Quantum channel estimation and asymptotic bound

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

Abstract. Quantum channel estimation in the one-parameter case is treated with the
asymptotic setting. This problem sometimes has square speedup. A sufficient condition for
this speedup is obtained. In this case, the existing results in Cramér-Rao approach seem to
contradict the existing results in covariant approach. In order to resolve this problem, we
introduce two kinds of asymptotic bounds.

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]. We also assume that use of entanglement
with reference system is available in the measurement process.

The main topic is the following. In the state estimation, when the number $n$ of prepared
states goes to infinity, the mean square error behaves as the order $O(\frac{1}{n})$ as in the estimation
of probability distribution. However, in the estimation of quantum channel, two different cases
were reported concerning asymptotic behavior of mean square error. As the first case, in the
estimations of depolarizing channels and Pauli channels, the optimal mean square error behaves
as $O(\frac{1}{n})$[1, 3]. As the second case, in the estimation of unitary, the optimal mean square error
behaves as $O(\frac{1}{n^2})$[4, 12, 6, 7, 5, 9, 8]. In order to clarify this point, we consider what case has
the order $O(\frac{1}{n^2})$ concerning the mean square error. As the first result, we derive a necessary
condition for the order $O(\frac{1}{n})$ by using Cramér-Rao approach for quantum state estimation.
While Fujiwara and Imai[2] derived this kind of condition, our condition is much simpler than
their condition. That is, it is easier to check our condition than to check their condition.

Next, we consider the meaning of Cramér-Rao approach when the order is $O(\frac{1}{n^2})$. In this
case, two different results concerning estimation of unitary were reported. One is based on the
Cramér-Rao approach[4, 12]. The other is based on the group covariant approach[6, 7, 5, 9, 8].
In particular, Imai and Hayashi [9] treated this problem in the case of phase estimation. They
seem to contradict with each other. In this paper, in order to resolve this problem, we introduce
another asymptotic bound for general channel estimation. Using this bound and the Cramér-Rao
bound, we consider the relation between these results and find a relation with superefficiency.
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 a superefficient estimator[23]. As the second result, in the estimation of phase action, we find that the bound by Cramér-Rao approach can be attained only by a quantum channel version of a superefficient estimator that works at specific points and the bound by group covariant approach can be attained by a suitable estimator that works at all points. Cramér-Rao approach is based on the asymptotically locally unbiased condition. So, we can conclude that the asymptotically locally unbiased condition is too weak for deriving a bound that can be attained in all points. Indeed, a similar phenomena happens in quantum state estimation when we use the large deviation criterion[15].

This paper is organized as follows. Some of obtained results are based on quantum state estimation with Cramér-Rao approach. Section 2 is devoted to a review of 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 bound based on Cramér-Rao approach 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. Section 4 is devoted to the global attainability of the Cramér-Rao bound in the channel estimation and 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 \} \}. 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 mean square error decreases in proportion to \( n^{-1} \) as in the estimation of probability distribution. So, we focus on the first order coefficient of the mean square error concerning \( n^{-1} \). In the most general setting, when a positive operator valued measure \( M^n \) on the total system \( \mathcal{H} \otimes^n \) takes values in the parameter space \( \Theta \subset \mathbb{R}^d \), it is allowed as an estimator.

For simplicity, in the following, we consider the case when the number \( d \) of parameters is one. So, the mean square error is given as

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

The optimal first order coefficient of this formulation is described by \( C_\theta \) and is called Quantum Cramér-Rao bound.

In the quantum case, there are several quantum extensions of Fisher information. 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 \( L_\theta^R \) and the SLD \( L_\theta^S \) as the operators satisfying

\[
\frac{d\rho_\theta}{d\theta} = \rho_\theta L_\theta^R, \quad \frac{d\rho_\theta}{d\theta} = \frac{1}{2} (L_\theta^S \rho_\theta + \rho_\theta L_\theta^S).
\]

Then, the RLD and SLD Fisher informations are given by [16, 17, 18]

\[
J_\theta^R := \text{Tr} \rho_\theta L_\theta^R (L_\theta^R)^\dagger, \quad J_\theta^S := \text{Tr} \rho_\theta (L_\theta^S)^2.
\]
When the range of $\rho_\theta$ contains the range of $\left(\frac{d\rho_\theta}{d\theta}\right)^2$, the RLD Fisher information has another expression:

$$J_\theta^R = \text{Tr}(\frac{d\rho_\theta}{d\theta})^2 \rho_\theta^{-1}.$$  
(1)

