Uncertainty relation revisited from quantum estimation theory

Yu Watanabe,[1,*] Takahiro Sagawa,[2,3] and Masahito Ueda[1,4]

[1] *Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[2] *The Hakubi Center, The Kyoto University, Yoshida-Ushinomiya-cho, Sakyo-ku, Kyoto 606-8302, Japan*

[3] *Yukawa Institute for Theoretical Physics, The Kyoto University, Kitashirakawa Oiwake-cho, Kyoto 606-8502, Japan*

[4] *ERATO Macroscopic Quantum Control Project, JST, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

(Received 27 October 2010; revised manuscript received 2 February 2011; published 28 October 2011)

We use quantum estimation theory to formulate bounds of errors in quantum measurement for arbitrary quantum states and observables in a finite-dimensional Hilbert space. We prove that the measurement errors of two noncommuting observables satisfy Heisenberg-type uncertainty relation, find the achievable bound, and propose a strategy to achieve it.

DOI: 10.1103/PhysRevA.84.042121 PACS number(s): 03.65.Ta, 02.50.Tt, 03.65.Aa, 03.65.Fd

I. INTRODUCTION

Quantum theory features two types of uncertainty: indeterminacy of observables and complementarity of quantum measurements. The indeterminacy [1] reflects the inherent nature of a quantum system [2–4], whereas the complementarity [5] involves quantum measurement, and estimation of a quantum state from the measurement outcomes is essential [6–8]. However, how to optimize the measurement and estimation for a given quantum system has remained an outstanding issue. The purpose of this paper is to report the resolution to this problem.

The complementarity implies that we cannot simultaneously perform precise measurements of noncommuting observables. There must exist trade-off relations between the measurement errors of the noncommuting observables. Whereas a number of trade-off relations have been found, they are neither achievable for all quantum states and observables [8–12] nor applicable to all quantum systems [13–15]. Due to advances in controlling quantum states, it is now possible to implement a scheme that performs a projection measurement on a part of samples and another projection measurement on the rest [16–19]. However, the achievable bound of the measurement errors for such a scheme is yet to be identified.

In this paper, we report the following three results. First, we prove that for all measurements the measurement errors of noncommuting observables are bounded from below by their commutation relation. This implies that not only quantum fluctuations but also measurement errors are bounded by the same commutation relation. However, the bound cannot be achieved for all quantum states and observables. Second, we find the achievable bound for the measurements that perform a projection measurement with or without noise on a part of samples and another measurement on the rest. We propose a scheme of the experimental setup that achieves the bound. Third, we numerically show that the bound is satisfied for all measurements. Therefore, we conjecture that all measurements satisfy the proposed trade-off relation, and that the measurements that achieve the bound are optimal for obtaining information about two noncommuting observables.

The complementarity in quantum measurement has often been discussed in terms of the variance of the measurement outcomes [9,10]. However, the variance of the measurement outcomes per se does not necessarily give a quantitative error concerning the measurement. To quantify this error in the measurement, it is essential to invoke quantum estimation theory (see Fig. 1). The measurement error is quantified by the difference between the information obtained by the measurement and that of the precise measurement concerning the observable. The information content corresponding to quantum estimation theory is the Fisher information that gives precision of the estimated value calculated from measurement outcomes. However, it is challenging to find the achievable bound for the Fisher information. Several bounds of the uncertainty relation have been derived by using the Fisher information [8,11,12]. However, those bounds, in general, cannot be achieved. The crucial point of our successful finding of the achievable bound is that we express the relevant operators in terms of generators of the Lie algebra su(d), where d is the dimension concerning Hilbert space $\mathcal{H}$. This greatly facilitates the analysis of our results.

This paper is organized as follows. In Sec. II, we introduce quantum estimation theory and define the measurement error based on it. In Sec. III, we prove the trade-off relation between the measurement errors of two noncommuting observables for all quantum measurement. The obtained relation clarifies that the errors in the measurement cannot violate Heisenberg-type uncertainty relation. In Sec. IV, we derive the achievable bound of the product of two measurement errors. This bound is tighter than Heisenberg-type uncertainty relation and shown to be achieved for a certain set of measurements that can be experimentally implemented. In Sec. V, we show the numerical evidence that the bound proved in Sec. IV is satisfied by all measurements. In Sec. VI, we summarize the main results of this paper and discuss some future problems.

II. QUANTUM ESTIMATION THEORY OF ERROR IN MEASUREMENT

A. Definition of error in measurement

Given n independent and identically distributed (i.i.d.) unknown quantum states $\hat{\rho}$ on the d-dimensional Hilbert space

*watanabe@cat.phys.s.u-tokyo.ac.jp

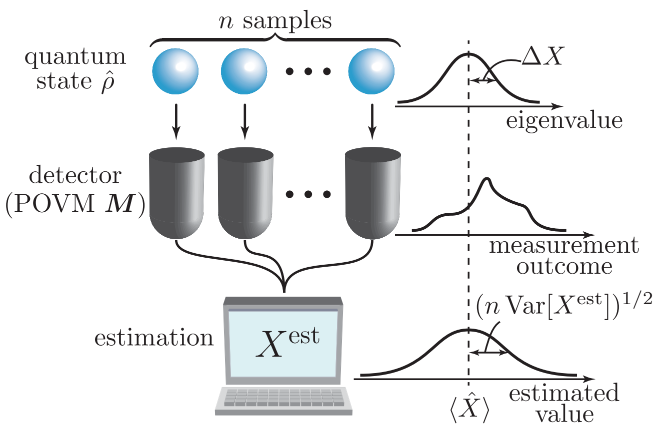

FIG. 1. (Color online) A measurement described by the positive operator valued measure (POVM) measurement $\boldsymbol{M}$ is performed to retrieve the information about the expectation value of an observable, $\langle \hat{X} \rangle$, for a quantum state $\hat{\rho}$. Since $\langle \hat{X} \rangle$ does not, in general, coincide with the value that is calculated from the distribution of the measurement outcomes due to noises introduced in the measurement process, it is necessary to estimate $\langle \hat{X} \rangle$ from the outcomes. The distribution of the estimated values is broader than the original one for $\hat{\rho}$ due to the measurement errors.

