in quantum estimation theory is the quantum Cramér-Rao inequality [4] [6], (cf., Appendix A): when the true value of the parameter is θ_0, the covariance matrix $V_{\theta_0}[M]$ of an arbitrary estimator M for the parameter θ that is locally unbiased at θ_0 is bounded from below as

$$V_{\theta_0}[M] \geq J_{\theta_0}(\sigma)^{-1}.$$

Here $J_\theta(\sigma)$ denotes the symmetric logarithmic derivative (SLD) Fisher information matrix with respect to the coordinate system θ of the output family $\{\Lambda_{U_\theta}(\sigma); \theta \in \Theta\}$, and the inequality means that the matrix $V_{\theta_0}[M] - J_{\theta_0}(\sigma)^{-1}$ is positive semidefinite.

Lemma 1 *For all $\sigma, \tau \in \mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$ and $\lambda \in [0, 1]$,*

$$J_\theta(\lambda\sigma + (1-\lambda)\tau) \leq \lambda J_\theta(\sigma) + (1-\lambda) J_\theta(\tau).$$

Proof This follows immediately from the convexity of the SLD Fisher metric [7]. □

In contrast to the one dimensional parameter case [7], we cannot conclude directly from Lemma 1 that the optimal input is a pure state. This is partly because the matrices $J_\theta(\sigma)$ and $J_\theta(\tau)$ appeared in the right-hand side do not always comparable with each other. However we have the following

Lemma 2 *Let ψ be the maximally entangled state (1). Then for all $\sigma \in \mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2)$,*

$$J_\theta(\sigma) \leq J_\theta(|\psi\rangle\langle\psi|). \tag{3}$$

The equality holds if and only if σ is a maximally entangled state.

Proof We first prove the inequality (3) for all pure states σ of the form $\sigma = |\psi(x)\rangle\langle\psi(x)|$, where

$$\psi(x) := \sqrt{1-x} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sqrt{x} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad \left(0 \leq x \leq \frac{1}{2}\right).$$

The corresponding parametric family of output states is $\{|\psi_\theta(x)\rangle\langle\psi_\theta(x)|; \theta \in \Theta\}$, where

$$\psi_\theta(x) := (U_\theta \otimes I)\psi = \sqrt{1-x} \begin{bmatrix} e^{i\alpha}\cos\phi \\ e^{-i\beta}\sin\phi \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sqrt{x} \begin{bmatrix} -e^{i\beta}\sin\phi \\ e^{-i\alpha}\cos\phi \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{4}$$

By a direct computation, the SLD Fisher information matrix $J_\theta(x) := J_\theta(\sigma)$ for the family (4) is

$$J_\theta(x) = 4 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos^2\phi - (1-2x)^2\cos^4\phi & (1-2x)^2\cos^2\phi\sin^2\phi \\ 0 & (1-2x)^2\cos^2\phi\sin^2\phi & \sin^2\phi - (1-2x)^2\sin^4\phi \end{bmatrix}. \tag{5}$$

It suffices to show that $J_\theta(x)$ is monotone increasing in x. In fact, for $0 \le y \le x \le 1/2$

$$J_\theta(x) - J_\theta(y) = 16(x-y)(1-x-y)\begin{bmatrix} 0 & 0 & 0 \\ 0 & \cos^4\phi & -\cos^2\phi\sin^2\phi \\ 0 & -\cos^2\phi\sin^2\phi & \sin^4\phi \end{bmatrix},$$

which is positive semidefinite, and equals zero if and only if $x = y$.

We next prove the inequality (3) for any pure state σ. Let $\sigma = |\psi(x,V,W)\rangle\langle\psi(x,V,W)|$ where $\psi(x,V,W) := (V\otimes W)\psi(x)$, $(V,W \in SU(2))$, and let $J_\theta(x,V,W) := J_\theta(\sigma)$ be the corresponding SLD Fisher information. Since we are dealing with the operations of the form $U_\theta\otimes I$, the SLD Fisher information matrix is invariant under the transformation W of the second frame, i.e., $J_\theta(x,V,W) = J_\theta(x,V,I)$ for all W. On the other hand, the transformation V of the first frame induces the coordinate transform $\theta\mapsto\theta' := (\phi',\alpha',\beta')$ of $SU(2)$ as

$$U_\theta V = \begin{bmatrix} e^{i\alpha'}\cos\phi' & -e^{i\beta'}\sin\phi' \\ e^{-i\beta'}\sin\phi' & e^{-i\alpha'}\cos\phi' \end{bmatrix}.$$

Then the above proof applies to the new coordinate system θ', to obtain the monotonicity of $J_{\theta'}(x)$ in x. (Note that $J_{\theta'}(x) \ne J_\theta(x,V,I)$ in general: $J_{\theta'}(x)$ is the list of components of the SLD Fisher metric with respect to the coordinate system θ', while $J_\theta(x,V,I)$ the list with respect to θ.) Since the monotonicity is purely a geometric property and is invariant under a coordinate transform, we conclude that $J_\theta(x,V,I)$ also exhibits monotonicity in x (with V fixed). Now we claim $J_\theta(1/2) = J_\theta(1/2,V,I)$ for all $V\in SU(2)$. In fact, it is a well known fact that given V, there is a $V'\in SU(2)$ such that $(V\otimes I)\psi = (I\otimes V')\psi$ for a maximally entangled ψ, (cf., Appendix B, Lemma 11). Therefore $J_\theta(1/2,V,I) = J_\theta(1/2,I,V') = J_\theta(1/2,I,I) = J_\theta(1/2)$. In summary, $J_\theta(x,V,W) \le J_\theta(1/2) = J_\theta(|\psi\rangle\langle\psi|)$ for all $V,W\in SU(2)$ and $0\le x\le 1/2$, with equality if and only if $x = 1/2$.

Now we prove the inequality (3) for any state $\sigma\in\mathcal{S}(\mathbb{C}^2\otimes\mathbb{C}^2)$. Let $\sigma = \sum_i\lambda_i\sigma_i$ be a pure state decomposition in which $\lambda_i > 0$ and $\sum_i\lambda_i = 1$. Then by Lemma 1 and the above fact, we conclude that

$$J_\theta(\sigma) \le \sum_i \lambda_i J_\theta(\sigma_i) \le J_\theta(|\psi\rangle\langle\psi|). \tag{6}$$

Finally, observe that the inequalities in (6) hold for any pure state decomposition of σ. As a consequence, $J_\theta(\sigma) = J_\theta(|\psi\rangle\langle\psi|)$ implies that every pure state decomposition of σ comprizes only maximally entangled pure states. This is the case only if σ itself is a maximally entangled pure state, (cf., [9]). The lemma was verified. □

Theorem 3 *For each value of the parameter θ, the Cramér-Rao lower bound $J_\theta(\sigma)^{-1}$ takes the minimum if and only if σ is a maximally entangled state.*

Proof This follows immediately from Lemma 2 and the fact that the function $f(t) = -1/t$ is operator monotone on $(0,\infty)$, [10]. □

Note that the inequalities derived in Lemmas 1, 2 are intrinsic properties of the SLD Fisher metric and are independent of a particular choice of the coordinate system θ. As a consequence, Theorem 3 holds for any parametrization of $SU(2)$.

