Comparison between the Cramer-Rao and the mini-max approaches in quantum channel estimation

Masahito Hayashi

Graduate School of Information Sciences, Tohoku University, Sendai, 980-8579, Japan.
Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117542
E-mail: hayashi@math.is.tohoku.ac.jp

Received:

Abstract: In a unified viewpoint in quantum channel estimation, we compare the Cramér-Rao and the mini-max approaches, which gives the Bayesian bound in the group covariant model. For this purpose, we introduce the local asymptotic mini-max bound, whose maximum is shown to be equal to the asymptotic limit of the mini-max bound. It is shown that the local asymptotic mini-max bound is strictly larger than the Cramér-Rao bound in the phase estimation case while the both bounds coincide when the minimum mean square error decreases with the order $O(\frac{1}{n})$. We also derive a sufficient condition for that the minimum mean square error decreases with the order $O(\frac{1}{n})$.

1. Introduction

In quantum information technology, it is usual to use quantum channel for sending quantum state. Since a quantum channel has noise, it is important to identify quantum channel. In this paper, we consider theoretical optimal performance of quantum channel estimation when we can apply the same unknown channel several times. In order to treat this problem, we employ quantum state estimation theory. In our setting, we can optimize our input state and our measurement[1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14]. As is illustrated in Fig. 1 with $n = 4$, it is assumed to be possible to use entanglement with reference system in the measurement process when the channel $\Lambda_\theta$ with the unknown parameter $\theta$ is applied $n$ times. This setting is mathematically equivalent with the setting given in Fig. 2 with $n = 4$, which has a single input state in the large input system and a single measurement in the large output system. In this paper, we consider quantum channel estimation with the formulation given by Fig. 2.

In the state estimation, when the number $n$ of prepared states goes to infinity, the mean square error (MSE) behaves as the order $O(\frac{1}{n})$ as in the estimation of probability distribution. However, in the estimation of quantum channel, two
different analyses were reported concerning asymptotic behavior of MSE. As the first case, in the estimations of depolarizing channels and Pauli channels, the optimal MSE behaves as $O\left(\frac{1}{n}\right)\[1,3\]$. As the second case, in the estimation of unitary, the optimal MSE behaves as $O\left(\frac{1}{n^2}\right)\[4,5,6,7,8,9,10,11\]$. In the second case, two different types of results were reported: One is based on the Cramér-Rao approach\[4,5\]. The other is based on the mini-max approach\[6,7,8,9,10,11\].

The Cramér-Rao approach is based on the notion of locally unbiased estimator, and allows one to give a simple lower bound (the Cramér-Rao bound) to the mean square error (MSE) at a given point. The mini-max approach aims to minimize the maximum of the MSE over all possible values of the parameter. The mini-max approach is more meaningful than the Cramér-Rao approach because the true value of the parameter is unknown. So, the Cramér-Rao bound is just
Comparison between Cramer-Rao and mini-max approaches

a lower bound in general, while it can be asymptotically achieved in the case of quantum state estimation with the $n$ copies of the unknown state. However, the Cramér-Rao bound has been considered so many times in the literatures\cite{13,15}. The reason seems that computing the mini-max bound is a much harder problem. Indeed, there are only very few examples of calculations of the mini-max bound, and most of them are in the compact group covariant setting, in which, as was shown by Holevo\cite{15}, the mini-max bound coincides with the Bayesian average of the MSE over the normalized invariant measure. So, many researchers\cite{6,7,8,9,10,11} applied this approach to quantum channel estimation under the group covariance. As the most simple case for unitary estimation, phase estimation has been treated with the mini-max approach by several papers\cite{9,10,11}, and it has been shown that the minimum MSE that behaves as $O\left(\frac{1}{n}\right)$. On the other hand, the Cramér-Rao approach suggests the noon state as the optimal input\cite{36,37} in phase estimation. The later estimation scheme was experimentally demonstrated in the case of $n = 4$\cite{32,33} and $n = 10$\cite{34}. Also, another group\cite{31} experimentally demonstrated an estimation protocol concerning the group covariant approach proposed by\cite{35}. In phase estimation with group covariant framework, the asymptotic minimum MSE behaves as $O\left(\frac{1}{n}\right)$\cite{9,10}. When we focus on the difference $\hat{\theta} - \theta$ between the true parameter $\theta$ and the estimate $\hat{\theta}$, the limiting distribution concerning the random variable $n(\hat{\theta} - \theta)$ can be obtained through Fourier transform of a function with a finite domain\cite{11}. However, the Cramér-Rao bound is different from the asymptotic limit of the mini-max bound. So, these results seem to contradict with each other. No existing research compares both approaches in a unified viewpoint.

The manuscript consists of two parts. In the first part, we discuss the Cramér-Rao approach for channel estimation, and give a simple formula to compute a bound based on the right logarithmic derivative introduced by Holevo\cite{20}. This formula is based on the Choi-Jamiolkowski representation matrix\cite{39,40} of the quantum channel, and holds under the condition that, at the given point, the support of the Choi-Jamiolkowski representation matrix contains the support of its derivative. Under this condition, we prove the additivity of the RLD Fisher information for quantum channels.

The second part is about the local asymptotic mini-max bound\cite{38}, in which the maximum of the MSE is taken over an interval, the width of which is sent to zero after computing the asymptotic limits. The obtained results are summarized as follows: (1) The local asymptotic mini-max bound is strictly larger than the Cramér-Rao bound in the case of phase estimation. (2) Both bounds coincide with each other when the asymptotic minimum MSE behaves as $O\left(\frac{1}{n}\right)$. (3) The local asymptotic mini-max bound is achievable (Proposition 3), using a variation of the two-step strategy\cite{23,22}. That is, there is an optimal sequence of estimators that achieves the local asymptotic mini-max bound at any point. As consequence, the maximum of the local asymptotic mini-max bound is shown to be equal to the asymptotic limit of the mini-max bound.

However, the Cramér-Rao bound has a different type of achievability. That is, there is an optimal sequence of estimators that achieves the Cramér-Rao at a given point. This characteristic is a curious quantum analogue of super efficiency in classical statistics. In estimation of probability distribution, if we assume a weaker condition for our estimator, there exists an estimator that surpasses the Cramér-Rao bound only in measure zero points. Such an estimator is called
In the estimation of phase action, we point out that the Cramér-Rao bound can be attained only by a quantum channel version of a superefficient estimator that works at specific points. Since the Cramér-Rao approach is based on the asymptotically locally unbiased condition, we can conclude that the asymptotically locally unbiased condition is too weak for deriving the local asymptotic mini-max bound, which can be attained in all points. Indeed, a similar phenomenon happens in quantum state estimation when we use the large deviation criterion.

This paper is organized as follows. Some of obtained results are based on quantum state estimation with the Cramér-Rao approach. Section 2 is devoted to a review of the Cramér-Rao approach in quantum state estimation. In this section, the symmetric logarithmic derivative (SLD) Fisher information and the right logarithmic derivative (RLD) Fisher information are explained. However, the Cramér-Rao bound is obtained only by the locally unbiased condition. So, it is needed to discuss its relation with the estimator that works globally. In Section 3, we treat SLD Fisher information and RLD Fisher information in the quantum channel estimation, and discuss the increasing order of SLD Fisher information. In Section 4, we give several examples where the maximum SLD Fisher information increases with $O(n^2)$. In Section 5, we compare the local asymptotic mini-max bound and the Cramér-Rao bound. We also show the global attainability of the local asymptotic mini-max bound in the channel estimation. It is also shown that the Cramér-Rao bound is closely related to a quantum channel version of superefficiency in the phase estimation.

2. Cramér-Rao bound in quantum state estimation

In quantum state estimation, we estimate the true state through the quantum measurement under the assumption that the true state of the given quantum system $\mathcal{H}$ belongs to a certain parametric state family $\{\rho_\theta | \theta \in \Theta \subset \mathbb{R}^d\}$. In the following, we consider the case when the number $d$ of parameters is one. Usually, we assume that $n$ quantum systems are prepared in the state $\rho_\theta$. Hence, the total system is described by the tensor product space $\mathcal{H}^\otimes n$, and the state of the total system is given by $\rho_\theta^\otimes n$.

In this case, when we choose a suitable measurement, the MSE decreases in proportion to $n^{-1}$ as in the estimation of probability distribution. So, we focus on the first order coefficient of the MSE concerning $\frac{1}{n}$. In the most general setting, any positive operator valued measure (POVM) $M^n$ on the total system $\mathcal{H}^\otimes n$ is allowed as an estimator when it takes values in the parameter space $\Theta \subset \mathbb{R}$. The MSE is given as

$$\text{MSE}_\theta(M^n) := \int (\hat{\theta} - \theta)^2 \text{Tr} \rho_\theta^\otimes n M^n(d\hat{\theta}).$$

In the quantum case, there are several quantum extensions of Fisher information when the state $\rho_\theta$ is differentiable at $\theta$ and $(I - P)\frac{\partial}{\partial \theta}(I - P) = 0$, where $P$ is the projection to the support of $\rho_\theta$. The largest one is the right logarithmic derivative (RLD) Fisher information $J_\theta^R$, and the smallest one is symmetric logarithmic derivative (SLD) Fisher information $J_\theta^S$. For these definitions, we
define the RLD $L^R_\theta$ and the SLD $L^S_\theta$ as the operators satisfying
\[
\frac{d\rho_\theta}{d\theta} = \rho_\theta L^R_\theta, \quad \frac{d\rho_\theta}{d\theta} = \frac{1}{2} \left( L^S_\theta \rho_\theta + \rho_\theta L^S_\theta \right).
\]
Then, the RLD and SLD Fisher informations are given by \[19,20,21\]
\[
J^R_\theta := \text{Tr} \rho_\theta (L^R_\theta)^\dagger (L^R_\theta), \quad J^S_\theta := \text{Tr} \rho_\theta (L^S_\theta)^2.
\]
When the range of $\rho_\theta$ contains the range of $(\frac{d\rho_\theta}{d\theta})^2$, the RLD Fisher information has another expression:
\[
J^R_\theta = \text{Tr} \left( \frac{d\rho_\theta}{d\theta} \right)^2 \rho_\theta^{-1}.
\] (1)