$\mathcal{H}$, we perform the same measurement on each of them. Here, we assume that the space of the possible quantum states is that of all density matrices on $\mathcal{H}$. If we want to know the expectation value $\langle \hat{X} \rangle := \mathrm{Tr}[\hat{\rho}\hat{X}]$ of a single observable $\hat{X}$, the optimal strategy (optimality is shown in Ref. [20]) is to perform the projection measurement $\boldsymbol{P} = \{\hat{P}_i\}_{i=1}^d$ corresponding to the spectral decomposition of $\hat{X} = \sum_i \alpha_i \hat{P}_i$ and then to calculate the estimated value of $\langle \hat{X} \rangle$ as

$$X^{\mathrm{est}}(\boldsymbol{n}) = \sum_{i=1}^{d} \alpha_i \frac{n_i}{n}, \tag{1}$$

where $\boldsymbol{n} = (n_1, \ldots, n_d)$ with n_i giving the number of times the outcome i is obtained ($\sum_i n_i = n$). The expectation value $\mathbb{E}[X^{\mathrm{est}}]$ of the estimator X^{est} is equivalent to $\langle \hat{X} \rangle$,

$$\mathbb{E}[X^{\mathrm{est}}] := \sum_{\boldsymbol{n}} X^{\mathrm{est}}(\boldsymbol{n}) p(\boldsymbol{n}) = \langle \hat{X} \rangle, \tag{2}$$

where the summation is taken over all non-negative integers $n_i \geqslant 0$ that satisfy $\sum_i n_i = n$, and

$$p(\boldsymbol{n}) = n! \prod_{i=1}^{d} \frac{p_i^{n_i}}{n_i!} \tag{3}$$

is the probability that the outcome i is obtained n_i times with $p_i = \mathrm{Tr}[\hat{\rho}\hat{P}_i]$. The estimators, such as (1), that satisfy (2) for all quantum states are called unbiased estimators. The variance of the estimator X^{est}, which quantifies the error in the whole measurements and the estimation process, is calculated to be

$$\mathrm{Var}[X^{\mathrm{est}}] := \mathbb{E}[(X^{\mathrm{est}})^2] - \mathbb{E}[X^{\mathrm{est}}]^2 = \frac{1}{n}(\Delta X)^2, \tag{4}$$

where

$$(\Delta X)^2 := \langle \hat{X}^2 \rangle - \langle \hat{X} \rangle^2 \tag{5}$$

characterizes the quantum fluctuation of observable $\hat{X}$. Since the variance decreases as n^{-1}, we can estimate $\langle \hat{X} \rangle$ from the measurement outcomes of $\boldsymbol{P}$ if n is sufficiently large.

When we perform the positive operator-valued measure (POVM) measurement $\boldsymbol{M} = \{\hat{M}_i\}_{i=1}^m$, the estimator cannot, in general, be written in the form of (1), and the variance is asymptotically greater than that of the optimal one:

$$\lim_{n\to\infty} n\mathrm{Var}[X^{\mathrm{est}}] \geqslant (\Delta X)^2, \tag{6}$$

where the left-hand side (LHS) and the right-hand side (RHS) show the variance of the concerned estimator and that of the optimal case per sample, respectively. The variance $\mathrm{Var}[X^{\mathrm{est}}]$ is caused by three different types of errors: quantum fluctuations, errors in the n identical measurements, and errors in the estimation process (see Fig. 1). The estimation error arises unless we use optimal estimators that minimize $\mathrm{Var}[X^{\mathrm{est}}]$ such as the maximum likelihood estimator. To quantify the error in single-shot measurements, it is necessary to use the estimator that minimizes the variance. We define the measurement error as

$$\varepsilon(\hat{X}; \boldsymbol{M}) := \min_{X^{\mathrm{est}}} \lim_{n\to\infty} n\mathrm{Var}[X^{\mathrm{est}}] - (\Delta X)^2, \tag{7}$$

where the minimization is taken over all consistent estimators [21] that asymptotically converge to $\langle \hat{X} \rangle$:

$$\lim_{n\to\infty} \mathrm{Prob}(|X^{\mathrm{est}} - \langle \hat{X} \rangle| < \delta) = 1 \tag{8}$$

for all quantum states $\hat{\rho}$ and arbitrary $\delta > 0$. The condition (8) implies that the POVM measurement does not involve any systematic error and we can quantify the measurement precisely. Examples of the consistent estimator include the average of eigenvalues (1) for the projection measurement, and the maximum likelihood estimator for the POVM measurement. These quantities also minimize $\lim_{n\to\infty} n\mathrm{Var}[X^{\mathrm{est}}]$. If there exists no consistent estimator of $\langle \hat{X} \rangle$, we define $\varepsilon(\hat{X}; \boldsymbol{M}) = +\infty$. Such a situation occurs, for example, when the projection measurement of an observable that does not commute with $\hat{X}$ is performed.