Theorem 3 hints that the optimal input will be a maximally entangled state. However, it alone does not lead to a decisive conclusion, because the Cramér-Rao lower bound $J_\theta(\sigma)^{-1}$ is not always achievable for a multi parameter quantum statistical model. Here we say that the Cramér-Rao lower bound is *achievable* at θ if there is a locally unbiased estimator M that satisfies $V_\theta[M] = J_\theta(\sigma)^{-1}$. In this sense the next theorem is the key to the conclusion that a maximally entangled state is in fact the optimal one.

Theorem 4 *The Cramér-Rao lower bound $J_\theta(x)^{-1}$ for the family (4) is achievable if and only if $x = 1/2$.*

Proof The Cramér-Rao lower bound is achievable at θ if and only if

$$\{\langle L_{\theta,i}\psi_\theta(x)|L_{\theta,j}\psi_\theta(x)\rangle\}_{1\le i,j\le 3}$$

are all real, where $\{L_{\theta,i}\}_{i=1}^3$ are SLDs, (cf., Appendix A, Corollary 10). By a direct computation, we have

$$\mathrm{Im}\,\langle L_{\theta,1}\psi_\theta(x)|L_{\theta,2}\psi_\theta(x)\rangle = \mathrm{Im}\,\langle L_{\theta,1}\psi_\theta(x)|L_{\theta,3}\psi_\theta(x)\rangle = 2(2x-1)\sin 2\phi,$$

and

$$\mathrm{Im}\,\langle L_{\theta,2}\psi_\theta(x)|L_{\theta,3}\psi_\theta(x)\rangle = 0.$$

The assertion immediately follows. □

The implication of Theorem 4 is profound. The existence of an estimator that achieves the Cramér-Rao lower bound implies the existence of compatible observables that correspond to the parameters of $SU(2)$. Theorem 4 thus asserts that the noncommutative nature of the $SU(2)$ parameters is "suppressible" (at least locally) by using a maximally entangled input. Moreover, the achievability condition used in the proof is an intrinsic property of the tangent space and hence is independent of a particular choice of the coordinate system θ. As a consequence, the local suppression of noncommutativity is also a parametrization independent (i.e., geometric) property.

In summary, for estimating the extended $SU(2)$ operation $\Lambda_{U_\theta} : \sigma \mapsto (U_\theta \otimes I)\sigma(U_\theta \otimes I)^*$, the optimal input σ is a maximally entangled state. This gives an estimation theoretic verification for the optimality of the use of a maximally entangled state. In the subsequent sections, we explore a differential geometric interpretation of this result to obtain a deeper insight into the role of entanglement.

3 Information geometry of pure states

It is well known that the parameter estimation theory for a classical statistical manifold is closely related to an information geometric structure of the manifold [11]. Such a geometric structure has been successfully extended to a quantum regime, i.e., to manifolds of faithful quantum states on a finite dimensional complex Hilbert space $\mathcal{H}$ [11, Chap. 7]. In this section we further extend an information geometric structure to the manifold $\mathcal{M} := \partial_e \mathcal{S}(\mathcal{H})$ of pure quantum states that is identified with the projective Hilbert space $P(\mathcal{H})$. For more information, see [12] [13], where a relation to Berry's phase and extensions to manifolds of generic quantum states are also presented.

Let $\mathcal{B}(\mathcal{H})$ and $\mathcal{B}_h(\mathcal{H})$ denote the sets of linear operators and Hermitian operators on $\mathcal{H}$. In order to introduce an information geometric structure on $\mathcal{M}$, the following lemma is useful.

Lemma 5 *For $\rho \in \mathcal{M}$ and $D \in \mathcal{B}_h(\mathcal{H})$, the following conditions are equivalent.*
(a) *There exists a unique tangent vector $X \in T_\rho\mathcal{M}$ that satisfies $D = X\rho$.*
(b) *There exists an operator $L \in \mathcal{B}_h(\mathcal{H})$ that satisfies $D = \frac{1}{2}(\rho L + L\rho)$ and $\mathrm{Tr}\rho L = 0$.*

Proof Let $\rho = |\psi\rangle\langle\psi|$ and let $\{\psi_i\}_{i=1}^n$ be an orthonormal (moving) frame of $\mathcal{H}$ with $\psi_1 = \psi$. We introduce $\mathbb{R}$-linear spaces $T_\rho^{(i)}$, $(i = 1, 2, 3)$, as follows. The $T_\rho^{(1)}$ is the set of Hermitian operators L satisfying $\mathrm{Tr}\rho L = 0$, the $T_\rho^{(2)}$ is the set $\{X\rho\,;\, X \in T_\rho\mathcal{M}\}$, and the $T_\rho^{(3)}$ is the set of Hermitian operators whose matrix representation with respect to the frame $\{\psi_i\}_{i=1}^n$ is of the form

$$\begin{bmatrix} 0 & \overline{a_2} & \cdots & \overline{a_n} \\ a_2 & & & \\ \vdots & & O & \\ a_n & & & \end{bmatrix}, \qquad (a_i \in \mathbb{C}).$$

It is easily shown that $\tilde{f}(L) := \frac{1}{2}(\rho L + L\rho)$ defines a surjective linear map $\tilde{f} : T_\rho^{(1)} \to T_\rho^{(3)}$. On the other hand, the tangent space $T_\rho\mathcal{M}$ is clearly isomorphic to $T_\rho^{(2)}$, and since $X\rho = |X\psi\rangle\langle\psi| + |\psi\rangle\langle X\psi|$, the space $T_\rho^{(2)}$ is obviously identical to $T_\rho^{(3)}$. (Note that $\mathrm{Tr}(X\rho) = X(\mathrm{Tr}\rho) = 0$.) Thus there is a surjective linear map $f : T_\rho^{(1)} \to T_\rho\mathcal{M}$. □

The operator L in Lemma 5 (b) is uniquely determined [14] only up to

$$\ker f = \{K \in \mathcal{B}_h(\mathcal{H})\,;\, K\rho = 0\}.$$

Because of this ambiguity, we must arbitrarily choose a representative of the SLD in order to define a one-one homomorphism $\mathcal{L}_\rho : T_\rho\mathcal{M} \to \mathcal{B}_h(\mathcal{H})$ which satisfies

$$d\rho = \frac{1}{2}(\rho\mathcal{L}_\rho + \mathcal{L}_\rho\rho).$$

In addition we assume that $\mathcal{L}_\rho$ is smooth in ρ. Such an operator-valued one-form $\mathcal{L}_\rho$ is called an *SLD representation*. When no confusion is likely to arise, we simply denote $\mathcal{L}_\rho(X)$ as L_X for each $X \in T_\rho\mathcal{M}$.

