Geometry of optimal estimation scheme for $SU(D)$ channels

Hiroshi Imai[*] and Akio Fujiwara[†]
Department of Mathematics, Osaka University
Toyonaka, Osaka 560-0043, Japan

Abstract

The problem of estimating an unknown $SU(D)$ channel $\Gamma_U : \rho \mapsto U\rho U^*$ is studied based on the quantum Cramér-Rao inequality. It is shown that the minimum estimation error is of $O(1/n^2)$, where n is the degree of extension of the channel. The mechanism behind this asymptotic behavior is investigated from a differential geometrical point of view.

1 Introduction

This paper deals with the problem of estimating an unknown unitary channel Γ_U acting on the set $\mathcal{S}(\mathcal{H})$ of density operators on a Hilbert space $\mathcal{H} \simeq \mathbb{C}^D$ as $\Gamma_U : \rho \mapsto U\rho U^*$, where $U \in SU(D)$. In particular, we investigate the optimal estimation scheme using the extension $(\mathrm{id} \otimes \Gamma_U)^{\otimes n} : \mathcal{S}((\mathcal{H} \otimes \mathcal{H})^{\otimes n}) \to \mathcal{S}((\mathcal{H} \otimes \mathcal{H})^{\otimes n})$, where n is an arbitrary positive integer.

Due to its obvious group covariant structure, the problem has been studied in a Bayesian framework [1, 2, 3, 4], using a covariant cost function averaged over $SU(D)$ with respect to the uniform prior distribution (i.e., the Haar measure). In contrast, our approach is a local one based on the quantum Cramér-Rao inequality. Such a local approach, the validity of which has been established in [5], has an advantage that it allows a direct comparison of estimation performances among various classes of quantum channels which do not necessarily possess *a priori* distributions such as the generalized Pauli channels [6]. It also allows us to invoke differential geometrical methods [7] in studying the roles of the quantum entanglement and the degree n of extension.

The paper is organized as follows. We summarize the main results in Section 2, and prove them in Section 3. In Section 4, we recast the main results from a differential geometrical point of view. In Section 5, we give brief concluding remarks, and further remarks on the admissibility of an input state are presented in Appendix A.

[*] himai@gaia.math.wani.osaka-u.ac.jp
[†] fujiwara@math.wani.osaka-u.ac.jp

2 Main Results

Let us introduce a local coordinate system $\theta = (\theta^1, \ldots, \theta^{D^2-1})$ of $SU(D)$ around a point U_0 by the exponential map:

$$U_\theta = U_0 \,\mathrm{Exp}\left(\sqrt{-1}\sum_{i=1}^{D^2-1}\theta^i X_i\right), \tag{1}$$

where $\{\sqrt{-1}\,X_i\}_{1\le i\le D^2-1}$ is a basis of Lie algebra $su(D)$ satisfying $\mathrm{Tr}X_iX_j = \frac{1}{2}\delta_{ij}$. By a suitable rearrangement of the constituent Hilbert spaces $\mathcal{H}$, we identify $(\mathrm{id}\otimes\Gamma_{U_\theta})^{\otimes n}$ with $\mathrm{id}^{\otimes n}\otimes\Gamma_{U_\theta}^{\otimes n}$. Once an input state $\psi^{(n)} \in \mathcal{H}^{\otimes n}\otimes\mathcal{H}^{\otimes n}$ is fixed, we have a quantum statistical model

$$\rho_\theta := (\mathrm{id}^{\otimes n}\otimes\Gamma_{U_\theta}^{\otimes n})\,(|\psi^{(n)}\rangle\langle\psi^{(n)}|),$$

and the problem of estimating the unknown unitary operation $U_\theta \in SU(D)$ is reduced to estimating the parameter θ of the model ρ_θ.

Let us decompose $\mathcal{H}^{\otimes n}$ into irreducible subspaces under the $SU(D)$ action as follows:

$$\mathcal{H}^{\otimes n} = \bigoplus_{\lambda}\left(\bigoplus_{[\lambda]\in\mathrm{STab}\,(\lambda)}\mathcal{H}^{[\lambda]}\right),$$

where λ runs over all possible Young frames (or Dynkin indices) and $\mathrm{STab}\,(\lambda)$ stands for the set of standard tableaux on λ. Then

$$\mathcal{H}^{\otimes n}\otimes\mathcal{H}^{\otimes n} = \bigoplus_{\lambda}\left(\bigoplus_{[\lambda]\in\mathrm{STab}\,(\lambda)}\mathcal{H}^{\otimes n}\otimes\mathcal{H}^{[\lambda]}\right).$$

