Quantum estimation by local observables

Masahiro Hotta[1,*] and Masanao Ozawa[2,†]
[1]*Department of Physics, Faculty of Science, Tohoku University, Sendai, 980-8578, Japan*
[2]*Graduate School of Information Sciences, Tohoku University, Sendai, 980-8579, Japan*
(Received 30 January 2004; published 30 August 2004)

Quantum estimation theory provides optimal observations for various estimation problems for unknown parameters in the state of the system under investigation. However, the theory has been developed under the assumption that every observable is available for experimenters. Here, we generalize the theory to problems in which the experimenter can use only locally accessible observables. For such problems, we establish a Cramér-Rao-type inequality by obtaining an explicit form of the Fisher information as a reciprocal lower bound for the mean-square errors of estimations by locally accessible observables. Furthermore, we explore various local quantum estimation problems for composite systems, where nontrivial combinatorics is needed for obtaining the Fisher information.

DOI: 10.1103/PhysRevA.70.022327 PACS number(s): 03.67.−a

I. INTRODUCTION

In many experimental situations, we are not allowed to have a large number of data enough to determine unknown parameters such as coupling constants of hypothetical interactions. In some cases, the number may be fairly small and it is crucial to theoretically explore the best estimator for the parameter from the small number of our available data. The problem becomes prominent for quantum systems, since optimal estimation must be well reconciled with the inevitable quantum uncertainty arising from available observables and unknown parameters. In such situations, the quantum estimation theory can play a significant role; for detailed reviews, we refer to Helstrom [1] and Holevo [2]. The theory provides the best observation on the system for the estimation with the minimum value of the estimate error.

Although the ordinary quantum estimation theory is certainly powerful for many estimation problems, the theory includes an implicit assumption which is not realistic in some practical experiments. The assumption is that every observable of the system is available for the observer or the experimenter. Contrary to this assumption, practically available observables are often restricted. For instance, it is a common situation in experiments that a particle is contained inside the laboratory at the origin of the time and that the experimenter can only use measuring devices inside the laboratory. However, according to the time evolution the particle may go out of the laboratory, so that the ability to estimate the state parameter is restricted to measuring devices inside the laboratory for the later time. Another example is found in elementary particle physics. It usually happens due to the limits of the present technology of measurement that our apparatus can probe only low-energy portions of the total Hilbert space with visible signals. Thus the observables are certainly restricted. In such situations, the observable provided by the ordinary theory for the best estimate may in general not be available. Then, the question becomes relevant as to what is the best estimate among those which are accessible *only by use of restricted observables*. Let us generally call such estimations local quantum estimations.

In this paper, we elaborate on the formulation of quantum estimation theory for local quantum estimation problems on an unknown parameter g. For the restricted density operators measured by our apparatus, a Fisher information is introduced. Then, we prove the quantum Cramér-Rao-type inequality for the local quantum estimation for g. The observable is specified that attains equality and yields the best local quantum estimate for g by its measurement. It is also pointed out that there exist nontrivial aspects in the analysis of the local quantum estimation for the composite system of identical subsystems. In that case we have two natural estimations and the corresponding two Fisher informations for the unknown parameter g. The first alternative takes a simple form to apply, but may give a smaller value of the Fisher information. The second alternative is able to give a larger value of the Fisher information and generates a better estimate for g, but have a pretty complicated form to deal with, compared to the first alternative. Especially, calculation of the second Fisher information requires solving independently the evolutions of many descendant operators.

In Sec. II, a brief review of the standard quantum estimation theory is given. In Sec. III, we discuss more physically the quantum estimation problem, including the biased-estimator case. Several expected advantages of the quantum estimation are also reviewed. In Sec. IV, we introduce the notion of local quantum estimation problems. In Sec. V, a quantum Cramér-Rao inequality for local quantum estimations is established. In Sec. VI, Fisher information is discussed for unnormalized pure states. In Sec. VII, we reveal nontrivial aspects of local quantum estimations for composite systems. In Sec. VIII, two general formulations are proposed for local quantum estimations for the composite system. In Sec. IX, a formula which is useful for the evaluation of one of two sorts of Fisher information for composite systems is given. In Sec. X, we apply our formulation for local quantum estimations to a decaying two-level system with a

*Electronic address: hotta@tuhep.phys.tohoku.ac.jp
†Electronic address: ozawa@math.is.tohoku.ac.jp

small unknown parameter. In the final section, we summarize our results of this paper.

II. CRAMÉR-RAO BOUND FOR QUANTUM ESTIMATORS

Let us briefly review quantum estimation theory in this section. For detailed reviews we refer the reader to Helstrom [1] and Holevo [2]

Let **S** be a closed quantum system described by a Hilbert space $\mathcal{H}$. We assume that the Hamiltonian H_{tot} has a constant real parameter $g \in G$—i.e., $H_{tot}=H_{tot}(g)$, where G is the set of possible values of g. The evolution equation for the density operator ρ_{tot} of **S** is given by

$$i\hbar\partial_t\rho_{tot} = [H_{tot}, \rho_{tot}]. \tag{1}$$

Then, the density operator of **S** at a given time t depends on the time t and the parameter g—i.e.,

$$\rho_{tot} = \rho_{tot}(g,t). \tag{2}$$

We shall consider the following quantum estimation problem for the parameter g. In order to estimate the parameter g in $H_{tot}(g)$, we assume that one measures an observable A at time t, and the output $\overline{g}$ is taken as the estimate for g. Thus, the observable A plays the role of the estimator of this statistical estimation problem.

By the Born statistical formula, the expectation value of the measurement output $\overline{g}$ in the state $\rho_{tot}(g)=\rho_{tot}(g,t)$ is given by

$$E_g[A] = \mathrm{Tr}[A\rho_{tot}(g)]. \tag{3}$$

Then, $E_g[A]$ is the mean of one's estimate $\overline{g}$ for the given true value $g(\in G)$. The variance of the estimate $\overline{g}$ for the true value $g(\in G)$ is given by

$$V_g[A] = (\Delta_g A)^2 = \mathrm{Tr}[A^2\rho_{tot}(g)] - \mathrm{Tr}[A\rho_{tot}(g)]^2, \tag{4}$$

where $\Delta_g A$ is the uncertainty of observable A in the state $\rho_{tot}(g)$. The mean-square error $E_g^2[A]$of the estimate $\overline{g}$ for the true value g is defined by

$$E_g^2[A] := E_g[(A-g)^2]. \tag{5}$$

By a simple manipulation, we obtain the relation

$$E_g^2[A] = V_g[A] + (E_g[A]-g)^2. \tag{6}$$

The estimator A is called *unbiased* if the mean estimate is correct for any possible values $g(\in G)$—i.e.,

$$E_g[A] = g \tag{7}$$

for any $g(\in G)$. In this section, we shall confine our attention to unbiased estimators A. As seen in Eq. (6), for unbiased A, the variance $V_g[A]$ represents the mean-square error of the estimate $\overline{g}$ for the true value $g(\in G)$:

$$E_g^2[A] = V_g[A]. \tag{8}$$

The lower bound for $V_g[A]$ is given by the well-known quantum Cramér-Rao inequality as follows. The symmetric logarithmic derivative (SLD) $L(g)$ for $\rho_{tot}(g)$ is defined as a self-adjoint operator satisfying

$$\partial_g\rho_{tot}(g) = \frac{1}{2}[\rho_{tot}(g)L(g) + L(g)\rho_{tot}(g)]. \tag{9}$$

Note here that

$$E_g[L(g)] = \mathrm{Tr}[\rho_{tot}(g)L(g)] = 0, \tag{10}$$

due to the normalization condition of ρ_{tot}. By the above relation, the SLD may not be determined uniquely; however, any two solutions $L_1(g)$ and $L_2(g)$ satisfy the relation [3]

$$L_1(g)\rho_{tot}(g) = L_2(g)\rho_{tot}(g). \tag{11}$$

Thus, the operator $L(g)\rho_{tot}(g)$ is uniquely determined. The Fisher information J_g of the parameter g in $\rho_{tot}(g)$ is uniquely defined by

$$J_g = \mathrm{Tr}[L^2(g)\rho_{tot}(g)]. \tag{12}$$

Then, every unbiased estimator A satisfies the quantum Cramér-Rao inequality [1,2]

$$E_g^2[A] = V_g[A] \geqslant \frac{1}{J_g}. \tag{13}$$

A simple proof is given in the Appendix.

III. MORE PHYSICAL REVIEW OF QUANTUM ESTIMATION

In this section, we shall discuss real experimental procedures for estimating the unknown parameter g in the framework of quantum estimation theory given above. In real experiments, rigorous unbiased estimators that satisfy Eq. (7) globally in the parameter space of g are usually not available for experimenters from technical reasons. Instead, only biased estimators are available. Even in such real situations, as we shall show in the following, the quantum estimation theory described in the previous section plays an active role. We also discuss some advantages of quantum estimation theory to provide more efficient methods in several estimation problems in physics.