We note that the measurement error $\varepsilon(\hat{X}; \boldsymbol{M})$ is defined as the limit of n going to infinity. If n is finite and not sufficiently large, we cannot use the statistical analysis to evaluate the errors. Therefore, the definition (7) is meaningful for the case in which n is infinite or, at least, sufficiently large.

B. Measurement error in terms of Fisher information

To express $\varepsilon(\hat{X}; \boldsymbol{M})$ in terms of the Fisher information, we expand the Hermitian operators on the d-dimensional Hilbert space by the generators of the Lie algebra su(d). The generators $\hat{\boldsymbol{\lambda}} = \{\hat{\lambda}_i\}_{i=1}^{d^2-1}$ are traceless, Hermitian, and orthonormal with respect to the trace-norm:

$$\lambda_i^\dagger = \lambda_i, \quad \mathrm{Tr}[\lambda_i] = 0, \quad \mathrm{Tr}[\hat{\lambda}_i \hat{\lambda}_j] = \delta_{ij}. \tag{9}$$

In terms of them, an arbitrary quantum state $\hat{\rho}$ can be expanded as

$$\hat{\rho} = \frac{1}{d}\hat{I} + \boldsymbol{\theta} \cdot \hat{\boldsymbol{\lambda}} = \frac{1}{d}\hat{I} + \sum_{i=1}^{d^2-1} \theta_i \hat{\lambda}_i, \tag{10}$$

where $\hat{I}$ is the identity operator and $\boldsymbol{\theta} \in \mathbb{R}^{d^2-1}$ is a (d^2-1)-dimensional real vector. Since $\hat{\rho}$ is unknown, $\boldsymbol{\theta}$ is unknown, and the dimension of the space Θ of all possible parameters $\boldsymbol{\theta}$ is d^2-1. An arbitrary observable can also be expanded in terms of the same set of generators as

$$\hat{X} = x_0\hat{I} + \boldsymbol{x}\cdot\hat{\boldsymbol{\lambda}}. \tag{11}$$

Then, the expectation value can be evaluated as

$$\langle\hat{X}\rangle = x_0 + \boldsymbol{x}\cdot\boldsymbol{\theta}. \tag{12}$$

Therefore, estimating $\langle\hat{X}\rangle$ amounts to estimating $\boldsymbol{x}\cdot\boldsymbol{\theta}$. For any consistent estimator X^{est} of $\langle\hat{X}\rangle$, the variance $\mathrm{Var}[X^{\mathrm{est}}]$ satisfies the following Cramér-Rao inequality [22]:

$$\lim_{n\to\infty} n\mathrm{Var}[X^{\mathrm{est}}] \geqslant \begin{cases} \boldsymbol{x}\cdot[J(\boldsymbol{M})]^{-1}\boldsymbol{x}, & \boldsymbol{x}\in\mathrm{supp}[J(\boldsymbol{M})],\\ +\infty, & \text{otherwise,}\end{cases} \tag{13}$$

where $J(\boldsymbol{M})$ is the Fisher information matrix whose ij component is defined as

$$[J(\boldsymbol{M})]_{ij} := \sum_{k=1}^{m} p_k(\partial_i \ln p_k)(\partial_j \ln p_k), \tag{14}$$

where $\partial_i = \partial/\partial\theta_i$ and $p_k = \mathrm{Tr}[\hat{\rho}\hat{M}_k]$. In general, $J(\boldsymbol{M})$ does not have the inverse. The inverse in (13) is defined by the Moore-Penrose pseudoinverse [23,24]. In the rest of this paper, the inverses of nonsquare matrices and matrices that have zero eigenvalue are defined by the Moore-Penrose pseudoinverse. If the RHS of (13) is finite, there exists some estimator, for example, the maximum likelihood estimator, that achieves the equality of (13). If the RHS of (13) is infinite, there exists no consistent estimator of $\langle\hat{X}\rangle$. Therefore, the measurement error can be written as

$$\varepsilon(\hat{X};\boldsymbol{M}) = \begin{cases} \boldsymbol{x}\cdot J(\boldsymbol{M})^{-1}\boldsymbol{x} - (\Delta X)^2, & \boldsymbol{x}\in\mathrm{supp}[J(\boldsymbol{M})],\\ +\infty, & \text{otherwise.}\end{cases} \tag{15}$$

The matrix $J(\boldsymbol{M})$ varies with varying the POVM, but it is bounded from above by the quantum Cramér-Rao inequality [25]:

$$J(\boldsymbol{M}) \leqslant J_Q, \tag{16}$$

$$\Leftrightarrow \quad J(\boldsymbol{M})^{-1} \geqslant J_Q^{-1}, \tag{17}$$

where J_Q is the quantum Fisher information matrix [26], which is a monotone metric on the quantum state space with the coordinate system $\boldsymbol{\theta}$. The quantum Fisher information matrix is not uniquely determined, but from the monotonicity there exist the minimum J_S and the maximum J_R [27], where J_S (J_R) is the symmetric (right) logarithmic derivative Fisher information matrix. Their ij elements are defined as