Given an input state $\psi^{(n)} \in \mathcal{H}^{\otimes n}\otimes\mathcal{H}^{\otimes n}$, let us decompose it as

$$\psi^{(n)} = \sum_{\lambda}\sum_{\mathrm{STab}(\lambda)} a^{[\lambda]}\,\psi^{[\lambda]}, \tag{2}$$

where $\psi^{[\lambda]}$ is a unit vector on the invariant subspace $\mathcal{H}^{\otimes n}\otimes\mathcal{H}^{[\lambda]}$, and the coefficients $a^{[\lambda]}$ satisfy the normalization

$$\sum_{\lambda}\sum_{\mathrm{STab}(\lambda)}|a^{[\lambda]}|^2 = 1.$$

Associated with the quantum statistical model ρ_θ is the symmetric logarithmic derivative (SLD) Fisher metric g [7], which will also be denoted as $g_{\psi^{(n)}}$ when the input state $\psi^{(n)}$ needs to be specified. The SLD Fisher metric g is a measure of statistical distinguishability, and is one of the most fundamental quantity in quantum estimation theory. In fact, it is related to the quantum Cramér-Rao inequality [8, 9]

$$V_\theta[M^{(n)}|\psi^{(n)}] \ge \left(J_\theta[\psi^{(n)}]\right)^{-1}, \tag{3}$$

where $V_\theta[M^{(n)}|\psi^{(n)}]$ is the covariance matrix of the locally unbiased estimator (POVM) $M^{(n)}$ for the parameter θ when the input state is $\psi^{(n)}$, and $J_\theta[\psi^{(n)}]$ is the SLD Fisher information matrix, i.e., the representation of the SLD Fisher metric g by components with respect to the coordinate system θ.

In view of the Cramér-Rao inequality (3), the way of finding an optimal estimation scheme is twofold. First, we optimize the input state $\psi^{(n)}$ to make the lower bound $\left(J_\theta[\psi^{(n)}]\right)^{-1}$ as small as possible, that is, to make the SLD Fisher metric g as large as possible. Second, we investigate if the corresponding lower bound is achievable, that is, if there is a locally unbiased estimator $M^{(n)}$ for which the equality holds in (3).

Motivated by the decomposition (2), let us first mention the problem of maximizing the SLD Fisher metric $g^{[\lambda]} := g_{\psi^{[\lambda]}}$ for the model $\rho_\theta^{[\lambda]} := (\mathrm{id}^{\otimes n} \otimes \Gamma_{U_\theta}^{\otimes n})\,(|\psi^{[\lambda]}\rangle\langle\psi^{[\lambda]}|)$ on the invariant subspace $\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$. Unfortunately, the set $\{J_\theta[\psi] \mid \psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}\}$ does not have the maximal element in general (see Appendix A). In other words, there is no input state $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ that maximizes the metric $g^{[\lambda]}$ itself. Hence, we must introduce a weaker optimality criterion.

Definition 1. *A state $\phi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ is called* admissible *in the component $\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ if*

$$\mathrm{tr}\, J_0[\phi] = \max \left\{ \mathrm{tr}\, J_0[\psi] \,\middle|\, \psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]} \right\}.$$

Suppose $\phi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ is admissible in $\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$. Then it is easily seen that there is no $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ that satisfies $J_0[\psi] \geq J_0[\phi]$ and $J_0[\psi] \neq J_0[\phi]$ simultaneously. Stated otherwise, there is no $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ that satisfies $g_\psi^{[\lambda]} \geq g_\phi^{[\lambda]}$ and $g_\psi^{[\lambda]} \neq g_\phi^{[\lambda]}$ at $U = U_0$. Moreover, since $J_0[\psi]$ is independent of the choice of $U_0 \in SU(D)$ (see the proof of Theorem 1), it follows that there is no $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ that satisfies $g_\psi^{[\lambda]} \geq g_\phi^{[\lambda]}$ and $g_\psi^{[\lambda]} \neq g_\phi^{[\lambda]}$ anywhere on $SU(D)$. This observation justifies the notion of admissibility as an alternative optimality criterion for input states.

The admissibility of the input state is closely related to the achievability of the Cramér-Rao inequality. In fact, we can prove the following.

Theorem 1. *For $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$, the following are equivallent:*

(a) *There is a locally unbiased estimator M on $\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ that satisfies $V_0[M|\psi] = (J_0[\psi])^{-1}$.*

(b) *ψ is admissible.*

As to a general input of the form (2), we have the following.

