Uncertainty principle in view of quantum estimation theory

Keiji Matsumoto

METR 97-08 October 1997

Uncertainty principle in view of quantum estimation theory

Keiji Matsumoto [1]

Abstract

Position-momentum uncertainty relation is examined in the light of quantum estimation theory, and some counterintuitive results are obtained. One important conclusion is that the conclusion that the so-called 'minimum-uncertainty state' is actually is the maximal uncertainty state.

Keywords: the uncertainty relation, quantum estimation theory, pure state model, Cramer-Rao type bound

[1]Department of Mathematical Engineering and Information Physics
University of Tokyo, Bunkyo-ku,Tokyo 113, Japan

1 Introduction

In this paper, we examine the position-momentum uncertainty in view of quantum estimation theory.

First, it must be emphasized that so-called 'Heisenberg's uncertainty',

$$\langle(\Delta X)^2\rangle\langle(\Delta P)^2\rangle \geq \frac{\hbar^2}{4}, \tag{1}$$

where $\langle(\Delta X)^2\rangle$ stands for

$$\langle\phi|(X - \langle\phi|X|\phi\rangle)^2|\phi\rangle,$$

has nothing to do with the Heisenberg's gedanken experiment which deals with the simultaneous measurement of the position and the momentum.

$\langle(\Delta X)^2\rangle$ (, or $\langle(\Delta P)^2\rangle$) in (1) is the variance of the data when only position (, or momentum) is measured. Therefore, (1) corresponds to the experiment in which the position is measured for the one of the group of identical particles and the momentum for the other group of identical particles. As a matter of fact, the inequality (1), is derived by H. P. Robertson [12] and some careful authors call the inequality the Robertson's uncertainty (Heisenberg himself had nothing to do with this inequality, as a matter of fact).

Second, it should be emphasised that, since the measurement obtained by the spectral decomposition of the position operator differs from that of the momentum operator, the simultaneous measurement in exact sense is impossible.

Though 'the simultaneous measurement in the weak sense' can be defined in several reasonable manners, here, we formulate the problem as a estimation of the *shift parameters* x_0 and p_0, in the *position-momentum shifted model*

$$\mathcal{M}_{xp} = \{\rho(\theta) \mid \rho(x_0, p_0) = \pi(D(x_0, p_0)|\phi_0\rangle),\ \theta = (x_0, p_0) \in \mathbf{R}^2\},$$

where

$$D(x_0, p_0) = \exp\frac{i}{\hbar}(p_0 X - x_0 P) \tag{2}$$

and $|\phi_0\rangle = \phi_0(x)$ is a member of $L^2(\mathbf{R}, \mathbf{C})$ such that

$$\langle\phi_0|X|\phi_0\rangle = \langle\phi_0|P|\phi_0\rangle = 0.$$

2 Preliminaries

2.1 Classical estimation theory

In this subsection, we briefly review the classical estimation theory (Throughout the paper, the usual estimation theory, or the estimation theory of the probability distribution is called classical estimation theory, in the sense that the theory is not quantum mechanical).

The theme of the classical estimation theory is identification of the probability distribution from which the N data $x_1, x_2, ..., x_N$ is produced. Usually, the probability distribution is assumed to be a member of a *model*, or a family

$$\mathcal{M} = \{p(x|\theta) | \theta \in \Theta \subset \mathbf{R}^m\}$$

of probability distributions and that the finite dimensional parameter $\theta \in \Theta \subset \mathbf{R}^m$ is to be estimated statistically.

Unbiased estimator $\hat{\theta} = \hat{\theta}(x_1, x_2, ..., x_N)$ of parameter θ the estimate which satisfies

$$\begin{aligned} E_\theta[\hat{\theta}] &\equiv \int dx_1 dx_2, ..., dx_N \hat{\theta}(x_1, x_2, ..., x_N) \prod_{i=1}^{N} p(x_i|\theta) \\ &= \theta, \end{aligned} \tag{3}$$

that is, the estimate which gives the true value of parameter in average. For the technical reason, we also define *locally unbiased estimator* $\hat{\theta}$ *at* θ_0 by

$$\begin{aligned} &E_{\theta_0}[\hat{\theta}] = \theta_0, \\ &\partial_j E_\theta[\hat{\theta}^i]\Big|_{\theta=\theta_0} = \delta^i_j. \end{aligned}$$

The estimator is unbiased iff it is locally unbiased at every $\theta \in \Theta$.

