DISTURBANCE OF OPERATION IN QUANTUM ESTIMATION FOR THE GAUSSIAN P-FUNCTION

Yoshiyuki Tsuda*, Keiji Matsumoto* and Masahito Hayashi**

For the quantum Gaussian state family, Hayashi proposed a quantum mechanical operation using beam splitters to estimate the location and scale parameters of the P-function, and he showed that it is asymptotically optimal. In this paper, we analyze the effect of disturbance of his operation caused by the randomness of the transparency of the beam splitters. It is shown that even if the variance of the random transparency is small, Hayashi’s estimators are improper in a sense that they are biased and asymptotically inconsistent. In such a case, we propose to stop the operation and correct the biases of estimators.

Key words and phrases: coherent state, counting measurement, quantum Gaussian state, heterodyne measurement.

1. Introduction

The quantum estimation is an application of the statistical estimation theory to the quantum mechanics. The general theory on the quantum mechanics is described by the linear algebra on a Hilbert space, and it might require some specific setup in order to consider the quantum estimation. However, our present problem on the Gaussian P-function is a special case in which only the basic analytical methods are used. This paper is unfolding our problem without linear algebraic premises.

The quantum estimation theory was first considered by Helstrom [2]. The main problem is to find an optimal scheme to estimate unknown parameter of the quantum state. Here, the scheme of estimation is composed of (i) physical operation and measurement in order to obtain data and (ii) the estimator based on the obtained data. Let $\rho$ denote a state of the system to be observed. Let $M$ denote a measurement on the system. If one carries out a measurement $M$, then he will obtain data $x$ as an observation of a random variable $X$ whose probability distribution is determined by $\rho$ and $M$. Suppose that we know the true state is an element of the set $S := \{\rho_{\theta}\}$ parameterized by $\theta \in \Theta$, and also suppose that we can select a measurement in the set $M = \{M_{\lambda}\}$ parameterized by $\lambda \in \Lambda$. Based on the obtained data $x$, one will estimate $\theta \in \Theta$ by an estimator $\hat{\theta}$.

In the standard statistical estimation problem, a statistician is allowed to select only the estimator $\hat{\theta}$. One features of the quantum estimation is that one could improve the accuracy of the estimation by interaction between samples. In the present paper, we consider the effect of disturbance of physical operation in a problem of quantum estimation for the Gaussian P-function model.

The models of Gaussian P-function are typical and important examples of the quantum estimation and it can be applied to the optical experiments. Our model has two-dimensional location parameters and one scale parameter. For the estimation problem
of the location parameters of the Gaussian $P$-function, Yuen and Lax [5] and Holevo [4] have given a lower bound of the variance of unbiased estimators and an optimal estimation scheme that attains the bound. For the problem of the location and scale parameters, Hayashi [1] has given a lower bound of the variance of unbiased estimators and an optimal estimation scheme that asymptotically attains the bound. His result gave the first practical example of the advantage of interaction of samples to the quantum estimation of the Gaussian model concerning the statistical estimation theory. We consider the model of the location and scale parameters and analyze the effect of disturbance of the asymptotically optimal scheme given by Hayashi [1]. He pointed out that his scheme is realized in the optical experiment by the physical operation using beam splitters. For example, if we have given $n = 2^m$ samples, then an experimenter uses beam splitters whose transparency is $1/2$. However, in practical, the transparency of prepared beam splitters is not always exactly the ideal quantity $(1/2)$, and it is often slightly different from the ideal one. If the transparency is not correct, then the physical operation proposed by Hayashi [1] is disturbed and it might affect the estimation. We analyze the effect of the disturbance to the estimation, and it turns out that the estimator is not even asymptotically consistent to the true parameter. In such a case, a naive scheme not using interactions between samples is better than that of Hayashi [1] since the disturbance could be avoided. In order to avoid the inconsistency of Hayashi [1]'s estimator, we propose a new scheme to estimate location and scale parameters by correcting that of Hayashi [1]. We will see that our scheme gives an asymptotically consistent estimator, and it is better than the naive estimator.