We shall first consider a quantum estimation for a parameter g in $H_{tot}(g)$ performed by a measurement of a general observable A for a single sample at time t. Here A is not assumed unbiased. Assume that we get an outcome $\overline{a}$ in the measurement and we make the estimate $\overline{g}$ for g as a function of the outcome $\overline{a}$. Usually, this function $\overline{g}=\phi(\overline{a})$ is obtained by the following way. The relation

$$E_g[A] = a \tag{14}$$

between the true value g and the mean output a from the measurement can be often solved theoretically as a function

$$g = \phi(a). \tag{15}$$

Usually, sensible experiments are designed to possess a suitable domain $\mathcal{G}(\subseteq G)$, which includes interesting values of g, and to allow the relation

$$\overline{g} = \phi(\overline{a}), \tag{16}$$

applying the above function ϕ to the measurement output $\overline{a}$, which gives a good estimate $\overline{g}$ for the given true value

$g(\in\mathcal{G})$ from the output of the single A measurement. Note that, due to the quantum nature, the observable A generally possesses nonzero values of the variance $V_g[A]$. The variance has a close relationship to the problem: to what extent the estimate $\overline{g}$ can be trusted. For example, if $V_g[A]$ is extremely small, then, even in the single measurement, the observed value $\overline{a}$ must be almost equal to the expectation value $E_g[A]$ for the correct value of g. Consequently, the estimate $\overline{g}$ has to almost coincide with the correct g value. From the viewpoint of the dimensional analysis, it is rather straightforward to introduce an expected error $\delta g_g[A]$ of the estimate $\overline{g_a}[A]$ for the true value of g as follows:

$$(\delta g_g[A])^2 := \frac{V_g[A]}{(\partial_g E_g[A])^2}. \tag{17}$$

Now, we shall give a justification of the above estimate and error evaluation from the viewpoint of quantum estimation theory. When $V_g[A]$ is small enough, a domain $\mathcal{G}$ in the space G of possible values can be chosen so narrow that $E_g[A]$ is linearly expanded around a physically interesting value $g_o(\in\mathcal{G})$:

$$E_g[A] = E_{g_o}[A] + \partial_g E_g[A]\big|_{g=g_o}(g-g_o) + O((g-g_o)^2). \tag{18}$$

Here it is quite useful to remind that a single measurement of the observable A simultaneously implies a single measurement of an observable $f(A)$, where $f(x)$ is an arbitrary real function of x. If an output $\overline{a}$ of the observable A is obtained, it is interpreted that an output $f(\overline{a})$ is observed for the observable $f(A)$ in the same measurement. In what follows, in order to make a useful choice of f, let us impose on the function f the locally unbiased condition:

$$E_g[f(A)] = g + O((g-g_o)^2). \tag{19}$$

Then it is noticed that Eq. (19) is satisfied for $g\in\mathcal{G}$ by a linearized function $f(x)$ such that

$$f(x) = \frac{x}{\partial_g E_g[f(A)]|_{g=g_o}} + g_o - \frac{E_{g_o}[f(A)]}{\partial_g E_g[f(A)]|_{g=g_o}}. \tag{20}$$

By linearity of f, for the mean output $a=E_g[A]$ we have

$$f(a) = f(E_g[A]) = E_g[f(A)], \tag{21}$$

so that from Eq. (15) we have

$$\phi(\overline{a}) = f(\overline{a}) + O((g-g_o)^2). \tag{22}$$

Thus, $\overline{g}=\phi(\overline{a})$ is now reproduced by substituting the output $\overline{a}$ into the function $f(x)$. By use of Eqs. (6) and (19), the mean-square error of the estimate $\overline{g}$ as the output $f(\overline{a})$ of the $f(A)$ measurement is evaluated as

$$E_g^2[A] = V_g[f(A)] + O((g-g_o)^4). \tag{23}$$

From Eq. (20), the variance of $f(A)$ is evaluated as

$$V_g[f(A)] = \frac{V_g[A]}{(\partial_g E_g|_{g=g_o}[A])^2} = (\delta g_g[A])^2[1+O(g-g_o)]; \tag{24}$$

thus, it is verified due to Eqs. (23) and (24) that

$$E_g^2[A] = (\delta g_g[A])^2 + [1+O(g-g_o)]. \tag{25}$$

Hence the validity of the error evaluation by $(\delta g_g[A])^2$ definition has been shown.

It is a significant result from quantum estimation theory that even for the biased observable A, the quantum Cramér-Rao inequality can be proven:

$$(\delta g_g[A])^2 \geqslant \frac{1}{J_g}. \tag{26}$$

Here the Fisher information is defined by Eq. (12) and the equality can be achieved by taking $A\propto L(g)$ for each value of g. This can be shown by adopting not A but the local unbiased operator $f(A)$ and returning to the general argument for the unbiased case in Sec. II. It is also possible to prove by using the biased observable A straightforwardly. The proof can be seen in the Appendix.

Next we shall discuss physically relevant cases, the N-sample systems, by naturally extending the single-sample argument. Let us take a composite system which consists of N identical **S** subsystems. Assume here that the density operator is independent and identically distributed:

$$\rho_{tot}^{(N)}(g,t) = \rho(g,t)^{\otimes N}. \tag{27}$$

Now the average estimators $\overline{A^{(N)}}$ defined by

$$\overline{A^{(N)}} := \frac{1}{N}\sum^{N} \mathbf{1}\otimes\cdots\otimes A\otimes\mathbf{1}\cdots\otimes\mathbf{1} \tag{28}$$

are available. According to the quantum law of large numbers [4], the measurement data for $\overline{A^{(N)}}$ are going to be normally distributed with the average $E_g[A]$ and the standard deviation $(V_g[A]/N)^{1/2}$ when the number N becomes large. Since the expected error is solely a pull back of the quantum deviation of the observable, we can trust the estimate

$$g = \overline{g} \pm \delta g_{\overline{g}}[\overline{A^{(N)}}] \tag{29}$$

to 1σ precision for the large-number cases.

Now let us discuss the Fisher information for the N-sample cases. The SLD for the composite system $L^{(N)}$ is defined for general density operators $\rho_{tot}^{(N)}$ by

$$\partial_g \rho_{tot}^{(N)} = \frac{1}{2}[\rho_{tot}^{(N)} L^{(N)} + L^{(N)}\rho_{tot}^{(N)}]. \tag{30}$$

For independent and identically distributed (i.i.d.) density operators, it is easily derived that the composite SLD $L^{(N)}$ is given by

$$L^{(N)} = \sum^{N} \mathbf{1}\otimes\cdots\otimes\mathbf{1}\otimes L\otimes\mathbf{1}\otimes\cdots\otimes\mathbf{1}, \tag{31}$$

where $\mathbf{1}$ is the identity operation and L is the SLD for the subsystem defined by Eq. (9). This result yields the follow-

ing simple relation for the Fisher information of the composite system $J_g^{(N)}=\mathrm{Tr}[(L^{(N)})^2\rho^{(N)}]$:

$$J_g^{(N)} = NJ_g^{(1)}. \tag{32}$$

By virtue of the Cramér-Rao inequality, it is easily noticed that the optimized estimation for g in a single measurement of the composite system is achieved by adopting the average estimator $\overline{L^{(N)}}=(1/N)L^{(N)}$. The expected error is given by

$$\delta g_g[\overline{L^{(N)}}] = \frac{1}{\sqrt{NJ_g^{(1)}}}. \tag{33}$$

This coincides with the usual error of the estimation for g based upon N independent data of measurements of $L^{(1)}$ for N subsystems. However, stress that there is no need to measure N times the estimator $L^{(1)}$ for each subsystem **S** to achieve the estimate. In the quantum estimation, just one measurement of the single observable $\overline{L^{(N)}}$ yields the best estimate. Other relative-difference components like $L\otimes\mathbf{1}\cdots\otimes\mathbf{1}-\mathbf{1}\otimes L\cdots\otimes\mathbf{1}$ remain unmeasured. This saves effectively the number of processes in the estimation and exposes an advantage of the quantum estimation.

When the entanglement between the subsystems is available, it is possible [5] that the large-N behavior of $\delta g_g[\overline{L^{(N)}}]$ can be improved beyond the $1/\sqrt{N}$ factor as

$$\delta g_g[\overline{L^{(N)}}] \propto \frac{1}{N}. \tag{34}$$

IV. QUANTUM ESTIMATION BY LOCAL OBSERVABLES

We shall now consider the following constraints on the quantum estimation problem discussed above. In the above general formulation, we have assumed that every observable A of the system **S** is available for our measurement to fix the g value. However, in practice the available observables are restricted. For instance, it is a common situation in experiments that a particle described as the system **S** is contained inside the laboratory at the origin of the time and that we can only use measuring devices inside our laboratory. However, according to the time evolution the particle may go out of the experimental apparatus or our laboratory, so that for the general t, our ability of estimating the parameter g is restricted by the measuring devices inside the laboratory.

Let $\mathcal{M}$ be a subspace of $\mathcal{H}$. The projection of $\mathcal{H}$ onto $\mathcal{M}$ is denoted by P. In this paper, we consider the following two constraints.

(i) The initial state is supposed to be supported by P—i.e.,

$$\rho_{tot}(t=0) = P\rho_{tot}(t=0)P. \tag{35}$$

(ii) The available observables $\mathcal{O}$ for our measurements are restricted to those of the form

$$\mathcal{O} = PXP + y(I-P), \tag{36}$$

where X is an arbitrary observable on $\mathcal{H}$ and y is an arbitrary real number.

Let $\{|a\rangle\}$ be the orthonormal basis of $\mathcal{M}$ and $\{|a\rangle\}$ be the orthonormal basis of $\mathcal{H}$ extending $\{|a\rangle\}$—i.e., $\{|a\rangle\}\subseteq\{|a\rangle\}$. Then, condition Eq. (35) is equivalent to the relation

$$\rho_{tot}(t=0) = \sum_{a,a'} |a\rangle\langle a|\rho_{tot}(t=0)|a'\rangle\langle a'|. \tag{37}$$

Thus, the density operator initially has only matrix elements inside $\mathcal{M}$, and according to the time evolution, the density operator ρ_{tot} may have matrix elements outside of $\mathcal{M}$.

In the case of estimating the parameter g by observing a particle **S** initially localized in a box using the measuring devices effective only inside the box, the subspace $\mathcal{M}$ corresponds to the space of wave functions localized in the box. In this case, the assumption that the particle is initially localized inside the box is represented by condition (i). Since we assume that we know that the particle inside the box at the origin of the time, by measuring later, for instance, the weight of the box, we can measure I-P via a negative result of the measurement. Since the measuring devices are only effective inside the box, the measuring interaction couples only with the observable of the form $P\times P$ so that it is natural to assume that they can measure only observables of the form $P\times P$. Therefore, the set of available observables are considered to be restricted to those given by condition (ii).

Initially the density operator ρ_{tot} have only matrix elements in $\{|a\rangle\}$. However, in the course of the time evolution, ρ_{tot} can have matrix elements outside of $\{|a\rangle\}$. For any nonnegative time $t\geqslant 0$, we define the accessible density operator $\rho_{\|}(t)$ for the subspace $\mathcal{M}$ by

$$\rho_{\|}(t) = P\rho_{tot}(t)P. \tag{38}$$

Obviously, $\rho_{\|}$ has the matrix representation

$$\rho_{\|}(t) = [\langle a|\rho_{tot}(t)|b\rangle]. \tag{39}$$

Then, by the corresponding properties of $\rho_{tot}(t)$, the operator $\rho_{\|}(t)$ is positive and satisfies

$$0 \leqslant \mathrm{Tr}\rho_{\|}(t) \leqslant 1. \tag{40}$$

In what follows, we shall consider the time domain of t from $t=0$ to the time just before $t=t_*$ such that $\mathrm{Tr}\rho_{\|}(t=t_*)=0$, where we allow the case $t_*=\infty$, so that we have

$$0 < \mathrm{Tr}\rho_{\|}(t) \leqslant 1, \tag{41}$$

for $t\in[0,t_*)$.