When the state family $\{\rho_\theta\}$ is given by $\rho_\theta := e^{\theta X}|u\rangle\langle u| e^{-\theta X}$, the condition (1) does not hold, where $X^\dagger = -X$. In this case, the SLD Fisher information is calculated as follows[14].

$$4(|\langle u| iX|u\rangle|^2) - (\langle u| iX|u\rangle)^2).$$  
(2)

Now, we introduce the unbiased condition by

$$\int \hat{\theta} \text{Tr} \rho_\theta^{\otimes n} M^n(d\hat{\theta}) = \theta, \ \forall \theta \in \Theta.$$ 

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 $\theta_0$:

$$\int \hat{\theta} \text{Tr} \rho_{\theta_0}^{\otimes n} M^n(d\hat{\theta}) = \theta_0, \ \frac{d}{d\theta} \int \hat{\theta} \text{Tr} \rho_{\theta}^{\otimes n} M^n(d\hat{\theta})|_{\theta=\theta_0} = 1.$$ 

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.

$$\text{MSE}_\theta(M^n) \geq \frac{1}{n} (J_\theta^R)^{-1}$$  
(3)

$$\text{MSE}_\theta(M^n) \geq \frac{1}{n} (J_\theta^S)^{-1}.$$  
(4)

Since $J_\theta^R$ is greater than $J_\theta^S$, the inequality (4) is more informative than the inequality (3). When the estimator $M^n$ is the spectral decomposition of the operator $\theta_0 I + \frac{1}{nJ_\theta} (L^{S(1)}_\theta + \cdots + L^{S(n)}_\theta)$, 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

$$J_\theta^R \geq J_\theta^S.$$ 

In fact, in the asymptotic setting, a suitable estimator usually satisfies the asymptotic locally unbiased condition:

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

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

$$\lim_{n \to \infty} n \text{MSE}_\theta(M^n) \geq (J_\theta^S)^{-1}.$$  
(5)

Further, by using the two-step method, the bound $(J_\theta^S)^{-1}$ can be universally attained for any true parameter $\theta$ [19, 20, 21]. So, defining the Cramér-Rao bound:

$$C_\theta := \inf_{\{M^n\}} \left\{ \lim_{n \to \infty} n \text{MSE}_\theta(M^n) \mid \{M^n\} \text{ satisfies the asymptotic locally unbiased condition. } \right\},$$

we obtain

$$C_\theta = (J_\theta^S)^{-1}.$$
3. Maximum SLD Fisher information in quantum channel estimation

In this section, we discuss applicability of 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 of \( \mathcal{H} \otimes \mathcal{R} \) as follows.

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

(6)

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

\[
B \otimes C |A\rangle = |BAC^T\rangle
\]

(7)

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 \( \rho[\Lambda_\theta] := (\Lambda_\theta \otimes \text{id})(|I\rangle\langle I|) \).

Then, when the input state is the maximally entangled state \( |\frac{1}{\sqrt{d}} I\rangle = \sum_{j=1}^{d} \frac{1}{\sqrt{d}} |j\rangle \otimes |j\rangle \), the output state is \( \frac{1}{d} \rho[\Lambda_\theta] = (\Lambda_\theta \otimes \text{id})(|\frac{1}{\sqrt{d}} I\rangle\langle I|) = |\frac{1}{\sqrt{d}} I\rangle\langle \frac{1}{\sqrt{d}} I| \). When the matrix \( \mathcal{A} A^T \) is a density matrix on \( \mathcal{H} \), \( |A\rangle\langle A| \) is a pure state on the product system \( \mathcal{H} \otimes \mathcal{R} \). Thus, the output state is given as

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

(8)

In the one-parameter case, we express the derivative \( \frac{d\rho(\Lambda_\theta)}{d\theta} \) by \( D[\Lambda_\theta] \).