Theorem 2. *If the input $\psi^{(n)}$ is a superposition of admissible states $\psi^{[\lambda]} \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$ as (2), the lower bound of (3) is achievable. Moreover, the SLD Fisher metric is decomposed as*

$$g = \sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\, g^{[\lambda]}. \tag{4}$$

All in all, it is reasonable to restrict ourselves to inputs $\psi^{(n)} \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{\otimes n}$ that are superpositions of admissible states $\psi^{[\lambda]} \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$. (A further discussion is given in the proof of Theorem 2.) In what follows, as a canonical choice of admissible states on $\mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$, we focus on maximally entangled inputs:

$$\psi_{\mathrm{ME}}^{[\lambda]} := \frac{1}{\sqrt{\dim \mathcal{H}^{[\lambda]}}} \sum_{\ell=1}^{\dim \mathcal{H}^{[\lambda]}} e_\ell \otimes f_\ell, \tag{5}$$

where $\{e_k\}_k$ and $\{f_\ell\}_\ell$ are arbitrary orthonormal bases of $\mathcal{H}^{\otimes n}$ and $\mathcal{H}^{[\lambda]}$. (For the admissibility of $\psi_{\rm ME}^{[\lambda]}$, see the proof of Theorem 1, and for a statistical meaning of this choice, see Appendix A.)

Now that the SLD Fisher metric g is given by a convex combination of the components $g^{[\lambda]}$ as (4), the problem amounts to finding the index λ that maximizes the SLD Fisher information matrix $J_0[\psi_{\rm ME}^{[\lambda]}]$. This is completely solved by the following.

Theorem 3. *For irreducible representations specified by the Dynkin index $\lambda = [n_1, n_2, \ldots, n_{D-1}]$, the SLD Fisher information matrix $J_0[\psi_{\rm ME}^{[\lambda]}]$ is given by*

$$\left(J_0[\psi_{\rm ME}^{[\lambda]}]\right)_{ij} = \frac{4\,c^{[\lambda]}}{D^2-1}\,\delta_{ij},$$

where

$$c^{[\lambda]} := \frac{1}{2D}\left[D^2\sum_{\mu=1}^{D-1} p_\mu + D\left(\sum_{\mu=1}^{D-1} p_\mu + \sum_{\mu=1}^{D-1} p_\mu^2 - 2\sum_{\mu=1}^{D-1}\mu\, p_\mu\right) - \left(\sum_{\mu=1}^{D-1} p_\mu\right)^2\right] \tag{6}$$

with

$$p_\mu := \sum_{\nu=\mu}^{D-1} n_\nu$$

the length of the μth row of the corresponding Young frame. In particular,

$$J_0[\psi_{\rm ME}^{[\lambda]}] \le J_0[\psi_{\rm ME}^{[n,0,\ldots,0]}] = \frac{2}{D(D+1)}\,n(n+D),$$

and the maximum is attained only if $\lambda = [n, 0, \ldots, 0]$.

3 Proof of Theorems

3.1 Proof of Theorem 1

We prove a more detailed assertion.

Lemma 4. *Let $\tau : SU(D) \to \mathcal{B}(\mathcal{H}^{[\lambda]})$ be an irreducible representation. For $\psi \in \mathcal{H}^{\otimes n}\otimes\mathcal{H}^{[\lambda]}$, the following are equivalent:*

(a) *There is a locally unbiased estimator M on $\mathcal{H}^{\otimes n}\otimes\mathcal{H}^{[\lambda]}$ that satisfies $V_0[M|\psi] = (J_0[\psi])^{-1}$.*

(b) *$\langle\psi|I\otimes[\tau_*(Y), \tau_*(Z)]\psi\rangle = 0$ for all $Y, Z \in su(D)$.*

(c) *$\langle\psi|I\otimes\tau_*(Y)\psi\rangle = 0$ for all $Y \in su(D)$.*

(d) *ψ is admissible.*

Proof. We first prove (a) $\Leftrightarrow$ (b). According to [10], (a) occurs if and only if

$$\{\langle L_{i,\theta}\,\psi_\theta | L_{j,\theta}\,\psi_\theta\rangle\}_{1\le i,j\le D^2-1}$$

are all real at $\theta = 0$, where $\psi_\theta := (I \otimes \tau(U_\theta))\psi$, and $L_{i,\theta}$ is an ith SLD of the pure state model $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. (See also [11].) By direct computation using the coordinate system (1) and the canonical representation $L_{i,\theta} = 2\,\partial_i \rho_\theta$ for pure state models [12], we have

$$L_{i,0}\,\psi_0 = 2\sqrt{-1}\,(I \otimes \tau(U_0))\,(I - |\psi\rangle\langle\psi|)\,(I \otimes \tau_*(X_i))\,\psi$$

and

$$\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle = 4\,\langle\psi| I \otimes \tau_*(X_i)\tau_*(X_j)\psi\rangle - 4\,\langle\psi| I \otimes \tau_*(X_i)\psi\rangle\,\langle\psi| I \otimes \tau_*(X_j)\psi\rangle. \tag{7}$$

As a consequence

$$\mathrm{Im}\,\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle = \frac{2}{\sqrt{-1}}\,\langle\psi| I \otimes [\tau_*(X_i), \tau_*(X_j)]\psi\rangle,$$

and the assertion immediately follows.