$$[J_S]_{ij} := \tfrac{1}{2}\langle\{\hat{L}_i,\hat{L}_j\}\rangle, \tag{18}$$

$$[J_R]_{ij} := \langle\hat{L}_j^{\prime\dagger}\hat{L}_i^{\prime}\rangle, \tag{19}$$

where the curly brackets {,} denote the anticommutator, and a Hermitian $\hat{L}_i$ and a non-Hermitian $\hat{L}_i'$ are defined to be the solution to

$$\partial_i\hat{\rho} = \tfrac{1}{2}\{\hat{\rho},\hat{L}_i\}, \tag{20}$$

$$\partial_i\hat{\rho} = \hat{\rho}\hat{L}_i'. \tag{21}$$

It can be shown (see Appendix A) that

$$\left[J_S^{-1}\right]_{ij} = \mathcal{C}_s(\hat{\lambda}_i,\hat{\lambda}_j), \tag{22}$$

$$\left[J_R^{-1}\right]_{ij} = \mathcal{C}(\hat{\lambda}_i,\hat{\lambda}_j), \tag{23}$$

where

$$\mathcal{C}_s(\hat{X},\hat{Y}) := \tfrac{1}{2}\langle\{\hat{X},\hat{Y}\}\rangle - \langle\hat{X}\rangle\langle\hat{Y}\rangle, \tag{24}$$

$$\mathcal{C}(\hat{X},\hat{Y}) := \langle\hat{X}\hat{Y}\rangle - \langle\hat{X}\rangle\langle\hat{Y}\rangle \tag{25}$$

are the symmetrized and unsymmetrized correlation functions of two observables. From the classical and quantum Cramér-Rao inequalities and

$$\boldsymbol{x}\cdot J_S^{-1}\boldsymbol{x} = \boldsymbol{x}\cdot J_R^{-1}\boldsymbol{x} = (\Delta X)^2, \tag{26}$$

we find that the inequality

$$\varepsilon(\hat{X};\boldsymbol{M}) = \boldsymbol{x}\cdot\left[J(\boldsymbol{M})^{-1} - J_Q^{-1}\right]\boldsymbol{x} \geqslant 0 \tag{27}$$

is satisfied for any quantum Fisher information.

We note that if the density matrix does not have full rank, for example, if the state is pure, the Cramér-Rao inequality (13) does not hold and needs a correction for the unbiasedness of the estimator due to the semi-positivity of the quantum states [28]. The measurement error defined in (7) satisfies (15) only for those states whose density matrices have full rank. However, all non-full-rank states can be approximated by full-rank states with arbitrary precision, and all states that can be generated realistically are mixed. Therefore, in the following, we consider only the case in which the density matrix has full rank.

III. TRADE-OFF RELATIONS ON MEASUREMENT ERROR FOR AN ARBITRARY MEASUREMENT

The first result in this paper is the following theorem:

Theorem 1. For all observables $\hat{X}_1$, $\hat{X}_2$ and quantum states $\hat{\rho}$, an arbitrary POVM $\boldsymbol{M}$ satisfies

$$\varepsilon(\hat{X}_1;\boldsymbol{M})\varepsilon(\hat{X}_2;\boldsymbol{M}) \geqslant \tfrac{1}{4}|\langle[\hat{X}_1,\hat{X}_2]\rangle|^2, \tag{28}$$

where the square brackets [,] denote the commutator.

Proof. From the quantum Cramér-Rao inequality [25], we have

$$J(\boldsymbol{M})^{-1} - J_R^{-1} \geqslant 0. \tag{29}$$

The matrix $J(\boldsymbol{M})$ is real and symmetric, and J_R is Hermitian. Thus, the following inequality is satisfied for all observables $\hat{X}_\mu = x_{0,\mu}\hat{I} + \boldsymbol{x}_\mu\cdot\hat{\boldsymbol{\lambda}}$ and $k\in\mathbb{R}$:

$$(\boldsymbol{x}_1 + ik\boldsymbol{x}_2)^\dagger[J(\boldsymbol{M})^{-1} - J_R^{-1}](\boldsymbol{x}_1 + ik\boldsymbol{x}_2) \geqslant 0$$
$$\Rightarrow\ k^2\varepsilon(\hat{X}_2;\boldsymbol{M}) + k|\langle[\hat{X}_1,\hat{X}_2]\rangle| + \varepsilon(\hat{X}_1;\boldsymbol{M}) \geqslant 0. \tag{30}$$

It follows from this inequality that the discriminant of the quadratic polynomial on the LHS of (30) is always negative. Therefore, (28) is proved. ■

Heisenberg originally discussed the trade-off relation between the measurement error of an observable and the disturbance in another noncommuting observable caused by the measurement. From this argument, it can be expected that the trade-off relation between measurement errors exists. We