From condition (ii), the available estimators A on $\mathcal{M}$ are naturally restricted and satisfy the relation

$$A = A_{\|} + a_{\perp}(\mathbf{1}-P) = A^{\dagger}, \tag{42}$$

where

$$A_{\|} = PA_{\|}P. \tag{43}$$

Using the definitions in Eqs. (38), (42), and (43), it can be shown that the expectation value of the available estimator A is given by

$$\langle A\rangle = \mathrm{Tr}[\rho_{tot}(t)A] = \mathrm{Tr}[\rho_{\|}(t)A_{\|}] + a_{\perp}[1-\mathrm{Tr}\rho_{\|}(t)]. \tag{44}$$

In order to define rigorously the notion of the "local" estimators $\tilde{A}$ corresponding to the available estimators A in

the restricted situation, let us extend the space $\mathcal{M}$ to a one-dimension-larger Hilbert space $\tilde{\mathcal{M}}$ by adding to the basis of $\mathcal{M}$ a normal vector $|B\rangle$ orthogonal to every $|a\rangle$— i.e., $\tilde{\mathcal{M}}:=\mathcal{M}\oplus\mathbf{C}|B\rangle$. Then, the local estimators $\tilde{A}$ acting on $\tilde{\mathcal{M}}$, which corresponds to the available estimator in Eq. (42), are defined by

$$\tilde{A}=A_{\|}+a_{\perp}|B\rangle\langle B|. \tag{45}$$

In particular, note that

$$\widetilde{1-P}=|B\rangle\langle B|. \tag{46}$$

Since the state $|B\rangle$ represents the inaccessible states by our local observation as seen in Eq. (46), we call $|B\rangle$ the blank state.

Further let us introduce the local density operator ρ acting on $\tilde{\mathcal{M}}$ and corresponding to $\rho_{\|}(=P\rho_{tot}P)$ by

$$\rho=\rho_{\|}+(1-\mathrm{Tr}\rho_{\|})|B\rangle\langle B|. \tag{47}$$

It is easily seen that ρ is positive and of unit trace. By a simple manipulation, the expectation value of the available estimator A in Eq. (44) can be reexpressed by use of the local estimator $\tilde{A}$ and the local density operator ρ as

$$\langle A\rangle=\mathrm{Tr}(\rho\tilde{A}). \tag{48}$$

V. CRAMÉR-RAO BOUND FOR LOCAL QUANTUM ESTIMATORS

In what follows, we shall consider the quantum Cramér-Rao inequality for the quantum estimation problem for the coupling constant g in the Hamiltonian $H_{tot}(g)$ by using *only* local measuring devices. By the time evolution, the local density operator $\rho=\rho(t,g)$ introduced in the previous section depends on the time t and the parameter g. Now we assume that one measures a local estimator $\tilde{A}$ at time t, and the output $\overline{a}$ determines the estimate $\overline{g_{\overline{a}}}[A]$ for g via the relation

$$E_{\overline{g_{\overline{a}}}[A]}[A]=\mathrm{Tr}\{\tilde{A}\rho(t,\overline{g_{\overline{a}}}[A])\}=\overline{a}.$$

We stress that, in this estimating process of g, we are allowed to use only local estimators $\tilde{A}$ instead of arbitrary observables in the theory.

The variance of the local observable $\tilde{A}$ for the correct g value is certainly given by

$$V_g[\tilde{A}]=\mathrm{Tr}[\tilde{A}^2\rho(t,g)]-\{\mathrm{Tr}[\tilde{A}\rho(t,g)]\}^2. \tag{49}$$

Then it is required at a given time t to find the minimum value of the expected error defined by

$$\delta g_g[\tilde{A}]:=\sqrt{\frac{V_g[\tilde{A}]}{(\partial_g E_g[\tilde{A}])^2}}. \tag{50}$$

It is shown that this problem is resolved by use of a solution of the problem on the estimate for the parameter g by arbitrary observables $\tilde{\mathcal{O}}$ on $\tilde{\mathcal{M}}$ as follows.

We define the local SLD $\tilde{L}(g)$ on $\tilde{\mathcal{M}}$ for an arbitrary local density operator $\rho(g)$ in Eq. (47) as a self-adjoint operator satisfying

$$\partial_g\rho(g)=\frac{1}{2}[\tilde{L}(g)\rho+\rho\tilde{L}(g)], \tag{51}$$

$$\tilde{L}(g)^{\dagger}=\tilde{L}(g). \tag{52}$$

Since $\mathrm{Tr}\,\rho(g)=1$ for any $g\in G$, we have

$$\mathrm{Tr}[\tilde{L}(g)\rho(g)]=0. \tag{53}$$

It is easy to construct a solution of Eq. (51) by introducing a SLD operator $L(g)$ on $\mathcal{M}$ for the accessible density operator $\rho_{\|}$. The SLD $L(g)$ on $\mathcal{M}$ is defined by

$$\partial_g\rho_{\|}=\frac{1}{2}[L(g)\rho_{\|}+\rho_{\|}L(g)], \tag{54}$$

$$L(g)^{\dagger}=L(g), \tag{55}$$

$$PL(g)P=L(g). \tag{56}$$

Due to the fact that $P\rho_{\|}P=\rho_{\|}$, we can find, at least, a solution of Eq. (54) for the SLD with $PL(g)P=L(g)$. Once the SLD $L(g)$ is given, then it is proved by a simple algebra that the operator defined by

$$\tilde{L}(g)=L(g)+\partial_g\ln[1-\mathrm{Tr}\rho_{\|}(g)]|B\rangle\langle B| \tag{57}$$

satisfies Eqs. (51) and (52); thus, it is a SLD on $\tilde{\mathcal{M}}$ for $\rho(g)$.

Here it is carefully noted that we may have

$$\mathrm{Tr}[L(g)\rho_{\|}(g)]\neq 0, \tag{58}$$

since the trace of $\rho_{\|}(g)$ is not necessarily normalized.

The operator $\tilde{L}(g)$ is determined uniquely up to the support of $\rho(g)$; any two solutions $\tilde{L}_1(g),\tilde{L}_2(g)$ satisfy $\tilde{L}_1(g)\rho(g)=\tilde{L}_2(g)\rho(g)$. The Fisher information J_g of the parameter g in $\rho(g)$ is uniquely defined by

$$J_g=\mathrm{Tr}[\tilde{L}(g)^2\rho(g)]=\mathrm{Tr}[L(g)^2\rho_{\|}(g)]+\frac{\{\mathrm{Tr}[L(g)\rho_{\|}(g)]\}^2}{1-\mathrm{Tr}[\rho_{\|}(g)]}, \tag{59}$$

where we have used Eqs. (47) and (57). Then, for the arbitrary observables $\tilde{\mathcal{O}}$ on $\tilde{\mathcal{M}}$ we have the quantum Cramér-Rao inequality

$$(\delta g_g[\tilde{\mathcal{O}}])^2:=\frac{V_g[\tilde{\mathcal{O}}]}{(\partial_g E_g[\tilde{\mathcal{O}}])^2}\geqslant\frac{1}{J_g}, \tag{60}$$

where J_g is given by Eq. (59). In order to apply the result to our local estimator problem, it is crucial to notice that the SLD in Eq. (57) takes the precise form of the local estimator on $\tilde{\mathcal{M}}$ in Eq. (45). Therefore, the equality can be attained by a local estimator. This indicates that the following quantum Cramér-Rao inequality for arbitrary local estimators $\tilde{A}$ on $\tilde{\mathcal{M}}$

really holds for the local density operators ρ corresponding to the accessible density operators $\rho_{\|}(g)$ $[=P\rho_{tot}(g)P]$:

$$(\delta g_g[\tilde{A}])^2 = \frac{V_g[\tilde{A}]}{(\partial_g E_g[\tilde{A}])^2} \geqslant \frac{1}{J_g}, \tag{61}$$

where the Fisher information J_g is given by Eq. (54). For a given $g \in G$, the equality is attained by a local estimator $\tilde{A}_o(g)$ such that

$$\tilde{A}_o(g) \propto \tilde{L}(g) = L(g) - \frac{\mathrm{Tr}[L(g)\rho_{\|}(g)]}{1-\mathrm{Tr}\rho_{\|}(g)}|B\rangle\langle B|. \tag{62}$$

Note that the local estimators which give the minimum expected error such as $\tilde{A}_o$ are unique only up to a factor and an additive term proportional to the identity operator. For instance, an estimator such that

$$\tilde{A}'_o(g) \propto L(g) + \frac{\mathrm{Tr}[L(g)\rho_{\|}(g)]}{1-\mathrm{Tr}\rho_{\|}(g)}P, \tag{63}$$

which has no matrix element for the blank state, also attains the equality.

VI. FISHER INFORMATION FOR UNNORMALIZED PURE STATES

In physics, it often happens that the measurement device is able to probe only a small part of the physical states of the total system. Even in such situations, nonunitary formulations are sometimes available. The state vectors $|\Psi(t)\rangle$ are governed by equations of motion with non-Hermitian Hamiltonians and evolve deterministically in the subspace $\mathcal{M}$, which is accessible by the experimental devices. Such examples are found in the various fields of physics including the scattering problems with weak absorption of quanta in the nuclear physics and the quantum optics, the flavor-oscillation studies in the elementary particle physics, and so on. The information about the coupling constant g in the equations of motion is imprinted on the state vectors $|\Psi(t,g)\rangle$ during the time evolution.