Let us introduce an information geometric structure on $\mathcal{M}$. We first define a Riemannian metric by the SLD Fisher metric:

$$g(X, Y) := \frac{1}{2}\mathrm{Tr}\rho(L_X L_Y + L_Y L_X) = \mathrm{Tr}(X\rho)L_Y.$$

It is invariant under the arbitrariness $\ker f$ of SLD representations. Moreover, it is shown that g is identical to the Fubini-Study metric [14]. We next introduce a pair of affine connections that are mutually dual with respect to the SLD Fisher metric. One is defined by

$$(\nabla_X Y)\,\rho := \frac{1}{2}\{\rho(XL_Y - \mathrm{Tr}\rho(XL_Y)) + (XL_Y - \mathrm{Tr}\rho(XL_Y))\rho\},$$

and is called the *exponential* connection. It is well defined because the right-hand side uniquely defines a derivative of ρ by Lemma 5. The other connection is defined via duality:

$$g(\nabla^*_X Y, Z) := Xg(Y, Z) - g(Y, \nabla_X Z) = \mathrm{Tr}(X(Y\rho))L_Z,$$

and is called the *mixture* connection. Note that in contrast to a quantum statistical manifold of faithful states, the mixture connection cannot be defined by $(\nabla^*_X Y)\rho = X(Y\rho)$, since $X(Y\rho)$ does not correspond to a derivative of ρ in general.

By a direct computation, the torsions T and T^* which correspond to ∇ and ∇^* are

$$T(X,Y)\rho = \frac{1}{4}[\,[L_X, L_Y], \rho], \qquad T^*(X,Y) = 0.$$

The Riemannian curvatures do not vanish in general. Thus one cannot expect the existence of the divergence on the space $(\mathcal{M}, g, \nabla, \nabla^*)$ in general.

Here is a special but important example: by differentiating the relation $\rho = \rho^2$ valid for pure states, we have a canonical choice $L_X := 2(X\rho)$ of the SLD representation [14]. Interestingly, the corresponding dualistic structure is reduced to the standard Riemannian structure of the projective Hilbert space $P(\mathcal{H})$ in which $\nabla = \nabla^* =$ the Levi-Civita connection of the Fubini-Study metric g. In fact, by using a (real) local coordinate system $\zeta = (\zeta^i)$ of $P(\mathcal{H})$, the components of the SLD Fisher metric g are given by

$$g_{ij} := g(\partial_i, \partial_j) = 2\mathrm{Tr}(\partial_i \rho)(\partial_j \rho), \qquad (\partial_i := \partial/\partial\zeta^i),$$

and the components of the mixture connection ∇^* are

$$\Gamma^*_{ij,k} := g(\nabla^*_{\partial_i}\partial_j, \partial_k) = \mathrm{Tr}(\partial_i\partial_j\rho)(2\partial_k\rho) = \frac{1}{2}(\partial_i g_{jk} + \partial_j g_{ki} - \partial_k g_{ij}).$$

Clearly the torsion T vanishes in this case. In what follows, we will work with this special differential geometric structure.

4 Information geometry of $SU(2)$ orbits on $\mathbb{C}P^3$

In this section, we regard $\mathbb{C}P^3$ as a real Riemannian manifold equipped with the SLD Fisher metric g, and explore the geometry of orbits of $SU(2)$ action $\psi \mapsto (U \otimes I)\psi$ on $\mathbb{C}P^3$. We say that unit vectors ψ and $\hat{\psi}$ on $\mathbb{C}^2 \otimes \mathbb{C}^2$ are equivalent (and denote $\psi \sim \hat{\psi}$) if they lie on the same $SU(2)$ orbit on $\mathbb{C}P^3$, i.e., if there is a $U \in SU(2)$ such that $|\hat{\psi}\rangle\langle\hat{\psi}| = |(U\otimes I)\psi\rangle\langle(U\otimes I)\psi|$. Let their Schmidt decompositions be

$$\psi = \sqrt{1-x}\, e_1 \otimes f_1 + \sqrt{x}\, e_2 \otimes f_2, \qquad \left(0 \le x \le \frac{1}{2}\right),$$

and

$$\hat{\psi} = \sqrt{1-\hat{x}}\, \hat{e}_1 \otimes \hat{f}_1 + \sqrt{\hat{x}}\, \hat{e}_2 \otimes \hat{f}_2, \qquad \left(0 \le \hat{x} \le \frac{1}{2}\right).$$

It is shown that $\psi \sim \hat{\psi}$ if and only if either $x = \hat{x} = 1/2$, or $x = \hat{x}$ and $\hat{f}_i = \lambda_i f_i$ for some $\lambda_i (\in \mathbb{C})$ of unit modulus, $(i = 1, 2)$, (cf., Appendix B, Lemma 11). As a consequence, $\mathbb{C}P^3$ is partitioned into disjoint $SU(2)$ orbits as

$$\mathbb{C}P^3 = \bigcup_{\psi \in \mathcal{I}} M_\psi,$$

where M_ψ denotes the orbit that passes through ψ. The orbit space $\mathcal{I} := \mathbb{C}P^3/SU(2)$ is identified with a complete list of initial points that generates disjoint orbits, and is explicitly given, for example, by

$$\psi = \sqrt{1-x}\begin{bmatrix}\cos\gamma \\ e^{i\delta}\sin\gamma\end{bmatrix} \otimes \begin{bmatrix}\cos\gamma \\ e^{-i\delta}\sin\gamma\end{bmatrix} + \sqrt{x}\begin{bmatrix}-e^{-i\delta}\sin\gamma \\ \cos\gamma\end{bmatrix} \otimes \begin{bmatrix}-e^{i\delta}\sin\gamma \\ \cos\gamma\end{bmatrix},$$

where $0 \leq x \leq 1/2$, $0 \leq \gamma \leq \pi/2$, and $0 \leq \delta < 2\pi$, (cf., Appendix B, Eq. (12)). Note that the parametrization $(x, \gamma, \delta) \mapsto \psi$ degenerates at $x = 1/2$ and at $\gamma = 0, \pi/2$. In [15] [16], other stratifications of $\mathbb{C}P^3$ based on different $SU(2)$ actions are presented.

We are interested in the relation between entanglement and the geometry of $SU(2)$ orbits as Riemannian submanifolds of $\mathbb{C}P^3$. Since the orbits that correspond to the same degree x of entanglement are isometric to each other, we choose representative orbits by setting $\gamma = 0$ in the orbit space $\mathcal{I}$. The corresponding orbits are given by Eq. (4).