2. Problem

When we discuss a general problem of quantum estimation, we need to consider any measurement $M$ which may be carried out to the quantum system. However, our interest in this paper is devoted to the problem of disturbance in the experiment proposed by Hayashi [1]. His optimal method uses only two kinds of measurements called the 'heterodyne measurement' and the 'counting measurement' and the state is denoted by a probabilistic superposition of the coherent states. It enables us to describe the problem by a Bayesian model in which the prior distribution of parameters of basic states means the state, and selecting some of random variables corresponds to the selecting measurements. This is why the linear algebraic preparation is not needed.

Let $(A_j, B_j)$ denote the complex amplitude $A_j + iB_j$ ($i = \sqrt{-1}$) of a coherent state, $(X_j, Y_j)$ denote the observed amplitude $X_j + iY_j$ by the heterodyne measurement and $Z_j$ denote the counted numbers by the counting measurement. Suppose that $W_j := (A_j, B_j, X_j, Y_j, Z_j)$ ($j = 1, ..., n$) are random variables. Suppose that we can always not observe $A_j$ or $B_j$ for any $j$, and we can selectively observe $(X_j, Y_j)$ or $Z_j$ for each $j$, so we do not observe $Z_j$ if we select to observe $(X_j, Y_j)$ and we do not observe $(X_j, Y_j)$ if we select $Z_j$. Let $S \subseteq \{1, ..., n\}$ be a set of indices of which we observe $(X_j, Y_j)$, hence we observe $Z_k$ if $k \notin S$. Let $(x_j, y_j)$ ($j \in S$) and $z_k$ ($k \notin S$) be observations of $(X_j, Y_j)$ and $Z_j$. For each $j \in S$, we assume that $X_j$ and $Y_j$ are continuous random variables on the set $R$ of real numbers, and, for each $k \notin S$ $Z_k$ is a discrete random variable on the set $N_0$ of non-negative integers. We also assume that the joint probability (density) function of $(X_j, Y_j)$ ($j \in S$) and $Z_k$ ($k \notin S$) is denoted as

$$E \left[ \prod_{j \in S} \left( \frac{\exp(-(x_j - A_j)^2 - (y_j - B_j)^2)}{\pi} \right) \prod_{k \notin S} \left( \frac{\exp(-A_k^2 - B_k^2)(A_k^2 + B_k^2)^z_k}{z_k!} \right) \right]$$

where the expectation is taken for $A_j$ and $B_j$ ($j = 1, ..., n$).
**REMARK 1.** Suppose that \( A_1, ..., A_n, B_1, ..., B_n \) are independently distributed. For \( j \in S \), \( A_j \) and \( B_j \) are distributed according to the normal distributions \( N(\theta, \nu/2) \) and \( N(\eta, \nu/2) \) of means \( \theta \) and \( \eta \) and variance \( \nu/2 \), respectively. For \( k \notin S \), \( A_k \) and \( B_k \) are distributed according to \( N(0, \nu/2) \). Then, all selected variables \((X_j, Y_j) (j \in S)\) and \( Z_k (d \notin S)\) are independently distributed. The distributions of \( X_j \) and \( Y_j \) are \( N(\theta, (\nu+1)/2) \) and \( N(\eta, (\nu+1)/2) \), respectively, and that of \( Z_k \) is the geometric distribution \( Geo(\nu) \) of mean \( \nu \), that has the probability function

\[
\frac{1}{\nu + 1} \left( \frac{\nu}{\nu + 1} \right)^z_k.
\]

Not only we can select observed variables, but we can also select pairs of variables \((A_j, B_j)\) and \((A_k, B_k)\) and transform them as

\[
\begin{align*}
A_j &\mapsto A_j \cos \tau + A_k \sin \tau, \\
B_j &\mapsto B_j \cos \tau + B_k \sin \tau, \\
A_k &\mapsto -A_k \sin \tau + A_k \cos \tau, \\
B_k &\mapsto -B_k \sin \tau + B_k \cos \tau,
\end{align*}
\]

where \( \tau \in \mathbb{R} \) is arbitrarily selected. We call this transformation \( g^{(j,k)} \). Note that the transformations has the group structure in the algebraic sense, and we write \( g^{(j_2,k_2)}_2 g^{(j_1,k_1)}_1 \) to denote that we first carry out \( g^{(j_1,k_1)}_1 \) and then next do \( g^{(j_2,k_2)}_2 \).