Let us evaluate the Fisher information for the pure state $|\Psi(t,g)\rangle$. The accessible density operator for the pure state reads

$$\rho_{\|}(t,g) = |\Psi(t,g)\rangle\langle\Psi(t,g)|, \tag{64}$$

where $\mathrm{Tr}[\rho_{\|}(0,g)]=1$ and at an advanced time $t\,(>0)$ the following relation holds:

$$0 < \mathrm{Tr}[\rho_{\|}(t,g)] \leqslant 1. \tag{65}$$

We define the SLD operator L on $\mathcal{M}$, in the same way discussed in the previous section, for the accessible density operator $\rho_{\|}$. Note that the operator L is not uniquely determined due to the purity of $\rho_{\|}$; however, the ambiguity is not relevant at all for the Fisher information, as commented on in the previous section. It is shown that we have a SLD

$$L = \frac{2}{\mathrm{Tr}[\rho_{\|}]}\partial_g\rho_{\|} - \frac{\mathrm{Tr}[\partial_g\rho_{\|}]}{(\mathrm{Tr}[\rho_{\|}])^2}\rho_{\|} \tag{66}$$

as a simple representative and the Fisher information itself is uniquely evaluated by

$$J = 4\left(\langle\partial_g\Psi|\partial_g\Psi\rangle - \frac{|\mathrm{Im}\langle\Psi|\partial_g\Psi\rangle|^2}{\langle\Psi|\Psi\rangle}\right) + 4\frac{|\mathrm{Re}\langle\Psi|\partial_g\Psi\rangle|^2}{1-\langle\Psi|\Psi\rangle}, \tag{67}$$

where $|\partial_g\Psi\rangle := \partial_g|\Psi(t,g)\rangle$. This result is an extension of that in Ref. [3], where the normalized pure-state theory is analyzed. The relation enables us to evaluate easily the Fisher information for many unnormalized pure state theories.

In Eq. (67), one may worry about the apparent divergence of the third term at $\langle\Psi|\Psi\rangle=1$, because the state evolves initially from the normalized state. However, for ordinary physical systems, the early behavior of the norm $\langle\Psi|\Psi\rangle$ is given by

$$\langle\Psi(t,g)|\Psi(t,g)\rangle \sim 1 - \alpha(g)t^2, \tag{68}$$

where α is a positive function of g. Thus the third term is evaluated in the early era as

$$4\frac{|\mathrm{Re}\langle\Psi|\partial_g\Psi\rangle|^2}{1-\langle\Psi|\Psi\rangle} \sim \frac{[\partial_g\alpha(g)]^2}{\alpha(g)}t^2. \tag{69}$$

Hence, the limit $t\to+0$ of Eq. (67) exists without any problems.

VII. PROBLEMS OF THE COMPOSITE SYSTEM

In the local estimation problem, some nontrivial aspects appear in the composite system analysis. Suppose a system **S**. Let us assume our measuring device for **S** is able to access only a subspace $\mathcal{M}$ of the Hilbert space of **S**. Later let P denote the projection operator onto $\mathcal{M}$ and $D_{\mathcal{M}}$ denote the dimension of the subspace $\mathcal{M}$. The accessible density operator on $\mathcal{M}$ is denoted by $\rho_{\|}$. The operator $\rho_{\|}(t,g)$ evolves in the subspace $\mathcal{M}$ and becomes dependent on the coupling constant g in the equation of motion. Let us consider a composite system $\mathbf{S}^{\otimes N}$ composed of N identical **S** subsystems. For instance, suppose that an independent and identically distributed (i.i.d.) initial condition is set for the total density operator $\rho_{tot}(0)$ of the composite system. Also assume that the unitary evolution of the total system is factorized—i.e., $U^{(N)}(t)=U(t)^{\otimes N}$. Even in such a simple situation, it can be pointed out that we have, at least, two natural alternatives for the estimation of g as follows.

The first alternative is rather simple. In the procedure, one first calculates the accessible density operator $\rho_{\|}^{(N)} = P^{\otimes N}\rho_{tot}P^{\otimes N}$ for $\mathbf{S}^{\otimes N}$ which is reduced to a direct product defined by

$$\rho_{\|}^{(N)}(t,g) := \rho_{\|}^{\otimes N}(t,g). \tag{70}$$

A local density operator for the accessible density operator $\rho_{\|}^{(N)}$ can be defined straightforwardly by

$$\rho_1^{(N)} := \rho_{\|}^{(N)} + (1 - \mathrm{Tr}\rho_{\|}^{(N)})|B\rangle\langle B|, \tag{71}$$

where $|B\rangle$ is the blank state. Let $j^{(N)}$ denote the Fisher information based upon the first local density operator $\rho_1^{(N)}$.

The estimation problem in the composite systems is nontrivial because we may have a construction of another local density operator for $\mathbf{S}^{\otimes N}$. We are able to define at first the local density operator for each subsystem $\mathbf{S}$. For the ith subsystem $\mathbf{S}_i$, the local density operator ρ_i corresponding to $\rho_{\|i}$ is written as

$$\rho_i = \rho_{\|i} + (1 - \mathrm{Tr}\rho_{\|i})|B_i\rangle\langle B_i|, \tag{72}$$

where $|B_i\rangle$ is the blank vector for the ith subsystem $\mathbf{S}_i$. Then we can define naturally the second local density operator $\rho_2^{(N)}$ for the composite system $\mathbf{S}^{\otimes N}$ by a direct product as follows:

$$\rho_2^{(N)} := \prod_{i=1}^{N} \otimes \rho_i. \tag{73}$$

Let $J^{(N)}$ denote the Fisher information based upon $\rho_2^{(N)}$. By construction, the Fisher information $J^{(N)}$ for the i.i.d. density operator is calculated as

$$J^{(N)} = NJ^{(1)}. \tag{74}$$

As seen above, there exist two independent Fisher informations for the composite system. Then, it is an important question: which alternative of the formulation gives us a more precise estimate for g, that is, which Fisher information, $j^{(N)}$ or $J^{(N)}$, is larger than another. The problem should be addressed for the general initial conditions for the density operators, beyond the above i.i.d. situations. Note first that the operators $\rho_1^{(N)}$ act on the Hilbert space $\mathcal{M}^{\otimes N} \oplus \mathbf{C}|B\rangle$ and the dimension of $\mathcal{M}^{\otimes N} \oplus \mathbf{C}|B\rangle$ is given by $D_1 = (D_{\mathcal{M}})^N + 1$. On the other hand, the operators $\rho_2^{(N)}$ act on the Hilbert space $\tilde{\mathcal{M}}^{\otimes N}$ and the dimension of $\tilde{\mathcal{M}}^{\otimes N}$ is given by $D_2 = (D_{\mathcal{M}} + 1)^N$. Since $D_2 > D_1$ always holds, it is naively expected that the second Fisher information $J^{(N)}$ is not less than the first Fisher information $j^{(N)}$. This guess can be proved affirmatively by use of the monotonicity argument for the Fisher information as will be mentioned later.

The above argument has been limited to the i.i.d. cases. In order to analyze the composite-system estimation generally, we must extend the above two formulations. Especially, nontrivial analyses are required to define $J^{(N)}$. These are formulated in Sec. VIII.

In Sec. IX, it is also pointed out that evaluation of the larger Fisher information $J^{(N)}$ requires solving time evolutions of various density operators corresponding to different initial conditions. Such a feature does not appear in the evaluation of both the standard Fisher information in the usual cases and the smaller Fisher information $j^{(N)}$ in the local estimation.

VIII. GENERAL FORMULATION FOR THE COMPOSITE SYSTEM

The available estimators for the composite system $\mathbf{S}^{\otimes N}$ now reads

$$A^{(N)} = \sum_{k_1 \cdots k_N} \omega_{k_1 \cdots k_N} A_{k_1} \otimes \cdots \otimes A_{k_N}, \tag{75}$$

which is just a natural extension of Eq. (42). Here A_k denote the available estimators for the subsystem $\mathbf{S}$, which take the form in Eq. (42) and $\omega_{k_1 \cdots k_N}$ are real coefficients. The corresponding extension of Eq. (45) is also possible. The local estimator $\tilde{A}^{(N)}$ corresponding to the available estimator $A^{(N)}$ is defined by

$$\tilde{A}^{(N)} = \sum_{k_1 \cdots k_N} \omega_{k_1 \cdots k_N} \tilde{A}_{k_1} \otimes \cdots \otimes \tilde{A}_{k_N}, \tag{76}$$

where $\tilde{A}_k$ are the local estimators corresponding to $\tilde{A}$ in Eq. (45).

In order to define the two Fisher informations $j^{(N)}$ and $J^{(N)}$ beyond the i.i.d. condition, let us consider the most general local density operator $\rho_{tot}^{(N)}(0) = P^{\otimes N}\rho_{tot}^{(N)}(0)P^{\otimes N}$ as the initial total density operator. In the unitary time evolution of the total system,

$$\rho_{tot}^{(N)}(t,g) = U^{(N)}(t,g)\rho_{tot}^{(N)}(0)U^{(N)\dagger}(t,g), \tag{77}$$

the total density operator becomes to have matrix elements between the inaccessible states.

Even for the general initial conditions, the definition of the first Fisher information $j^{(N)}$ is essentially unchanged. Let us introduce the accessible operators $\rho_{\|}^{(N)}$ by reducing the total density operator $\rho_{tot}^{(N)}$ as

$$\rho_{\|}^{(N)} = P^{\otimes N}\rho_{tot}^{(N)}P^{\otimes N}. \tag{78}$$

For the accessible density operator $\rho_{\|}^{(N)}$, a SLD operator $L^{(N)}$ is defined by

$$\partial_g \rho_{\|}^{(N)} = \frac{1}{2}[L^{(N)}\rho_{\|}^{(N)} + \rho_{\|}^{(N)}L^{(N)}], \tag{79}$$

$$(L^{(N)})^{\dagger} = L^{(N)}, \tag{80}$$

$$P^{\otimes N}L^{(N)}P^{\otimes N} = L^{(N)}. \tag{81}$$

According to Eq. (59), the Fisher information $j^{(N)}$ is defined straightforwardly as follows:

$$j^{(N)} := \mathrm{Tr}[(L^{(N)})^2\rho_{\|}^{(N)}] + \frac{(\mathrm{Tr}[L^{(N)}\rho_{\|}^{(N)}])^2}{1 - \mathrm{Tr}[\rho_{\|}^{(N)}]}. \tag{82}$$

Next, in order to define the second Fisher information $J^{(N)}$, what we want is a proper definition of a local density operator $\rho^{(N)}$ acting on $\tilde{\mathcal{M}}^{\otimes N}$ such that the total density operator $\rho_{tot}^{(N)}$ is reduced into $\rho^{(N)}$. Here, it is quite natural to impose the condition that expectation values of all the available observables $A^{(N)}$ for $\rho_{tot}^{(N)}$ be equivalent to those of the corresponding local observables $\tilde{A}^{(N)}$ for $\rho^{(N)}$:

$$\mathrm{Tr}[A^{(N)}\rho_{tot}] = \mathrm{Tr}[\tilde{A}^{(N)}\rho^{(N)}]. \tag{83}$$

By some manipulations it is soon noticed that the above constraint is really satisfied by defining the local density operator $\rho^{(N)}$ as follows. Let index α_j for $j = 1, \ldots, N$ below take index a_j for states in $\mathcal{M}$ or the index B for the blank state.