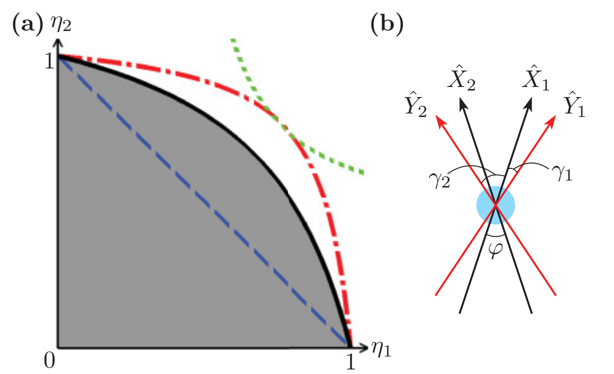

FIG. 2. (Color online) (a) Plots of measurement errors of 10^9 randomly chosen POVMs for $\dim\mathcal{H} = 4$ $(S = 3/2)$, $\hat{\rho} = \{\hat{I}/(2S+1) + |S\rangle\langle S|\}/2$, $\hat{X}_1 = \hat{S}_x$, and $\hat{X}_2 = (\sqrt{3}\hat{S}_x + \hat{S}_y)/2$, where $\eta_\mu := [\varepsilon(\hat{X}_\mu;\boldsymbol{M})/(\Delta X_\mu)^2 + 1]^{-1}$ with $0 \leqslant \eta_\mu \leqslant 1$. The red dash-dotted, blue dashed, black solid, and green dotted curves show the bounds set by (28), (34), (42), and the inequality obtained in Ref. [12], respectively. (b) The directions of spin observables $\hat{Y}_1$ and $\hat{Y}_2$ that achieve the bound set by (42) for two spin observables $\hat{X}_1$ and $\hat{X}_2$. The projection measurement of $\hat{Y}_\nu$ can be implemented, for example, by using cold atoms and a linearly polarized laser whose propagation direction is specified by γ_ν.

have proved this in the form of (28). Holevo proved a similar trade-off relation for position and momentum for the coherent state [8].

Equation (28) is satisfied for all quantum states and observables on any finite-dimensional Hilbert space. However, the equality in (28) cannot be achieved for all quantum states and observables [see the dash-dotted curve in Fig. 2(a)]. For example, for $\hat{\rho} = \hat{I}/d$,

$$\langle[\hat{X}_1,\hat{X}_2]\rangle = 0 \tag{31}$$

for all observables $\hat{X}_1$ and $\hat{X}_2$, and thus the RHS of (28) vanishes. For a POVM $\boldsymbol{M}$ to achieve the equality of (28), the POVM must satisfy

$$\varepsilon(\hat{X}_1;\boldsymbol{M}) = 0, \tag{32}$$
$$\varepsilon(\hat{X}_2;\boldsymbol{M}) < +\infty. \tag{33}$$

To satisfy (32), the POVM $\boldsymbol{M}$ must be the projection measurement of $\hat{X}_1$. However, in this case, the error for $\hat{X}_2$ diverges and the conditions (33) cannot be satisfied unless $\hat{X}_2$ commutes with $\hat{X}_1$.

IV. ATTAINABLE BOUND OF THE MEASUREMENT ERRORS

A simple but not optimal way to estimate $\langle\hat{X}_1\rangle$ and $\langle\hat{X}_2\rangle$ is to perform one projection measurement $\boldsymbol{P}_1$ on n_1 samples and another projection measurement $\boldsymbol{P}_2$ on $n_2 = n - n_1$ samples. This measurement scheme is asymptotically equivalent to the POVM measurement that randomly performs those two projection measurements with probabilities $q_\mu = n_\mu/n$ $(\nu = 1,2)$. If $\hat{X}_1$ and $\hat{X}_2$ cannot be simultaneously block-diagonalizable, the measurement errors satisfy

$$\varepsilon(\hat{X}_1;\boldsymbol{M})\varepsilon(\hat{X}_2;\boldsymbol{M}) = (\Delta X_1)^2(\Delta X_2)^2. \tag{34}$$

However, this measurement scheme does not exploit possible correlations between the observables $\hat{X}_1$ and $\hat{X}_2$. To utilize them, it is sufficient to perform projection measurements of two observables $\hat{Y}_\nu$ with probability q_ν $(\nu = 1,2)$ with $q_1 + q_2 = 1$, where $\hat{Y}_1$ and $\hat{Y}_2$ are linear combinations of $\hat{X}_1$ and $\hat{X}_2$. Therefore, we consider the following classes of POVM.

First, we define a set of projection measurements $\mathcal{P}_{\hat{X}_1,\hat{X}_2}$ as that of all projection measurements corresponding to the spectral decompositions of the observables that are linear combinations of $\hat{X}_1$ and $\hat{X}_2$:

$$\mathcal{P}_{\hat{X}_1,\hat{X}_2} = \left\{\boldsymbol{P} = \{\hat{P}_i\}_{i=1}^d \,\middle|\, {}^\exists a_1,a_2,\beta_i \in \mathbb{R},\ a_1\hat{X}_1 + a_2\hat{X}_2 = \sum_i \beta_i\hat{P}_i\right\}. \tag{35}$$

The set of the measurement schemes that probabilistically perform projection measurements with $\mathcal{P}_{\hat{X}_1,\hat{X}_2}$ is defined as

$$\mathcal{M}_{\rm random} := \{q_1\boldsymbol{P}_1 + q_2\boldsymbol{P}_2 \mid \boldsymbol{P}_1,\boldsymbol{P}_2 \in \mathcal{P}_{\hat{X}_1,\hat{X}_2},\ q_1,q_2 \geqslant 0,\ q_1 + q_2 = 1\}, \tag{36}$$

where

$$q_1\boldsymbol{P}_1 + q_2\boldsymbol{P}_2 = \{q_\nu\hat{P}_{\nu,i}\}_{\nu=1,2}^{i=1,\ldots,d} \tag{37}$$

for $\boldsymbol{P}_\nu = \{\hat{P}_{\nu,i}\}_{i=1}^d$.