**REMARK 2.** Suppose that \( A_j, B_j, A_k, B_k \) are independently and identically distributed. Let \( A_j \) and \( A_k \) are distributed according to \( N(\theta, \nu/2) \) and \( B_j \) and \( B_k \) are distributed according to \( N(\eta, \nu/2) \). If we carry out the transformation \( g^{(j,k)}_\tau \), then the distributions of \( A_j, B_j, A_k, B_k \) are given by

\[
\begin{align*}
N(\theta(\cos \tau + \sin \tau), \nu/2), & \quad N(\eta(\cos \tau + \sin \tau), \nu/2), \\
N(\theta(-\sin \tau + \cos \tau), \nu/2), & \quad N(\eta(-\sin \tau + \cos \tau), \nu/2),
\end{align*}
\]

respectively.

Suppose that \( A^1, ..., A^n, B_1, ..., B_n \) are independently distributed, and \( A_j \) \((j = 1, ..., n)\) is distributed according to \( N(\theta, \nu/2) \) and \( B_j \) \((j = 1, ..., n)\) is distributed according to \( N(\eta, \nu/2) \), where \( \theta, \eta \in \mathbb{R} \) and \( \nu > 0 \) are unknown. Under the rules of selecting observed variables and transformations, we consider to estimate \( \theta, \eta \) and \( \nu \) based on observed variables.

### 3. Naive estimation without transformation

In this section, we consider one of the most naive method to estimate unknown parameters, that is, we do not carry out any transformations and we let \( S := \{1, ..., n\} \). Then, from Remark \( \Box \) \( X_1, ..., X_n, Y_1, ..., Y_n \) are independently distributed, and \( X_j \) is distributed according to \( N(\theta, (\nu+1)/2) \) and \( Y_j \) is distributed according to \( N(\eta, (\nu+1)/2) \) for \( j = 1, ..., n \). Let

\[
\bar{\theta} := \frac{1}{n} \sum_{j=1}^{n} X_j, \quad \bar{\eta} := \frac{1}{n} \sum_{j=1}^{n} Y_j,
\]

\[
\bar{\nu} := \frac{1}{n-1} \sum_{j=1}^{n} (X_j - \bar{\theta})^2 + \frac{1}{n-1} \sum_{j=1}^{n} (Y_j - \bar{\eta})^2 - 1.
\]
be estimators for $\theta$, $\eta$ and $\nu$. Then, we have
\[
E(\bar{\theta}) = \theta, \quad E(\bar{\eta}) = \eta, \quad E(\bar{\nu}) = \nu,
\]
so that they are unbiased. Moreover, since the covariance matrix of $(\bar{\theta}, \bar{\eta}, \bar{\nu})$ is
\[
V_1 := \begin{pmatrix}
\frac{\nu+1}{2n} & 0 & 0 \\
0 & \frac{\nu+1}{2n} & 0 \\
0 & 0 & \frac{(\nu+1)^2}{n-1}
\end{pmatrix},
\]
we can see that they are asymptotically consistent.

4. Hayashi’s estimation using transformation

In this section, we consider Hayashi’s method to estimate $\theta$, $\eta$ and $\nu$. Note first that, by some transformations of $A_1, ..., A_n, B_1, ..., B_n$ by $G := g_{\tau}^{(j_1,k_1), ..., g_{\tau}^{(j_n,k_n)}$, we obtain random variables $A'_1, ..., A'_n, B'_1, ..., B'_n$, which are mutually independent and $A_1, B_1, A_j, B_j (j \geq 2)$ are distributed according to
\[
N(\sqrt{n}\theta, \nu/2), \quad N(\sqrt{n}\eta, \nu/2), \quad N(0, \nu/2), \quad N(0, \nu/2),
\]
respectively.

**REMARK 3.** For $t = 1, ..., n-1$, let
\[
\tau_t := \tan^{-1} t^{-1/2}.
\]
Then, a transformation defined by
\[
G_1 := g_{\tau_{n-1}}^{(1,n)} \cdots g_{\tau_1}^{(1,2)}
\]
generates $A'_1, ..., A'_n, B'_1, ..., B'_n$ which are independently distributed according to (4.1).