Then the matrix elements of $\rho^{(N)}$ on $\tilde{\mathcal{M}}^{\otimes N}$ are given by

$$\begin{aligned}\langle\alpha_1\alpha_2\cdots\alpha_N|\rho^{(N)}|\alpha_1'\alpha_2'\cdots\alpha_N'\rangle &= \prod_{j=1}^{N}[\delta_{\alpha_j B}\delta_{\alpha_j' B}+(1-\delta_{\alpha_j B})(1-\delta_{\alpha_j' B})] \\ &\times\sum_{x_1\cdots x_N}\prod_{k=1}^{N}[\delta_{\alpha_k B}\delta_{x_k 1}+(1-\delta_{\alpha_k B})\delta_{x_k 0}] \\ &\times\prod_{k'=1}^{N}[\delta_{\alpha_{k'}' B}\delta_{x_{k'} 1}+(1-\delta_{\alpha_{k'}' B})\delta_{x_{k'} 0}] \\ &\times\mathrm{Tr}[(P_{1,x_1}\otimes P_{2,x_2}\otimes\cdots\otimes P_{N,x_N})\rho_{tot}], \end{aligned} \quad (84)$$

where Tr stands for the trace operation on $\mathcal{H}^{\otimes n}$. For $m=1,\ldots,N$, the subscript x_m takes 0 or 1 and the operator P_{m,x_m} is defined by

$$P_{m,x_m}=|a_m'\rangle\langle a_m|, \quad \text{if} \quad x_m=0 \quad (85)$$

$$P_{m,x_m}=\mathbf{1}-P, \quad \text{if} \quad x_m=1. \quad (86)$$

By construction the Hermicity of the operator $\rho^{(N)}$ is trivial. Further, taking $A^{(N)}=\mathbf{1}^{\otimes N}$ in Eq. (83) yields the normalization condition

$$\mathrm{Tr}[\rho^{(N)}]=1. \quad (87)$$

The positivity of $\rho^{(N)}$ is also proved as follows. Suppose an arbitrary vector $|\Psi\rangle$ on $\tilde{\mathcal{M}}^{\otimes N}$:

$$|\Psi\rangle=\sum_{\alpha_1\cdots\alpha_N}C_{\alpha_1\cdots\alpha_N}|\alpha_1\cdots\alpha_N\rangle=\sum_{(i_1\cdots i_k)}|\Psi_{[i_1\cdots i_k]}\rangle, \quad (88)$$

where

$$|\Psi_{[\emptyset]}\rangle=\sum_{a_1\cdots a_N}C_{a_1\cdots a_N}|a_1\cdots a_N\rangle, \quad (89)$$

$$|\Psi_{[1]}\rangle=\sum_{a_2\cdots a_N}C_{Ba_2\cdots a_N}|Ba_2\cdots a_N\rangle, \quad (90)$$

and so on. Then, using the definition of $\rho^{(N)}$ in Eq. (84), the expectation values of $\rho^{(N)}$ for the arbitrary state vectors $|\Psi\rangle$ are evaluated as follows:

$$\begin{aligned}\langle\Psi|\rho^{(N)}|\Psi\rangle &= \sum_{i_1\cdots i_k}\langle\Psi_{[i_1\cdots i_k]}|\rho^{(N)}|\Psi_{[i_1\cdots i_k]}\rangle \\ &= \sum_{i_1\cdots i_k}\mathrm{Tr}[\tilde{P}_{[i_1\cdots i_k]}\rho^{(N)}] \\ &= \sum_{i_1\cdots i_k}\mathrm{Tr}[P_{[i_1\cdots i_k]}\rho_{tot}], \end{aligned} \quad (91)$$

where $\tilde{P}_{[i_1\cdots i_k]}=|\Psi_{[i_1\cdots i_k]}\rangle\langle\Psi_{[i_1\cdots i_k]}|$ and $P_{[i_1\cdots i_k]}$ are defined by replacing $|B\rangle\langle B|$'s in the operator $\tilde{P}_{[i_1\cdots i_k]}$ to $\mathbf{1}-P$. Noting that the operators $P_{[i_1\cdots i_k]}$ can be expressed as $P_{[i_1\cdots i_k]}=\sum_\beta|\Phi_{\beta,[i_1\cdots i_k]}\rangle\langle\Phi_{\beta,[i_1\cdots i_k]}|$ by use of the vectors $|\Phi_{\beta,[i_1\cdots i_k]}\rangle$ in the total Hilbert space, it is proved that

$$\langle\Psi|\rho^{(N)}|\Psi\rangle=\sum_{\beta,i_1,\cdots i_k}\langle\Phi_{\beta,[i_1\cdots i_k]}|\rho_{tot}^{(N)}|\Phi_{\beta,[i_1\cdots i_k]}\rangle\geqslant 0. \quad (92)$$

Taking account of the normalization condition in Eq. (87), this implies the positivity of the operator $\rho^{(N)}$.

Since $\mathrm{Tr}[\rho^{(N)}]=1$, we can define in the usual way a SLD operator $\mathcal{L}$ for the local density operator $\rho^{(N)}$:

$$\partial_g\rho^{(N)}=\frac{1}{2}(\mathcal{L}\rho^{(N)}+\rho^{(N)}\mathcal{L}). \quad (93)$$

Then, the Fisher information $J^{(N)}$ is defined by

$$J^{(N)}=\mathrm{Tr}[\rho^{(N)}\mathcal{L}^2]. \quad (94)$$

Now let us comment on the inequality $J^{(N)}\geqslant j^{(N)}$, using the monotonicity of the Fisher information. The point is that there exists a mapping R of the density operators defined in $\tilde{\mathcal{M}}^{\otimes N}$ onto the density operators defined in $\mathcal{M}^{\otimes N}\oplus\mathbf{C}|B\rangle$. Let us denote $P_\parallel$ a projection operator onto the subspace of vectors in $\tilde{\mathcal{M}}^{\otimes N}$ that do not include the blank states at all. Denote $P_\perp$ a projection operator onto the subspace of vectors that include more than one sub-blanck vectors $|B_i\rangle$. It should be noted that

$$P_\parallel\rho^{(N)}P_\parallel=\rho_\parallel^{(N)}. \quad (95)$$

Let us define the mapping R as follows.

$$\begin{aligned}R[\rho^{(N)}] &= P_\parallel\rho^{(N)}P_\parallel+\mathrm{Tr}[P_\perp\rho^{(N)}]|B\rangle\langle B| \\ &= \rho_\parallel^{(N)}+\mathrm{Tr}[P_\perp\rho^{(N)}]|B\rangle\langle B|. \end{aligned} \quad (96)$$

By definition, it is clear that the mapping is linear and of unit trace:

$$\mathrm{Tr}[R[\rho^{(N)}]]=\mathrm{Tr}[P_\parallel\rho^{(N)}]+\mathrm{Tr}[P_\perp\rho^{(N)}]=1. \quad (97)$$

It is also easily seen that this mapping is completely positive, since so are $\rho^{(N)}\mapsto P_\parallel\rho^{(N)}P_\parallel$ and $\rho^{(N)}\mapsto\mathrm{Tr}[P_\perp\rho^{(N)}]$.

Using the relation $\mathrm{Tr}[\rho^{(N)}]=1$, we obtain

$$\varrho:=R[\rho^{(N)}]=\rho_\parallel^{(N)}+(1-\mathrm{Tr}[\rho_\parallel^{(N)}])|B\rangle\langle B|. \quad (98)$$

Then, the first Fisher information $j^{(N)}$ is given by $\mathrm{Tr}[\varrho^{(N)}(\tilde{L})^2]$, where $\tilde{L}$ is the SLD operator corresponding to $\varrho^{(N)}$. According to the monotonicity theorem for the Fisher information [6], it must be satisfied under the projective mapping R that $J^{(N)}\geqslant j^{(N)}$. This result does not depend on whether the total density operators $\rho_{tot}^{(N)}$ are factorized or entangled.