When the state family $\{\rho_\theta\}$ is given by $\rho_\theta := e^{i\theta H} |u\rangle \langle u| e^{-i\theta H}$, the condition (1) does not hold, where $H$ is an Hermitian matrix. In this case, the SLD Fisher information is calculated as follows \[17\].
\[
4(\langle u|H^2|u\rangle - \langle u|H|u\rangle^2).
\] (2)

Now, we introduce the unbiased condition by
\[
\int \hat{\theta} \text{Tr} \rho_\theta^\otimes_n M^n(d\hat{\theta}) = \theta, \quad \forall \theta \in \Theta.
\]
However, this condition is sometimes too restrictive in the asymptotic setting. So, we consider the Taylor expansion at a point $\theta_0$ and focus on the first order. Then, we obtain the locally unbiased condition at $\theta_0$:
\[
\int \hat{\theta} \text{Tr} \rho_\theta^\otimes_n M^n(d\hat{\theta}) = \theta_0, \quad \frac{d}{d\theta} \int \hat{\theta} \text{Tr} \rho_\theta^\otimes_n M^n(d\hat{\theta})|_{\theta=\theta_0} = 1.
\]
Under the locally unbiased condition at $\theta_0$, an application of Schwarz inequality similar to the classical case yields the quantum Cramér-Rao inequalities for both quantum Fisher information.
\[
\text{MSE}_{\theta_0}(M^n) \geq \frac{1}{n} (J^R_{\theta_0})^{-1}
\] (3)
\[
\text{MSE}_{\theta_0}(M^n) \geq \frac{1}{n} (J^S_{\theta_0})^{-1}.
\] (4)

Since $J^R_{\theta}$ is greater than $J^S_{\theta}$, the inequality (4) is more informative than the inequality (3). When the estimator $M^n$ is the spectral decomposition of the operator $\theta_0 I + \frac{1}{n} \rho_\theta (L^S_{\theta_0}^{(1)} + \cdots + L^S_{\theta_0}^{(n)})$, the equality in (4) holds, where $X^{(j)}$ is given as $I^{\otimes j-1} \otimes X \otimes I^{\otimes n-j}$. Then, we obtain the following inequality
\[
J^R_{\theta} \geq J^S_{\theta}.
\]
This inequality seems to imply that $J^R_{\theta}$ is not as meaningful as $J^S_{\theta}$ in the one-parametric case. However, as will be explained latter, $J^R_{\theta}$ provides a meaningful bound for MSE in the case of channel estimation.
In fact, in the asymptotic setting, a suitable estimator usually satisfies the asymptotic locally unbiased condition:

$$\lim_{n \to \infty} \int \theta \text{Tr} \rho_\theta \otimes^n M^n (d\theta) = \theta_0,$$

for all points \(\theta_0\). Under the above condition, using (4), we obtain the inequality

$$\limsup_{n \to \infty} n \text{MSE}_\theta (M^n) \geq (J_S^\theta)^{-1}.$$

Further, by using the two-step method, the bound \((J_S^\theta)^{-1}\) can be universally attained for any true parameter \(\theta\). So, defining the Cramér-Rao bound:

$$C_\theta := \inf \left\{ \text{MSE}_\theta (M^n) \left| \begin{array}{c} \text{\{M^n\}} \\
\text{satisfies the asymptotic locally unbiased condition.} \end{array} \right. \right\},$$

we obtain

$$C_\theta = (J_S^\theta)^{-1}.$$

On the other hand, its multi-parameter case is much complicated even in the asymptotic setting [24,25,26,27,28]. So, this paper does not treat the multi-parameter case.

3. Maximum SLD and RLD Fisher informations in quantum channel estimation

In this section, we apply the Cramér-Rao approach to estimation of channel. In the quantum system, the channel is given by a trace preserving completely positive (TP-CP) map \(\Lambda\) from the set of densities on the input system \(\mathcal{H} := \mathbb{C}^d\) to the set of densities on the output system \(\mathcal{K} := \mathbb{C}^{d'}\). By using \(dd'\) linear maps \(F_i\) from \(S(\mathcal{H})\) to \(S(\mathcal{K})\), any TP-CP map \(\Lambda\) can be described by \(\Lambda(\rho) = \sum_{i=1}^{dd'} F_i \rho F_i^\dagger\).

Hence, our task is to estimate the true TP-CP map under the assumption that the true TP-CP map belongs to a certain family of TP-CP maps \(\{\Lambda_\theta\}\).

In order to characterize a TP-CP map \(\Lambda_\theta\), we formulate the notation concerning states on the tensor product system \(\mathcal{H} \otimes \mathcal{R}\), where \(\mathcal{R}\) is a system of the same dimensionality as \(\mathcal{H}\) and is called the reference system. Using a linear map \(A\) from \(\mathcal{R}\) to \(\mathcal{H}\), we define an element \(\langle A \rangle\) of \(\mathcal{H} \otimes \mathcal{R}\) as follows.

$$\langle A \rangle := \sum_{j,k} A_{j,k} |j\rangle_H \otimes |k\rangle_R,$$

where \(\{|j\rangle_H\}_{j=1,\ldots,d}\) and \(\{|k\rangle_R\}_{k=1,\ldots,d}\) are complete orthonormal systems (CONSs) of \(\mathcal{H}\) and \(\mathcal{R}\). Hence, the relation

$$B \otimes C |A \rangle = |BAC^T \rangle$$

holds. This notation is applied to the cases of \(\mathcal{K} \otimes \mathcal{H}\) and \(\mathcal{K} \otimes \mathcal{R}\).

Now, we focus on the matrix \(\rho_\theta[A_\theta] := (A_\theta \otimes \text{id})(|I\rangle\langle I|),\) which is called the Choi-Jamiolkowski representation matrix [39,40]. Then, when the input state is
the maximally entangled state $|\frac{1}{\sqrt{d}}I\rangle\rangle = \sum_{j=1}^{d} |\frac{1}{\sqrt{d}}j\rangle \otimes |j\rangle$, the output state is $\frac{1}{2}\rho[A\theta] = (A\theta \otimes \text{id})(|\frac{1}{\sqrt{d}}I\rangle\rangle\langle\frac{1}{\sqrt{d}}I|)$. When the matrix $A\theta$ is a density matrix on $\mathcal{H}$, $|\langle A|\rangle\langle A|$ is a pure state on the product system $\mathcal{H} \otimes \mathcal{R}$. Thus, the output state is given as

$$(A\theta \otimes \text{id})(|\langle A|\rangle\langle A|) = (I \otimes A^T)(A\theta \otimes \text{id})(|I\rangle\rangle\langle I|)(I \otimes \mathcal{A}) = (I \otimes A^T)\rho[A\theta](I \otimes \mathcal{A}).$$

In the one-parameter case, we express the derivative \(\frac{d\rho[A\theta]}{d\theta}\) by $D[A\theta]$. When the input state is the product state $|\psi\rangle\langle\psi| \otimes |u\rangle\langle u|$, the output state is $A\theta(|\psi\rangle\langle\psi| \otimes |u\rangle\langle u|) = (I \otimes u \cdot v^T)\rho[A\theta](I \otimes \mathcal{A} \cdot u^T)$. Since

$$\langle\mathcal{A}|\text{Tr}_K\rho[A\theta]|\mathcal{A}\rangle|\psi\rangle\langle u| = u \cdot v^T \text{Tr}_K\rho[A\theta]|\mathcal{A}\rangle|\psi\rangle,$$

we have

$$\langle\mathcal{A}|\text{Tr}_K\rho[A\theta]|\mathcal{A}\rangle = 1.$$ 

Thus, we obtain

$$\text{Tr}_K\rho[A\theta] = I.$$ (6)

Taking the derivative in (6), we obtain

$$\text{Tr}_K D[A\theta] = 0.$$ (7)

Now, we back to our estimation problem. In this problem, our choice is given by a pair of the input state $\rho$ and the quantum measurement $M$. When we fix the input state, our estimation problem can be reduced to the state estimation with the state family $\{A\theta(\rho) | \theta \in \Theta\}$. In the one-parameter case, we focus on the suprema

$$J^R[A\theta] := \sup_{\rho} J^R[A\theta, \rho], \quad J^S[A\theta] := \sup_{\rho} J^S[A\theta, \rho],$$

where $J^S[A\theta, \rho]$ and $J^R[A\theta, \rho]$ are the SLD and RLD Fisher informations when the input state is $\rho$. In particular, it is important to calculate the supremum $J^S[A\theta]$, which is smaller than $J^R[A\theta]$.