If we draw the transformation $g_{\tau}^{(j,k)}$ as
\[
(A_j, B_j) \quad \tau \quad (A_j \cos \tau + A_k \sin \tau, B_j \cos \tau + B_k \sin \tau)
\]
\[
(A_k, B_k) \quad (A_J \sin \tau + A_k \cos \tau, -B_j \sin \tau + B_k \cos \tau)
\]
then, $G_1$ can be expressed as
\[
\begin{align*}
(A_1, B_1) & \\
(A_2, B_2) & \quad \tau \quad 2^{-1/2} \\
(A_3, B_3) & \quad \tan^{-1} 3^{-1/2} \\
(A_4, B_4) & \quad \tan^{-1} 2^{-1/2}
\end{align*}
\]
for $n = 4$.

The transformation generating $A'_1, ..., A'_n, B'_1, ..., B'_n$ which are which are distributed according to (4.1) is not unique as the follows.

**REMARK 4.** Suppose that $n = 2^m$ for a natural number $m$. For $t = 1, ..., m$, let $J_t := \{(j,k)\}$ be the set of pairs of natural numbers $j$ and $k$ satisfying
\[
j \equiv 1 \pmod{2^t}, \quad k - j = 2^{t-1}.
\]
Then, a transformation defined by
\[ G_2 := \prod_{(j,k) \in J_m} g_{\pi/4}^{(j,k)} \cdots \prod_{(j,k) \in J_t} g_{\pi/4}^{(j,k)} \cdots \prod_{(j,k) \in J_1} g_{\pi/4}^{(j,k)} \]
also generates \( A'_1, \ldots, A'_n, B'_1, \ldots, B'_n \) which are independently distributed according to (4.1).

In the same way as (4.2), \( G_2 \) can be expressed as
\[
\begin{align*}
(A_1, B_1) & \quad \pi/4 \quad (A'_1, B'_1) \\
(A_2, B_2) & \quad \pi/4 \quad (A'_2, B'_2) \\
(A_3, B_3) & \quad \pi/4 \quad (A'_3, B'_3) \\
(A_4, B_4) & \quad (A'_4, B'_4)
\end{align*}
\]
for \( n = 2^2 = 4 \).

After such transformations obtaining \( A'_1, \ldots, A'_n, B'_1, \ldots, B'_n \), we let \( S = \{1\} \) to select observed variables \((X_1, Y_1), Z_2, \ldots, Z_n\). Let
\[
\hat{\theta} := X_1/\sqrt{n}, \quad \hat{\eta} := Y_1/\sqrt{n}, \quad \hat{\nu} := \frac{1}{n-1} \sum_{j=2}^{n} Z_j,
\]
then from Remark 1 we have
\[
E(\hat{\theta}) = \theta, \quad E(\hat{\eta}) = \eta, \quad E(\hat{\nu}) = \nu,
\]
so that they are unbiased. Moreover, since the covariance matrix of \((\hat{\theta}, \hat{\eta}, \hat{\nu})\) is
\[
V_2 := \begin{pmatrix}
\frac{\nu+1}{2n} & 0 & 0 \\
0 & \frac{\nu+1}{2n} & 0 \\
0 & 0 & \nu \frac{\nu+1}{n+1}
\end{pmatrix},
\]
we can see that they are asymptotically consistent and this Hayashi’s estimators \( \hat{\theta}, \hat{\eta}, \hat{\nu} \) dominates the naive estimators \( \bar{\theta}, \bar{\eta}, \bar{\nu} \) in a sense that the difference \( V_1 - V_2 \) of the covariance matrices is non-negative definite.

Hayashi [1] also proved that his estimators are asymptotically optimal in the quantum mechanical setup which is more general than that of here.

5. Noisy transformation

The transformations \( G \) of \( A_1, \ldots, A_n, B_1, \ldots, B_n \) to \( A'_1, \ldots, A'_n, B'_1, \ldots, B'_n \) is closely related to physical operation of in quantum optics. Actually, each \( g_{\pi/4}^{(j,k)} \) corresponds to the interference of two light beams using a beam splitter of transparency \( \cos^2 \tau \). For example, if an experimenter tries to realize the transformation \( G_2 \), then he has to prepare beam splitters of transparency \( 1/2 \), since \( G_2 \) is constructed by \( g_{\pi/4}^{(j,k)} \)’s.