The components g_{ij} of the metric g on the orbit (4) with respect to the coordinate system θ are given by the SLD Fisher information matrix $J_\theta(x)$, Eq. (5), and the volume element is

$$\omega := \sqrt{\det J_\theta(x)}\, d\phi d\alpha d\beta = 8\sqrt{x(1-x)} \sin 2\phi \; d\phi d\alpha d\beta.$$

This simple formula already offers some information about the relation between entanglement and the geometry of orbits: an orbit maximally inflates at $x = 1/2$, and collapses as $x \to 0$. Note that the scaling factor $\sqrt{x(1-x)}$ is identical, up to a constant factor, to the concurrence [9] [15].

In order to get full information about the global structure of the orbits, we compute the Riemannian curvature R of the Levi-Civita connection. Let $R_{ijkl} := g(R(\partial_i, \partial_j)\partial_k, \partial_l)$ denote the components. (The readers may be warned not to confuse the order of indices with that used in a standard book of differential geometry such as [17]. We follow the book [11].) Due to the symmetries $R_{ijkl} = -R_{jikl} = -R_{ijlk} = R_{klij}$, there are at most 36 nonvanishing components: six independent components R_{1212}, R_{1313}, R_{2323}, R_{1213}, R_{1223}, R_{1323}, and those which are obtained by permuting indices. By a direct computation, they are given by

$$\begin{aligned}
R_{1212} &= -4\cos^2\phi\,[1 + (1-2x)^2\{3 - 4(1 + x - x^2)\cos^2\phi\}], \\
R_{1313} &= -4\sin^2\phi\,[1 + (1-2x)^2\{3 - 4(1 + x - x^2)\sin^2\phi\}], \\
R_{2323} &= -64x^2(1-x)^2\cos^2\phi\sin^2\phi, \\
R_{1213} &= -16(1-2x)(1 + x - x^2)\cos^2\phi\sin^2\phi, \\
R_{1223} &= R_{1323} = 0.
\end{aligned}$$

The components $R_{jk} := R_{ijk}{}^{i} = R_{ijkl}g^{li}$ of the Ricci curvature Ric then becomes

$$\begin{aligned}
R_{11} &= 4(1 - 2x + 2x^2), \\
R_{22} &= 2\cos^2\phi\,[1 + (1-2x)^2\{1 - 2(1 + 2x - 2x^2)\cos^2\phi\}], \\
R_{33} &= 2\sin^2\phi\,[1 + (1-2x)^2\{1 - 2(1 + 2x - 2x^2)\sin^2\phi\}], \\
R_{12} &= R_{21} = R_{13} = R_{31} = 0, \\
R_{23} &= R_{32} = 4(1-2x)^2(1 + 2x - 2x^2)\cos^2\phi\sin^2\phi.
\end{aligned}$$

It is easy to show that the orbit is Einstein (i.e., $Ric = \lambda g$ for a constant λ) if and only if $x = 1/2$ or $x = 0$. The scalar curvature $\rho := R_j{}^{j} = R_{jk}g^{kj} = 2(1 - x + x^2)$ indicates that the larger the parameter x ($\in [0, 1/2]$) is, the "flatter" the orbit becomes on average. Let us take a closer look at this point.

The sectional curvature with respect to the subspace spanned by $\{\partial_1, \partial_2\}$ is given by

$$\frac{R_{1212}}{(g_{12})^2 - g_{11}g_{22}} = 1 + x - x^2 - \frac{8x(1-x)}{1 + 4x(1-x) - (1-2x)^2\cos 2\phi}.$$

This is independent of ϕ if and only if $x = 1/2$ or $x = 0$. When $x = 1/2$, the orbit turns out to be a space of constant positive curvature 1/4, in that

$$R_{ijkl} = \frac{1}{4}(g_{jk}\, g_{il} - g_{ik}\, g_{jl}),$$

for all $i, j, k, l = 1, 2, 3$. (This is confirmed either by a direct computation, or by the fact that the orbit is a 3-dimensional Einstein manifold [17, p. 293].) Since the fundamental group of the orbit is $\mathbb{Z}_2$, it is the quotient $S^3(2)/\{\pm I\}$, (cf., [18]), i.e., the 3-dimensional real projective space $\mathbb{R}P^3(2)$ of radius 2. It is also important to observe that for $0 < x < 1/2$, the orbit is not of constant curvature and hence is *not isometric* (though diffeomorphic) to $S^3(r)/\{\pm I\}$ for any $r > 0$. Since the manifold $SU(2)$ equipped with the Cartan-Killing metric is isometric to S^3, these facts could be paraphrased by saying that the "shape" of the Riemannian manifold $SU(2)/\{\pm I\} \cong SO(3)$, the coordinates of which are to be estimated, comes into full view only through the $SU(2)$ action $\psi \mapsto (U \otimes I)\psi$ on a maximally entangled ψ. This gives a geometric insight into Theorems 3 and 4.

When $x = 0$, on the other hand, the orbit collapses to a lower dimensional manifold in which $\partial_2 = \partial_3$. In this case, the only independent component R_{1212} of the Riemannian curvature tensor satisfies $R_{1212} = (g_{12})^2 - g_{11}g_{22}$. Namely, the collapsed manifold is a space of constant positive curvature 1. Since the manifold is simply connected, it is the 2-dimensional sphere S^2 of unit radius. This is, of course, in accordance with the known isomorphism between $\mathbb{C}P^1$ and S^2.

In summary we have

Theorem 6 *The Riemannian manifold $SU(2)/\{\pm I\} \cong SO(3)$ is isometrically embedded into $\mathbb{C}P^3$ as an $SU(2)$ orbit M_ψ if and only if ψ is a maximally entangled state.*

5 Discussions

We have studied a quantum statistical estimation problem of operators $U \in SU(2)$ acting on $\partial_e\mathcal{S}(\mathbb{C}^2 \otimes \mathbb{C}^2) \simeq \mathbb{C}P^3$ as $\psi \mapsto (U \otimes I)\psi$. It was shown that the quantum Cramér-Rao lower bound takes the minimum, and is achievable, if and only if ψ is a maximally entangled state (Theorems 3, 4), and that an $SU(2)$ orbit on $\mathbb{C}P^3$ equipped with the SLD Fisher metric is isometric to $SU(2)/\{\pm I\} \cong SO(3)$ if and only if ψ is a maximally entangled state (Theorem 6).