For the variance of locally unbiased estimator at θ, the following theorem gives bound of efficiency of the estimation.

Theorem 1 *(Cramer-Rao inequality) For any locally unbiased estimator $\hat{\theta}$ at θ,*

$$V_\theta[\hat{\theta}] \geq \frac{1}{N} J^{-1}(\theta). \tag{4}$$

Here, N is the number of the data and $J(\theta)$ is $m \times m$ real symmetric matrix defined by

$$J(\theta) \equiv \left[\int dx p(x|\theta) \partial_i \ln p(x|\theta) \partial_j \ln p(x|\theta) \right], \tag{5}$$

where ∂_i stands for $\partial / \partial \theta^i$.

The best estimator, or the estimator $\hat{\theta}$ satisfying (6)*, is given by*

$$\begin{aligned} \hat{\theta}^i(x_1, ..., x_N) &= \hat{\theta}^i_{(\theta)}(x_1, ..., x_N) \\ &\equiv \theta^i + \sum_{j=1}^{m} [J^{-1}(\theta)]^{ij} \partial_j \ln \prod_{k=1}^{N} p(x_k|\theta). \end{aligned}$$

$J(\theta)$ is called *Fisher information matrix*, because the larger the $J(\theta)$ is, the more precise estimate can be done with the same number of data. Metaphorically speaking, we obtain as much information as $J(\theta)$ per data. Actually, as easily seen by putting $N = 1$ in Cramer-Rao (CR) inequality, we can obtain $J(\theta)$ as the minimum variance of locally unbiased estimate when only one data is given.

$$V_\theta[\hat{\theta}] \geq J^{-1}(\theta). \tag{6}$$

The trouble with the (6) is that the best estimator $\hat{\theta}_{(\theta)}$ is dependent on the true value of the parameter θ, which is unknown to us. To avoid this dilemma, we loosen the unbiasedness conditions, and consider *consistent estimators*, which is defined by,

$$\lim_{N \to \infty} E_\theta[\hat{\theta}(x_1, x_2, ..., x_N)] = \theta.$$

Theorem 2 *If the estimator is consistent,*

$$V_\theta[\hat{\theta}] \geq \frac{1}{N} J^{-1}(\theta) + o\left(\frac{1}{N}\right) \tag{7}$$

holds true.

The maximum likelihood estimator $\hat{\theta}_{MLE}$, which is defined by,

$$\hat{\theta}_{MLE} \equiv \mathrm{argmax} \left\{ \sum_{j=1}^{N} \ln p(x_i|\theta) \,\middle|\, \theta \in \Theta \subset \mathbf{R}^m \right\}.$$

is consistent and achieves the equality in (7).

Notice that to obtain $\hat{\theta}_{MLE}$, we need no information about the true value of the parameter beforehand. Hence, the Fisher information matrix is a good measure of the efficiency of the optimal consistent estimator.

2.2 Locally unbiased measurement

In this subsection and the next, conventional theory of quantum estimation is reviewed briefly. Suppose that a physical state belongs to a certain manifold $\mathcal{M} = \{\rho(\theta)|\theta \in \Theta \subset \mathbf{R}^m\} \subset \mathcal{P}_n$, and that the true value of the parameter θ is not known. In this paper, we restrict ourselves to the case of *pure state model*, or the case where any member of of the model $\mathcal{M}$ is pure state,

$$\begin{aligned} \rho(\theta) &= \pi(|\phi(\theta)\rangle) \\ &\equiv |\phi(\theta)\rangle\langle\phi(\theta)|. \end{aligned} \tag{8}$$