Next, (b) $\Leftrightarrow$ (c) is a direct consequence of the fact that Lie algebra $su(D)$ is simple [13]. In fact, $[su(D), su(D)] = su(D)$, so that $[\tau_*(su(D)), \tau_*(su(D))] = \tau_*(su(D))$.

Finally we prove (c) $\Leftrightarrow$ (d). Since the (i,j)th entry of $J_0[\psi]$ is given by $\mathrm{Re}\,\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle$, we have from (7) that

$$\mathrm{tr}\,J_0[\psi] = 4\,\langle\psi| I \otimes C^{[\lambda]}\,\psi\rangle - 4\sum_{i=1}^{D^2-1} \left|\langle\psi| I \otimes \tau_*(X_i)\psi\rangle\right|^2, \qquad C^{[\lambda]} := \sum_{i=1}^{D^2-1} \tau_*(X_i)^2.$$

Since $\{X_i\}_i$ are chosen to be Killing orthonormal up to scaling, the operator $C^{[\lambda]}$ is the second order Casimir operator [13] for the representation τ, and is a scalar multiple of the identity: $C^{[\lambda]} = c^{[\lambda]} I$. The coefficient $c^{[\lambda]}$ is explicitly given by (6), see [14]. As a consequence,

$$\mathrm{tr}\,J_0[\psi] = 4c^{[\lambda]} - 4\sum_{i=1}^{D^2-1} \left|\langle\psi| I \otimes \tau_*(X_i)\psi\rangle\right|^2 \le 4c^{[\lambda]}$$

for all ψ. Now observe that the upper bound $4c^{[\lambda]}$ is achievable. In fact, let ψ be a maximally entangled state $\psi_{\mathrm{ME}}^{[\lambda]}$, then

$$\langle\psi_{\mathrm{ME}}^{[\lambda]} | I \otimes \tau_*(X_i)\,\psi_{\mathrm{ME}}^{[\lambda]}\rangle = \frac{1}{\dim \mathcal{H}^{[\lambda]}}\,\mathrm{Tr}\,\tau_*(X_i) = 0,$$

because elements of $\tau_*(su(D))$ have trace zero. Therefore

$$\mathrm{tr}\,J_0[\psi_{\mathrm{ME}}^{[\lambda]}] = 4c^{[\lambda]} = \max\left\{\mathrm{tr}\,J_0[\psi] \,\middle|\, \psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}\right\}.$$

The equivalence (c) $\Leftrightarrow$ (d) now follows immediately. □

3.2 Proof of Theorem 2

Let $\tau^{[\lambda]} : SU(D) \to \mathcal{B}(\mathcal{H}^{[\lambda]})$ be irreducible representations, and let $\psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{\otimes n}$ be decomposed into

$$\psi = \sum_\lambda \sum_{\mathrm{STab}(\lambda)} a^{[\lambda]}\,\psi^{[\lambda]}, \tag{8}$$

where $\psi^{[\lambda]} \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}$. Further let

$$\psi_\theta := (I \otimes U_\theta^{\otimes n})\psi = \sum_\lambda \sum_{\mathrm{STab}(\lambda)} a^{[\lambda]}\ (I \otimes \tau^{[\lambda]}(U_\theta))\psi^{[\lambda]},$$

and let $L_{i,\theta}$ be an ith SLD of the corresponding model $\rho_\theta = |\psi_\theta\rangle\langle\psi_\theta|$. Then by an evaluation similar to (7), we have

$$\begin{aligned}&\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle \\ &\quad = 4\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ \langle \psi^{[\lambda]} | I \otimes \tau_*^{[\lambda]}(X_i)\tau_*^{[\lambda]}(X_j)\psi^{[\lambda]}\rangle \\ &\qquad -4\left(\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ \langle \psi^{[\lambda]} | I \otimes \tau_*^{[\lambda]}(X_i)\psi^{[\lambda]}\rangle\right)\left(\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ \langle \psi^{[\lambda]} | I \otimes \tau_*^{[\lambda]}(X_j)\psi^{[\lambda]}\rangle\right). \quad (9)\end{aligned}$$

Now suppose that $\psi^{[\lambda]}$ are all admissible. It then follows from Lemma 4 (c) that

$$\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle = 4\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ \langle \psi^{[\lambda]} | I \otimes \tau_*^{[\lambda]}(X_i)\tau_*^{[\lambda]}(X_j)\psi^{[\lambda]}\rangle.$$

As a consequence

$$\mathrm{Im}\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle = \frac{2}{\sqrt{-1}}\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ \langle \psi^{[\lambda]} | I \otimes [\tau_*^{[\lambda]}(X_i), \tau_*^{[\lambda]}(X_j)]\psi^{[\lambda]}\rangle = 0,$$

which follows from Lemma 4 (b). This proves the achievability of (3). On the other hand,

$$J_0[\psi] = [\mathrm{Re}\langle L_{i,0}\psi_0 | L_{j,0}\psi_0\rangle]_{ij} = \sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ J_0[\psi^{[\lambda]}].$$

This proves the decomposition (4).