The information geometric study of $SU(2)$ orbits presented in Section 4 has clarified what happens when the degree x of entanglement varies: as x increases toward 1/2, the orbit inflates and hence points on the orbit are getting separated from each other. This is the geometric mechanism behind the estimation theoretic Theorem 3. In fact, the larger the SLD distance of two nearby quantum states becomes, the easier one can distinguish these states, as the quantum Cramér-Rao inequality asserts. Theorem 4, on the other hand, concerns the existence of a set of simultaneously measurable observables as an estimator for the 3-dimensional parameter of $SU(2)$. More precisely, it asserts that the noncommutative nature of the $SU(2)$ parameters "disappears" when (and only when) we use a maximally entangled state as the input. Theorem 4 can also be viewed as providing an "operational" characterization of the otherwise inaccessible quantity of the Fubini-Study metric tensor.

Finally we touch upon a generalization to $SU(n)$. For the achievability of the Cramér-Rao lower bound, a result analogous to Theorem 4 holds for all n. In fact, the only essential ingredient of the proof is that elements of the Lie algebra $su(n)$ have trace zero. On the other hand, the maximality of the SLD Fisher metric analogous to Lemma 2 does not hold for $n \geq 3$. This fact suggests an essential role of the dimensionality. A detailed analysis of the statistical estimation of $SU(n)$, as well as the proofs of the above facts, will be presented in a subsequent paper.

Appendices

A Estimation of pure states

This appendix gives a brief account of the parameter estimation theory for a finite dimensional pure state model. For more information, see [12] [13] [14] [19] [20]. Suppose an unknown quantum state lies in a parametric family $\{\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta| \ ; \ \|\psi_\theta\| = 1, \ \theta = (\theta^1, ..., \theta^d) \in \Theta \subset \mathbb{R}^d\}$ of pure states on a finite dimensional complex Hilbert space $\mathcal{H}$. The problem is to estimate, by means of a certain measurement, the true value of the parameter θ. We assume that the parametrization $\theta \mapsto \rho_\theta$ is smooth and nondegenerate. An estimator for the parameter θ is given by a pair $(M, \hat{\theta})$, where $M = \{M(x); x \in \mathcal{X}\}$ is a positive operator valued measure that takes values on a finite set $\mathcal{X}$, and $\hat{\theta} : \mathcal{X} \to \Theta$ is a map that gives an estimate of θ from a measurement outcome x. In the quantum estimation theory, we often assume the local unbiasedness condition on estimators [6]: an estimator $(M, \hat{\theta})$ for the parameter θ is called *locally unbiased* at $\theta = \theta_0$, or θ_0-*unbiased* for short, if the unbiasedness condition

$$\sum_{x \in \mathcal{X}} \hat{\theta}^i(x) \ \mathrm{Tr}\rho_\theta M(x) = \theta^i, \qquad (i = 1, ..., d),$$

and its differentiation

$$\sum_{x \in \mathcal{X}} \hat{\theta}^i(x) \ \mathrm{Tr}(\partial_j \rho_\theta) M(x) = \delta^i_j, \qquad (i, j = 1, ..., d),$$

hold at $\theta = \theta_0$. Clearly an estimator is unbiased if and only if it is locally unbiased at every $\theta \in \Theta$.

The performance of an estimator is usually evaluated by the covariance matrix. When the actual quantum state is $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, the covariance matrix $V_\theta[M, \hat{\theta}] = [v_\theta^{ij}]$ for a θ-unbiased estimator $(M, \hat{\theta})$ is defined by

$$v_\theta^{ij} := \sum_{x \in \mathcal{X}} (\hat{\theta}^i(x) - \theta^i)(\hat{\theta}^j(x) - \theta^j) \ \mathrm{Tr}\rho_\theta M(x).$$

The smaller the covariance matrix is, the more accurately one can estimate the parameter θ.

One of the most important notion in the quantum estimation theory is the *symmetric logarithmic derivative* (SLD): given a model $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$, the SLD with respect to θ^i, $(i = 1, ..., d)$, is defined by the Hermitian operator $L_{\theta,i}$ that satisfies the equation

$$\partial_i \rho_\theta = \frac{1}{2}(L_{\theta,i}\rho_\theta + \rho_\theta L_{\theta,i}).$$

Since the SLD is not unique for a pure state model, it is convenient to work with the vector $\ell_{\theta,i} := 2(\partial_i \rho_\theta)\psi_\theta$, which is identical to $L_{\theta,i}\psi_\theta$ for any representative $L_{\theta,i}$ of the SLD. While the vectors $\{\ell_{\theta,1}, ..., \ell_{\theta,d}\}$ are not always $\mathbb{C}$-linearly independent, they are $\mathbb{R}$-linearly independent (due to the nondegeneracy of the parametrization $\theta \mapsto \rho_\theta$). Moreover, since $\langle\psi_\theta|\ell_{\theta,i}\rangle = 0$, the vectors $\{\psi_\theta, \ell_{\theta,1}, ..., \ell_{\theta,d}\}$ are also $\mathbb{R}$-linearly independent. The positive definite matrix $J_\theta := [\mathrm{Re}\,\langle\ell_{\theta,i}|\ell_{\theta,j}\rangle]$ is called the *SLD Fisher information matrix*. For later convenience, we introduce the dual vectors

$\{\ell_\theta^i\}_{i=1}^d$ of $\{\ell_{\theta,i}\}_{i=1}^d$ by $\ell_\theta^i := J_\theta^{ij}\ell_{\theta,j}$, where J_θ^{ij} is the (i,j)th entry of the inverse matrix J_θ^{-1}, and Einstein's summation convention is used. Note that $\mathrm{Re}\,\langle \ell_\theta^i | \ell_\theta^j \rangle = J_\theta^{ij}$.

Associated with a θ-unbiased estimator $(M,\hat{\theta})$ are the vectors

$$\xi_\theta^i := \sum_{x\in\mathcal{X}} (\hat{\theta}^i(x) - \theta^i)\, M(x)\psi_\theta, \qquad (i = 1, ..., d).$$

Due to the θ-unbiasedness, they satisfy

$$\langle \xi_\theta^i | \psi_\theta \rangle = 0, \qquad \mathrm{Re}\,\langle \xi_\theta^i | \ell_{\theta,j} \rangle = \delta_j^i, \qquad (i,j = 1, ..., d). \tag{7}$$

We denote by $\boldsymbol{X}_\theta$ the ordered list $[\xi_\theta^1, ..., \xi_\theta^d]$ of vectors, and by $\boldsymbol{X}_\theta^*\boldsymbol{X}_\theta$ the $d\times d$ matrix whose (i,j)th entry is $\langle x_\theta^i | x_\theta^j \rangle$. (We take $\boldsymbol{X}_\theta$ for a "matrix" whose ith column is the vector $|\xi_\theta^i\rangle$.) Now by a standard argument [6, p.88, p.274], we have

$$V_\theta[M,\hat{\theta}] \geq \boldsymbol{X}_\theta^*\boldsymbol{X}_\theta. \tag{8}$$

and

$$V_\theta[M,\hat{\theta}] \geq \mathrm{Re}\,\boldsymbol{X}_\theta^*\boldsymbol{X}_\theta \geq J_\theta^{-1}. \tag{9}$$

Here $\mathrm{Re}\,\boldsymbol{X}_\theta^*\boldsymbol{X}_\theta$ is the matrix whose (i,j)th entry is $\mathrm{Re}\,\langle x_\theta^i | x_\theta^j \rangle$. (The matrix $\mathrm{Im}\,\boldsymbol{X}_\theta^*\boldsymbol{X}_\theta$ is defined likewise.) The inequality (9) is called the *SLD Cramér-Rao inequality*.