When the input state is the product state \( |u\rangle\langle v| \otimes |u\rangle\langle u| \), the output state is \( (I \otimes u \cdot v^T)\rho[\Lambda_\theta](I \otimes \overline{v} \cdot u^\dagger) \). Since

\[
\text{Tr}_K(I \otimes u \cdot v^T)\rho[\Lambda_\theta](I \otimes \overline{v} \cdot u^\dagger) = |u\rangle\langle u|,
\]

(9)

we have

\[
(I \otimes u \cdot v^T)\rho[\Lambda_\theta](I \otimes \overline{v} \cdot u^\dagger) = (\text{Tr}_R(I \otimes u \cdot v^T)\rho[\Lambda_\theta](I \otimes \overline{v} \cdot u^\dagger)) \otimes |u\rangle\langle u|.
\]

(10)

This relation implies that the system \( \mathcal{R} \) has no information concerning \( \Lambda_\theta \). In this case, it is enough to treat the state \( \text{Tr}_R(I \otimes u \cdot v^T)\rho[\Lambda_\theta](I \otimes \overline{v} \cdot u^\dagger) \) for estimation of \( \Lambda_\theta \). Since this state is independent of \( |u| \), it is abbreviated by \( \langle \overline{v}|\rho[\Lambda_\theta]|\overline{v} \rangle \). Thus,

\[
\langle \overline{v}|(\text{Tr}_K \rho[\Lambda_\theta])(\overline{v}) = \text{Tr}_K(\overline{v}|\rho[\Lambda_\theta]|\overline{v}) 1,
\]

(11)

which implies that \( \langle \overline{v}|(\text{Tr}_K \rho[\Lambda_\theta])(\overline{v}) = 1 \). Thus, we obtain

\[
\text{Tr}_K \rho[\Lambda_\theta] = I.
\]

(12)

Taking the derivative in (12), we obtain

\[
\text{Tr}_K D[\Lambda_\theta] = 0.
\]

(13)
Now, we back to our estimation problem. In this problem, our choice is given by a pair of the input state $\rho$ and quantum measurement $M$. When we fix the input state, our estimation problem can be reduced to the state estimation with the state family $\{\Lambda_\theta(\rho) | \theta \in \Theta\}$. In the one-parameter case, we focus on the suprema

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

(14)

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

When $n$ applications of the unknown channel $\Lambda_\theta$ 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:

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

for all points $\theta_0$, and denotes the mean square error of $(\rho_n, M^n)$ by $\text{MSE}_n(\rho_n, M^n)$. Assume that $J^S[\Lambda_\theta^n]$ 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

$$\limsup_{n \to \infty} n^\alpha \text{MSE}_n(\rho_n, M^n) \geq \limsup_{n \to \infty} \frac{n^\alpha}{J^S[\Lambda_\theta^n]}.$$

We define the Cramér-Rao bound:

$$\tilde{C}_\alpha[\Lambda_\theta] := \inf_{\{(\rho_n, M^n)\}} \left\{ \limsup_{n \to \infty} n^\alpha \text{MSE}_n(\rho_n, M^n) \left| \{\rho_n, M^n\}\text{ satisfies the asymptotic locally unbiased condition.} \right. \right\}.$$

Thus, we obtain

$$\tilde{C}_\alpha[\Lambda_\theta] = \limsup_{n \to \infty} \frac{n^\alpha}{J^S[\Lambda_\theta^n]}.$$ (15)

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

(C) The range of $\rho[\Lambda_\theta]$ contains the range of $D[\Lambda_\theta]^2$.

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

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

$$J^R[\Lambda_\theta] = \| \text{Tr}_\mathcal{K} D[\Lambda_\theta] \rho[\Lambda_\theta]^{-1} D[\Lambda_\theta] \|.$$ (16)

**Proof:** Assume that the input state is given by $|A\rangle\langle A|$ and $A$ is an invertible matrix. Then, the range of $(I \otimes A^T)\rho[\Lambda_\theta](I \otimes A)$ contains the range of $((I \otimes A^T)D[\Lambda_\theta](I \otimes A))^{-1}$.