Whatever measuring apparatus is used to produce the estimate $\hat{\theta}$ of the true value of the parameter θ, the probability that the estimate $\hat{\theta}$ lie in a particular subset B of the space $\mathbf{R}^m$ of the results will be given by

$$\Pr\{\hat{\theta} \in B|\theta\} = \mathrm{tr}\rho(\theta)M(B) \tag{9}$$

when θ represents the true value of parameter. Here M is a mapping of subsets $B \subset \mathbf{R}^m$ to non-negative Hermitian operators on $\mathcal{H}$, such that

$$M(\phi) = O, M(\mathbf{R}^m) = I, \tag{10}$$

$$M(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} M(B_i) \ \ (B_i \cap B_j = \phi, i \neq j), \tag{11}$$

(see Ref.[5],p.53 and Ref.[6],p.50.). M is called a *generalized measurement* or *measurement*, because there is a corresponding measuring apparatus to any M satisfying (11) [11][15]. A measurement E is said to be *simple* if E is projection valued.

A generalized measurement M is called an *unbiased measurement* in the model $\mathcal{M}$, if $E_\theta[M] = \theta$ holds for all $\theta \in \Theta$, i.e.,

$$\int \hat{\theta}^j \mathrm{tr}\rho(\theta)\, M((d\hat{\theta}) = \theta^j, \quad (j = 1, \cdots, m). \tag{12}$$

Differentiation yields

$$\int \hat{\theta}^j \mathrm{tr}\frac{\partial \rho(\theta)}{\partial \theta^k}\, M(d\hat{\theta}) = \mathrm{tr}\frac{\partial \rho(\theta)}{\partial \theta^k}\, X^j = \delta^j_k, \quad (j, k = 1, \cdots, m). \tag{13}$$

If (12) and (13) hold at a some θ, M is said to be *locally unbiased* at θ. Obviously, M is unbiased iff M is locally unbiased at every $\theta \in \Theta$. For simplicity, only locally unbiased estimators are treated from now on.

As a measure of error of a locally unbiased measurement M, we employ the covariance matrix with respect to M at the state ρ_θ, $V_\theta[M] = [v_\theta^{jk}] \in \mathbf{R}^{m\times m}$, where

$$v_\theta^{jk} = \int (\hat{\theta}^j - \theta^j)(\hat{\theta}^k - \theta^k)\mathrm{tr}\rho(\theta) M(d\hat{\theta}). \tag{14}$$

2.3 SLD CR bound and the attainable CR type bound

Analogically to the classical estimation theory, in the quantum estimation theory, we have the following *SLD CR inequality*, which is proved for the exact state model by Helstrom [4][5], and is proved for the pure state model by Fujiwara and Nagaoka [3]:

$$V_\theta[M] \geq (J^S(\theta))^{-1}, \tag{15}$$

i.e., $V_\theta[M] - (J^S(\theta))^{-1}$ is non-negative definite. Here $V_\theta[M]$is a covariance matrix of an unbiased measurement M, and $J^S(\theta)$ is called *SLD Fisher information matrix*, and is defined by

$$J^S(\theta) \equiv [\mathrm{Re}\langle l_i(\theta)|l_j(\theta)\rangle], \tag{16}$$

where $|l_i(\theta)\rangle$ $(i = 1, ..., m)$ are defined afterward.

The inequality (15) is of special interest, because J^{S-1} is the one of the best bounds in the sense explained later.

First, we set some notations. The *horizontal lift* $|l_X\rangle$ of a tangent vector $X \in \mathcal{T}_{\rho(\theta)}(\mathcal{M})$ to $|\phi(\theta)\rangle$, is an element of $\mathcal{H}$ which satisfies

$$X\rho(\theta) = \frac{1}{2}(|l_X\rangle\langle\phi(\theta)| + |\phi(\theta)\rangle\langle l_X|), \tag{17}$$

and

$$\langle l_X|\phi(\theta)\rangle = 0. \tag{18}$$

Here, X in the left hand side of (17) is to be understood as a differential operator, and $\pi(*)$ is defined as $\pi(|\phi(\theta)\rangle) = \rho(\theta)$. $|l_i(\theta)\rangle$ is defined to be a horizontal lift of $\partial_i \in \mathcal{T}_{\rho(\theta)}(\mathcal{M})$.