We now focus on achievability of the SLD Cramér-Rao lower bound at a given θ. In what follows, we shall drop the subscript θ for notational simplicity.

Lemma 7 $\mathrm{Re}\,\boldsymbol{X}^*\boldsymbol{X} = J^{-1}$ *if and only if* $\xi^i = \ell^i$ *for all* i.

Proof The second relation in (7) with $i = j$, which is equivalent to $\mathrm{Re}\,\langle \xi^i - \ell^i | \ell^i \rangle = 0$, implies that for each i, the vectors $\{\xi^i, \ell^i, \xi^i - \ell^i\}$ form a right triangle (ξ^i being the hypotenuse) with respect to the real inner product $\mathrm{Re}\,\langle\cdot|\cdot\rangle$. Then by the Pythagoras theorem, we see that $\mathrm{Re}\,\langle \xi^i | \xi^i \rangle = \mathrm{Re}\,\langle \ell^i | \ell^i \rangle$ if and only if $\xi^i = \ell^i$. □

Lemma 8 $V[M,\hat{\theta}] = \mathrm{Re}\,\boldsymbol{X}^*\boldsymbol{X}$ *implies* $\mathrm{Im}\,\boldsymbol{X}^*\boldsymbol{X} = 0$.

Proof If $V[M,\hat{\theta}] = \mathrm{Re}\,\boldsymbol{X}^*\boldsymbol{X}$, then $\mathrm{Re}\,\boldsymbol{X}^*\boldsymbol{X} \geq \boldsymbol{X}^*\boldsymbol{X}$ by (8) and (9), that is, $0 \geq i\,\mathrm{Im}\,\boldsymbol{X}^*\boldsymbol{X}$. Since $\mathrm{Im}\,\boldsymbol{X}^*\boldsymbol{X}$ is a real skew-symmetric matrix, this implies that $\mathrm{Im}\,\boldsymbol{X}^*\boldsymbol{X} = 0$. □

Motivated by (7), we say that a collection of vectors $\{\eta^1, ..., \eta^d\}$ in $\mathcal{H}$ is *θ-unbiased* if

$$\langle \eta^i | \psi \rangle = 0, \qquad \mathrm{Re}\,\langle \eta^i | \ell_j \rangle = \delta_j^i, \qquad (i,j = 1, ..., d).$$

In this case, the vectors $\{\eta^1, ..., \eta^d\}$ are necessarily $\mathbb{R}$-linearly independent. The next lemma, due to Matsumoto [20], subsumes the converse of Lemma 8.

Lemma 9 *Suppose a collection of vectors* $\{\eta^1, ..., \eta^d\}$ *is θ-unbiased and satisfies the condition*

$$\mathrm{Im}\,\boldsymbol{Y}^*\boldsymbol{Y} = 0,$$

where $\boldsymbol{Y} = [\eta^1, ..., \eta^d]$. *Then there is a projection valued measure* $E = \{E(x); x \in \mathcal{X}\}$ *and real numbers* $\{a^i(x); x \in \mathcal{X}, i = 1, ..., d\}$ *such that*

$$\eta^i = \sum_{x \in \mathcal{X}} a^i(x) E(x) \psi, \qquad (i = 1, ..., d). \tag{10}$$

In particular, letting $\hat{\theta}^i(x) := \theta^i + a^i(x)$, *the pair* $(E, \hat{\theta})$ *forms a* θ*-unbiased estimator that satisfies* $V[E, \hat{\theta}] = \operatorname{Re} \boldsymbol{Y}^* \boldsymbol{Y}$.

Proof Since $\operatorname{Im} \boldsymbol{Y}^*\boldsymbol{Y} = 0$, the Gram matrix $\boldsymbol{Y}^*\boldsymbol{Y}$ with respect to the complex inner product $\langle \cdot | \cdot \rangle$ is identical to the Gram matrix $\operatorname{Re} \boldsymbol{Y}^*\boldsymbol{Y}$ with respect to the real inner product $\operatorname{Re} \langle \cdot | \cdot \rangle$, and is positive definite because $\{\eta^1, ..., \eta^d\}$ are $\mathbb{R}$-linearly independent. This implies that $\{\eta^1, ..., \eta^d\}$ are $\mathbb{C}$-linearly independent. Moreover, since $\langle \eta^i | \psi \rangle = 0$, the vectors $\{\psi, \eta^1, ..., \eta^d\}$ are also $\mathbb{C}$-linearly independent.

Let $V := \operatorname{Span}_{\mathbb{C}} \{\psi, \eta^1, ..., \eta^d\}$. Since $\langle \eta^i | \psi \rangle$ and $\langle \eta^i | \eta^j \rangle$ are all real, there is an orthonormal basis $\{e_1, ..., e_{d+1}\}$ of V such that $\langle e_k | \psi \rangle$ and $\langle e_k | \eta^i \rangle$ are all real, and that $\langle e_k | \psi \rangle \neq 0$ for all k. (To find such a basis, one first performs the Gram-Schmidt procedure on $\{\psi, \eta^1, ..., \eta^d\}$, and then rotates the basis slightly to meet the condition $\langle e_k | \psi \rangle \neq 0$.) Then letting $E(k) := |e_k\rangle\langle e_k|$ and $a^i(k) := \langle e_k | \eta^i \rangle / \langle e_k | \psi \rangle$, we have

$$\eta^i = \sum_{k=1}^{d+1} a^i(k) E(k) \psi.$$

If $V = \mathcal{H}$ then let $\mathcal{X} = \{1, ..., d+1\}$, otherwise let $\mathcal{X} = \{0, 1, ..., d+1\}$, $E(0)$ the projection onto $V^\perp$, and $a^i(0) = 0$. Then the projection valued measure $E = \{E(x)\}_{x \in \mathcal{X}}$ and real numbers $\{a^i(x); x \in \mathcal{X}, i = 1, ..., d\}$ satisfy (10).