Using the formula (1), we obtain

$$J^R[\Lambda_\theta, |A\rangle\langle A|] = \text{Tr}((I \otimes A^T)D[\Lambda_\theta](I \otimes A))^{-1}(I \otimes A^T)\rho[\Lambda_\theta](I \otimes A)^{-1}$$

$$= \text{Tr}(I \otimes A^T)D[\Lambda_\theta](I \otimes A)(I \otimes A^T)D[\Lambda_\theta](I \otimes A)^{-1}\rho[\Lambda_\theta]^{-1}(I \otimes A)^{-1}$$

$$= \text{Tr}(I \otimes A^T)D[\Lambda_\theta]\rho[\Lambda_\theta]^{-1}D[\Lambda_\theta]$$

$$= \text{Tr} A^T \left( \text{Tr}_\mathcal{K} D[\Lambda_\theta] \rho[\Lambda_\theta]^{-1} D[\Lambda_\theta] \right),$$
where \((I \otimes A^T)[\rho|\Lambda\rangle|I \otimes A|)^{-1}\) is the inverse of \((I \otimes A^T)[\rho|\Lambda\rangle|I \otimes A|)\) on its range. So, the supremum of \(\text{Tr} A^T \rho|\Lambda\rangle|\Lambda\rangle \rho|\Lambda\rangle^{-1} \rho|\Lambda\rangle\) with the condition \(\text{rank } A = \dim \mathcal{H}\) equals \(\| \text{Tr} K \rho|\Lambda\rangle|\Lambda\rangle \rho|\Lambda\rangle^{-1} \rho|\Lambda\rangle\|\).

For a non-invertible matrix \(A\), the inequality \(J^R[\rho|\Lambda\rangle|I\rangle\langle A|\langle I| \leq \| \text{Tr} K \rho|\Lambda\rangle|\Lambda\rangle \rho|\Lambda\rangle^{-1} \rho|\Lambda\rangle\|\) can be shown by modifying the above discussion, which will be given in [25].

**Theorem 2** When the condition \((C)\) holds,

\[ J^R[\rho|\Lambda\rangle|I\rangle \langle I| = n J^R[\rho|\Lambda\rangle], \]  

(17)

Since \(n J^S[\rho|\Lambda\rangle] \leq J^S[\rho|\Lambda\rangle|I\rangle \langle I| \leq J^R[\rho|\Lambda\rangle|I\rangle \langle I|\), \(J^S[\rho|\Lambda\rangle|I\rangle \langle I|\) increases in order \(n\) under the assumption of Theorem 1, i.e., \(J^S[\rho|\Lambda\rangle|I\rangle \langle I| = O(n)\). When the rank of \(\rho|\Lambda\rangle\) is the maximum, i.e., \(\text{dd} \), this condition holds and \(J^S[\rho|\Lambda\rangle|I\rangle \langle I| = O(n)\). Even if the rank of \(\rho|\Lambda\rangle\) is not the maximum, we have an example satisfying the condition \((C)\) as follows.

A channel \(\Lambda\) 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

\[ \Lambda(\rho) = \sum_{k,l} d_{k,l} \rho_{k,l} |k\rangle \langle l|, \]

where \(\rho = \sum_{k,l} \rho_{k,l} |k\rangle \langle l|\). In this case, the state \(\rho|\Lambda\rangle\) is written as the following form

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

That is, its range is included by the space spanned by \(|k\rangle \langle k| \rangle \rangle\). When a channel family \(\{\rho|\Lambda\rangle\}\) consists of phase damping channels and the range of \(\rho|\Lambda\rangle\) is the space spanned by \(|k\rangle \langle k| \rangle \rangle\), the condition \((C)\) holds.

**Proof:** For any matrix \(X\), we define matrixes:

\[ [X]_{k} := \rho|\Lambda\rangle \otimes \cdots \otimes \rho|\Lambda\rangle \otimes X \otimes \rho|\Lambda\rangle \otimes \cdots \otimes \rho|\Lambda\rangle \]

(18)

\[ (X)_{k} := I \otimes \cdots \otimes I \otimes X \otimes I \otimes \cdots \otimes I \]

(19)

\[ [X]_{k,j} := \rho|\Lambda\rangle \otimes \cdots \otimes \rho|\Lambda\rangle \otimes X \otimes \rho|\Lambda\rangle \otimes \cdots \otimes \rho|\Lambda\rangle \otimes X \otimes \rho|\Lambda\rangle \otimes \cdots \otimes \rho|\Lambda\rangle \]

(20)

Here, (20) is the definition in the case of \(j > k\). In the case of \(j < k\), \([X]_{j,k}\) is defined to be \([X]_{j,k}\). Then,

\[ D[\rho|\Lambda\rangle] = \sum_{k=1}^{n} [D[\rho|\Lambda\rangle]_{k} \]