It is proved that SLD CR bound is attainable iff $\langle l_j|l_i\rangle$ is real for any i, j. When SLD-CR bound is attainable, that bound is achieved by a simple measurement, i.e., a projection valued measurement. Especially, when the model has only one parameter, SLD CR bound is always attainable.

Is there any better bound than SLD Fisher information matrix which is always attainable? The answer is negative: Letting A be a real hermitian matrix which is larger than J^{S-1}, that is, $A > J^{S-1}$, there exists such an unbiased estimator M that $V[M]$ is not smaller than A. In other words, there is no better bound than SLD CR bound.

In the general case, therefore, no matrix makes attainable lower bound of $V_\theta[M]$. Hence, instead, we determine the *attainable CR type bound* $\mathrm{CR}(G, \theta, \mathcal{M})$, which is defined by

$$\mathrm{CR}(G, \theta, \mathcal{M}) = \min\{\mathrm{Tr}GV_\theta[M] \mid M \text{ is locally unbiased at } \theta\} \tag{19}$$

for an arbitrary nonnegative symmetric real matrix G. G is called *weight matrix.* To make the estimational meaning of (19) clear, let us restrict ourselves to the case when G is $diag(g_1, g_2..., g_m)$. Letting v_{ii} is the (i, i)-th component of $V[M]$,

$$\mathrm{Tr}GV[M] = \sum_i g_i v_{ii}, \tag{20}$$

is the weighed sum of the covariance of the estimation of θ^i. If one wants to know, for example, θ^1 more precisely than other parameters, then he set g_1 larger than any other g_i, and pick up a measurement which minimize $\sum_i g_i v_{ii}$.

3 Estimation theory of the 2-parameter pure state model

In this section, some theorems about the attainable CR-type bound of 2-parameter pure state model is reviewed (see Ref. [10]).

We denote by $\mathcal{V}_\theta(\mathcal{M})$ the boundary of the region of the map $V_\theta[*]$ from an unbiased estimator to a 2×2 matrix. It is known that $\mathcal{V}_\theta(\mathcal{M})$ is convex.

As for the 2-parameter pure state model, the region $\mathcal{V}_\theta(\mathcal{M})$ (for notational simplicity, we often omit θ and $\mathcal{M}$) is already explicitly calculated. Define $\beta(\theta)$ by

$$\beta \equiv \frac{|\,\mathrm{Im}\langle l_1 | l_2 \rangle\,|}{\sqrt{\det J^S}}, \tag{21}$$

and the map $\mathbf{V}(*)$ from a matrix to a matrix by

$$\mathbf{V}(V) \equiv (J^S)^{1/2}(V - J^{S-1})(J^S)^{1/2}, \tag{22}$$

and $\mathcal{V}$ is a set of matrices which satisfy

$$V - J^{S-1} \geq 0, \tag{23}$$

and

$$\det \sqrt{\mathbf{V}(V)} + \left(\frac{1}{\beta^2} - 1 \right)^{1/2} \mathrm{Tr} \sqrt{\mathbf{V}(V)} \geq 1. \tag{24}$$

It should be remarked that β is geometrical scalar in the sense that it is unchanged by the change of the parameterization of the model, and that it takes value from 0 to 1. It is also remarkable that β is nothing but Berry's phase [1] [14] per unit area.

Theorem 3 *If $\beta(\theta', \mathcal{M}') \geq \beta(\theta, \mathcal{M})$ and $J^S(\theta', \mathcal{M}') = J^S(\theta, \mathcal{M})$, then for any weight matrix G,*

$$\mathrm{CR}(G, \theta', \mathcal{M}') \geq \mathrm{CR}(G, \theta, \mathcal{M}). \tag{25}$$

Proof The inequality (24) implies that if $\beta(\theta', \mathcal{M}') \geq \beta(\theta, \mathcal{M})$,

$$\mathcal{V}_\theta(\mathcal{M}) \subset \mathcal{V}_{\theta'}(\mathcal{M}'),$$

which directly leads to (25), for any weight matrix G. □

Theorem 4 *If $\beta(\theta', \mathcal{M}') = \beta(\theta, \mathcal{M})$ and $J^S(\theta', \mathcal{M}') \leq J^S(\theta, \mathcal{M})$, then for any weight matrix G, (25) holds true.*

The proof is omitted here for simplicity.