In real experimental setups, measurements always suffer from noises which deteriorate the precision of projection measurement $\boldsymbol{P}$. Such a noisy measurement can be expressed as

$$\boldsymbol{M} = F\boldsymbol{P} = \left\{\sum_j F_{ij}\hat{P}_j\right\}_i, \tag{38}$$

where F is the so-called information processing matrix or probability transition matrix whose elements satisfy

$$F_{ij} \geqslant 0, \quad \sum_i F_{ij} = 1. \tag{39}$$

The measurements described by $F\boldsymbol{P}$ cover a broad class of experimentally realizable measurements. For example, a typical scheme of quantum nondemolition (QND) measurement belongs to this class [16–18]. We note that the noise of a measurement in this class is described by a classical noise that is characterized by a classical noisy channel with F_{ij}. We define a set of measurements,

$$\mathcal{M}_{\rm noisy} := \left\{F\boldsymbol{M} \,\middle|\, \boldsymbol{M} \in \mathcal{M}_{\rm random}, F_{ij} \geqslant 0, \sum_i F_{ij} = 1\right\}, \tag{40}$$

which include random measurements consisting of noisy projection measurements. Note that the classes of measurements described above satisfy

$$\mathcal{M}_{\rm random} \subset \mathcal{M}_{\rm noisy} \subset \mathcal{M}_{\rm all}, \tag{41}$$

where $\mathcal{M}_{\rm all}$ denotes the totality of POVM measurements.

The second main result in this paper is the following theorem:

Theorem 2. For all observables $\hat{X}_1$, $\hat{X}_2$ and quantum states $\hat{\rho}$, an arbitrary POVM $\boldsymbol{M} \in \mathcal{M}_{\text{noisy}}$ satisfies

$$\varepsilon(\hat{X}_1;\boldsymbol{M})\varepsilon(\hat{X}_2;\boldsymbol{M}) \geqslant (\Delta_Q X_1)^2(\Delta_Q X_2)^2 - [\mathcal{C}_Q(\hat{X}_1,\hat{X}_2)]^2. \quad (42)$$

Moreover, the measurements that achieve the equality of (42) exist for all quantum states and observables.

Here Δ_Q and $\mathcal{C}_Q$ are defined as follows. Let $\mathcal{H}_a$ ($a = 1,2,\ldots$) be the simultaneous irreducible invariant subspace of $\hat{X}_1$ and $\hat{X}_2$ ($\mathcal{H} = \oplus_a \mathcal{H}_a$), and $\hat{P}_a$ the projection operator on $\mathcal{H}_a$. We define the probability distribution as $p_a := \langle \hat{P}_a \rangle$ and the post-measurement state of the projection measurement $\{\hat{P}_1,\hat{P}_2,\ldots\}$ as

$$\hat{\rho}_a := \frac{1}{p_a}\hat{P}_a\hat{\rho}\hat{P}_a. \quad (43)$$

Then, Δ_Q and $\mathcal{C}_Q$ are defined as

$$(\Delta_Q X_\mu)^2 := \sum_a p_a \left(\langle \hat{X}_\mu^2 \rangle_a - \langle \hat{X}_\mu \rangle_a^2 \right), \quad (44)$$

$$\mathcal{C}_Q(\hat{X}_1,\hat{X}_2) := \sum_a p_a \left(\frac{1}{2}\langle\{\hat{X}_1,\hat{X}_2\}\rangle_a - \langle \hat{X}_1\rangle_a \langle \hat{X}_2 \rangle_a \right), \quad (45)$$

where $\langle \hat{X} \rangle_a = \mathrm{Tr}[\hat{\rho}_a \hat{X}]$. If $\hat{X}_1$ and $\hat{X}_2$ are simultaneously block-diagonalizable, then quantum fluctuations and correlations of observables are determined by the diagonal blocks of $\hat{\rho}$. (Note that $\langle \hat{X}_\mu \rangle$ is independent of the off-diagonal blocks of $\hat{\rho}$.) If two observables commute with each other, the RHS of (42) vanishes.

For qubits ($\dim \mathcal{H} = 2$), (42) can be proven for all POVM measurements, as stated in the following theorem:

Theorem 3. For all quantum states $\hat{\rho}$ and observables $\hat{X}_\mu$ on the two-dimensional Hilbert space, (42) is satisfied for all POVMs $\boldsymbol{M} \in \mathcal{M}_{\text{all}}$.

Inequality (42) is stronger than (28) and the trade-off relations obtained by Nagaoka [12] [see Fig. 2(a)], and it reduces to the trade-off relation found in Ref. [13] for $\dim \mathcal{H} = 2$ and $\hat{\rho} = \hat{I}/2$. The optimal measurement of Englert's complementarity [14] for $\dim \mathcal{H} = 2$ achieves the bound set by (42).