(21)

\[ \| \sum_{k=1}^{n} (X)_{k} \| = n \|X\|. \]

(22)

Using (12), we have

\[ \text{Tr} (X)_{k} = (\text{Tr} X)_{k}. \]

(23)
For $k \neq j$, using (12) and (13), we have

$$\text{Tr}_{K^\otimes n}[D]\{[\Lambda_\theta]\}_{k}(\rho[\Lambda_\theta])^{-1}_k[D][\Lambda_\theta] = \text{Tr}_{K^\otimes n}[D][\Lambda_\theta]_{k,j} = 0. \quad (24)$$

For any $k$, using (23), we have

$$\text{Tr}_{K^\otimes n}[D]\{[\Lambda_\theta]\}_{k}(\rho[\Lambda_\theta])^{-1}_k[D][\Lambda_\theta] = \text{Tr}_{K^\otimes n}[D][\Lambda_\theta][\rho[\Lambda_\theta]^{-1}D][\Lambda_\theta]_{k} = \{\text{Tr}_{K} D}[\Lambda_\theta][\rho[\Lambda_\theta]^{-1}D][\Lambda_\theta]\}_{k} \quad (25)$$

Thus, combining (21), (22), (24), and (25), we obtain

$$J^R[\Lambda_\theta^\otimes n] = \| \text{Tr}_{K^\otimes n}\sum_{k,j} [D][\Lambda_\theta]_{k}(\rho[\Lambda_\theta])^{-1}_k[D][\Lambda_\theta]_{j} \|$$

$$= \| \text{Tr}_{K^\otimes n}\sum_{k=1}^{n} [D][\Lambda_\theta]_{k}(\rho[\Lambda_\theta])^{-1}_k[D][\Lambda_\theta]_{k} + \sum_{k \neq j} [D][\Lambda_\theta]_{k,j} \|$$

$$= \| \text{Tr}_{K^\otimes n}\sum_{k=1}^{n} (\text{Tr}_{K} D)[\Lambda_\theta][\rho[\Lambda_\theta]^{-1}D][\Lambda_\theta]_{k} \| = n\| \text{Tr}_{K} D[\Lambda_\theta][\rho[\Lambda_\theta]^{-1}D][\Lambda_\theta]\|.$$

Next, 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 $\Lambda_\theta(\rho) = e^{\theta X} \rho e^{-\theta X}$ and $X^\dagger = -X$. Using (2), we obtain

$$J^S[\Lambda_\theta] = (\lambda_{\max}(iX) - \lambda_{\min}(iX))^2, \quad (26)$$

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

$$J^S[\Lambda_\theta^\otimes n] = n^2(\lambda_{\max}(iX) - \lambda_{\min}(iX))^2. \quad (27)$$

4. Global attainability in channel estimation

The discussion in the previous section treats the bound only with the locally unbiased condition. In order to check the operational meaning, we have to consider the global attainability for this bound as in the case of state estimation. Here, remember that this bound can be attained by using the two-step method in the case of state estimation. For this purpose, we introduce another quantity. In the following definition, we focus on an $\epsilon$-neighborhood $U_{\theta,\epsilon}$ of $\theta$. Define

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

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

Concerning this quantity, we have the following two propositions, whose proofs will be given in [25].

**Proposition 3** When $\tilde{C}_\alpha[\Lambda_{\theta_0}]$ is continuous,

$$C_\alpha[\Lambda_{\theta_0}] \geq \tilde{C}_\alpha[\Lambda_{\theta_0}]. \quad (28)$$

**Proposition 4** Assume that $R := \sup_{\theta \in \Theta} |\theta| < \infty$. When the order parameter $\alpha$ equals 1 and $\tilde{C}_\alpha[\Lambda_{\theta_0}]$ is continuous,

$$\tilde{C}_1[\Lambda_{\theta_0}] = \tilde{C}_1[\Lambda_{\theta_0}].$$
By using the two-step method[22], the bound $C_{\alpha}[\Lambda_{\theta}]$ can be attained at all points $\theta$ as follows.