4 The estimation of the shift parameters

Our purpose is to examine how efficiently we can estimate the shift parameters $\theta = (x_0, p_0)$.

The horizontal lifts of $\partial/\partial x_0$, $\partial/\partial p_0$ are

$$\begin{aligned} \underset{|\phi(\theta)\rangle}{h}\left(\frac{\partial}{\partial x_0}\right) &= -\frac{i}{\hbar}\Delta P_\theta|\phi(\theta)\rangle, \\ \underset{|\phi(\theta)\rangle}{h}\left(\frac{\partial}{\partial p_0}\right) &= \frac{i}{\hbar}\Delta X_\theta|\phi(\theta)\rangle, \end{aligned}$$

where

$$\begin{aligned} \langle A\rangle_\theta &\equiv \langle\phi(\theta)|A|\phi(\theta)\rangle, \\ \Delta A_\theta &\equiv A - \langle A\rangle_\theta \end{aligned}$$

and the SLD Fisher information matrix $J^S(\theta)$ is,

$$\begin{aligned} [J^S(\theta)]_{x_0,x_0} &= \frac{1}{\hbar^2}\langle(\Delta P_\theta)^2\rangle, \\ &= \frac{1}{\hbar^2}\langle(\Delta P_{(0,0)})^2\rangle, \\ [J^S(\theta)]_{p_0,p_0} &= \frac{1}{\hbar^2}\langle(\Delta X_{(0,0)})^2\rangle, \\ [J^S(\theta)]_{x_0,p_0} &= \frac{1}{\hbar^2}Cov(X, P)_{(0,0)}, \end{aligned}$$

where $Cov(X,P)_\theta \equiv \frac{1}{2}\langle(\Delta X_\theta \Delta P_\theta + \Delta P_\theta \Delta X_\theta)$. The absolute value $\beta(\theta)$ of the eigenvalue of $\mathbf{D}$ at θ is calculated as,

$$\begin{aligned}\beta(\theta) &= \frac{1}{2i}\frac{\langle[P,X]\rangle_{(0,0)}/\hbar^2}{\hbar^2\sqrt{\langle(\Delta X_{(0,0)})^2\rangle\langle(\Delta P_{(0,0)})^2\rangle - (Cov(X,P)_{(0,0)})^2}} \\ &= \frac{1}{\sqrt{\langle(\Delta X_\theta)^2\rangle\langle(\Delta P_\theta)^2\rangle - (Cov(X,P)_\theta)^2}}.\end{aligned} \tag{26}$$

Notice that $J^S(\theta)$ and $\beta(\theta)$ are independent of the true value of parameters.

We are interested in the attainable CR type bound and in the index β of noncommutative nature of the model.

As for β, (26) indicates that the larger the formal 'covariance matrix' [2] of X and P

$$\begin{bmatrix} \langle(\Delta P_{(0,0)})^2\rangle & Cov(X,P)_{(0,0)} \\ Cov(X,P)_{(0,0)} & \langle(\Delta X_{(0,0)})^2\rangle \end{bmatrix}$$

is, the smaller the noncommutative nature between x_0 and p_0.

As for the attainable CR type bound, because the SLD Fisher information matrix is proportional to the formal 'covariance matrix' and β decreases as the determinant of the 'covariance matrix' increases, we can metaphorically say that the larger the 'covariance matrix' implies the possibility of more efficient estimate of x_0 and p_0, which is seemingly paradoxical.