However, in practical cases, it is difficult to prepare beam splitters of exactly the same transparency as the ideal quantity \((1/2)\), and in many cases, each quantity is slightly different. Hence, we consider, when \( n = 2^m \) and we try to carry out transformation \( G_2 \), but each \( \tau \) is independently and identically distributed according to \( N(\pi/4, \epsilon \log 2) \), and we calculate expectations and variances of Hayashi’s estimators \( \hat{\theta}, \hat{\eta}, \hat{\nu} \) in order to see the effect of the randomness of transformations.
Let $A':=(A_1',...,A_n',B_1',...,B_n')$. Then, note that Remark shows that

$$E(\hat{\theta}) = \frac{1}{\sqrt{n}} E(E(X_1 \mid A)) = E(A'_1),$$

$$E(\hat{\theta}_2) = \frac{1}{\sqrt{n}} E(E(X_2 \mid A)) = E(B'_1),$$

$$E(\hat{\nu}) = \frac{1}{n-1} \sum_{j=2}^{n} E(E(Z_j \mid A))$$

$$= \frac{1}{n-1} \sum_{j=2}^{n} E(A_{j}^{2} + B_{j}^{2}),$$

$$V(\hat{\theta}) = \frac{1}{n} E(V(X_1 \mid A))$$

$$= \frac{1}{n} E(A_1^2 + 1/2) - E(\hat{\theta})^2,$$

$$V(\hat{\eta}) = \frac{1}{n} E(V(Y_1 \mid A))$$

$$= \frac{1}{n} E(B_1^2 + 1/2) - E(\hat{\eta})^2,$$

$$V(\hat{\nu}) = \frac{1}{(n-1)^2} E \left( V \left( \sum_{j=2}^{n} Z_j \bigg| A \right) \right)$$

$$= \frac{1}{(n-1)^2} \sum_{j=2}^{n} (E((A'_j + B'_j)^4 + (A'_j + B'_j)^2) - E(A'_j^2 + B'_j^2)E((A'_j + B'_j)^2))$$

$$+ \frac{2}{(n-1)^2} \sum_{2 \leq j < k \leq n} (E((A'_j + B'_j)(A'_k + B'_k)^2))$$

$$- \frac{2}{(n-1)^2} \sum_{2 \leq j < k \leq n} (E((A'_j + B'_j)^2)E((A'_k + B'_k)^2)),$$

hold, so that these quantities depend only on bivariate, at most, fourth moments of $A$, and we can obtain the following results.

**THEOREM 1.** For any $\epsilon > 0$, we have

(5.1) \quad E(\hat{\theta}) = n^{-\epsilon/2} \theta, \quad E(\hat{\eta}) = n^{-\epsilon/2} \eta,

(5.2) \quad V(\hat{\theta}) = \frac{1 + \nu}{2n} + \frac{\theta^2}{n} - \frac{\theta^2}{n^\epsilon} - \frac{n^{-1} - n^{-\epsilon}}{2^{1+\epsilon} - 2\epsilon \theta^2},

(5.3) \quad V(\hat{\eta}) = \frac{1 + \nu}{2n} + \frac{\eta^2}{n} - \frac{\eta^2}{n^\epsilon} - \frac{n^{-1} - n^{-\epsilon}}{2^{1+\epsilon} - 2\epsilon \eta^2},

(5.4) \quad E(\hat{\nu}) = \nu + \frac{-1 + 2^{1+\epsilon} - 2^{1+\epsilon} \epsilon \eta^{4} + 2^{2^{1+\epsilon} - 2\epsilon} \eta^{2} + n^{1-\epsilon}}{2^{\epsilon}(2^{2^\epsilon} - 1)(1 + n)} (\theta^2 + \eta^2).