It should be noted that for any input ψ of the form (8) having a fixed set of coefficients $\{a^{[\lambda]}\}_\lambda$, we obtain from (9) that

$$\mathrm{tr}\, J_0[\psi] \le 4\sum_\lambda \sum_{\mathrm{STab}(\lambda)} |a^{[\lambda]}|^2\ c^{[\lambda]}.$$

Moreover, this upper bound is achievable if $\psi^{[\lambda]}$ are all admissible. This observation supports the validity of restricting inputs ψ to superpositions of admissible states.

3.3 Proof of Theorem 3

By direct calculation using (7), we have

$$\left(J_0[\psi_{\mathrm{ME}}^{[\lambda]}]\right)_{ij} = \frac{4}{\dim \mathcal{H}^{[\lambda]}}\, K_\tau(X_i, X_j),$$

where

$$K_\tau(Y, Z) := \mathrm{Tr}\, \tau_*(Y)\tau_*(Z).$$

Since, for each $U \in SU(D)$, the adjoint action $\mathrm{Ad}(U) : su(D) \to su(D) : Y \mapsto UYU^{-1}$ is K_τ-orthogonal, in that $K_\tau(\mathrm{Ad}(U)Y, \mathrm{Ad}(U)Z) = K_\tau(Y, Z)$, it follows from [15, Theorem VIII.2.4] that K_τ is identical, up to a constant multiple, to the Killing metric. In other words, there is a constant r_τ satisfying $K_\tau(Y, Z) = r_\tau \operatorname{Tr} YZ$, so that $K_\tau(X_i, X_j) = (r_\tau/2)\, \delta_{ij}$. Consequently,

$$(D^2-1)\,\frac{r_\tau}{2} = \sum_{i=1}^{D^2-1} K_\tau(X_i, X_i) = \sum_{i=1}^{D^2-1} \operatorname{Tr} \tau_*(X_i)^2 = \operatorname{Tr} C^{[\lambda]} = \dim \mathcal{H}^{[\lambda]}\, c^{[\lambda]}.$$

By using these relations, we obtain

$$\left(J_0[\psi_{\mathrm{ME}}^{[\lambda]}]\right)_{ij} = \frac{2\, r_\tau}{\dim \mathcal{H}^{[\lambda]}}\, \delta_{ij} = \frac{4\, c^{[\lambda]}}{D^2-1}\, \delta_{ij}. \tag{10}$$

We next show that (10) takes the maximum at $\lambda = [n, 0, \ldots, 0]$. Letting $M := \sum_{\mu=1}^{D-1} p_\mu$, the coefficient $c^{[\lambda]}$ is rewritten as

$$c^{[\lambda]} = \frac{1}{2D}\left[D^2 M + D\left(M + \sum_{\mu=1}^{D-1} p_\mu^2 - 2\sum_{\mu=1}^{D-1} \mu\, p_\mu\right) - M^2\right]. \tag{11}$$

The problem is thus reduced to maximizing (11) under the constraint that $M \le n$ and

$$p_1 \ge p_2 \ge \cdots \ge p_{D-1} \ge 0.$$

Since

$$\sum_{\mu=1}^{D-1} p_\mu^2 - 2\sum_{\mu=1}^{D-1} \mu\, p_\mu \le \sum_{\mu=1}^{D-1} p_\mu^2 - 2\sum_{\mu=1}^{D-1} p_\mu \le \left(\sum_{\mu=1}^{D-1} p_\mu\right)^2 - 2\sum_{\mu=1}^{D-1} p_\mu = M^2 - 2M,$$

we have

$$c^{[\lambda]} \le \frac{1}{2D}\left[D^2 M + D(M^2 - M) - M^2\right] = \frac{D-1}{2D}(M^2 + DM) \le \frac{D-1}{2D}(n^2 + Dn).$$

By checking the condition for each inequality to saturate, it is easily seen that this upper bound is attained if and only if $\lambda = [n, 0, \ldots, 0]$.