Let $\hat{\theta}^i(k) := \theta^i + a^i(k)$. Then the estimator $(E, \hat{\theta})$ is θ-unbiased, and

$$\langle \eta^i | \eta^j \rangle = \sum_k \sum_l a^i(k) a^j(l) \langle E(k)\psi | E(l)\psi \rangle = \sum_k (\hat{\theta}^i(k) - \theta^i)(\hat{\theta}^j(k) - \theta^j) \langle \psi | E(k)\psi \rangle$$

is the (i, j)th entry of the covariance matrix $V[E, \hat{\theta}]$, proving that $V[E, \hat{\theta}] = \boldsymbol{Y}^*\boldsymbol{Y} = \operatorname{Re} \boldsymbol{Y}^*\boldsymbol{Y}$. $\square$

We say that the SLD lower bound (9) is *achievable* at θ if there is a θ-unbiased estimator $(M, \hat{\theta})$ for which $V[M, \hat{\theta}] = J^{-1}$ holds. The next corollary is also due to Matsumoto [20].

Corollary 10 *The SLD lower bound is achievable if and only if* $\langle \ell_i | \ell_j \rangle$ *are all real.*

Proof We first note that $\langle \ell_i | \ell_j \rangle$ are all real if and only if $\langle \ell^i | \ell^j \rangle$ are all real. Assume first that $V[M, \hat{\theta}] = J^{-1}$ for a certain θ-unbiased estimator $(M, \hat{\theta})$. Then by Lemmas 7, 8, and (9), we have $\operatorname{Im} \langle \ell^i | \ell^j \rangle = 0$ for all i, j. Assume next that $\langle \ell^i | \ell^j \rangle$ are all real. Since the collection $\{\ell^1, ..., \ell^d\}$ is θ-unbiased, we see from Lemma 9 that there is an estimator $(E, \hat{\theta})$ that satisfies $V[E, \hat{\theta}] = [\operatorname{Re} \langle \ell^i | \ell^j \rangle] = J^{-1}$. $\square$

B Characterization of $SU(2)$ orbit space

Let ψ and $\hat{\psi}$ be unit vectors on $\mathbb{C}^2 \otimes \mathbb{C}^2$, and let their Schmidt decompositions be

$$\psi = \sqrt{1-x}\, e_1 \otimes f_1 + \sqrt{x}\, e_2 \otimes f_2, \qquad \left(0 \leq x \leq \frac{1}{2}\right),$$

and

$$\hat{\psi} = \sqrt{1-\hat{x}}\, \hat{e}_1 \otimes \hat{f}_1 + \sqrt{\hat{x}}\, \hat{e}_2 \otimes \hat{f}_2, \qquad \left(0 \le \hat{x} \le \frac{1}{2}\right).$$

These vectors are equivalent ($\psi \sim \hat{\psi}$) if there is a $U \in SU(2)$ such that $|\hat{\psi}\rangle\langle\hat{\psi}| = |(U \otimes I)\psi\rangle\langle(U \otimes I)\psi|$. We claim

Lemma 11 *$\psi \sim \hat{\psi}$ if and only if either $x = \hat{x} = 1/2$, or $x = \hat{x}$ and $\hat{f}_i = \lambda_i f_i$ for some $\lambda_i \in \mathbb{C}$ of unit modulus, $(i = 1, 2)$.*

Proof Since we are dealing with the $SU(2)$ action $\psi \mapsto (U \otimes I)\psi$, we set, without loss of generality, as

$$e_1 = \hat{e}_1 = f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad e_2 = \hat{e}_2 = f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

It suffices to show that for

$$\hat{f}_1 = \begin{bmatrix} e^{i\alpha_0} \cos\phi_0 \\ e^{-i\beta_0} \sin\phi_0 \end{bmatrix}, \qquad \hat{f}_2 = \begin{bmatrix} -e^{i\beta_0} \sin\phi_0 \\ e^{-i\alpha_0} \cos\phi_0 \end{bmatrix}, \qquad (0 \le \phi_0 \le \frac{\pi}{2},\ 0 \le \alpha_0, \beta_0 < 2\pi),$$

$\psi \sim \hat{\psi}$ if and only if either $x = \hat{x} = 1/2$, or $x = \hat{x}$ and $\phi_0 = 0$.

We first show the 'if' part. If $x = \hat{x} = 1/2$, let $(\phi, \alpha, \beta) = (\phi_0, \alpha_0 + \pi \pmod{2\pi}, 2\pi - \beta_0)$, else if $x = \hat{x}$ and $\phi_0 = 0$, let $(\phi, \alpha, \beta) = (0, \alpha_0, 0)$. Then by a routine calculation, we have $|(U \otimes I)\psi\rangle\langle(U \otimes I)\psi| = |\hat{\psi}\rangle\langle\hat{\psi}|$ for $U = U_{(\phi,\alpha,\beta)}$.

We next show the 'only if' part. Since the $SU(2)$ action does not change the singular values of a Schmidt decomposition, $\psi \sim \hat{\psi}$ implies $x = \hat{x}$. As we have already shown that the equation $|(U \otimes I)\psi\rangle\langle(U \otimes I)\psi| = |\hat{\psi}\rangle\langle\hat{\psi}|$ with $x = \hat{x} = 1/2$ has a solution $U \in SU(2)$ for any $\{\hat{f}_i\}_i$, we need only consider the case when $x = \hat{x} \neq 1/2$. By a direct computation, we see that the equation with $x = \hat{x} = 0$ implies $\phi = \phi_0 = 0$, while the equation with $0 < x = \hat{x} < 1/2$ implies $\phi = \phi_0 = 0$ and $\alpha = \alpha_0$. The claim was verified. □

Let us specify a complete list of equivalence classes explicitly. Let the orthonormal frames $\{e_i\}_i$, $\{\hat{e}_i\}_i$, and $\{\hat{f}_i\}_i$ be as in the above proof, and let

$$f_1 = \begin{bmatrix} e^{i\alpha_1} \cos\phi_1 \\ e^{-i\beta_1} \sin\phi_1 \end{bmatrix}, \qquad f_2 = \begin{bmatrix} -e^{i\beta_1} \sin\phi_1 \\ e^{-i\alpha_1} \cos\phi_1 \end{bmatrix}, \qquad (0 \le \phi_1 \le \frac{\pi}{2},\ 0 \le \alpha_1, \beta_1 < 2\pi).$$

We denote $\{f_1, f_2\} \sim \{\hat{f}_1, \hat{f}_2\}$ if $\psi \sim \hat{\psi}$. Then by Lemma 11, $\{f_1, f_2\} \sim \{\hat{f}_1, \hat{f}_2\}$ if and only if either $x = \hat{x} = 1/2$, or $x = \hat{x}$ and

$$[\hat{f}_1, \hat{f}_2] = [f_1, f_2] \begin{bmatrix} e^{i\mu} & 0 \\ 0 & e^{-i\mu} \end{bmatrix}, \qquad (\exists \mu \in \mathbb{R}). \tag{11}$$