If $\epsilon$ is positive and small enough, then we have

(5.5) \quad V(\hat{\nu}) = O(n^{-2\epsilon})

as $n$ goes to $\infty$. 
PROOF. First, for \( j = 1, \ldots, n (= 2^m) \) and \( t = 0, 1, \ldots, m \), let \( A_j^{(t)} \) and \( B_j^{(t)} \) denote \( A_j^{(t)} \) and \( B_j^{(t)} \) after the transformation \( \prod_{(j,k) \in J_t} g_{j(k)}^{\pi/4} \cdots \prod_{(j,k) \in J_t} g_{j(k)}^{\pi/4} \) which is on the way of \( G_2 \). Then, for any natural numbers \( r_A, r_B, s_A, s_B, t \) and \( j = 2^{t-1} + 1 \), the recurrent equations

\[
E(A_1^{(t)} r_A B_1^{(t)} r_B) = E((A_1^{(t-1)} \cos \tau + A_j^{(t-1)} \sin \tau)^r_A (B_1^{(t-1)} \cos \tau + B_j^{(t-1)} \sin \tau)^r_B),
\]

\[
E(A_1^{(t-1)} r_A A_1^{(t-1)} r_B) = E(A_j^{(t-1)} r_A B_j^{(t-1)} r_B),
\]

\[
E(A_1^{(t-1)} r_A A_1^{(t-1)} r_B) = E(A_j^{(t-1)} r_A B_j^{(t-1)} r_B)
\]

hold. Hence, we have (5.1) to (5.3). Next, for any \( r_A, r_B, t, j = 2^{t-1} + 1 \), and for any natural number \( k \) satisfying

\[
j' - j \equiv 0 \pmod{2^t},
\]

the recurrent equations

\[
E(A_j^{(t)} r_A B_j^{(t)} r_B) = E((-A_1^{(t-1)} \sin \tau + A_j^{(t-1)} \cos \tau)^r_A (-B_1^{(t-1)} \sin \tau + B_j^{(t-1)} \cos \tau)^r_B),
\]

\[
E(A_j^{(t)} r_A B_j^{(t)} r_B) = E(A_j^{(t-1)} r_A B_j^{(t-1)} r_B)
\]

hold. Hence, we have (5.4). Finally, for any \( r_A, r_B, s_A, s_B, t, \Delta = 1, \ldots, m-t, j = 2^{t-1} + 1 \) and \( k = 2^{t+\Delta-1} + 1 \), the recurrent equations

\[
E(A_1^{(t)} r_A B_1^{(t)} r_B A_1^{(t)} s_A B_j^{(t)} s_B) = E((A_1^{(t-1)} \cos \tau + A_j^{(t-1)} \sin \tau)^r_A (B_1^{(t-1)} \cos \tau + B_j^{(t-1)} \sin \tau)^r_B
\]

\[
\times (-A_1^{(t-1)} \sin \tau + A_j^{(t-1)} \cos \tau)^s_A (-B_1^{(t-1)} \sin \tau + B_j^{(t-1)} \cos \tau)^s_B),
\]

\[
E(A_1^{(t+\Delta)} r_A B_1^{(t+\Delta)} r_B A_1^{(t+\Delta)} s_A B_j^{(t+\Delta)} s_B) = E((A_1^{(t+\Delta-1)} \cos \tau + A_j^{(t+\Delta-1)} \sin \tau)^r_A
\]

\[
\times (B_1^{(t+\Delta-1)} \cos \tau + B_j^{(t+\Delta-1)} \sin \tau)^r_B
\]

\[
\times (-A_1^{(t-1)} \sin \tau + A_j^{(t-1)} \cos \tau)^s_A (-B_1^{(t-1)} \sin \tau + B_j^{(t-1)} \cos \tau)^s_B),
\]

\[
E(A_j^{(t+\Delta)} r_A B_j^{(t+\Delta)} r_B A_j^{(t+\Delta)} s_A B_j^{(t+\Delta)} s_B) = E((-A_1^{(t+\Delta-1)} \sin \tau + A_j^{(t+\Delta-1)} \cos \tau)^r_A
\]

\[
\times (-B_1^{(t+\Delta-1)} \sin \tau + B_j^{(t+\Delta-1)} \cos \tau)^r_B
\]

\[
\times (-A_1^{(t-1)} \sin \tau + A_j^{(t-1)} \cos \tau)^s_A (-B_1^{(t-1)} \sin \tau + B_j^{(t-1)} \cos \tau)^s_B)
\]