When $n$ applications of the unknown channel $A\theta$ are available, the input state $\rho_n$ and the measurement $M^n$ are given as a state on $(\mathcal{H} \otimes \mathcal{R})^{\otimes n}$ and a POVM on $(\mathcal{K} \otimes \mathcal{R})^{\otimes n}$. For a sequence of estimators $\{(\rho_n, M^n)\}$, we consider the asymptotic locally unbiased condition:

$$\lim_{n \to \infty} \int \hat{\theta} \text{Tr} A_{\theta_0}(\rho_n)M^n(d\hat{\theta}) = \theta_0, \quad \lim_{n \to \infty} \frac{d}{d\theta} \int \hat{\theta} \text{Tr} A_{\theta}(\rho_n)M^n(d\hat{\theta}) \bigg|_{\theta = \theta_0} = 1$$

for all points $\theta_0$, and denotes the MSE of $(\rho_n, M^n)$ by $\text{MSE}_n(\rho_n, M^n)$). Assume that $J^S[A_{\theta_0}^{\otimes n}]$ behaves as $O(n^\alpha)$ when $n$ goes to infinity. When $\{(\rho_n, M^n)\}$ satisfies the asymptotic locally unbiased condition, the inequality (5) yields that

$$\limsup_{n \to \infty} \frac{n^\alpha}{J^S[A_{\theta_0}^{\otimes n}]} \geq \limsup_{n \to \infty} \frac{\text{MSE}_n(\rho_n, M^n)}{J^S[A_{\theta_0}^{\otimes n}]}.$$
We define the Cramér-Rao bound:

\[ C_\alpha[A_B] = \inf_{\{\rho_n, M^n\}} \left\{ \limsup_{n \to \infty} n^n \text{MSE}_\theta(\rho_n, M^n) \middle| \{\rho_n, M^n\} \text{ satisfies the asymptotic locally unbiased condition.} \right\}. \]

Thus, we obtain

\[ C_\alpha[A_B] = \limsup_{n \to \infty} \frac{n^n}{J^S[A_B^\otimes n]} . \tag{8} \]

Since

\[ J^S[A_B^\otimes(n+m)] \geq J^S[A_B^\otimes n] + J^S[A_B^\otimes m], \tag{9} \]

the limit \( \lim_{n \to \infty} J^S[A_B^\otimes n] \) exists. (For example, see Lemma A.1 of [21].) Thus, \( C_\alpha[A_B] \) can be defined by \( \lim_{n \to \infty} \frac{n^n}{J^S[A_B^\otimes n]} \).

In order to treat the above values, we consider the following condition:

(C) The range of \( \rho_{[A_B]} \) contains the range of \( D[A_B]^2 \).

Assume that the condition (C) does not hold. When the input state is the maximally entangled state \( \frac{1}{\sqrt{n}} |\rangle \langle \rangle \rangle \), the RLD Fisher information diverges. So, \( J^R[A_B] \) is infinity.

**Theorem 1** When the condition (C) holds,

\[ J^R[A_B] = \| \text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]}^{-1} D[A_B] \|. \]

**Proof:** Assume that the input state is given by \( |\rangle \langle \rangle \rangle \) and \( A \) is an invertible matrix. Then, the range of \( (I \otimes A^T) \rho_{[A_B]}(I \otimes \overline{A}) \) contains the range of \( (I \otimes A^T) D[A_B](I \otimes \overline{A})^2 \).

Assume that \( \rho_{[A_B]} \) is invertible. Using the formula (1), we obtain

\[ J^R[A_B, |\rangle \langle \rangle \rangle] = \text{Tr}(I \otimes A^T) D[A_B](I \otimes \overline{A})^2((I \otimes A^T) \rho_{[A_B]}(I \otimes \overline{A}))^{-1} \]
\[ = \text{Tr}(I \otimes A^T) D[A_B](I \otimes \overline{A})(I \otimes A^T) D[A_B](I \otimes \overline{A})^{-1} \rho_{[A_B]}^{-1}(I \otimes A^T)^{-1} \]
\[ = \text{Tr}(I \otimes \overline{A} A^T) D[A_B] \rho_{[A_B]}^{-1} D[A_B] \]
\[ = \text{Tr}(\overline{A} A^T)(\text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]}^{-1} D[A_B]), \]

where \( ((I \otimes A^T) \rho_{[A_B]}(I \otimes \overline{A}))^{-1} \) is the inverse of \( (I \otimes A^T) \rho_{[A_B]}(I \otimes \overline{A}) \) on its range. So, the supremum of \( \text{Tr}(\overline{A} A^T)(\text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]}^{-1} D[A_B]) \) with the condition \( \text{rank} \overline{A} A^T = \dim \mathcal{H} \) equals \( \| \text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]}^{-1} D[A_B] \| \). Assume that \( \rho_{[A_B]} \) is non-invertible. We choose an arbitrary small real number \( \epsilon > 0 \). Similar calculations and the operator monotonicity of \( x \mapsto -x^{-1} \) yield that

\[ \text{Tr}(I \otimes A^T) D[A_B](I \otimes \overline{A})^2((I \otimes A^T)(\rho_{[A_B]} + \epsilon I)(I \otimes \overline{A}))^{-1} \leq \text{Tr}(\overline{A} A^T)(\text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]} + \epsilon I)^{-1} D[A_B]) \]
\[ \leq \text{Tr}(\overline{A} A^T)(\text{Tr}_\mathcal{K} D[A_B] \rho_{[A_B]}^{-1} D[A_B]), \tag{10} \]
which is a bounded value due to Condition (C). Now, we focus on two matrixes on the range of \((I \otimes A^T)\rho[\bar{A}]\) and \((I \otimes A^T)\rho[\bar{A}]\) and \((I \otimes A^T)D[\bar{A}]\). Since the right hand side of (10) is independent of \(\epsilon\), the range of \((I \otimes A^T)\rho[\bar{A}]\) contains that of \((I \otimes A^T)D[\bar{A}]\). So, taking the limit \(\epsilon \to 0\), we obtain

\[
J^R[A_0, |A\rangle \langle A|] = (I \otimes A^T)D[\bar{A}](I \otimes \bar{A})^{-1}((I \otimes A^T)\rho[\bar{A}])(I \otimes A^T)\rho[\bar{A}])^{-1}
\]

\[
= \lim_{\epsilon \to 0} Tr((I \otimes A^T)D[A_0](I \otimes \bar{A})^{-1}((I \otimes A^T)(\rho[A_0] + \epsilon I)(I \otimes \bar{A})^{-1})
\]

\[
\leq Tr \bar{A}A^T(Tr_{\bar{A}} D[A_0]\rho[A_0]^{-1}D[A_0])
\]

The remaining task is to show the inequality

\[
J^R[A_0, |A\rangle \langle A|] \leq \| Tr_{\bar{A}} D[A_0]\rho[A_0]^{-1}D[A_0]\|
\]

for a non-invertible matrix \(A\). Define \((I \otimes A^T)^{-1}\) and \((I \otimes \bar{A})^{-1}\) as the inverses of \(I \otimes A^T\) and \(I \otimes \bar{A}\) whose domains are the ranges of the matrixes \(I \otimes A^T\) and \(I \otimes \bar{A}\). Letting \(P\) and \(P'\) be the projections to the ranges of the matrixes \(I \otimes A^T\) and \(I \otimes \bar{A}\), we have \((I \otimes A^T) = P(I \otimes A^T) = (I \otimes A^T)P' = P(I \otimes A^T)P',\) and \((I \otimes \bar{A}) = P'(I \otimes \bar{A}) = (I \otimes \bar{A})P = P'(I \otimes \bar{A})P\). Then, \((I \otimes A^T)\rho[A_0](I \otimes \bar{A}) = P(I \otimes A^T)P'\rho[A_0](I \otimes \bar{A})P\).

In the following, we consider the case when \(\rho[A_0]\) is invertible. The matrix \((P'\rho[A_0]P')^{-1}\) is defined as the inverse of \(P'\rho[A_0]P'\) whose domain and range are the range of \(P'\). Thus,

\[
J^R[A_0, |A\rangle \langle A|] = Tr P((I \otimes A^T)D[A_0](I \otimes \bar{A})^{-1}((I \otimes A^T)\rho[A_0]P)\rho[A_0])^{-1}
\]

\[
= Tr(I \otimes A^T)D[A_0]P'(I \otimes \bar{A})^{-1}(P'\rho[A_0]P')^{-1}(I \otimes A^T)^{-1}P'I \otimes A^T)P'\rho[A_0]P'(I \otimes \bar{A})^{-1}
\]

\[
= Tr(P'\rho[A_0])^{-1}P'I \otimes A^T)D[A_0]P'(I \otimes \bar{A})^{-1}
\]

\[
\leq Tr \rho[A_0]^{-1}D[A_0](I \otimes \bar{A}A^T)D[A_0]
\]

\[
= Tr \bar{A}A^T(Tr_{\bar{A}} D[A_0]\rho[A_0]^{-1}D[A_0])
\]

Therefore, the inequality (11) holds.

Next, we consider the case when \(\rho[A_0]\) is non-invertible. We choose an arbitrary small real number \(\epsilon > 0\). Similar calculations and the operator monotonicity of \(x \mapsto x^{-1}\) yield that

\[
Tr P((I \otimes A^T)D[A_0](I \otimes \bar{A})^{-1}((I \otimes A^T)\rho[A_0] + \epsilon I)(I \otimes \bar{A})^{-1})
\]

\[
\leq Tr \bar{A}A^T(Tr_{\bar{A}} D[A_0]\rho[A_0] + \epsilon I^{-1}D[A_0])
\]

\[
\leq Tr \bar{A}A^T(Tr_{\bar{A}} D[A_0]\rho[A_0]^{-1}D[A_0]),
\]

which is a bounded value due to Condition (C). Now, we focus on two matrixes on the range of \(P, P((I \otimes A^T)\rho[A_0](I \otimes \bar{A})P\) and \(P((I \otimes A^T)D[A_0](I \otimes \bar{A})^{-2}P\). Since the right hand side of (13) is independent of \(\epsilon\), the range of \(P(I \otimes A^T)\rho[A_0](I \otimes \bar{A})P\)
contains that of $P((I \otimes A^T)D[A_\theta](I \otimes A))^2P$. So, taking the limit $\epsilon \to 0$, we obtain

$$J^R[A_\theta, |A|] = \text{Tr} P((I \otimes A^T)D[A_\theta](I \otimes A))^2P(I \otimes A^T)\rho[A_\theta](I \otimes A)^{-1} \rho[A_\theta]^{-1}(I \otimes A)P^{-1} \leq \text{Tr} A^T(\text{Tr}_\mathcal{K} D[A_\theta]|\rho[A_\theta]|^{-1}D[A_\theta]).$$

Therefore, the inequality \([11]\) holds.