It is worth noting that the information $J^{(N)}$ possesses a decomposition representation. Let us consider an arbitrary subsequence $(i_1,i_2,\ldots,i_n)$ of the sequence $(1,2,3,\ldots,N)$. Define that $\rho_{[i_1,i_2,\ldots,i_n]}$ is a $(\dim\mathcal{M})^{N-n}\times(\dim\mathcal{M})^{N-n}$ matrix which is composed of components of $\rho^{(N)}$ with $\alpha_{i_m}=\alpha_{i_m}'=B$ for $m=1,\ldots,n$. The following are examples:

$$\begin{aligned}&\langle a_2a_3\cdots a_N|\rho_{[1]}|a_2'a_3'\cdots a_N'\rangle: \\ &\quad=\langle Ba_2a_3\cdots a_N|\rho^{(N)}|Ba_2'a_3'\cdots a_N'\rangle, \end{aligned} \quad (99)$$

$$\begin{aligned}&\langle a_2a_4\cdots a_N|\rho_{[1,3]}|a_2'a_4'\cdots a_N'\rangle: \\ &\quad=\langle Ba_2Ba_4\cdots a_N|\rho^{(N)}|Ba_2'Ba_4'\cdots a_N'\rangle, \end{aligned} \quad (100)$$

$$\langle a_1|\rho_{[2,3,\ldots,N]}|a_1'\rangle := \langle a_1 BB\cdots B|\rho^{(N)}|a_1' BB\cdots B\rangle. \tag{101}$$

Note that the empty subsequence ø corresponds to the accessible density operator:

$$\langle a_1 a_2\cdots a_N|\rho_{[\emptyset]}|a_1' a_2'\cdots a_N'\rangle = \langle a_1 a_2\cdots a_N|\rho_{\|}^{(N)}|a_1' a_2'\cdots a_N'\rangle. \tag{102}$$

By definitions the local density operators $\rho_{[i_1,i_2,\ldots,i_n]}$ are non-negative—i.e., $\rho_{[i_1,i_2,\ldots,i_n]} \geqslant 0$. For each $\rho_{[i_1,i_2,\ldots,i_n]}$, we can introduce a partial SLD operator $\mathcal{L}_{[i_1,i_2,\ldots,i_n]}$ as

$$\partial_g \rho_{[i_1,i_2,\ldots,i_n]} = \frac{1}{2}(\mathcal{L}_{[i_1,i_2,\ldots,i_n]}\rho_{[i_1,i_2,\ldots,i_n]} + \rho_{[i_1,i_2,\ldots,i_n]}\mathcal{L}_{[i_1,i_2,\ldots,i_n]}). \tag{103}$$

Then it is possible to rewrite the second information as

$$J^{(N)} = \sum_{(i_1,i_2,\ldots,i_n)} J^{(N)}_{[i_1,i_2,\ldots,i_n]}, \tag{104}$$

$$J^{(N)}_{[i_1,i_2,\ldots,i_n]} = \mathrm{Tr}[\rho_{[i_1,i_2,\ldots,i_n]}\mathcal{L}^2_{[i_1,i_2,\ldots,i_n]}]. \tag{105}$$

Here $\Sigma_{(i_1,i_2,\ldots,i_n)}$ means the sum over all the subsequences $(i_1,i_2,\ldots,i_n)$ of $(1,2,3,\ldots,N)$, including the empty subsequence ø. The decomposition representation makes the evaluation of $J^{(N)}$ easier in many practical applications by using a useful formula for the operators $\rho_{[i_1,i_2,\ldots,i_n]}$ in the next section.

IX. EVALUATION OF THE LOCAL DENSITY OPERATOR

The accessible density operators $\rho_{\|}^{(N)}$ can be followed by our apparatus, since the operators $\rho_{\|}^{(N)}$ are completely local by definition. Meanwhile, the local density operator $\rho^{(N)}$ has been so far defined based upon the total density operator $\rho_{tot}^{(N)}$ in the previous section. We must say that the definition is too formal from the the practical viewpoint, because we seldom know global information about the total density $\rho_{tot}^{(N)}$ due to the limitation of our ability to measure the system. For the realistic evaluation of $J^{(N)}$, it is convenient to write down $\rho^{(N)}$ explicitly in terms of locally accessible quantities just as the operator $\rho_{\|}^{(N)}$. Such a reformulation can be realized for the cases with factorized evolutions—i.e., $U^{(N)}(t)=[U(t)]^{\otimes N}$—as follows. It should be emphasized that we do *not* need to assume the i.i.d. condition for the initial density operator.

Suppose that a composite system $\mathbf{S}^{\otimes N}$ of N identical $\mathbf{S}$ subsystems is governed by a unitary evolution and that the evolution is factorized for each subsystem $\mathbf{S}$:

$$\rho_{tot}^{(N)}(t) = [U(t)]^{\otimes N}\rho_{tot}^{(N)}(0)[U(t)^\dagger]^{\otimes N}, \tag{106}$$

where $U(t)$ is the unitary time evolution operator for $\mathbf{S}$ and $\rho_{tot}(0)$ is arbitrary initial density operator, which may have entanglement between the subsystems.

Let $\mathcal{O}_{\mathcal{M}} := \{e_a | e_a^\dagger = e_a, Pe_aP = e_a\}$ denote the complete basis of the available observables acting on $\mathcal{M}$ for each subsystem $\mathbf{S}$. Even in our local experiments, we are able to define and measure the projective evolutions for the available observables e_a. The projective evolutions are given by stochastic mappings $\Gamma(g,t)[e_a]$ which are defined by

$$\Gamma(g,t)[e_a] := PU(t)e_aU(t)^\dagger P. \tag{107}$$

In various physical systems, the dynamics is first given by not $\Gamma(g,t)$ but a Lindblad differential form given by

$$\partial_t \rho_{\|} = T_g[\rho_{\|}] \tag{108}$$

for arbitrary density operators $\rho_{\|}$ on $\mathcal{M}$. Here T_g is a time-independent Lindblad super-operator. Then the super operator T_g is related formally to the stochastic mapping $\Gamma(g,t)$ via

$$T_g = \partial_t \Gamma(g, t=0). \tag{109}$$

By integrating formally as $\Gamma(g,t)=e^{tT_g}$, we can recover the stochastic superoperators $\Gamma(g,t)$.

Stress that the operators $\Gamma(g,t)[e_a]$ are completely local quantities we can observe. Moreover, the projective evolutions for the composite available observables are also completely local quantities, which are written as

$$\Gamma(g,t)^{\otimes k}[e_{a_1}\otimes\cdots\otimes e_{a_k}] = P^{\otimes k}U(t)^{\otimes k}(e_{a_1}\otimes\cdots\otimes e_{a_k})(U(t)^{\otimes k})^\dagger P^{\otimes k}. \tag{110}$$

Our aim in this section is to express the operators $\rho_{[i_1,i_2,\cdots,i_n]}$ of the local density operator $\rho^{(N)}$ only in terms of the accessible operators like $\Gamma(g,t)^{\otimes k}[e_{a_1}\otimes\cdots\otimes e_{a_k}]$.

Since the initial density operator satisfies $\rho_{tot}^{(N)}(0) = P^{(N)}\rho_{tot}^{(N)}(0)P^{(N)}$, the operator $\rho_{tot}^{(N)}(0)$ can be expanded using the basis $\{e_a\}$:

$$\rho_{tot}^{(N)}(0) = \sum_{a_1}\sum_{a_2}\cdots\sum_{a_N} C_{a_1a_2\cdots a_N} e_{a_1}\otimes e_{a_2}\otimes\cdots\otimes e_{a_N}, \tag{111}$$

where the real coefficients $C_{a_1a_2\cdots a_N}$ is uniquely determined by $\rho_{tot}^{(N)}(0)$. After rather straightforward calculations, we argue that the following relation really holds:

$$\rho_{[i_1,i_2,\ldots,i_n]}(t) = \sum_{m=0}^{n}(-1)^{n-m}\sum_{(j_1,\ldots,j_m)\subseteq(i_1,\ldots,i_n)}\mathrm{Tr}_{(i_1,i_2,\ldots,i_n)-(j_1,\ldots,j_m)} \times[\Gamma^{\otimes(N-m)}_{[j_1,\ldots,j_m]}(g,t)[\mathrm{Tr}_{(j_1,\ldots,j_m)}[\rho_{tot}^{(N)}(0)]]]. \tag{112}$$

Here $\Sigma_{(j_1,\ldots,j_m)\subseteq(i_1,\ldots,i_n)}$ means the sum over all the subsequences $(j_1,\ldots,j_m)$ of the sequence $(i_1,\ldots,i_n)$. $\mathrm{Tr}_{(j_1,\ldots,j_m)}$ is a trace operation in terms of the $(j_1,\ldots,j_m)$ degrees of freedom. $\mathrm{Tr}_{(i_1,i_2,\ldots,i_n)-(j_1,\ldots,j_m)}$ means a trace operation in terms of the complementary subsequence to the subsequence $(j_1,\ldots,j_m)$ of $(i_1,\ldots,i_n)$. When $m=0$, $\mathrm{Tr}_{(j_1,\ldots,j_m)}$ is reduced into the identical operation. The operation $\Gamma^{\otimes(N-m)}_{[j_1,\ldots,j_m]}$ is the time-evolution operator for all degrees of freedom removing the $(j_1,\ldots,j_m)$ part. Here it is better to note again that even though the formula includes subtractions, all the operators $\rho_{[i_1,i_2,\ldots,i_n]}$ are non-negative $\rho_{[i_1,i_2,\ldots,i_n]} \geqslant 0$.

In Eq. (112), note that the operators $\rho_{[i_1,i_2,\ldots,i_n]}(t)$ cannot be evaluated only from the knowledge about the accessible density operator $\rho_{\|}^{(N)}(t)$. It is required to solve independently time evolutions of many descendant operators:

$$\Gamma^{\otimes(N-m)}_{[j_1,\ldots,j_m]}(g,t)[\mathrm{Tr}_{(j_1,\ldots,j_m)}[\rho^{(N)}_{tot}(0)]].$$

Here let us just draw the outline of the proof using the $\rho_{[1]}$ case. Substituting Eqs. (99) and (85), the following manipulation is possible:

$$\begin{aligned}\langle a_2 a_3 \cdots a_N|\rho_{[1]}|a_2' a_3' \cdots a_N'\rangle &= \langle Ba_2 a_3 \cdots a_N|\rho^{(N)}|Ba_2' a_3' \cdots a_N'\rangle \\ &= \langle a_2 a_3 \cdots a_N|\mathrm{Tr}_1[(\mathbf{1}-P) \\ &\quad \otimes P^{\otimes(N-1)}\rho^{(N)}_{tot}(\mathbf{1}-P) \\ &\quad \otimes P^{\otimes(N-1)}]|a_2' a_3' \cdots a_N'\rangle.\end{aligned}$$

Moreover, we can rewrite the equation as follows:

$$\rho_{[1]} = \mathrm{Tr}_1[\mathbf{1}\otimes P^{\otimes(N-1)}\rho^{(N)}_{tot}(t)\mathbf{1}\otimes P^{\otimes(N-1)}] - \mathrm{Tr}_1[\Gamma^{\otimes N}(g,t)[\rho^{(N)}_{tot} \times(0)]].$$

It is noticed that the first term of the right-hand-side is reduced using the expansion of $\rho^{(N)}_{tot}(0)$ as follows:

$$\begin{aligned}&\mathrm{Tr}_1[\mathbf{1}\otimes P^{\otimes(N-1)}\rho^{(N)}_{tot}(t)\mathbf{1}\otimes P^{\otimes(N-1)}] \\ &= \sum_{a_1\cdots a_N} C_{a_1 a_2\cdots a_N}\mathrm{Tr}_1[U(t)e_{a_1}U^\dagger(t)][PU(t)e_{a_2}U^\dagger(t)P]\otimes\cdots \\ &\quad \otimes [PU(t)e_{a_N}U^\dagger(t)P] \\ &= \Gamma^{\otimes(N-1)}(g,t)\left[\sum_{a_1\cdots a_N} C_{a_1 a_2\cdots a_N}\mathrm{Tr}_1[e_{a_1}]e_{a_2}\otimes\cdots\otimes e_{a_N}\right] \\ &= \Gamma^{\otimes(N-1)}(g,t)[\mathrm{Tr}_1[\rho^{(N)}_{tot}(0)]]. \qquad (113)\end{aligned}$$

Consequently we arrive at the relation for $\rho_{[1]}$ in Eq. (112):

$$\rho_{[1]} = \Gamma^{\otimes(N-1)}(g,t)[\mathrm{Tr}_1[\rho^{(N)}_{tot}(0)]] - \mathrm{Tr}_1[\Gamma^{\otimes N}(g,t)[\rho^{(N)}_{tot}(0)]]. \qquad (114)$$

The proofs for the other components in Eq. (112) can be achieved in the similar ways.