**Proposition 5** Assume that $R := \sup_{\theta \in \Theta} |\theta| < \infty$ and $C_1[\Lambda_{\theta}]$ is continuous. For any $\delta > 0$, there exists a sequence of estimators $\{(\rho_n, M^n)\}$ such that

$$C_{\alpha}[\Lambda_{\theta}], \{(\rho_n, M^n)\} \leq C_{\alpha}[\Lambda_{\theta}] + \delta$$  \hspace{1cm} \text{(29)}

for all points $\theta$.

A proof of this proposition will be given in [25].

As a typical example, we consider the case of $\Lambda_{\theta}(\rho) = e^{i\theta X} \rho e^{-i\theta X}$ with $X = \begin{pmatrix} i/2 & 0 \\ 0 & -i/2 \end{pmatrix}$.

In this case, we replace mean square error by $\min_k (\hat{\theta} + 2\pi k - \theta)^2$. The minimum average error behaves as $\pi^2$ [8, 9]. 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. Since the bound $C_2[\Lambda_{\theta}]$ can be globally attained, $C_2[\Lambda_{\theta}] = \pi^2$. Since $J^2[\Lambda_{\theta}^n] = n^2$, $\tilde{C}_2[\Lambda_{\theta}] = 1$. Hence, the bound $\tilde{C}_2[\Lambda_{\theta}]$ can be attained only in a specific point as follows.

**Proposition 6** Assume that $R := \sup_{\theta \in \Theta} |\theta| < \infty$ and $C_1[\Lambda_{\theta}]$ is continuous. For any $\delta > 0$ and any $\theta_0 \in \Theta$, there exists a sequence of estimators $\{(\rho_n, \theta_0), M_{\theta_0}^n\}$ satisfying the asymptotically locally unbiased condition and the relations:

$$\limsup_{n \to \infty} n^0 \text{MSE}_{\theta_0}(\rho_n, \theta_0), M_{\theta_0}^n \leq C_{\alpha}[\Lambda_{\theta}] + \delta, \quad \forall \theta \neq \theta_0$$  \hspace{1cm} \text{(30)}

$$\limsup_{n \to \infty} n^0 \text{MSE}_{\theta_0}(\rho_n, \theta_0), M_{\theta_0}^n \leq \tilde{C}_{\alpha}[\Lambda_{\theta_0}] + \delta.$$  \hspace{1cm} \text{(31)}

A proof of this proposition will be given in [25].

In estimation of probability distribution, there exists a superefficient estimator that has smaller error at a discrete set than Cramér-Rao bound[23]. Since such a superefficient estimator cannot be useful, it is thought to be better to choose a condition for our estimators in order to remove superefficient estimators in statistics. In this classical case, if we assume the asymptotic locally unbiased condition, we have no superefficient estimator. Proposition 6 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 $\{(\rho_n, M^n)\}$ is called $q$-channel-superefficient at $\theta$ with the order $1/n^2$ when $\lim_{n \to \infty} n^0 \text{MSE}_{\theta}(\rho_n, M^n) < C_{\alpha}[\Lambda_{\theta}]$.

Hence, in order to remove the q-channel-superefficiency problem, it is better to adopt the bound $C_{\alpha}[\Lambda_{\theta}]$ as the criterion instead of $\tilde{C}_{\alpha}[\Lambda_{\theta}]$.

Finally, we consider the relation with the adaptive method proposed by Nagaoka[13]. In this method, we apply our POVM to each single system $\mathcal{H}$, and we decide $k$-th POVM based on the knowledge of previous $k - 1$ outcomes. In this case, Fujiwara [24] analyzed the asymptotic behaviour of the MSE of this estimator. Now, we consider the case of $nn$ applications of the unknown channel $\Lambda_{\theta}$. In this case, we divide $nn$ applications into $n$ groups consisting of $m$ applications. When we apply the adaptive method mentioned in Fujiwara[24] to these groups, the MSE of this estimator behaves as $\frac{1}{nJ^2[\Lambda_{\theta}^n]}$, which is close to $\frac{C_{\alpha}[\Lambda_{\theta}]}{nm^2}$. So, when $\alpha > 1$, this method cannot realize the optimal order $O(\frac{1}{\alpha nm^2})$.