**Theorem 2** When two channel families \(\{A_\theta\}\) and \(\{\tilde{A}_\theta\}\) satisfy the condition (C), then the additivity

$$J^R[A_\theta \otimes \tilde{A}_\theta] = J^R[A_\theta] + J^R[\tilde{A}_\theta]$$

holds.

**Proof:** Let \(\mathcal{K}\) and \(\tilde{\mathcal{K}}\) be output systems of the channels \(A_\theta\) and \(\tilde{A}_\theta\). Then, the relation \([11]\) guarantees

$$\text{Tr}_{\mathcal{K}\otimes\tilde{\mathcal{K}}} D[A_\theta] \otimes D[\tilde{A}_\theta] = \text{Tr}_\mathcal{K} D[A_\theta] \text{Tr}_\tilde{\mathcal{K}} D[\tilde{A}_\theta] = 0. \quad (14)$$

Since \(D[A_\theta] \otimes D[\tilde{A}_\theta] = D[A_\theta] \otimes \rho[\tilde{A}_\theta] + \rho[A_\theta] \otimes D[\tilde{A}_\theta]\), Theorem \([11]\) and \([13]\) yield

$$J^R[A_\theta \otimes \tilde{A}_\theta] = \|\text{Tr}_{\mathcal{K}\otimes\tilde{\mathcal{K}}}(D[A_\theta] \otimes \rho[\tilde{A}_\theta] + \rho[A_\theta] \otimes D[\tilde{A}_\theta])(\rho[\tilde{A}_\theta]^{-1} \otimes \rho[A_\theta]^{-1})(D[A_\theta] \otimes \rho[\tilde{A}_\theta] + \rho[A_\theta] \otimes D[\tilde{A}_\theta])\|$$

$$= \|\text{Tr}_{\mathcal{K}\otimes\tilde{\mathcal{K}}}(D[A_\theta]\rho[A_\theta]^{-1} D[A_\theta] \otimes \rho[\tilde{A}_\theta] + \rho[A_\theta] \otimes D[\tilde{A}_\theta]\rho[\tilde{A}_\theta]^{-1} D[\tilde{A}_\theta] + 2D[A_\theta] \otimes D[\tilde{A}_\theta])\|$$

$$= \|\text{Tr}_{\mathcal{K}} D[A_\theta] \rho[A_\theta]^{-1} D[A_\theta] + \| \text{Tr}_{\tilde{\mathcal{K}}} D[\tilde{A}_\theta] \rho[\tilde{A}_\theta]^{-1} D[\tilde{A}_\theta]\|$$

$$= J^R[A_\theta] + J^R[\tilde{A}_\theta], \quad (15)$$

where the equation \([15]\) follows from the additivity property concerning matrix norm:

$$\|X \otimes I + I \otimes Y\| = \|X\| + \|Y\|$$

for any two Hermitian matrices \(X\) and \(Y\).

**Corollary 1** When a channel family \(\{A_\theta\}\) satisfies the condition (C), then

$$J^R[A_\theta^{\otimes n}] = nJ^R[A_\theta].$$

Since \(nJ^S[A_\theta] \leq J^S[A_\theta^{\otimes n}] \leq J^R[A_\theta^{\otimes n}] \leq J^S[A_\theta^{\otimes n}]\), \(J^S[A_\theta^{\otimes n}]\) increases in order \(n\) under the assumption of Theorem \([11]\) i.e., \(J^S[A_\theta^{\otimes n}] = O(n)\). When the rank of \(\rho[A_\theta]\) is the maximum, i.e., \(dd'\), this condition holds and \(J^S[A_\theta^{\otimes n}] = O(n)\). However, there is an example that does not satisfy the above condition but satisfies the condition (C) as follows. So, the condition (C) is weaker than the condition that \(\rho[A_\theta]\) has the maximum rank.
A channel $A$ is called a phase damping channel when the output system $\mathcal{K}$ equals the input system $\mathcal{H}$ and there exist complex numbers $d_{k,l}$ such that

$$A_{\{d_{k,l}\}}(\rho) = \sum_{k,l} d_{k,l} \rho_{k,l} |k\rangle \langle l|,$$

where $\rho = \sum_{k,l} \rho_{k,l} |k\rangle \langle l|$. In this case, the state $\rho[A]$ is written as the following form

$$\rho[A_{\{d_{k,l}\}}] = \sum_{k,l} d_{k,l} |k\rangle \langle l| \langle l| \langle k|.$$

That is, the range of $\rho[A]$ is included by the space spanned by $\{|k\rangle_K|k\rangle_R\}$. When a channel family $\{A_\theta\}$ is given as a one-parameter subfamily of $\{\Lambda_{\{d_{k,l}\}}|d_{k,l}|$ is strictly positive$\}$, the condition (C) holds. Therefore, there exists a channel family that satisfies the condition (C) but consists of non-full-rank channels.

Further, we have the following observation.

**Corollary 2** Assume that the condition (C) holds and there exists a normalized vector $u$ in the input system $\mathcal{H}$ such that $\|\text{Tr}_K D[A_\theta] \rho[A_\theta]^{-1} D[A_\theta]\| = \|\text{Tr}_K D[A_\theta] \rho[A_\theta]^{-1} D[A_\theta]\|_u$, $\langle u | \rho[A_\theta] \langle u | \rangle = 0$, and $\text{Tr}_K \otimes \rho[A_\theta] = 0$. Then, $J^S[A_\theta] = J^R[A_\theta]$, and this bound can be attained by the input pure state $|u\rangle \langle u|$ on $\mathcal{H}$. That is, it can be attained without use of the reference system.

**Proof:** Since $\langle u | \rho[A_\theta] \langle u | \rangle = 0$, $J^S[A_\theta] = J^R[A_\theta]$, and this bound can be attained by the input pure state $|u\rangle \langle u|$ on $\mathcal{H}$. That is, it can be attained without use of the reference system.

4. Examples

We consider the case where the condition (C) does not hold. As the simplest example, we consider the one-parameter unitary case, i.e., the case when $A_\theta(\rho) = e^{i\theta H} \rho e^{-i\theta H}$ with an Hermitian matrix $H$. Using (12), we obtain

$$J^S[A_\theta] = (\lambda_{\text{max}}(H) - \lambda_{\text{min}}(H))^2,$$

where $\lambda_{\text{max}}(H)$ and $\lambda_{\text{min}}(H)$ are the maximum and minimum of eigenvalues of $H$. So, we obtain

$$J^S[A_\theta^{\otimes n}] = n^2 (\lambda_{\text{max}}(H) - \lambda_{\text{min}}(H))^2.$$

In particular, in the two-dimensional case, when $H = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix}$, the optimal input is $\sqrt{\frac{1}{4}}(|0\rangle \otimes n + |1\rangle \otimes n)$, which is called the noon state. This unitary estimation is called phase estimation and this estimation with the noon state was experimentally realized with $n = 4$[32,33] and $n = 10$[34].
Next, we consider the $d$-dimensional system $\mathbb{C}^d$ spanned by $\{|j\rangle\}_{j=0}^{d-1}$ and the unitary matrix $X$ defined as $X|j\rangle := |j+1\rangle \mod d$. Using a distribution $\{p_j\}_{j=0}^{d-1}$ and a real diagonal matrix $H$ with diagonal elements $\{h_j\}_{j=0}^{d-1}$, we define the TP-CP map $A_0$ by

$$A_0(\rho) := \sum_{j=0}^{d-1} p_j X^j e^{i\theta H} \rho e^{-i\theta H} X^{-j}.$$ 

This TP-CP map can be regarded as the stochasitc application of the unitary $X^j$ after the application of the unitary $e^{i\theta H}$.

Let $h_a$ and $h_b$ be the maximum and the minimum eigenvalues of $H$. Using two-dimensional reference system spanned by $|0\rangle_R$ and $|1\rangle_R$, we choose the following input state:

$$|\Phi_n\rangle := \frac{1}{\sqrt{2}}(|a\rangle^\otimes n|0\rangle_R + |b\rangle^\otimes n|1\rangle_R).$$

In this case, as the first step, we apply the following measurement $\{M_j\}$:

$$M_j := |j + (a, \ldots, a)\rangle \langle j + (a, \ldots, a)| \otimes |0\rangle_R \langle 0| + |j + (b, \ldots, b)\rangle \langle j + (b, \ldots, b)| \otimes |1\rangle_R \langle 1|.$$ 

When the outcome of this measurement is $j$, the resulting state is the pure state

$$\frac{1}{\sqrt{2}}(e^{inh_a}|j + (a, \ldots, a)\rangle|0\rangle_R + e^{inh_b}|j + (b, \ldots, b)\rangle|1\rangle_R). \quad (16)$$

The SLD Fisher information of the above family is $n^2(h_a - h_b)^2$. Therefore, since the maximum SLD Fisher information behaves as $O(n^2)$ at most, $J_S[A_0^\otimes n]$ behaves as $O(n^2)$.