The relation in Eq. (112) makes the evaluation of $J^{(N)}$ possible, only based upon our local knowledge.

X. DECAYING TWO-LEVEL MODEL WITH A SMALL UNKNOWN PARAMETER

In order to demonstrate our formulation explicitly, let us consider a system including a small unknown parameter g. In many physical systems, the estimation of such a small parameter often provides significant physical information. For example, tiny coupling constants in the elementary particle interactions produce only quite rare processes; however, the analyses give a lot of important constraints of high-energy features beyond today's accelerator technology. For simplicity suppose a decaying two-level system including the small coupling g. The model has been frequently used in physics, for instance, to analyze the flavor-oscillating phenomena in the K_0-$\overline{K_0}$ system [7]. The Hamiltonian of the example is given as

$$H = -i\hbar\begin{bmatrix}\Gamma_+ & 0\\ 0 & \Gamma_-\end{bmatrix} + g\hbar\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \qquad (115)$$

where $\Gamma_\pm > 0$, $\Gamma_+ \neq \Gamma_-$ and $|g| \ll \Gamma_\pm, |\Gamma_+ - \Gamma_-|$. In the two-level subspace, time evolution of the density matrix $\rho_{\|}$ is governed by the following equation of motion:

$$i\hbar\partial_t\rho_{\|} = H\rho_{\|} - \rho_{\|}H^\dagger. \qquad (116)$$

Define the states $|\pm\rangle$ as

$$\sigma_z|\pm\rangle = \pm|\pm\rangle. \qquad (117)$$

Here σ_z is the z component of the Pauli matrix. Let us estimate the time evolution of $|\pm\rangle$ in the first order of g. It is solved as

$$|\pm(g,t)\rangle = e^{-\Gamma_\pm t}|\pm\rangle + igd(t)|\mp\rangle + O(g^2), \qquad (118)$$

where the function $d(t)$ is given as

$$d(t) = \frac{e^{-\Gamma_+ t} - e^{-\Gamma_- t}}{\Gamma_+ - \Gamma_-}. \qquad (119)$$

Then the information $J_\pm$ for the state $|\pm(t)\rangle$ is evaluated as

$$J_\pm(t) = 4d(t)^2 + O(g). \qquad (120)$$

The Cramér-Rao bound is always achieved by an observable

$$A(t_*) = \sigma_y + O(g), \qquad (121)$$

where σ_y is the y component of the Pauli matrix.

In this simple model we are able to optimize the measurement time t. The Fisher information takes its maximum value in the lowest order:

$$J_{max} = \frac{4}{(\Gamma_+ - \Gamma_-)^2}\left[\left(\frac{\Gamma_-}{\Gamma_+}\right)^{\Gamma_+/(\Gamma_+-\Gamma_-)} - \left(\frac{\Gamma_+}{\Gamma_-}\right)^{\Gamma_-/(\Gamma_--\Gamma_+)}\right]^2 + O(g), \qquad (122)$$

at

$$t_* = \frac{\ln\Gamma_+ - \ln\Gamma_-}{\Gamma_+ - \Gamma_-} + O(g). \qquad (123)$$

In Eq. (123), when Γ_- is much smaller than Γ_+:

$$\Gamma_- \ll \Gamma_+, \qquad (124)$$

$t_* \sim -(1/\Gamma_+)\ln(\Gamma_-/\Gamma_+)$ becomes larger logarithmically. In fact, the time t_* can be late until the first-order estimation breaks down, the time $t_g \sim -(1/\Gamma_+)\ln(g/\Gamma_+)$. Interestingly, at $t=t_*$, the survival probability for the state $|+(g,t)\rangle$ in the two-level subspace is estimated as

$$\langle +(g,t_*)|+(g,t_*)\rangle \sim \left(\frac{\Gamma_-}{\Gamma_+}\right)^2 \ll 1. \qquad (125)$$

Against naive expectations, this indicates that the best quantum estimation can be achieved at the time after the state has almost escaped from the two-level subspace.

Analysis of a composite system of the two subsystems may be also instructive. Let us first take the initial state as an i.i.d. state:

$$|\Psi(0)\rangle = |+\,+\rangle. \tag{126}$$

For the density matrix $\rho_{\|}^{(2)}(t)=|++(t)\rangle\langle++(t)|$, the Fisher information $j^{(2)}$ defined by

$$j^{(2)}: = \mathrm{Tr}[\rho_{\|}^{(2)}(L^{(2)})^2] + \frac{[\mathrm{Tr}[\rho_{\|}^{(2)}L^{(2)}]]^2}{1-\mathrm{Tr}\rho_{\|}^{(2)}} \tag{127}$$

is evaluated as

$$j^{(2)}(t) = 8d(t)^2 e^{-2\Gamma_+ t} + O(g). \tag{128}$$

Because $j^{(1)}=4d^2+O(g)$, the relation $j^{(2)}(t)=2j^{(1)}$ is not satisfied due to the $e^{-2\Gamma_+ t}$ factor and $j^{(2)}$ is exponentially smaller than $2j^{(1)}$ at $t>0$.

Let us compare the result with $J^{(2)}$ defined by Eq. (104). In this case each component of the local density operator is defined as

$$\rho_{[\emptyset]}(t) = \Gamma^{\otimes 2}(g,t)[\rho(0)] = \rho_{\|}^{(2)}(t), \tag{129}$$

$$\rho_{[1]}(t) = \Gamma(g,t)[\mathrm{Tr}_1[\rho(0)]] - \mathrm{Tr}_1[\Gamma^{\otimes 2}(g,t)[\rho(0)]], \tag{130}$$

$$\rho_{[2]}(t) = \Gamma(g,t)[\mathrm{Tr}_2[\rho(0)]] - \mathrm{Tr}_2[\Gamma^{\otimes 2}(g,t)[\rho(0)]], \tag{131}$$

$$\rho_{[1,2]}(t) = 1 - \mathrm{Tr}_2[\Gamma(g,t)[\mathrm{Tr}_1[\rho(0)]]] - \mathrm{Tr}_1[\Gamma(g,t) \times [\mathrm{Tr}_2[\rho(0)]]] + \mathrm{Tr}_{1,2}[\Gamma^{\otimes 2}(g,t)[\rho(0)]]. \tag{132}$$

As seen above, to calculate $\rho_{[1]}$, $\rho_{[2]}$ and $\rho_{[1,2]}$, we need the time evolution of the partial density matrices $\Gamma(g,t)[\mathrm{Tr}_1[\rho(0)]]$ and $\Gamma(g,t)[\mathrm{Tr}_2[\rho(0)]]$. It should be stressed that these evolutions cannot be obtained only from knowledge of $\rho_{\|}^{(2)}(t)$, for instance, by taking any traces for $\rho_{\|}^{(2)}$. They must be calculated independently by solving Eq. (116) for the initial density matrices $\mathrm{Tr}_1[\rho(0)]$ and $\mathrm{Tr}_2[\rho(0)]$. For the initial i.i.d. density matrix, each $J^{(2)}$ component is calculated as

$$J_{[\emptyset]}^{(2)} = j^{(2)} = 8d(t)^2 e^{-2\Gamma_+ t} + O(g), \tag{133}$$

$$J_{[1]}^{(2)} = 4d(t)^2(1-e^{-2\Gamma_+ t}) + O(g), \tag{134}$$

$$J_{[2]}^{(2)} = 4d(t)^2(1-e^{-2\Gamma_+ t}) + O(g), \tag{135}$$

$$J_{[1,2]}^{(2)} = O(g). \tag{136}$$

Thus the total information $J^{(2)}$ is precisely equal to twice $J^{(1)}$:

$$J^{(2)} = 8d(t)^2 + O(g) = 2J^{(1)}. \tag{137}$$

Next let us discuss an entangled case. Initially we take a state as

$$|\Phi(0)\rangle = \frac{1}{\sqrt{2}}[|+-\rangle + |-+\rangle]. \tag{138}$$

Calculation of $j^{(2)}(t)$ is easy and the results are as follows:

$$j^{(2)}(t) = 8d(t)^2[e^{-2\Gamma_+ t} + e^{-2\Gamma_- t}] + O(g). \tag{139}$$

Note that $j^{(2)}(t)/j^{(1)}(t)$ vanishes exponentially in time just as in the i.i.d. case.