We examine these points in the *shifted harmonic oscillator model* $\mathcal{M}_{xp,n}$, which is defined to be the position-momentum shifted model in which $\phi_0(x)$ is equal to the nth eigenstate $|n\rangle$ of the harmonic oscillator. For $\mathcal{M}_{xp,n}$, we have

$$\begin{aligned}[J^S(\theta)]_{x_0,x_0} &= \frac{4}{\hbar}\left(n+\frac{1}{2}\right), \\ [J^S(\theta)]_{p_0,p_0} &= \frac{4}{\hbar}\left(n+\frac{1}{2}\right), \\ [J^S(\theta)]_{x_0,p_0} &= [J^S(\theta)]_{(x_0,p_0)} = 0,\end{aligned}$$

[2]Note that this formal 'covariance matrix' is not the covariance matrix of any measurement related to position or momentum

and

$$\beta(\theta) = \frac{1}{2(n+1/2)}.$$

Hence, if n is large, 'noncommutative nature' of the parameters is small.

What about the efficiency of the estimate ? We define $\mathcal{M}'_{xp,n}$ by normalizing the parameters in $\mathcal{M}_{xp,n}$ as

$$\theta = (x_0, p_0) \longrightarrow \theta = (n+1/2)^{-1/2} x_0, (n+1/2)^{-1/2} p_0).$$

Then, directly from the definition,

$$\mathrm{CR}(G, \theta, \mathcal{M}_{xp,n}) = \frac{1}{n+1/2} \mathrm{CR}(G, \theta, \mathcal{M}'_{xp,n}), \tag{27}$$

and the SLD Fisher information matrix of $\mathcal{M}'_{xp,n}$ is equal to $\frac{4}{\hbar} I_m$ for any n.

Because the index β is unchanged by the change of the parameter,

$$\mathrm{CR}(G, \theta, \mathcal{M}'_{xp,0}) \geq \mathrm{CR}(G, \theta, \mathcal{M}'_{xp,1}) \geq ... \geq \mathrm{CR}(G, \theta, \mathcal{M}'_{xp,n}) \geq ...,$$

which, combined with (27) leads to

$$\mathrm{CR}(G, \theta, \mathcal{M}_{xp,0}) \geq \mathrm{CR}(G, \theta, \mathcal{M}_{xp,1}) \geq ... \geq \mathrm{CR}(G, \theta, \mathcal{M}_{xp,n}) \geq ...,$$

for any θ and any G. Therefore, if the 'covariance matrix' larger, the more efficient estimate of the parameter is possible. Especially, when $n = 0$, or in the case of the so-called 'minimum uncertainty state', the efficiency of the estimation is the lowest.

Especially, when $n = 0$, or in the case of the so-called 'minimum uncertain state', the position parameter x_0 and the momentum parameter p_0 are maximally 'noncommutative' in the sense β is larger than that of any other $\mathcal{M}_{xp,n}$ $(n \neq 0)$. It is easily shown that β is maximal, or coherent, iff $|\phi_0\rangle$ is in the squeezed state, or $|\phi_0\rangle = S(v, w)|0\rangle$, where, letting $a = (Q + iP)/\sqrt{2\hbar}$, $S(\xi)$ is the operator defined by

$$S(\xi) = \exp\left\{ \frac{1}{2} (\, \xi a^{\dagger 2} - \overline{\xi} a^2 \,) \right\}. \tag{28}$$

$n = 0$-case can be also said to be 'maximally uncertain', in the sense that the efficiency of the estimation is lower than any other $\mathcal{M}_{xp,n}$ $(n \neq 0)$.

If n is very large, how efficiently can we estimate the shift parameters? Given N particles, we divide them into two groups, to one of which we apply the best measurement for x_0 and to the other of which we apply the best measurement for p_0 . The parameter x_0 and p_0 is estimated only from the data from the first group and the second group, respectively. Then, the attained efficiency of the estimation of x_0 is

$$\frac{\hbar}{(N/2)\,4(n+1/2)}$$

and the efficiency of the estimation of p_0 is

$$\frac{\hbar}{(N/2)\,4(n+1/2)}$$

which are combined to yield the efficiency of this estimate par sample,

$$g_{x_0}[V[M]]_{x_0,x_0} + g_{p_0}[V[M]]_{p_0,p_0} = \frac{\hbar g_{x_0}}{2(n+1/2)} + \frac{\hbar g_{p_0}}{2(n+1/2)}, \tag{29}$$

in this estimation scheme. Therefore, we have

$$\mathrm{CR}(\mathrm{diag}(g_1, g_2), \mathcal{M}_{xp,n}) \leq \frac{\hbar g_{x_0}}{2(n+1/2)} + \frac{\hbar g_{p_0}}{2(n+1/2)},$$

which implies that arbitrarily precise estimate is possible if $\mathcal{M}_{xp,n}$ with large enough n is fortunately given.