5. Local asymptotic mini-max bound

In this section, we consider the relation between the discussion in the previous section and estimating protocols in a different viewpoint. Consider the phase estimation with inputing the noon state $\frac{1}{\sqrt{2}}(|0\rangle^\otimes n + |1\rangle^\otimes n)$. Then, the output state is $\frac{1}{\sqrt{2}}(e^{i\theta/2}|0\rangle^\otimes n + e^{-i\theta/2}|1\rangle^\otimes n)$. In this case, we cannot distinguish the parameters $\theta$ and $\theta + 2\pi/n$. For example, when we apply measurement $\{\frac{1}{\sqrt{2}}(|0\rangle^\otimes n \pm |1\rangle^\otimes n)\}$, the probability with the outcome $\frac{1}{\sqrt{2}}(|0\rangle^\otimes n \pm |1\rangle^\otimes n)$ equals $\cos^2 n\theta/2$, as is shown in Fig 3 and the Fisher information equals $n^2$. Even if the parameter $\theta$ is assumed to belong to $(0, \pi/n]$, we cannot distinguish the two parameters $\theta = \pi/3n$ and $\theta = 2\pi/3n$ with so high probability because we have only two outcomes. In order to distinguish two parameters $\theta = \pi/3n$ and $\theta = 2\pi/3n$ in this measurement, we need to repeat this measurement with several times, e.g., $k$. Since the number of application of the unknown unitary is $N := kn$, the error behaves as $\frac{1}{N} k$, which is different from $O(\frac{1}{Nk})$. So, we cannot conclude that the above method attains the order $O(\frac{1}{\sqrt{N}})$ concerning MSE. Therefore, we need to discuss what a bound can be attained globally, more carefully.
probability

Fig. 3. Phase estimation with noon state $n = 20$

For this purpose, we focus on an $\epsilon$-neighborhood $U_{\theta, \epsilon}$ of $\theta$ and define the local asymptotic mini-max risk:

$$C_\alpha[A_{\theta_0}, \{(\rho_n, M^n)\}] := \lim_{\epsilon \to 0} \limsup_{n \to \infty} \sup_{\theta \in U_{\theta_0, \epsilon}} n^\alpha \text{MSE}_\theta(\rho_n, M^n)$$

and the local asymptotic mini-max bound:

$$C_\alpha[A_{\theta_0}] := \inf_{\{(\rho_n, M^n)\}} C_\alpha[A_{\theta_0}, \{(\rho_n, M^n)\}].$$

Concerning the local asymptotic mini-max bound, we have the following two propositions.

**Proposition 1** When $\tilde{C}_\alpha[A_{\theta_0}]$ is continuous and the convergence (3) is compactly uniform,

$$C_\alpha[A_{\theta_0}] \geq \tilde{C}_\alpha[A_{\theta_0}].$$

**Proof:** For any $\delta > 0$, we choose an integer $N$ and an $\epsilon$-neighborhood $U_{\theta_0, \epsilon}$ satisfying

$$C_\alpha[A_{\theta_0}] + \delta \geq n^\alpha \text{MSE}_\theta(\rho_n, M^n), \quad \forall n \geq N, \forall \theta \in U_{\theta_0, \epsilon}.$$ (18)

We introduce two quantities

$$\eta_n(\theta) := \int \hat{\theta} \text{Tr} A_\theta(\rho_n) M^n(d\hat{\theta})$$

$$\nu_n(\theta) := \int (\hat{\theta} - \eta_n(\theta))^2 \text{Tr} A_\theta(\rho_n) M^n(d\hat{\theta}).$$
Then, we obtain
\[ \text{MSE}_\theta(\rho_n, M^n) = v_n(\theta) + (\eta_n(\theta) - \theta)^2. \]
Let \( J_{\theta,n} \) be the Fisher information of the distribution family \( \{ \text{Tr} \Lambda_\theta(\rho_n)M^n(d\tilde{\theta}) \} \theta \in \Theta \). This quantity is smaller than \( J^S[A^{\otimes n}_\theta] \). Deforming the classical Cramér-Rao inequality, we obtain
\[ v_n(\theta) \geq \left( \frac{d\eta_n(\theta)}{d\theta} \right)^2 J_{\theta,n}. \]
Thus, we obtain
\[ \text{MSE}_\theta(\rho_n, M^n) \geq \left( \frac{d\eta_n(\theta)}{d\theta} \right)^2 J_{\theta,n} + (\eta_n(\theta) - \theta)^2. \] (19)
As is shown later, for any \( \delta > 0 \), there exists a sufficiently large integer \( N \) satisfying the following. For any \( n \geq N \), there exists \( \theta_n \in U_{\theta_0, \epsilon} \) such that
\[ \frac{d\eta_n(\theta_n)}{d\theta} \geq 1 - \delta. \] (20)
Using (19), we obtain
\[ \text{MSE}_{\theta_n}(\rho_n, M^n) \geq \left( \frac{d\eta_n(\theta_n)}{d\theta} \right)^2 J_{\theta_n,n}. \] (21)
Take the limit \( n \to \infty \). Then, the continuity of \( \tilde{C}_{\alpha}[A_{\theta_0}] \), the compact uniformity of the convergence (8), (18), and (21) imply that
\[ C_{\alpha}[A_{\theta_0}] + \delta \geq (1 - \delta)^2 \tilde{C}_{\alpha}[A_{\theta_0}]. \]
Taking the limit \( \delta \to 0 \), we obtain (17).

Finally, we show the existence of the integer \( N \) satisfying the above condition given by (20) by using reduction to absurdity. Assume that for any \( \delta > 0 \), there exists a subsequence \( n_k \) such that \( \frac{d\eta_{n_k}(\theta)}{d\theta} < 1 - \delta \) for any \( \theta \in U_{\theta_0, \epsilon} \). Thus,
\[ [\eta_{n_k}(\theta + \epsilon/2) - (\theta + \epsilon/2)] - [\eta_{n_k}(\theta - \epsilon/2) - (\theta - \epsilon/2)] \]
\[ = \eta_{n_k}(\theta + \epsilon/2) - \eta_{n_k}(\theta - \epsilon/2) - \epsilon < -\epsilon \delta. \]
Then,
\[ \max\{ |\eta_{n_k}(\theta + \epsilon/2) - (\theta + \epsilon/2)|, |\eta_{n_k}(\theta - \epsilon/2) - (\theta - \epsilon/2)| \} > \frac{\epsilon \delta}{2}. \]
That is,
\[ \max\{ (\eta_{n_k}(\theta + \epsilon/2) - (\theta + \epsilon/2))^2, (\eta_{n_k}(\theta - \epsilon/2) - (\theta - \epsilon/2))^2 \} > \frac{\epsilon^2 \delta^2}{4}. \]
Using (19), we obtain
\[ \max\{ \text{MSE}_{\theta+\epsilon/2}(\rho_{n_k}, M^{n_k}), \text{MSE}_{\theta-\epsilon/2}(\rho_{n_k}, M^{n_k}) \} \geq \frac{\epsilon^2 \delta^2}{4}. \]
Since \( \text{MSE}_{\theta+\epsilon/2}(\rho_n, M^n) \) behaves as \( O\left(\frac{1}{n^2}\right) \), we obtain contradiction. \( \blacksquare \)
Proposition 2 Assume that $E := \sup_{\theta \in \Theta} |\theta| < \infty$. When the order parameter $\alpha$ equals 1 and $\tilde{C}_1[A_{\theta_0}]$ is continuous, 

$$C_1[A_{\theta_0}] = \tilde{C}_1[A_{\theta_0}] = \lim_{n \to \infty} \frac{n}{J^S[A_{\theta_0}^m]}.$$ 

Proof: As is mentioned in Section 3, (9) guarantees the convergence of $\lim_{n \to \infty} \frac{n}{J^S[A_{\theta_0}^m]}$. It is enough to show the inequality $\tilde{C}_1[A_{\theta_0}] \geq C_1[A_{\theta_0}]$. For an arbitrary real number $\delta > 0$ and an arbitrary integer $m$, let $\{\rho_m, M^m\}$ be a locally unbiased estimator at $\theta_0$ such that 

$$\frac{1}{J^S[A_{\theta_0}^m]} + \delta > \text{MSE}_{\theta_0}(\rho_m, M^m).$$

(22) 

We define another coordinate $\eta(\theta)$ by 

$$\eta(\theta) := \int \theta \text{Tr} \Lambda_{\theta_0} M^n(d\theta),$$

and denote the MSE concerning the parameter $\eta$ of an estimator $(\rho_m, M^m)$ by $\text{MSE}_{\eta}(\theta)$, $(\rho_m, M^m)$. For any $\delta > 0$, we choose an integer $m$, a sufficiently small $\epsilon'$-neighborhood $U_{\theta_0, \epsilon}$ such that 

$$\tilde{C}_1[A_{\theta_0}^m] + \delta \geq \frac{1}{J^S[A_{\theta_0}^m]}$$

(23) 

$$1 - \delta \leq \frac{\eta(\theta) - \eta(\theta')}{\theta - \theta'} \leq 1 + \delta$$

(24) 

for $\forall \theta, \theta' \in U_{\theta_0, \epsilon}$. 