hold. Hence, we have (5.5). □

By this theorem, we can see that Hayashi’s estimators \( \hat{\theta}, \hat{\eta} \) and \( \hat{\nu} \) are biased and asymptotically inconsistent.
Let us compare $\bar{\nu}$ and $\hat{\nu}$. Since $\hat{\nu}$ is biased, it is better to compare them by mean square errors, that is,

$$
\begin{align*}
\bar{M} &:= E((\hat{\nu} - \nu)^2) = V(\hat{\nu}), \\
\hat{M} &:= E((\hat{\nu} - \nu)^2) = V(\hat{\nu}) + E(\hat{\nu} - \nu)^2.
\end{align*}
$$

If $n = 2$, then we have

$$
\begin{align*}
\bar{M} &= (\nu + 1)^2, \\
\hat{M} &= \nu^2 + \nu + (2\nu + 1)(1 - 2^{-2\epsilon})(\theta^2 + \eta^2) + \frac{3 + 2^{-8\epsilon} - 4^{-1-\epsilon}}{2} (\theta^4 + \eta^4).
\end{align*}
$$

We can see that, if $\epsilon$ or $\theta^2 + \eta^2$ is small enough, then $\bar{M} < \hat{M}$. In case of $\eta = 0$, $(\epsilon, \theta)$ satisfying $\bar{M} = \hat{M}$ are numerically calculated for $\nu = 0.1, 1, 10, 100$ (solid lines) and for $\nu = 0, \infty$ (dashed lines) and plotted in Figure 1. If the parameters are in the lower-left side of the lines, then $\bar{M} < \hat{M}$, so that Hayashi’s estimator $\nu$ for $\nu$ has advantage in the sense of mean square error.

**Figure 1**

6. Correcting Hayashi’s method by stopping the transformation

We have seen that Hayashi’s estimators $\hat{\theta}$, $\hat{\eta}$ and $\hat{\nu}$ are not so good asymptotically if the transformation is noisy. Hence, in the case of $n = 2^m$, we consider to stop the transformation $G_2$ of Remark 4 at $m_0$-th step. Let $W_j := (A_j, B_j, X_j, Y_j, Z_j)$ ($j = 1, ..., n$) be random variables obtained by the transformation

$$
\prod_{(j,k) \in J_{m_0}} g^{(j,k)}_{t_j} \prod_{(j,k) \in J_1} g^{(j,k)}_{t_j}
$$

of $W := (A_j, B_j, X_j, Y_j, Z_j)$ ($j = 1, ..., n$). Let

$$
s := \{2^{m_0} j + 1 \mid j = 0, ..., 2^{m-m_0} - 1\}
$$

be a set of indices of which we observe $(X_j, Y_j)$, hence we observe $Z_k$ if $k \notin S$. Let

$$
\begin{align*}
\tilde{\theta}_{m_0} &:= \frac{1}{2^{m-m_0/2-\epsilon m_0/2}} \sum_{j \in S} X_j, \\
\tilde{\eta}_{m_0} &:= \frac{1}{2^{m-m_0/2-\epsilon/2}} \sum_{j \in S} Y_j, \\
\tilde{\nu}_{m_0} &:= \frac{1}{2^{m-1}} \sum_{k \notin S} Z_k + \frac{2^{m-m_0} - 1}{(2^m - 1)2^{m-m_0}} \sum_{j \in S} (X_j^2 + Y_j^2) \\
&- \frac{\alpha}{\beta} \sum_{j,k \in S, j \neq k} (X_j X_k + Y_j Y_k) - \frac{2^m - 2^{m_0}}{2^{m_0}(2^m - 1)}
\end{align*}
$$

where

$$
\begin{align*}
\alpha &= 2^{-\epsilon - m - m_0} \left(2(2+\epsilon)m_0 - 2^{1+\epsilon + 2m_0 + \epsilon m_0} + 2^{1+\epsilon + m + 2m_0 + \epsilon m_0} \\
&+ 2^{2m_0 + \epsilon(2 + m_0) - 2^{m + 2m_0 + \epsilon(2 + m_0)} - 8m_0}\right), \\
\beta &= (2 + 2^\epsilon)(-2 + 2^m)(-2^m + 2^{m_0}).
\end{align*}
$$
Figure 1. Pairs of $(\epsilon, \theta)$ satisfying $\bar{M} = \hat{M}$ for $\nu = 0, 1, 10, 100$ (solid lines) and for $\nu = 0, \infty$ (dashed lines) in case of $\eta = 0$. 
Then, by the same argument of the proof of Theorem 1, we have