We emphasize that the bound set by (42) can be achieved for all quantum states and observables, whereas the bound set by (28) cannot. For example, for

$$\hat{\rho} = \frac{r}{(2S+1)}\hat{I} + (1-r)|S\rangle\langle S|, \quad (46)$$

$$\hat{X}_1 = \hat{S}_x, \quad \hat{X}_2 = \hat{S}_x \cos\varphi + \hat{S}_y \sin\varphi, \quad (47)$$

$$q_1 = q_2 = 1/2, \quad (48)$$

the measured observable

$$\hat{Y}_\nu = \hat{S}_x \cos\gamma_\nu + \hat{S}_y \sin\gamma_\nu \quad (49)$$

is determined by the solution to

$$\cos\varphi + \cos\varphi\cos^2(\gamma_1-\gamma_2) - 2\cos(\gamma_1+\gamma_2-\varphi) \times \cos(\gamma_1-\gamma_2) = 0, \quad (50)$$

where $\hat{S}_i$ is the spin operator of total spin S in the i $(= x,y,z)$ direction, and $|S\rangle$ is the eigenstate of $\hat{S}_z$ with eigenvalue S. The RHS of (28) and that of (42) are given by $[\frac{1}{2}(1-r)S\sin\varphi]^2$ and $[rS(2S-1)/6 + S/2]^2\sin^2\varphi$, respectively. Such an optimal measurement can be implemented, for example, by using cold atoms [16–19]. By letting an ensemble of atoms interact with a linearly polarized off-resonant laser whose propagation direction is parallel to the direction specified by γ_ν in $\hat{Y}_\nu$ [see Fig. 2(b)], the angle of a paramagnetic Faraday rotation of the laser polarization carries information about $\langle \hat{Y}_\nu \rangle$. The rotation angle can be detected by a polarimeter using a polarization-dependent beam splitter. If the intensity of the laser is sufficiently strong, this scheme achieves the projection measurement of $\hat{Y}_\nu$.

In the following, we prove Theorems 2 and 3:

Proof of Theorem 2. If two POVMs satisfy $\boldsymbol{M}' = F\boldsymbol{M}$ with an information processing matrix F, they satisfy

$$J(\boldsymbol{M}') \leqslant J(\boldsymbol{M}). \quad (51)$$

Hence, we have only to consider the case in which $\boldsymbol{M} \in \mathcal{M}_{\text{random}}$.

Let $\hat{X}_\mu = x_{\mu,0}\hat{I} + \boldsymbol{x}_\mu \cdot \hat{\boldsymbol{\lambda}}$ be a linear combination of $\hat{Y}_\nu = y_{\nu,0}\hat{I} + \boldsymbol{y}_\nu \cdot \hat{\boldsymbol{\lambda}}$ ($\nu = 1,2$), and $A = (a_{\mu\nu})$ be its coefficient:

$$\hat{X}_\mu = \sum_\nu a_{\mu\nu}\hat{Y}_\nu. \quad (52)$$

We consider the POVM measurement $\boldsymbol{M} = q_1\boldsymbol{P}_1 + q_2\boldsymbol{P}_2 \in \mathcal{M}_{\text{random}}$, where $\boldsymbol{P}_\nu = \{\hat{P}_{\nu,i}\}$ corresponds to the spectral decompositions of the observables $\hat{Y}_\nu = \sum_i \beta_{\nu,i}\hat{P}_{\nu,i}$.

As shown in Appendix A, the inverse of $J(\boldsymbol{M})$ can be obtained as

$$\boldsymbol{y}_\nu \cdot J(\boldsymbol{M})^{-1}\boldsymbol{y}_\nu = (\Delta Y_\nu)^2 + (q_\nu^{-1}-1)(\Delta_Q Y_\nu)^2, \quad (53)$$

$$\boldsymbol{y}_1 \cdot J(\boldsymbol{M})^{-1}\boldsymbol{y}_2 = \mathcal{C}_s(\hat{Y}_1,\hat{Y}_2) - \mathcal{C}_Q(\hat{Y}_1,\hat{Y}_2). \quad (54)$$

Let

$$\tilde{J}(\boldsymbol{M}) := [R^{\mathrm{T}}J(\boldsymbol{M})^{-1}R]^{-1}, \quad (55)$$

$$\tilde{J}_S := \left[R^{\mathrm{T}}J_S^{-1}R\right]^{-1}, \quad (56)$$

$$\tilde{\varepsilon}(\boldsymbol{M}) := \tilde{J}(\boldsymbol{M})^{-1} - \tilde{J}_S^{-1} \quad (57)$$

be 2×2 matrices, where

$$R := (\, \boldsymbol{x}_1 \;\; \boldsymbol{x}_2 \,) \quad (58)$$

is a $(d^2-1)\times 2$ matrix. From (52), (53), and (54), we obtain

$$\tilde{J}_S = \begin{pmatrix} (\Delta X_1)^2 & \mathcal{C}_s(\hat{X}_1,\hat{X}_2) \\ \mathcal{C}(\hat{X}_1,\hat{X}_2) & (\Delta X_2)^2 \end{pmatrix} = A\begin{pmatrix} (\Delta Y_1)^2 & \mathcal{C}_s(\hat{Y}_1,\hat{Y}_2) \\ \mathcal{C}(\hat{Y}_1,\hat{Y}_2) & (\Delta Y_2)^2 \end{pmatrix}A^{\mathrm{T}}, \quad (59)$$

$$\tilde{\varepsilon}(\boldsymbol{M}) = A\begin{pmatrix} \frac{q_2}{q_1}(\Delta_Q Y_1)^2 & -\mathcal{C}_Q(\hat{Y}_1,\hat{Y}_2) \\ -\mathcal{C}_Q(\hat{Y}_1,\hat{Y}_2) & \frac{q_1}{q_2}(\Delta_Q Y_2)^2 \end{pmatrix}A^{\mathrm{T}}. \quad (60)$$

The measurement error of the observable $\hat{X}_\mu$ can be written as

$$\varepsilon(\hat{X}_\mu;\boldsymbol{M}) = [\tilde{\varepsilon}(\boldsymbol{M})]_{\mu\mu}. \quad (61)$$