Next, let $(\rho_{nm}^n, M_{nm}^n)$ be the estimator given as the average value of $n$ times applications of the estimator $(\rho_m, M^m)$ concerning the original parameter $\theta$. Then, we can choose a sufficiently large number $n$ satisfying the following: When the true parameter is $\theta_0$, the estimate of $(\rho_{nm}^n, M_{nm}^n)$ belongs to $U_{\theta_0, \epsilon'}$ with the probability higher than $1 - \delta$. The second inequality in (24) guarantees that 

$$\frac{1}{n} \text{MSE}_{\theta_0}(\rho_m, M^m) + \delta E = \text{MSE}_{\theta_0}(\rho_{nm}^n, M_{nm}^n) + \delta E \geq \frac{1}{(1 + \delta)^2} \text{MSE}_{\eta}(\rho_{nm}^n, M_{nm}^n).$$

(25) 

Thus, (22), (23), and (25) imply that 

$$(1 + \delta)^2 \left( \frac{\tilde{C}_1[A_{\theta_0}^m]}{nm} + \frac{2\delta}{n} + \delta E \right) \geq \text{MSE}_{\eta}(\theta_0)(\rho_{nm}^n, M_{nm}^n).$$

Since $\text{MSE}_{\eta}(\theta_0)(\rho_{nm}^n, M_{nm}^n)$ is continuous and $(1 + \delta)^2 \left( \frac{\tilde{C}_1[A_{\theta_0}^m]}{nm} + \frac{2\delta}{n} + \delta E \right) > (1 + \delta)^2 \left( \frac{\tilde{C}_1[A_{\theta_0}^m]}{nm} + \frac{2\delta}{n} + \delta E \right)$, we can choose a sufficiently small number $0 < \epsilon' < \epsilon$ such that 

$$(1 + \delta)^2 \left( \frac{\tilde{C}_1[A_{\theta_0}^m]}{nm} + \delta(E + 2) \right) \geq \text{MSE}_{\eta}(\theta_0)(\rho_{nm}^n, M_{nm}^n)$$

(26)
for $\theta \in U_{\theta_0, \epsilon'}$.

Let $(\rho_{\ln m}^\prime, M_{\ln m}^\prime)$ be the estimator given as the average value of $l$ times applications of the estimator $(\rho_{\ln m}^\prime, M_{\ln m}^\prime)$ concerning $\eta$. Since the estimator $(\rho_{\ln m}^\prime, M_{\ln m}^\prime)$ is unbiased concerning the parameter $\eta$,\

$$
\text{MSE}_{\eta}(\rho_{\ln m}^\prime, M_{\ln m}^\prime) = \frac{\text{MSE}_{\eta}(\rho_{\ln m}^\prime, M_{\ln m}^\prime)}{l},
$$

(27)

We choose a sufficiently large number $l$ satisfying the following: When the true parameter is $\theta \in U_{\theta_0, \epsilon}/2$, the estimate $\eta$ of $(\rho_{\ln m}^\prime, M_{\ln m}^\prime)$ belongs to $U_{\theta_0, \epsilon}/2$ with the probability $1 - p_l$, where the probability $p_l$ exponentially goes to 0 as $l \to \infty$.

The first inequality in (24) guarantees that\

$$
\frac{1}{(1 - \delta)^2} \text{MSE}_{\eta}(\rho_{\ln m}^\prime, M_{\ln m}^\prime) + E_p l \geq \text{MSE}_{\eta}(\rho_{\ln m}^\prime, M_{\ln m}^\prime), \quad \forall \theta \in U_{\theta_0, \epsilon}/2.
$$

(28)

Thus, the relations (26), (27), and (28) imply\

$$
\frac{(1 + \delta)^2}{l(1 - \delta)^2} (\tilde{C}_1[A_{\theta_0}] + nm\delta(E + 2)) \geq \lim sup_{k \to \infty} \text{MSE}_{\eta}(\rho_k^m, M_k^m), \quad \forall \theta \in U_{\theta_0, \epsilon}/2.
$$

(29)

Taking the limit $l \to \infty$, we obtain\

$$
\frac{(1 + \delta)^2}{(1 - \delta)^2} (\tilde{C}_1[A_{\theta_0}] + nm\delta(E + 2)) \geq \lim_{l \to \infty} \text{MSE}_{\eta}(\rho_{\ln m}^\prime, M_{\ln m}^\prime), \quad \forall \theta \in U_{\theta_0, \epsilon}/2.
$$

Finally, using the above $n$ and $m$, we define a sequence of estimators $\{(\rho_k^m, M_k^m)\}$ by the following way. For given $k$, we choose maximum $l$ such that $nm \leq k$. Then, the estimator $(\rho_k^m, M_k^m)$ defined as $(\rho_{\ln m}^\prime, M_{\ln m}^\prime)$. In this definition, we only use $nm$ applications, and the remaining $k - nm$ applications are discarded. So, we obtain\

$$
\frac{(1 + \delta)^2}{(1 - \delta)^2} (\tilde{C}_1[A_{\theta_0}] + nm\delta(E + 2)) \geq \limsup_{k \to \infty} k \text{MSE}_{\eta}(\rho_k^m, M_k^m), \quad \forall \theta \in U_{\theta_0, \epsilon}/2,
$$

which implies that\

$$
\frac{(1 + \delta)^2}{(1 - \delta)^2} (\tilde{C}_1[A_{\theta_0}] + nm\delta(E + 2)) \geq C_1[A_{\theta_0}, \{(\rho_k^m, M_k^m)\}_{k=1}^{\infty}] \geq C_1[A_{\theta_0}].
$$

Since $\delta > 0$ is arbitrary, $\tilde{C}_1[A_{\theta_0}] \geq C_1[A_{\theta_0}]$. 

Now, remember that the Cramér-Rao bound can be attained by using the two-step method in the case of state estimation. By using the two-step method\[22, 23\], the local asymptotic mini-max bound $C_\alpha[A_{\theta_0}]$ can be attained at all points $\theta$ as follows.
Proposition 3 Assume that $E := \sup_{\theta \in \Theta} |\theta| < \infty$ and $C_1[A_{\theta}]$ is continuous. For any $\delta > 0$, there exists a sequence of estimators $\{(\rho_n, M^n)\}$ such that

$$C_\alpha[A_{\theta}, (\rho_n, M^n)] \leq C_\alpha[A_{\theta}] + \delta$$

(30)

for all points $\theta$. Further, when the parameter space $\Theta$ is compact,

$$\lim_{n \to \infty} n^\alpha \min_{(\rho_n, M^n)} \max_{\theta \in \Theta} \text{MSE}_\theta(\rho_n, M^n) = \max_{\theta \in \Theta} C_\alpha[A_{\theta}].$$

(31)

Proof: We use a two-step method slightly different from [23]. Before applying the unknown channel $A_{\theta}$, for any real number $\delta > 0$, we choose an $\epsilon_i$-neighborhood $U_{\theta_i, \epsilon_i}$ satisfying the following three conditions: (1) $\cup_i U_{\theta_i, \epsilon_i} = \Theta$. (2) For any $\theta_i$, there exists a sequence of estimators $\{(\rho_n(\theta_i), M^n(\theta_i))\}$ such that $C_\alpha[A_{\theta_i}] + \delta/2 \geq \limsup_{n \to \infty} n^\alpha \sup_{\theta \in U_{\theta_i, \epsilon_i}} \text{MSE}_\theta(\rho_n(\theta_i), M^n(\theta_i))$. (3) $\sup_{\theta \in U_{\theta_i, \epsilon_i}} C_\alpha[A_{\theta}] \leq C_\alpha[A_{\theta}] + \delta/2$.

We divide $n$ applications of the unknown channel $A_{\theta}$ to two groups: The first group consists of $\sqrt{n}$ applications and the second group consists of $n - \sqrt{n}$ applications. In the first step, we apply a POVM $M$ to the first group. This POVM $M$ is a POVM on the single system $H$ and is required to satisfy that $J_\theta$ is non-degenerate at all points $\theta$. Based on $\sqrt{n}$ obtained data, we estimate which $\epsilon_i$-neighborhood $U_{\theta_i, \epsilon_i}$ contains the true parameter, and obtain the first step estimate $\theta_i$. The error probability $P_{\theta_i}$ of this step goes to 0 exponentially, i.e., $P_{\theta_i}$ behaves as $e^{-c\sqrt{n}}$, where $c$ depends on $\theta$.

In the second step, we apply the estimator $\rho_n - \sqrt{n}(\theta_i)$ to the second group, and obtain our final estimate from the outcome of the estimator $\rho_n - \sqrt{n}(\theta_i)$. We express this estimator by $\rho_n$ and $M^n$. Its MSE is evaluated as

$$\text{MSE}_\theta(\rho_n, M^n) \leq EP_{\theta_i} + (1 - P_{\theta_i}) \sup_{\theta \in U_{\theta_i, \epsilon_i}} \text{MSE}_\theta(\rho_n - \sqrt{n}(\theta_i), M^n - \sqrt{n}(\theta_i)).$$

(32)

Since $n^\alpha EP_{\theta_i}$ goes to 0, we obtain

$$\limsup_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n) \leq \sup_{\theta \in U_{\theta_i, \epsilon_i}} C_\alpha[A_{\theta}] + \delta/2 \leq C_\alpha[A_{\theta}] + \delta.$$

Thus, we obtain (30). Further, the relation (32) yields that

$$\sup_{\theta \in \Theta} \text{MSE}_\theta(\rho_n, M^n) \leq n^\alpha EP_{\theta_i} + \sup_{\theta \in \Theta} (1 - P_{\theta_i}) \sup_{\theta \in U_{\theta_i, \epsilon_i}} \text{MSE}_\theta(\rho_n - \sqrt{n}(\theta_i), M^n - \sqrt{n}(\theta_i)).$$

The compactness of $\Theta$ guarantees that $\sup_{\theta \in \Theta} n^\alpha EP_{\theta_i} \to 0$. Thus,

$$\limsup_{n \to \infty} n^\alpha \sup_{\theta \in \Theta} \text{MSE}_\theta(\rho_n, M^n) \leq \sup_{\theta \in \Theta} \sup_{\theta \in U_{\theta_i, \epsilon_i}} C_\alpha[A_{\theta}] + \delta/2 \leq \sup_{\theta \in \Theta} C_\alpha[A_{\theta}] + \delta.$$