4 Geometry of $SU(D)$ estimation

Theorem 3 implies that, for each n, the optimal input is $\psi_{\mathrm{ME}}^{[n,0,\ldots,0]}$, and that the optimal strategy for estimating an unknown $SU(D)$ channel exhibits

$$\min_{M^{(n)}, \psi^{(n)}} V_\theta\left[M^{(n)} \middle|\, \psi^{(n)}\right] = \frac{D+1}{n(n+D)}\left(J_\theta[\psi_{\mathrm{ME}}^{[1]}]\right)^{-1}. \tag{12}$$

The implication of this result is profound[1]. In the standard (classical) statistics, it is commonly believed that the estimation error approaches zero in the rate of $O(1/n)$. In contrast, for estimating

[1] Although in a different setting, an analogous asymptotic property has been obtained in [1, 2, 3, 4]; see also [16]

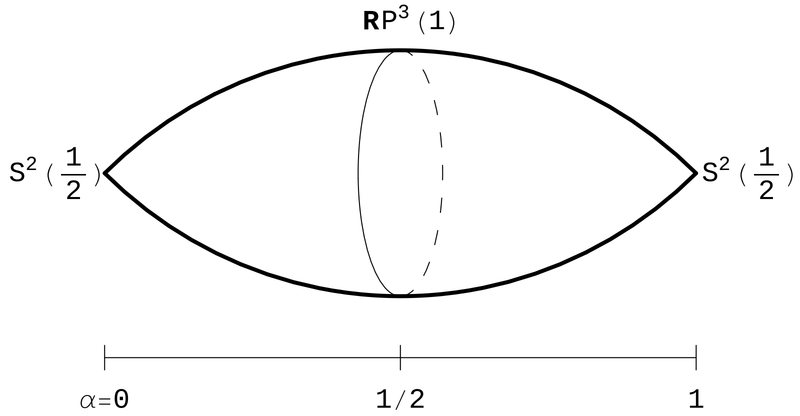

Figure 1: The global structure of the manifold of output states for $n = 1$. When $\alpha = 0$ or 1, it collapses to 2-dimensional sphere S^2 of radius 1/2; when $\alpha = 1/2$, it is isometric to 3-dimensional real projective space $\mathbb{R}P^3$ of unit radius; otherwise it is diffeomorphic, but is not isometric, to $\mathbb{R}P^3$ of any radius.

an unknown $SU(D)$ channel, the estimation error approaches zero asymptotically in the rate of $O\left(1/n^2\right)$ as (12) asserts.

Let us recast this result in terms of differential geometry. Theorem 3 asserts that the output manifold

$$\mathcal{M}^{[\lambda]} := \left\{ (\mathrm{id}^{\otimes n} \otimes \Gamma_U^{\otimes n})(|\psi_{\mathrm{ME}}^{[\lambda]}\rangle\langle\psi_{\mathrm{ME}}^{[\lambda]}|) \;\middle|\; U \in SU(D) \right\}$$

for a maximally entangled input $\psi_{\mathrm{ME}}^{[\lambda]}$ is *locally isometric*, up to a scaling factor $\sqrt{c^{[\lambda]}}$, to the Riemannian manifold $SU(D)$ equipped with the Cartan-Killing metric. On the other hand, it is easily seen that $\mathcal{M}^{[\lambda]}$ is diffeomorphic to $SU(D)/\mathbb{Z}_D$. As a consequence, we have the following.

Theorem 5. *The output manifold $\mathcal{M}^{[\lambda]}$ is isometric to $SU(D)/\mathbb{Z}_D$ up to a scaling factor $\sqrt{c^{[\lambda]}}$.*

In order to get a better perspective on Theorem 5, let us study the simplest case $SU(2)$ in detail. When $n = 1$, an input $\psi \in \mathcal{H} \otimes \mathcal{H}$ is decomposed into the following Schmidt form:

$$\psi = \sqrt{1-\alpha}\, e_1 \otimes f_1 + \sqrt{\alpha}\, e_2 \otimes f_2,$$

where $\alpha \in [0, 1]$ describes the degree of entanglement. The structure of the corresponding output manifold $\{(\mathrm{id} \otimes \Gamma_U)(|\psi\rangle\langle\psi|) \mid U \in SU(2)\}$ was studied in detail in [11], and is illustrated in Fig. 1. When $\alpha = 0$ or 1, the output manifold degenerates to a 2-dimensional sphere $\mathbb{C}P^1 \cong S^2$ of radius 1/2, which is nothing but the Bloch sphere. When $0 < \alpha < 1$, on the other hand, the global topology of the output manifold completely changes into one which is diffeomorphic to the 3-dimensional real projective space $SU(2)/\{\pm I\} \cong SO(3) \cong \mathbb{R}P^3$. Moreover, as the degree α of entanglement approaches 1/2, the manifold gradually inflates and hence points on the manifold are getting separated from each other. Finally when α reaches 1/2 (i.e., when the input is maximally entangled), the maximally inflated output manifold becomes *isometric* to $\mathbb{R}P^3$ of unit radius. This is the underlying differential geometrical mechanism for the admissibility of a maximally entangled input. In fact, the larger the SLD Fisher distance of two nearby quantum states becomes, the easier one can distinguish these states, as the quantum Cramér-Rao inequality asserts. For general n, the situation is similar: the output manifold inflates maximally (on average) when the input is a maximally entangled state $\psi_{\mathrm{ME}}^{[n]}$ on the invariant subspace specified by the Dynkin index $\lambda = [n]$, and it becomes isometric to $\mathbb{R}P^3$ of radius $r_n = \sqrt{(n^2+2n)/3}$. In summary, the degree of entanglement controls the "shape" of the output manifold, while the degree n of extension controls its maximal "radius."

As to $SU(D)$ for $D \geq 3$, on the other hand, the output manifold is not of constant curvature even for a maximally entangled input $\psi_{\mathrm{ME}}^{[\lambda]}$. In fact, the dimension of a Cartan subalgebra of $su(D)$ is greater than one, and the sectional curvature vanishes there. Thus the notion of radius is not relevant for the case $D \geq 3$. However the situation is analogous: if the input is taken to be a maximally entangled state $\psi_{\mathrm{ME}}^{[\lambda]}$, then as the degree n of extension increases, the "size" of output manifold increases in the rate $\sqrt{c^{[\lambda]}}$ which is asymptotically linear in n, while the "shape" is kept unchanged.

5 Concluding remarks

The problem of estimating an unknown $SU(D)$ channel $\Gamma_U : \rho \mapsto U\rho U^*$ was studied based on the quantum Cramér-Rao inequality. By invoking extensions $(\mathrm{id} \otimes \Gamma_U)^{\otimes n}$, it was shown that there was a sequence of input states $\psi^{(n)}$ and estimators $M^{(n)}$ on $(\mathcal{H} \otimes \mathcal{H})^{\otimes n}$ that exhibited

$$\min_{M^{(n)},\psi^{(n)}} V_\theta \left[M^{(n)} \middle| \, \psi^{(n)} \right] = O\left(\frac{1}{n^2}\right).$$

The optimal coefficient was also determined explicitly. Further, the mechanism behind this asymptotic behavior was investigated from a differential geometrical point of view.

Combining this result with the former one obtained in [6], we can conclude that there are at least two classes of quantum channels that exhibit essentially different asymptotic behaviors: the minimal estimation error is of $O(1/n^2)$ for $SU(D)$ channels, while it is of $O(1/n)$ for generalized Pauli channel[2]. It is an open problem whether there is a quantum channel that exhibits an asymptotic rate $O(1/n^s)$ with $s \neq 1, 2$.

Appendix

A Nonexistence of maximal SLD metric

In this appendix, we demonstrate that the set $\{J_0[\psi] \,|\, \psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}\}$ does not in general have the maximal element. Let us consider the irreducible representation $\lambda = [n_1]$ of $SU(2)$, which is also called the highest weight $j := n_1/2$ representation. We take $X_i := \sigma_i/2$ to be the basis of $su(2)$, where $\{\sigma_i\}_i$ are Pauli matrices. We further take an input of the form

$$\psi = \psi^{[n_1]} := \sum_{k=0}^{n_1} \sqrt{\alpha_k}\, e_k \otimes f_k,$$

[2]Recently, it was shown that low-noise channels also exhibited the same asymptotic behavior $O(1/n)$ [17].

where $(\alpha_i)_i$ is a probability vector, and $f_k := |j, m\rangle$ is the standard orthonormal basis of $\mathcal{H}^{[\lambda]}$ with $m := j - k$, $(0 \le k \le n_1)$, satisfying

$$\begin{aligned} \hat{S}_{\pm}|j,m\rangle &= \sqrt{j(j+1) - m(m \pm 1)}\,|j, m \pm 1\rangle, \\ \hat{S}_3\,|j,m\rangle &= m\,|j,m\rangle. \end{aligned}$$

with $\hat{S}_i := \tau_*^{[n_1]}(X_i)$, and $\hat{S}_{\pm} := \hat{S}_1 \pm \sqrt{-1}\,\hat{S}_2$. The corresponding SLD Fisher information matrix $J_0[\psi]$ is given by

$$(J_0[\psi])_{ij} = 4 \sum_k \alpha_k \operatorname{Re}\langle f_k | \hat{S}_i \hat{S}_j f_k\rangle - 4 \left(\sum_k \alpha_k \langle f_k | \hat{S}_i f_k\rangle \right) \left(\sum_k \alpha_k \langle f_k | \hat{S}_j f_k\rangle \right).$$