\[ E(\tilde{\theta}_{m_0}) = \theta, \quad E(\tilde{\eta}_{m_0}) = \eta, \]

\[ V(\tilde{\theta}_{m_0}) = \frac{1 + \nu}{2^{1+m-m_0}} - \frac{(2^\nu - 1)^2(2^{m_0} - 2^{m_0})}{2^{2+2m}(2^\nu - 2)} \theta^2, \]

\[ V(\tilde{\eta}_{m_0}) = \frac{1 + \nu}{2^{1+m-m_0}} - \frac{(2^\nu - 1)^2(2^{m_0} - 2^{m_0})}{2^{2+2m}(2^\nu - 2)} \eta^2, \]

and, if \( m_0 \) is fixed, then

\[ V(\tilde{\nu}_{m_0}) = O(n^{-1}) \]

as \( n \) goes to \( \infty \).

7. Numerical comparison and conclusion

For \( n = 2^6, 2^8, 2^{10}, 2^{12} \) and for \( \epsilon = 0.0001, \tilde{\theta}_{m_0}, \tilde{\eta}_{m_0} \) and \( \tilde{\nu}_{m_0} \) are constructed, and, for \( \theta = 0, 2, 4, 6, 8, 10, \eta = 0 \) and \( \nu = 0, 2, 4, 6, 8, 10 \), the sum of mean square errors

\[ M_{m_0} := E((\tilde{\theta}_{m_0} - \theta)^2 + (\tilde{\eta}_{m_0} - \eta)^2 + (\tilde{\nu}_{m_0} - \nu)^2) \]

are numerically calculated and their relative errors

\[ 1 - M_{m_0}/M_0 \]

are shown in Table 1. The Hayashi’s estimators or the corrected estimators are better than the naive estimators if the relative error is positive. The corrected estimators \( \tilde{\theta}_{m/2}, \tilde{\eta}_{m/2} \) and \( \tilde{\nu}_{m/2} \) are better than the Hayashi’s estimators if the relative error of the corrected ones is larger.

When the location parameter \( \theta \) is close to the origin or the scale parameter \( \nu \) is large, the effect of the randomness of the transformation is small and Hayashi’s estimators have the good performance. However, when the location is far from the origin or the scale is small, the randomness of the transformations significantly influences Hayashi’s estimators and its performance is inferior to the naive estimator. Our corrected estimators also loses the accuracy when the location is far from the origin or the scale is small, but such bad effects are relatively smaller than that of Hayashi’s estimators. Indeed, in the range of our simulation, the corrected estimators are always better than the naive estimators (see Figure 2 for \( n = 2^{10} \)).

Table 1

| \( n \) | \( M_{m_0} \) |
|-------|--------|
| 2^6   |       |
| 2^8   |       |
| 2^{10} |       |
| 2^{12} |       |

Acknowledgments

The authors wish to thank Professor M. Akahira of the University of Tsukuba and Professor H. Imai of the University of Tokyo for their support and encouragement. The authors wish to thank Dr. A. Tomita and Dr. F. Yura of IMAI Quantum Computation and Information Project for useful comments from physical points of view.
References

[1] Hayashi, M. (2000). Asymptotic quantum theory for the thermal state family. Quantum communication, computing and measurement 2 (edited by Kumar, P. D’ariano, G. M. and Hirota, O.), Plenum, New York, 99-104.

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

[3] Hirota, O. (1985). Optical Communication Theory (in Japanese), Morikita Shuppan.

[4] Holevo, A. S. (1982). Probabilistic and Statistical Aspect of Quantum Theory, North-Holland.

[5] Yuen, H. P., and Lax, M. (1973). Multiple-parameter Quantum Estimation and Measurement of Nonsymmetric Observables, Trans. IEEE, IT-19, 740-750.
Figure 2. The goodness of estimators for $n = 2^{10}$. 