Since the part $\geq$ of (31) is trivial, we obtain (31).
Proposition 3 holds even when we replace the MSE by a general error function $R(\theta, \hat{\theta})$ for one-parametric family satisfying the following conditions: (1) the relation $R(\theta, \hat{\theta}) \leq (\hat{\theta} - \theta)^2$ holds with a local coordinate when $\hat{\theta}$ is close to $\theta$. (2) the maximum of $R(\theta, \hat{\theta})$ exists. Therefore, we can apply Proposition 3 to the following case: Assume that the one-parameter channel family $\{A_\theta\}$ has a compact group covariant structure, that is, its parameter space is given as an interval $[a, b]$ and there is a unitary representation $U_\rho$ of $\mathbb{R}$ such that $U_\rho A_\theta(\rho) U_\rho^\dagger = A_{\theta + \rho}$. The error is given by $\min_{k \in \mathbb{Z}} \{(\hat{\theta} + k(b - a) - \theta)^2\}$ instead of the square error $(\hat{\theta} - \theta)^2$. In this case, due to the group covariance, $C_\alpha[A_\theta]$ does not depend on the true parameter $\theta$. Application of Proposition 3 implies that the global min-max error behaves $C_\alpha[A_\theta] \frac{1}{n^{\alpha}}$.

In the phase estimation case, the unknown parameter $\theta$ belongs to $[0, 2\pi)$, and the minimum of the worst value of the average error $\max_{\rho} \text{MSE}_\theta(\rho, n^{\alpha})$ behaves as $\frac{1}{n^{\alpha}} \Theta[0, 1]$. That is, the leading decreasing order is $O(1/n^2)$ and the leading decreasing coefficient is $\pi^2$ when we apply the optimal estimator. Proposition 3 implies that $\max_{\rho} C_2[A_\theta] = \pi^2$. Since $C_2[A_\theta]$ does not depend on $\theta$ due to the homogenous structure, we can conclude that $C_2[A_\theta] = \pi^2$. So, the equation $J^2[A_\theta] = n^2$ implies the equation $C_2[A_\theta] = 1$. Hence, the Cramér-Rao bound $C_2[A_\theta]$ cannot be attained globally in this model. However, it can be attained in a specific point in the following sense.

**Proposition 4** Assume that $E := \sup_{\theta \in \Theta} |\theta| < \infty$ and $C_1[A_{\theta_0}]$ is continuous. For any $\delta > 0$ and any $\theta_0 \in \Theta$, there exists a sequence of estimators $\{(\rho_n, m_n^{\alpha})\}$ satisfying the asymptotically locally unbiased condition and the relations:

$$
\limsup_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, \theta_0, m_n^{\alpha}) \leq C_\alpha[A_\theta] + \delta, \quad \forall \theta \neq \theta_0
$$

$$
\limsup_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, \theta_0, m_n^{\alpha}) \leq \tilde{C}_\alpha[A_{\theta_0}] + \delta.
$$

In estimation of probability distribution, there exists a superefficient estimator that has smaller error at a discrete set than the Cramér-Rao bound [20]. Since such a superefficient estimator cannot be useful, many statisticians think that it is better to impose a condition for our estimators for removing superefficient estimators. In this classical case, if we assume the asymptotic locally unbiased condition, we have no superefficient estimator. Proposition 4 means that even if the asymptotic locally unbiased condition is assumed, there exists an estimator that behaves in the similar way to a superefficient estimator in the case of unitary estimation. So, we call such an estimator a $q$-channel-superefficient estimator.

That is, a sequence of estimators $\{(\rho_n, m_n^{\alpha})\}$ is called $q$-channel-superefficient at $\theta$ with the order $n^{\alpha}$ when $\limsup_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, m_n^{\alpha}) < C_\alpha[A_\theta]$. Hence, in order to remove the $q$-channel-superefficiency problem, it is better to adopt the bound $C_\alpha[A_\theta]$ as the criterion instead of $\tilde{C}_\alpha[A_\theta]$.

**Proof:** We choose $\epsilon_n$-neighborhoods $U_{\theta_0, \epsilon_n}$ in the same way, and define the neighborhood $U_{\theta_0, \epsilon_n}$. We apply the same first step as Proposition 3 to neighborhoods $\{U_{\theta_0, \epsilon_n}\}$, and obtain the first step estimate $\hat{\theta}_1$. When the first step estimate $\hat{\theta}_1$ is not $\theta_0$, we apply the same method as Proposition 3 in the second
Comparison between Cramer-Rao and mini-max approaches

step. When the first step estimate $\theta_i$ is $\theta_0$, we apply the asymptotically locally unbiased estimator whose MSE behaves as $(\tilde{C}_\alpha[A_\theta]+\delta)/n^\alpha$ asymptotically.

Further, since there exists an asymptotically locally unbiased estimator that surpasses the bound $C_\alpha[A_\theta]$, the asymptotically locally unbiased condition is too weak for deriving the local asymptotic mini-max bound, which is more meaningful. In order to avoid this problem, it is sufficient to impose the following condition:

(CU) The limit $\lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n)$ exists for all $\theta$ and this convergence is compactly uniform concerning $\theta$.

Under the condition (CU), $\lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n)$ is continuous concerning $\theta$, and

$\lim_{n \to \infty} n^\alpha \sup_{\theta \in U_{\theta_0}, \epsilon} \text{MSE}_\theta(\rho_n, M^n) = \sup_{\theta \in U_{\theta_0, \epsilon}} \lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n)$.

Thus,

$\lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n) = \lim_{\epsilon \to 0} \sup_{\theta \in U_{\theta_0, \epsilon}} \lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n) = \lim_{\epsilon \to 0} \lim_{n \to \infty} n^\alpha \sup_{\theta \in U_{\theta_0, \epsilon}} \text{MSE}_\theta(\rho_n, M^n) = C_\alpha[A_\theta, \{(\rho_n, M^n)\}]$.

Therefore, we obtain the following corollary.

**Corollary 3** When a sequence of estimators $\{(\rho_n, M^n)\}$ satisfies the condition (CU),

$\lim_{n \to \infty} n^\alpha \text{MSE}_\theta(\rho_n, M^n) \geq C_\alpha[A_\theta]$.

Therefore, the condition (CU) is better in estimation of quantum channel than the asymptotically locally unbiased condition.

Finally, we consider the relation with the adaptive method proposed by Nagaoka [16]. In this method, we apply our POVM to each single system $\mathcal{H}$, and we decide the $k$-th POVM based on the knowledge of previous $k - 1$ outcomes. In this case, Fujiwara [30] analyzed the asymptotic behavior of the MSE of this estimator. Now, we consider the case of $nm$ applications of the unknown channel $A_\theta$. In this case, we divide $nm$ applications into $n$ groups consisting of $m$ applications. When we apply the adaptive method mentioned in Fujiwara [30] to these groups, the MSE of this estimator behaves as $\frac{n^J S[\Lambda_\theta^m]}{nm}$, which is close to $\frac{C_\alpha[A_\theta]}{nm^\alpha}$. So, when $\alpha > 1$, this method cannot realize the optimal order $O(\frac{1}{(nm)^\alpha})$.

**6. Discussion**

We have compared the Cramér-Rao bound $\tilde{C}_\alpha[A_\theta]$ and the local asymptotic mini-max bound $C_\alpha[A_\theta]$ in quantum channel estimation, which contains quantum state estimation. When the model has group covariant structure, the local asymptotic mini-max bound $C_\alpha[A_\theta]$ coincides with the limit of the global mini-max bound. We have also shown that both bounds $\tilde{C}_\alpha[A_\theta]$ and $C_\alpha[A_\theta]$ coincide in quantum channel estimation when the maximum of SLD Fisher information...
$J^S[\Lambda_g^\otimes n]$ behaves as $O(n)$. The case of state estimation can be regarded as a special case of this case. That is, the conventional state estimation has no difference between both bounds. However, we have shown that the Cramér-Rao bound $\tilde{C}_\alpha[A_0]$ is different from the local asymptotic mini-max bound $C_\alpha[A_0]$ in the phase estimation. So, we can conclude that the local asymptotic mini-max bound $C_\alpha[A_0]$ is more meaningful and does not necessarily coincide with the Cramér-Rao bound $\tilde{C}_\alpha[A_0]$.

In order to clarify the asymptotic leading order of $J^S[\Lambda_g^\otimes n]$, we have derived the condition (C) as a sufficient condition for $J^S[\Lambda_g^\otimes n] = O(n)$. That is, the condition $\| \text{Tr}_K D[A_0]|\rho[A_0]\|^{-1}D[A_0]\| = \infty$ is a necessary condition for square speedup. This condition has been derived from the following two facts. One is the supremum of the RLD Fisher information satisfies the additive property. The other is the RLD Fisher information is an upper bound of the SLD Fisher information. This, the supremum of the RLD Fisher information is the upper bound of the regularized supremum of the SLD Fisher information, which equals the inverse of the Cramér-Rao bound. However, it is an open problem to clarify whether this upper bound can be attained by the regularized SLD Fisher information.