The efficiency (29) is achievable by the following maximum likelihood estimator up to the first order of $1/N$:

$$\begin{aligned} \hat{x}_0 &= \mathrm{argmax}_{x_0} \sum_{j=1}^{N/2} \ln p(x_j - x_0), \\ \hat{p}_0 &= \mathrm{argmax}_{p_0} \sum_{j=1}^{N/2} \ln \tilde{p}(p_j - p_0), \end{aligned} \tag{30}$$

where $x_1, x_2, ..., x_{N/2}$ and $p_1, p_2, ..., p_{N/2}$, be is data produced by the measurement of the position and the momentum of the given states, and their probability distribution is denoted by $p(x)$ and $\tilde{p}(p)$, respectively.

5 Planck's constant and Uncertainty

In this section, we focus on Planck's constant. Let

$$P' = \hbar^{-1/2}P,\ X' = \hbar^{-1/2}X,$$
$$p_0' = \hbar^{-1/2}p_0,\ x_0' = \hbar^{-1/2}x_0,$$

and define

$$\mathcal{M}'_{xp} = \{\rho(p_0', x_0) \,|\, \rho(p_0', x_0') = \pi(\exp i(p_0'X' - x_0'P')|\phi_0\rangle\,)\}.$$

Since Planck's constant does not appear in the commutation relation $[P', X'] = -i$, the attainable CR type bound of $\mathcal{M}'_{xp}$ is not dependent on $\hbar$ if definition of $|\phi_0\rangle$ does not include $\hbar$.

If $\mathrm{CR}(\mathcal{M}'_{xp})$ has some finite value, the identity

$$\mathrm{CR}(\mathcal{M}_{xp}) = \hbar\mathrm{CR}(\mathcal{M}'_{xp})$$

implies

$$\lim_{\hbar\to 0}\mathrm{CR}(\mathcal{M}_{xp}) = 0.$$

Therefore, in the limit of $\hbar \to 0$, position and momentum can be simultaneously measured as precisely as needed.

However, it must be noticed that the attainable CR type bound of the *position shifted model*

$$\mathcal{M}_x = \{\rho(x_0) \,|\, \rho(x_0) = \pi(\exp(-ix_0P)|\phi_0\rangle\,)\}$$

and of the *momentum shifted model*

$$\mathcal{M}_p = \{\rho(p_0) \,|\, \rho(p_0) = \pi(\exp(ip_0P)|\phi_0\rangle\,)\}$$

also tends to zero as $\hbar \to 0$, and that the ratio

$$\frac{\mathrm{CR}(\mathcal{M}_{xp})}{\sqrt{\mathrm{CR}(\mathcal{M}_x)\,\mathrm{CR}(\mathcal{M}_p)}} = \frac{\mathrm{CR}(\mathcal{M}'_{xp})}{\sqrt{\mathrm{CR}(\mathcal{M}'_x)\,\mathrm{CR}(\mathcal{M}'_p)}}$$

is independent of $\hbar$, where $\mathcal{M}'_x$ and $\mathcal{M}'_p$ are defined in the same manner as $\mathcal{M}'_{xp}$. Therefore, noncommutative nature of the model is unchanged even if $\hbar$ tends to 0. Actually, as in (26)the index β of the noncommutative nature of the model is independent of $\hbar$.