5. Discussion

We have treated quantum channel estimation as a general framework containing quantum state estimation. We have introduced two kinds of asymptotic bound in quantum channel estimation.
One is given from the limit of maximum of SLD Fisher information $J^S[\Lambda^{\otimes n}_\theta]$. The other is given by more operational quantities. The relation between both bounds is closely related to the asymptotic behavior of $J^S[\Lambda^{\otimes n}_\theta]$. We have shown that both bounds coincide when $J^S[\Lambda^{\otimes n}_\theta]$ behaves as $O(n)$. The case of state estimation can be regarded as a special case of the above case. That is, the conventional state estimation has no difference between both quantities. However, in the case of unitary estimation, $J^S[\Lambda^{\otimes n}_\theta]$ behaves as $O(n^2)$, and there exists a difference between both bounds.

So, the asymptotic behavior of $J^S[\Lambda^{\otimes n}_\theta]$ is quite important. In order to clarify this behavior, we have derived the condition (C) as a sufficient condition for $J^S[\Lambda^{\otimes n}_\theta] = O(n)$. Indeed, Fujiwara and Imai [2] also obtained another sufficient condition. Since the relation with their condition is not clear, it is an open problem. Our condition (C) trivially contains the case when the state $\rho[\Lambda_\theta]$ is a full rank state on the tensor product system while it is not so easy to derive the above full rank condition from the above and Imai’s condition. Further, we have also obtained another example for $J^S[\Lambda^{\otimes n}_\theta] = O(n^2)[25]$ under the condition (C). This example is a larger class than the unitary model. So, we can expect that $J^S[\Lambda^{\otimes n}_\theta]$ 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 Center 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.

References

[1] Fujiwara A 2001 Phys. Rev. A 63 042304
[2] Fujiwara A and Imai H 2008 J. Phys. A: Math. Theor. 41 255304
[3] Fujiwara A and Imai H 2003 J. Phys. A: Math. Gen. 36 8093–103
[4] Fujiwara A 2002 Phys. Rev. A 65 012316
[5] Hayashi M 2006 Phys. Lett. A 354 183–9
[6] Chiribella G, D’Ariano G M, Perinotti P and Sacchi M F 2004 Phys. Rev. Lett. 93 180503
[7] Bagan E, Baig M and Mimoz-Tapia R 2004 Phys. Rev. A 70 030301
[8] Buzek V, Derka R and Massar S 1999 Phys. Rev. Lett. 82 2267–10
[9] Imai H and Hayashi M 2009 Fourier Analytic Approach to Phase Estimation in Quantum Systems J. Phys. A 41 043054
[10] Kahn J 2007 Phys. Rev. A 75 022326
[11] Imai H and Fujiwara F 2007 J. Phys. A: Math. Theor. 40 4391–400
[12] Hayashi M 2005 Phys. Lett. A 354 183–9
[13] Hotta M, Karasawa T and Ozawa M 2006 J. Phys. A: Math. Gen. 39 14465–70
[14] Nagaoka H 1989 Proc. 12th Symp. on Inform. Theory and its Appl. p 577; Nagaoka H 2005 Fourier Analytic Approach to Phase Estimation in Quantum Systems J. Phys. A 35 7689–727
[15] Holevo A S 1982 Probabilistic and Statistical Aspects of Quantum Theory (Amsterdam: North-Holland); (Originally published in Russian 1980)
[16] Holevo A S 2007 Quantum Information: An Introduction (Berlin: Springer)
[17] Holevo A S 1999 Phys. Lett. A 201 119–24
[18] Hayashi M 2002 J. Phys. A: Math. Gen. 35 7689–727
[19] Holevo A S 1982 Probabilistic and Statistical Aspects of Quantum Theory (Amsterdam: North-Holland); (Originally published in Russian 1980)
[20] Holevo A S 2006 Quantum Information: An Introduction (Berlin: Springer)
[21] Holevo A S 2009 Selected Papers on Probability and Statistics (American Mathematical Society Translations Series 2 Vol. 277) ed American Mathematical Society Staff (American Mathematical Society) p 95 (Preprint quant-ph/0608198)
[23] LeCam L 1986 Asymptotic Methods in Statistical Decision Theory (New York: Springer)
[24] Fujiwara A 2006 J. Phys. A: Math. Gen. 39 12489–504
[25] Hayashi M 2010 Two kinds of asymptotic bounds in quantum channel estimation Preprint arXiv:1003.4575v1 [quant-ph]