Let $n_1 = 2$ for definiteness. Then after some calculation, we have

$$(J_0[\psi])_{11} = (J_0[\psi])_{22} = 2\,(\alpha_0 + 2\alpha_1 + \alpha_2), \quad (J_0[\psi])_{33} = 4\,[\alpha_0 + \alpha_2 - (\alpha_0 - \alpha_2)^2],$$

and the off-diagonal elements are all zero. Consequently, the input ψ is admissible if and only if $\alpha_0 = \alpha_2$, for which we have

$$J_0[\psi] = 4 \begin{bmatrix} 1-\alpha & & \\ & 1-\alpha & \\ & & 2\alpha \end{bmatrix}, \tag{13}$$

where $\alpha_0 = \alpha_2 = \alpha$ and $\alpha_1 = 1 - 2\alpha$, with $0 \le \alpha \le 1/2$.

Now suppose there is an input ϕ which gives the maximal Fisher information matrix. Let us denote the matrix as

$$J_0[\phi] = 4 \begin{bmatrix} a & * & * \\ * & b & * \\ * & * & c \end{bmatrix},$$

where the off-diagonal elements are suppressed. Since ϕ is necessarily admissible,

$$a + b + c = (1 - \alpha) + (1 - \alpha) + 2\alpha = 2.$$

On the other hand, since $J_0[\phi] \ge J_0[\psi]$ for all α, it holds that

$$a \ge 1 - \alpha, \quad b \ge 1 - \alpha, \quad c \ge 2\alpha$$

for all α. As a consequence, $a, b, c \ge 1$, so that

$$a + b + c \ge 3.$$

This is a contradiction, proving that no such a ϕ exists.

Incidentally, the formula (13) demonstrates what happens when the entanglement parameter α is changed. In order to get better distinguishability in the first (and the second) direction of the parameter, we need to make α as small as possible. But accordingly, we lose distinguishability in the third direction. In general, if one tries to get more information about some directions, then he loses information about the other directions, as long as input states are chosen among admissible ones. This is the statistical, as well as the geometrical, meaning of the fact that no maximal element exists in $\{J_0[\psi]\,;\, \psi \in \mathcal{H}^{\otimes n} \otimes \mathcal{H}^{[\lambda]}\}$. In a sense, a maximally entangled input (e.g., $\alpha = 1/3$ in the above example) gives an estimation scheme "impartial" to all directions of the parameter.

References

[1] G. Chiribella, G.M. D'Ariano, P. Perinotti, and M.F. Sacchi, "Efficient use of quantum resources for the transmission of a reference frame," Phys. Rev. Lett. **93**, 180503 (2004).

[2] E. Bagan, M. Baig, and R. Muñoz-Tapia, "Quantum reverse-engineering and reference frame alignment without non-local correlations," Phys. Rev. A **70**, 030301(R) (2004).

[3] M. Hayashi, "Parallel treatment of estimation of $SU(2)$ and phase estimation," Phys. Lett. A **354**, 183-189 (2006).

[4] J. Kahn, "Fast rate estimation of an unitary operation in SU(d)," quant-ph/0603115.

[5] A. Fujiwara, "Strong consistency and asymptotic efficiency for adaptive quantum estimation problems," J. Phys. A: Math. Gen. **39**, 12489-12504 (2006).

[6] A. Fujiwara and H. Imai, "Quantum parameter estimation of a generalized Pauli channel," J. Phys. A: Math. Gen. **36**, 8093-8103 (2003).

[7] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Transl. Math. Monographs, vol. 191 (AMS, Providence, 2000).

[8] C.W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, NY, 1976).

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

[10] K. Matsumoto, "A new approach to the Cramer-Rao type bound of the pure state model," J. Phys. A: Math. Gen. **35**, 3111-3124 (2002).

[11] A. Fujiwara, "Estimation of SU(2) operation and dense coding: An information geometric approach," Phys. Rev. A **65**, 012316 (2002).

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

[13] D.H. Sattinger and O.L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics*, Applied Mathematical Sciences 61 (Springer, New York, 1986).

[14] A.O. Barut and R. Raczka, *Theory of Group Representations and Applications* (Polish Scientific Publishers, Warszawa, 1977).

[15] B. Simon, *Representaions of Finite and Compact Groups*, Graduate Studies in Mathematics 10 (AMS, Providence, 1996).

[16] M. Balllester, "Estimation of SU(d) using entanglement," quant-ph/0507073.

[17] M. Hotta, T. Karasawa, and M. Ozawa, "N-body-extended channel estimation for low-noise parameters," J. Phys. A: Math. Gen. **39**, 14465-14470 (2006).