Remark Notice the discussion in this section is essentially valid for the mixed position-momentum shifted model,

$$\{\rho(p_0, x_0) \,|\, \rho(p_0, x_0) = D(x_0,\, p_0\,)\rho_0 D^*(x_0,\, p_0\,),\, (x_0,\, p_0) \in \mathbf{R}^2\},$$

where the state ρ_0 is mixed, and $D(x_0,\, p_0\,)$ is the operator defined by (2).

6 Semiparametric estimation of the shift parameters

In section 4, our conclusion is that if we are fortunate enough, we can estimate the average of the position and the momentum with arbitrary accuracy at the same time.

One may argue that this is because we make full use of knowledge about the shape of the wave function of the given state. However, this argument is not thoroughly true.

In the classical estimation theory, we have the following very strong result. Suppose that we are intersected in the mean value $\theta \in \mathbf{R}$ of the probability distribution, and that the shape of the probability distribution is unknown except it is symmetric around θ. In other words, we set up the *semiparametric model* such that,

$$\{p(x \,|\, \theta, g) \,|\, p(x \,|\, \theta, g) = g(x - \theta),\ \theta \in \mathbf{R},\ g(x) \text{ is symmetric around } 0\}, \tag{31}$$

and estimate the parameter $\theta \in \mathbf{R}$ from the data $x_1, x_2, ..., x_N \in \mathbf{R}$.

If $g(x)$ is known, the variance of the best consistent estimator is given by

$$\frac{1}{NJ} + o\left(\frac{1}{N}\right) \tag{32}$$

where J is the Fisher information,

$$J = \int \left(\frac{d}{dx} \ln g(x) \right)^2 g(x) dx. \tag{33}$$

In the case where $g(x)$ is not known, the *theorem* 2.2 in the Ref. [2] insists that the bound (32) is attainable:

Theorem 5 *If $g(x)$ is absolutely continuous, the bound* (32) *is attainable by some consistent estimate (see pp. 649-650 in th Ref* [2]*).*

By the use of this theorem, if $p(x)$ and $\tilde{p}(p)$ defined in the end of section 4 are symmetric about x_0 and p_0 respectively, we can use the semiparametric estimates, instead of the maximum likelihood estimates (30), and can achieve the same efficiency as (30). Then, if we are so fortunate that the $|\phi_0\rangle$ is happen to be $|n\rangle$ with very large n, our estimate is quite accurate.

Acknowledgment

The author is grateful to Dr. A. Fujiwara, for introducing the quantum estimation theory to the author. The author is also thankful to Mr. M. Hayasi for inspiring discussions.

References

[1] M. V. Berry, "Quantal phase factors accompanying adiabatic changes," Proc. Roy. Soc. London **A392**, 45–57 (1984).

[2] P. J. Bickel, "The 1980 Wald Memorial Lectures On Adaptive Estimation," The Annals of Statistics, Vol.10, No.3, pp. 647-671 (1982).

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

[4] C. W. Helstrom, "Minimum Mean-Square Error Estimation in Quantum Statistics," Phys. Lett., **25A**, 101-102(1967).

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

[6] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982) (in Russian, 1980).

[7] E. L. Lehmann, "Theory of Point Estimation," Jhon Wiley (1983).

[8] E. L. Lehmann, "Testing Statistical Hypothesis," 2nd ed., Jhon Wiley (1986).

[9] K. Matsumoto, "A new approach to the Cramer-Rao type bound of the pure state model," METR 96-09 (1996).

[10] K. Matsumoto, "Uhlmann's parallelity in quantum estimation theory ," METR 97-03 (1997).

[11] M. Ozawa, "Quantum measuring processes of continuous observables," J. Math. Phys. **25**, 79-87 (1984).

[12] H. P. Robertson, Phys. Rev., **34**(1), 163-164, (1924).

[13] J. J. Sakurai, "Modern Quantum Mechanics," Benjamin/Cummings Publishing Company,Inc,(1985).

[14] A.Shapere,F.Wilczek, "GEOMETRIC PHASES IN PHYSICS," Advanced Series in Mathematical Physics,vol. 5, World Scientific (1989).

[15] W.F.Steinspring, "Positive functions on C^*-algebras," Proc. Am. Math. Soc. **6**, 211-216(1955).