Evaluation of $J^{(2)}(t)$ needs not only the density matrix

$$\rho_{\|}(t) = |\Phi(t)\rangle\langle\Phi(t)|, \tag{140}$$

but also another density matrix

$$\Gamma(g,t)[\mathrm{Tr}_1[\rho(0)]] = \Gamma(g,t)[\mathrm{Tr}_2[\rho(0)]] = \Gamma(g,t)\left[\frac{1}{2}\mathbf{1}\right]. \tag{141}$$

After some manipulations the form of $J^{(2)}(t)$ results in

$$J^{(2)}(t) = 8d(t)^2[e^{-2\Gamma_+ t} + e^{-2\Gamma_- t}] + 4d(t)^2[1+2e^{-(\Gamma_+ + \Gamma_-)t}]^2 \times \frac{[e^{-\Gamma_+ t} - e^{-\Gamma_- t}]^2}{e^{-2\Gamma_+ t}(1-e^{-2\Gamma_- t}) + e^{-2\Gamma_- t}(1-e^{-2\Gamma_+ t})} + O(g). \tag{142}$$

Note that at the early era ($t\sim 0$), both $j^{(2)}(t)$ and $J^{(2)}(t)$ have 4 times the information compared with the single system:

$$j^{(2)}(t\sim 0) \sim 4j^{(1)}, \tag{143}$$

$$J^{(2)}(t\sim 0) \sim 4J^{(1)}. \tag{144}$$

Thus the information is twice larger than the above i.i.d. case. Obviously this advantage arises due to the entanglement between subsystems.

For the entangled case, $J^{(2)}/J^{(1)}$ becomes smaller than the value of the i.i.d. case (equal to 2) in the late time. Hence, the i.i.d. density operator becomes more relevant than the entangled density operator for the estimation of g. In the limit of $t\to\infty$, the value of $J^{(2)}/J^{(1)}$ for the entangled case approaches to the single-system value:

$$\lim_{t\to\infty}\frac{J^{(2)}(t)}{J^{(1)}(t)} \sim 1. \tag{145}$$

This is due to contributions of the one-blank states ($|B\pm\rangle$ and $|\pm B\rangle$). Consequently, it can be said that the measurement should be at the early times in order to utilize enhancement of the Fisher information by the entanglement.

So far we have discussed only systems with small numbers of samples. For the practical estimation of the small parameter g, the many-sample estimation is inevitable beyond the above simple examples. For instance, the minimized expected error δg is given by

$$\delta g = \frac{1}{\sqrt{NJ^{(1)}}} \tag{146}$$

for the i.i.d. cases of N-sample systems. Then, in order to get a meaningful estimate, the number of the samples must be, at least, $O(1/(g^2 J^{(1)}))$ for the correct value g. It is expected that large entanglement between many samples may extremely improve the estimation for g and make the number of the samples enough for the estimation much smaller.

XI. SUMMARY

We have investigated deeply the local quantum estimation problem of an unknown parameter. The practical restriction of experimental observables takes place in various situations of the physical experiments. For a typical example, in particle physics we can probe only low-energy visible sectors of the whole system by our present devices. Such obstacles appear because of the limit of the present technology and so on. Moreover, observation of quantum phenomena, which happen only at quite small rates, often becomes the crucial target of experiments, which may derive some profound results of physics like, for instance, CPT violation [8]. In such situations, it is generally difficult to take a large number of data as one wishes, at least in the first stage of the experimental studies. Hence the local quantum estimation becomes really important when the experimental arrangements are designed, because the estimation theory provides among our available probes the optimized observable which quantum fluctuation is most suppressed in the estimation based upon a limited number of the data.

In spite of such relevance of the local quantum estimation, the problem has never been discussed in detail, as far as the authors know. In this paper, the detailed analysis and formulations based upon the Fisher informations have been completed. After a brief review on the standard quantum estimation theory, the local quantum estimator for the local estimation has been defined by Eq. (42). The notion of the local density operators was clearly introduced in Eq. (47), and the Cramér-Rao inequality in the local quantum estimate theory [Eq. (61)] has been proven by taking the local Fisher information defined by Eq. (59). The inequality is a fundamental tool in the theory and will play a significant role in the local estimation in various physical applications. In Sec. VI, the Fisher information for the unnormalized pure state was commented on. The formula in Eq. (67) is an extension of that derived by Fujiwara and Nagaoka, who discussed the Fisher information for normalized pure states. It is known that in many physical systems nonunitary theories of pure states also are available and that the validity is well verified by the experiments. In such systems with nonunitary evolution, Eq. (67) is quite useful to evaluate the Fisher information for an unknown parameter. In Sec. VII, it was pointed out that the local quantum estimation in the composite system has two independent formulations, using the i.i.d. cases. In Sec. VIII, two general formulations of the local quantum estimation for the composite system were proposed. For the composite system of N identical subsystems, we have two Fisher informations, $j^{(N)}$ and $J^{(N)}$. The information $j^{(N)}$ takes a simple form to define, but gives, in general, much smaller values than $J^{(N)}$. The theory of information $J^{(N)}$ can generate a more precise estimate for g, but has a pretty complicated form to deal with, compared to the $j^{(N)}$ case. In order to avoid the troublesome procedures in evaluation of $J^{(N)}$, we showed in Sec. IX the formula in Eq. (112), which makes the evaluation tractable. As seen in Eq. (112), calculation of the Fisher information $J^{(N)}$ requires solving evolutions of many descendant operators, $\Gamma^{\otimes(N-m)}_{[j_1,\ldots,j_m]}(g,t)[\mathrm{Tr}_{(j_1,\ldots,j_m)}[\rho^{(N)}_{tot}(0)]]$, independently of solving the accessible density operator $\rho^{(N)}_{\|}(t)$ itself. Such processes never appear in the ordinary quantum estimation theory, where the Fisher information can be evaluated by using only a time-evolved density operator. In Sec. X, we demonstrated explicitly our formulation of the local quantum estimation by applying to a decaying two-level system with a small unknown parameter.

We hope that the analysis in this paper enables the quantum estimation theory to take a more active part in the real experimental studies, which suffer from the restriction of available observables and the practical limitation of the number of the data.

ACKNOWLEDGMENTS

The authors thank M. Hayashi for helpful commens. M.O. is supported by the Strategic Information and Communications R&D Promotion Scheme of the MPHPT of Japan, by the CREST project of the JST, and by the Grant-in-Aid for Scientific Research of the JSPS.

APPENDIX

In this appendix, the Cramér-Rao inequality is proved. Let us write the triangular inequality relation as

$$\mathrm{Tr}(X^\dagger X)\mathrm{Tr}(Y^\dagger Y) \geqslant |\mathrm{Tr}(X^\dagger Y)|^2, \tag{A1}$$

where X and Y are arbitrary operators acting on the Hilbert space. Decomposing the operator $X^\dagger Y$ into the sum of the real and imaginary parts as

$$X^\dagger Y = \frac{1}{2}(X^\dagger Y + Y^\dagger X) + \frac{1}{2}(X^\dagger Y - Y^\dagger X), \tag{A2}$$

another inequality relation arises:

$$\begin{aligned}\mathrm{Tr}(X^\dagger X)\mathrm{Tr}(Y^\dagger Y) &\geqslant \frac{1}{4}|\mathrm{Tr}(X^\dagger Y + Y^\dagger X)|^2 + \frac{1}{4}|\mathrm{Tr}(X^\dagger Y - Y^\dagger X)|^2 \\ &\geqslant \frac{1}{4}|\mathrm{Tr}(X^\dagger Y + Y^\dagger X)|^2. \end{aligned}\tag{A3}$$

Here let us take

$$X = L(g)\sqrt{\rho_{tot}(g)}, \tag{A4}$$

$$Y = (A - E_g[A])\sqrt{\rho_{tot}(g)}. \tag{A5}$$

Then, from the inequality (A3), we can derive that

$$\begin{aligned}&\mathrm{Tr}[\rho_{tot}(g)L(g)^2]\mathrm{Tr}[\rho_{tot}(g)(A - E_g[A])^2] \\ &\quad\geqslant \frac{1}{4}|\mathrm{Tr}[\rho_{tot}(g)\{L(g)(A - E_g[A]) + (A - E_g[A])L(g)\}]|^2.\end{aligned}\tag{A6}$$

The right-hand-side term in the above inequality is able to be calculated using Eqs. (10), (9), and (3) successively as follows:

$$\frac{1}{4}|\mathrm{Tr}[\rho_{tot}(g)\{L(g)(A-E_g[A])+(A-E_g[A])L(g)\}]|^2$$

$$=\frac{1}{4}|\mathrm{Tr}[\rho_{tot}(g)[L(g)A+AL(g)]]|^2$$

$$=\frac{1}{4}|\mathrm{Tr}[A[\rho_{tot}(g)L(g)+L(g)\rho_{tot}(g)]]|^2$$

$$=\{\mathrm{Tr}[A\partial_g\rho_{tot}(g)]\}^2=(\partial_g E_g[A])^2. \quad \text{(A7)}$$

Consequently the relation (A6) implies the following inequality:

$$\frac{V_g[A]}{(\partial_g E_g[A])^2}\geqslant\frac{1}{J_g}; \quad \text{(A8)}$$

thus, the inequality (26) is proved. For the unbiased case with $E_g[A]=g$, the inequality (A8) is reduced to Eq. (13). The equality is trivially attained when $X\propto Y$ in Eq. (A3) and the relation $X\propto Y$ holds in Eqs. (A4) and (A5) when we set $A\propto L(g)$, because $E_g[L(g)]=0$.

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

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

[3] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).

[4] N. Giri and W. von Waldenfels, Z. Wahrscheinlichkeitstheor. Verwandte Geb. **42**, 129 (1978).

[5] V. Bužek, R. Derka, and S. Massar, Phys. Rev. Lett. **82**, 2207 (1999); M. Hayashi, IEICE Technical Report No. QIT 2003-61 (2003).

[6] E. A. Morozova and N. N. Chentsov, Itogi Nauki Tehnik **36**, 69 (1990); D. Petz, Linear Algebr. Appl. **244**, 81 (1996); D. Petz and C. Sudár, J. Math. Phys. **37**, 2662 (1996).

[7] For a review, see D. H. Perkins, *Introduction to High Energy Physics* (Addison-Wesley, Reading, MA, 1982).

[8] P. H. Eberhard (unpublished); W. C. Carithers, J. H. Christenson, P. H. Eberhard, D. R. Nygren, T. Modis, T. P. Pun, E. L. Schwartz, and H. Sticker, Phys. Rev. D **14**, 290 (1976); J. Ellis, J. S. Hagelin, D. V. Nanopoulos, and M. Srednicki, Nucl. Phys. B **241**, 381 (1984); P. Huet and M. E. Peskin, *ibid.* **434**, 3 (1995).