When $x = \hat{x} \neq 1/2$, the equation (11) characterizes all the frames $\{\hat{f}_1, \hat{f}_2\}$ that are equivalent to $\{f_1, f_2\}$, and the solution is as follows. If $\phi_1 = 0$, then $\phi_0 = 0$, $\alpha_0 = \alpha_1 + \mu \pmod{2\pi}$, and β_0 arbitrary; if $\phi_1 = \pi/2$, then $\phi_0 = \pi/2$, $\beta_0 = \beta_1 - \mu \pmod{2\pi}$, and α_0 arbitrary; if $0 < \phi_1 < \pi/2$, then $\phi_0 = \phi_1$, $\alpha_0 = \alpha_1 + \mu \pmod{2\pi}$, and $\beta_0 = \beta_1 - \mu \pmod{2\pi}$. A representative for the first case is given by $(\phi_0, \alpha_0, \beta_0) = (0, 0, 0)$, i.e.,

$$f_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad f_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

for the second, $(\phi_0, \alpha_0, \beta_0) = (\pi/2, 0, 0)$, i.e.,

$$f_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad f_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix},$$

and for the third, $(\phi_0, \alpha_0, \beta_0) = (\phi_1, 0, \beta_1 + \alpha_1 \pmod{2\pi})$, i.e.,

$$f_1 = \begin{bmatrix} \cos\gamma \\ e^{-i\delta}\sin\gamma \end{bmatrix}, \qquad f_2 = \begin{bmatrix} -e^{i\delta}\sin\gamma \\ \cos\gamma \end{bmatrix}, \qquad (0 < \gamma < \frac{\pi}{2},\ 0 \le \delta < 2\pi).$$

In summary, a complete list of representatives of equivalence classes is as follows: For $x = 1/2$,

$$\psi = \frac{1}{\sqrt{2}}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right),$$

for $0 \le x < 1/2$ and $0 \le \gamma < \pi/2$,

$$\psi = \sqrt{1-x}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} \cos\gamma \\ e^{-i\delta}\sin\gamma \end{bmatrix} + \sqrt{x}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -e^{i\delta}\sin\gamma \\ \cos\gamma \end{bmatrix}, \qquad (0 \le \delta < 2\pi)$$

and for $0 \le x < 1/2$ and $\gamma = \pi/2$

$$\psi = \sqrt{1-x}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \sqrt{x}\begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

The above family shows a discontinuity at $x = 1/2$ and at $\gamma = \pi/2$. By a slight modification, however, we obtain a complete list of representatives that forms a 3-dimensional smooth compact submanifold of $\mathbb{C}P^3$:

$$\psi = \sqrt{1-x}\begin{bmatrix} \cos\gamma \\ e^{i\delta}\sin\gamma \end{bmatrix} \otimes \begin{bmatrix} \cos\gamma \\ e^{-i\delta}\sin\gamma \end{bmatrix} + \sqrt{x}\begin{bmatrix} -e^{-i\delta}\sin\gamma \\ \cos\gamma \end{bmatrix} \otimes \begin{bmatrix} -e^{i\delta}\sin\gamma \\ \cos\gamma \end{bmatrix}, \tag{12}$$

where $0 \le x \le 1/2$, $0 \le \gamma \le \pi/2$, and $0 \le \delta < 2\pi$. Note that the parametrization degenerates at $x = 1/2$ and at $\gamma = 0, \pi/2$.

References

[1] C. H. Bennett and S. J. Wiesner, "Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states," Phys. Rev. Lett. **69**, pp. 2881-2884 (1992).

[2] S. Bose, M. B. Plenio, and V. Vedral, "Mixed state dense coding and its relation to entanglement measures," J. Mod. Opt. **47**, pp. 291-310 (2000); A. M. Childs, J. Preskill, and J. Renes, "Quantum information and precision measurement," J. Mod. Opt. **47**, pp. 155-176 (2000).

[3] A. S. Holevo, "The capacity of the quantum channel with general signal states," IEEE Trans. Inform. Theory **44**, pp. 269-273 (1998); B. Schumacher and M. D. Westmoreland, "Sending classical information via noisy quantum channels," Phys. Rev. A **56**, pp. 131-138 (1997); A. Winter, *Coding theorems of quantum information theory*, Ph. D. Thesis, University of Bielefeld (1999), quant-ph/9907077; T. Ogawa, *A study on the asymptotic property of the hypothesis testing and the channel coding in quantum mechanical systems*, Ph. D. Thesis, The University of Electro-Communications (2000).

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

[5] F. Hiai and D. Petz, “The proper formula for relative entropy and its asymptotics in quantum probablility,” Commun. Math. Phys. **143**, pp. 99-114 (1991); T. Ogawa and H. Nagaoka, “Strong converse and Stein’s lemma in quantum hypothesis testing,” IEEE Trans. Inform. Theory **46**, 2428-2433 (2000).

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

[7] A. Fujiwara, “Quantum channel identification problem,” Phys. Rev. A **63**, 042304 (2001).

[8] J. Fiurášek and Z. Hradil, “Maximum-likelihood estimation of quantum processes,” Phys. Rev. A **63**, 020101 (2001); M. F. Sacchi, “Maximum-likelihood reconstruction of completely positive maps,” Phys. Rev. A **63**, 054104 (2001).

[9] W. K. Wootters, “Entanglement of formation of an arbitrary state of two qubits,” Phys. Rev. Lett. **80**, pp. 2245-2248 (1998).

[10] R. Bhatia, *Matrix Analysis*, GTM 169 (Springer, New York, 1997).

[11] S. Amari and H. Nagaoka, *Methods of Information Geometry* (AMS, Providence, 2000).

[12] A. Fujiwara, *A geometrical study in quantum information systems*, Ph. D. Thesis, The University of Tokyo (1995).

[13] A. Fujiwara, “Geometry of Quantum Information Systems,” in *Geometry in Present Day Science*, edited by O.E. Barndorff-Nielsen and E.B.V. Jensen (World Scientific, Singapore, 1999), pp. 35-48.

[14] A. Fujiwara and H. Nagaoka, “Quantum Fisher metric and estimation for pure state models,” Phys. Lett. A **201**, pp. 119-124 (1995).

[15] M. Kuś and K. Życzkowski, “Geometry of entangled states,” Phys. Rev. A **63**, 032307 (2001).

[16] N. Barros e Sá, “Decomposition of Hilbert space in sets of coherent states,” J. Phys. A: Math. Gen. **34**, pp. 4831-4850 (2001).

[17] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry, vol. I* (Interscience, New York, 1963).

[18] J. A. Wolf, *Spaces of Constant Curvature*, 5th ed. (Publish or Perish, Wilmington, 1984).

[19] A. Fujiwara and H. Nagaoka, “An estimation theoretical characterization of coherent states,” J. Math. Phys. **40**, pp. 4227-4239 (1999).

[20] K. Matsumoto, “A new approach to the Cramer-Rao type bound of the pure state model,” quant-ph/9711008.