Further, Fujiwara and Imai [2] and Matsumoto [14] also obtained another sufficient condition. Since the relation with their conditions is not clear, its clarification is an open problem. Our condition (C) trivially contains the case when the state $|\rho[A_0]\rangle$ is a full rank state on the tensor product system while it is not so easy to derive the above full rank condition from Fujiwara and Imai’s condition. Further, we have also obtained another example for $J^S[\Lambda_g^\otimes n] = O(n^2)$ under the condition (C). This example is a larger class than the unitary model. So, we can expect that $J^S[\Lambda_g^\otimes n]$ behaves as $O(n^2)$ if the condition (C) does not hold. This is a challenging open problem.

Acknowledgment

The author was partially supported by a Grant-in-Aid for Scientific Research in the Priority Area ‘Deepening and Expansion of Statistical Mechanical Informatics (DEX-SMI)’, No. 18079014 and a MEXT Grant-in-Aid for Young Scientists (A) No. 20686026. The Centre for Quantum Technologies is funded by the Singapore Ministry of Education and the National Research Foundation as part of the Research Centres of Excellence programme. The author thanks Mr. Wataru Kumagai for helpful comments. He also thanks the referees for helpful comments concerning this manuscript. In particular, the first referee’s report was much help to improve the presentation in Section 1.

Lemmas needed for Theorem 1

Lemma 1 Any strictly positive definite matrix $A$ and any projection $P$ satisfy the inequality

$$A^{-1} \geq (PAP)^{-1},$$

where $(PAP)^{-1}$ is the inverse matrix with the domain $P$. 


Proof: Let \( R \) be the operator norm of the matrix \( PA(I-P) \). Lemma 2 guarantees the inequality
\[
PA(I-P) + (I-P)AP \leq \epsilon P + \frac{R^2}{\epsilon} (I-P)
\] (34)
for any \( \epsilon > 0 \). Thus,
\[
A \leq PAP + (I-P)A(I-P) + \epsilon P + \frac{R^2}{\epsilon} (I-P)
\]
\[
= P(A + \epsilon I)P + (I-P)(A + \frac{R^2}{\epsilon} I)(I-P)
\]
Since the function \( x \mapsto -x^{-1} \) is operator monotone,
\[
A^{-1} \geq (P(A + \epsilon I)P)^{-1} + ((I-P)(A + \frac{R^2}{\epsilon} I)(I-P))^{-1} \geq (P(A + \epsilon I)P)^{-1}.
\]
Taking the limit \( \epsilon \to 0 \), we obtain (33).

Lemma 2 Let \( A \) be a positive semi-definite matrix and \( P \) be a projection. Then, the inequality
\[
PA(I-P) + (I-P)AP \leq \epsilon P + \frac{R^2}{\epsilon} (I-P)
\] (35)
holds for any \( \epsilon > 0 \), where \( R \) is the operator norm of the matrix \( PA(I-P) \).

Proof: Choose an arbitrary normalized vector \( u \). Let \( t \) be \( \|Pu\|^2 \). Then,
\[
\langle u \ | PA(I-P) + (I-P)AP | u \rangle \leq 2\sqrt{t} \sqrt{1-tR}
\]
\[
\leq t\epsilon + (1-t) \frac{R^2}{\epsilon} = \langle u | \epsilon P + \frac{R^2}{\epsilon} (I-P) | u \rangle,
\]
which implies (35).

References
1. Fujiwara A.: Quantum channel identification problem. *Phys. Rev. A* 63 042304 (2001)
2. Fujiwara A., Imai H.: A fibre bundle over manifolds of quantum channels and its application to quantum statistics. *J. Phys. A: Math. Theor.* 41 255304 (2008)
3. Fujiwara A., Imai H.: Quantum parameter estimation of a generalized Pauli channel. *J. Phys. A: Math. Gen.* 36 8093 (2003)
4. Fujiwara A.: Estimation of SU(2) operation and dense coding: an information geometric approach. *Phys. Rev. A* 65 012316 (2002)
5. Imai H., Fujiiwara F.: Geometry of optimal estimation scheme for SU(D) channels. *J. Phys. A: Math. Theor.* 40 4391 (2007)
6. Hayashi M.: Parallel treatment of estimation of SU(2) and phase estimation. *Phys. Lett. A* 354 183 (2006)
7. Chiribella G., D’Ariano G. M., Perinotti P., Sacchi M. F.;: Efficient use of quantum resources for the transmission of a reference frame. *Phys. Rev. Lett.* 93 180503 (2004)
8. Bagan E., Baig M., Munoz-Tapia R: Quantum reverse-engineering and reference frame alignment without non-local correlations. *Phys. Rev. A* 70 030301 (2004)
9. A. Luis and J. Perina,: Optimum phase-shift estimation and the quantum description of the phase difference. *Phys. Rev. A*, 54, 4564 (1996).
10. Buzek V., Derka R., Massar S.: Optimal quantum clocks. *Phys. Rev. Lett.* 82 2207 (1999)
11. Imai H., Hayashi M.: Fourier Analytic Approach to Phase Estimation in Quantum Systems. New J. Phys. 11 043034 (2009)
12. Kahn J.: Fast rate estimation of an unitary operation in SU(d). Phys. Rev. A 75 022326 (2007)
13. Hotta M., Karasawa T., Ozawa M.: N-body-extended channel estimation for low-noise parameters. J. Phys. A: Math. Gen. 39 14465 (2006)
14. Matsumoto K.: On metric of quantum channel spaces. arXiv:1005.4759 (2010)
15. Holevo A. S.: Covariant measurements and uncertainty relations. Rep. Math. Phys., 16, 385–400, (1979).
16. Hotta M., Karasawa T., Ozawa M.: N-body-extended channel estimation for low-noise parameters. J. Phys. A: Math. Gen. 39 14465 (2006)
17. Matsumoto K.: On metric of quantum channel spaces. arXiv:1005.4759 (2010)
18. Hayashi M.: Two quantum analogues of Fisher information from a large deviation viewpoint of quantum estimation. J. Phys. A: Math. Gen. 35 7689 (2002)
19. Helstrom C. W.: Quantum Detection and Estimation Theory (New York: Academic Press) (1976)
20. Holevo A. S.: Probabilistic and Statistical Aspects of Quantum Theory (Amsterdam: North-Holland) (1982); (Originally published in Russian 1980)
21. Hayashi M.: Quantum Information: An Introduction (Berlin: Springer) (2006)
22. Hotta M., Matsumoto K.: Statistical model with measurement degree of freedom and quantum physics. RIMS kokyuroku No 1055 (Kyoto: Kyoto University) p 96 (1998) (In Japanese); Hayashi M., Matsumoto K.: Asymptotic Theory of Quantum Statistical Inference ed M Hayashi (Singapore: World Scientific) p 125 (2005) (reprinted)
23. Gill R., Massar S.: State estimation for large ensembles. Phys. Rev. A 61 042312 (2000)
24. Hayashi M.: Quantum estimation and the quantum central limit theorem. in Selected Papers on Probability and Statistics (American Mathematical Society Translations Series 2) Vol. 277, pp 95 - 123 (2009). (It was originally published in Japanese in Bulletin of Mathematical Society of Japan, Sugaku, Vol. 55, No. 4, 368–391 (2003).)
25. Guiţă M., Kahn J.: Local asymptotic normality for qubit states. Phys. Rev. A, 73, 052108 (2006)
26. Guiţă M., Jencova A.: Local asymptotic normality in quantum statistics. Commun. Math. Phys., 276, 341-379 (2007)
27. Guiţă M., Janssens B., Kahn J.: Optimal estimation of qubit states with continuous time measurements. Commun. Math. Phys., 277, 127-160 (2008)
28. Hayashi M., Matsumoto K.: Asymptotic performance of optimal state estimation in qubit system. Journal of Mathematical Physics, 49, 102101 (2008)
29. LeCam L.: Asymptotic Methods in Statistical Decision Theory. (New York: Springer) (1986)
30. Fujiwara A.: Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. J. Phys. A: Math. Gen. 39 12489 (2006)
31. Higgins B. L., Berry D. W., Bartlett S. D., Wiseman H. M., Pryde G. J.: Entanglement-free Heisenberg-limited phase estimation. Nature, 450, 393-396 (2007)
32. Nagata T., Okamoto R., O’Brien J. L., Sasaki K., Takeuchi S.: Beating the Standard Quantum Limit with Four-Entangled Photons. Science, 316, no. 5825, 726 (2007)
33. Okamoto R., Hofmann H. F., Nagata T., O’Brien J. L., Sasaki K., Takeuchi S.: Beating the standard quantum limit: phase super-sensitivity of N-photon interferometers. New J. Phys., 10 073033 (2008)
34. Jones J. A., Karlen S. D., Fitzsimons J., Aradian A., Benjamin S. C., Briggs G. A. D., Morton J. J. L.: Magnetic Field Sensing Beyond the Standard Quantum Limit Using 10-Spin NOON States. Science 324, 1166-1168 (2009)
35. Kitaev A. Y., Shen A. H., VyaIy M. N.: Classical and Quantum Computation, (Graduate Studies in Mathematics 47). American Mathematical Society, (2002)
36. Giovannetti V., Lloyd S., Maccone L.: Quantum-enhanced measurements: beating the standard quantum limit. Science, 306, 1330-1336 (2004)
37. Giovannetti V., Lloyd S., Maccone L.: Quantum-enhanced “Quantum metrology”. Phys. Rev. Lett., 96, 010401 (2006)
38. Hajek J.: Local asymptotic minimax and admissibility in estimation. Proc. Sixth Berkeley Symp. on Math. Statist. and Prob., Vol. 1 (Univ. of Calif. Press, 1972), 175-194.
39. Choi M.-D.: Completely Positive Linear Maps on Complex Matrices. *Lin. Alg. Appl.*, 10, 285–290 (1975)

40. Jamiołkowski A.: *Rep. Math. Phys.*, 3, 275 (1972)

Communicated by name