Because $\tilde{\varepsilon}(\boldsymbol{M})$ is symmetric, we have

$$\begin{aligned}\varepsilon(\hat{X}_1;\boldsymbol{M})\varepsilon(\hat{X}_2;\boldsymbol{M}) &\geqslant \det[\tilde{\varepsilon}(\boldsymbol{M})] \\ &= \det\begin{pmatrix} (\Delta_Q Y_1)^2 & \mathcal{C}_Q(\hat{Y}_1,\hat{Y}_2) \\ \mathcal{C}_Q(\hat{Y}_1,\hat{Y}_2) & (\Delta_Q Y_2)^2 \end{pmatrix}(\det A)^2 \\ &= (\Delta_Q X_1)^2(\Delta_Q X_2)^2 - [\mathcal{C}_Q(\hat{X}_1,\hat{X}_2)]^2.\end{aligned} \quad (62)$$

The condition for the equality to hold is that the off-diagonal elements of $\tilde{\varepsilon}(\boldsymbol{M})$ vanish. The observables $\hat{Y}_\nu$ that satisfy this condition exist for all $\hat{\rho}$. ■

Next we prove (42) for the two-dimensional Hilbert space. We first prove the following two lemmas:

Lemma 1 For all POVM $\boldsymbol{M} \in \mathcal{M}_{\text{all}}$,

$$\mathrm{Tr}\big[J(\boldsymbol{M})J_S^{-1}\big] \leqslant d-1 \tag{63}$$

is satisfied. This lemma was also shown in Ref. [29].

Proof. Let the spectral decomposition of each element of $\boldsymbol{M} = \{\hat{M}_i\}$ be

$$\hat{M}_i = \sum_j k_{ij}|\psi_{ij}\rangle\langle\psi_{ij}|, \tag{64}$$

and we define an associated POVM,

$$\boldsymbol{N} = \{\hat{N}_{ij} = k_{ij}|\psi_{ij}\rangle\langle\psi_{ij}|\}_{i,j}. \tag{65}$$

From the fact that there always exists an information processing matrix F such that $\boldsymbol{M} = F\boldsymbol{N}$,

$$J(\boldsymbol{M}) \leqslant J(\boldsymbol{N}) \tag{66}$$

is satisfied. By denoting

$$\hat{N}_{ij} = \frac{k_{ij}}{d}\hat{I} + \boldsymbol{v}_{ij}\cdot\hat{\boldsymbol{\lambda}}, \quad p_{ij} = \langle\hat{N}_{ij}\rangle, \tag{67}$$

from the facts that

$$\hat{N}_{ij}^2 = k_{ij}\hat{N}_{ij}, \quad \sum_{ij} k_{ij} = d, \tag{68}$$

we obtain

$$\begin{aligned}\mathrm{Tr}\big[J(\boldsymbol{M})J_S^{-1}\big] &\leqslant \mathrm{Tr}\big[J(\boldsymbol{N})J_S^{-1}\big] = \sum_{ij} p_{ij}^{-1}\boldsymbol{v}_{ij}\cdot J_S^{-1}\boldsymbol{v}_{ij}\\ &= \sum_{ij}\frac{(\Delta\hat{N}_{ij})^2}{\langle\hat{N}_{ij}\rangle} = \sum_{ij}\frac{k_{ij}\langle\hat{N}_{ij}\rangle - \langle\hat{N}_{ij}\rangle^2}{\langle\hat{N}_{ij}\rangle}\\ &= \sum_{ij}(k_{ij} - \langle\hat{N}_{ij}\rangle) = d-1.\end{aligned}$$

■

Lemma 2. Let $K := \tilde{J}_S^{-1/2}\tilde{J}(\boldsymbol{M})\tilde{J}_S^{-1/2}$. For the two-dimensional Hilbert space ($d=2$), the following inequalities hold:

$$\mathrm{Tr}[K] \leqslant 1 \quad \Leftrightarrow \quad \det[K^{-1} - I] \geqslant 1. \tag{69}$$

Proof. Because

$$P := J(\boldsymbol{M})^{-1/2}R[R^{\mathrm{T}}J(\boldsymbol{M})^{-1}R]^{-1}R^{\mathrm{T}}J(\boldsymbol{M})^{-1/2} \tag{70}$$

is a projection matrix ($P^2 = P \leqslant I$), we have

$$\begin{aligned}\mathrm{Tr}[K] &= \mathrm{Tr}\big[[R^{\mathrm{T}}J(\boldsymbol{M})^{-1}R]^{-1}R^{\mathrm{T}}J_S^{-1}R\big]\\ &= \mathrm{Tr}\big[PJ(\boldsymbol{M})^{1/2}J_S^{-1}J(\boldsymbol{M})^{1/2}\big]\\ &\leqslant \mathrm{Tr}\big[J(\boldsymbol{M})J_S^{-1}\big].\end{aligned} \tag{71}$$

Therefore, from Lemma 1, the statement is proved. ■

Proof of Theorem 3. If $\hat{X}_1$ and $\hat{X}_2$ commute with each other,

$$(\Delta_Q X_\mu)^2 = \mathcal{C}_Q(\hat{X}_1,\hat{X}_2) = 0, \tag{72}$$

and therefore the RHS of (42) vanishes. Hence, we have only to consider the case in which $\hat{X}_1$ and $\hat{X}_2$ do not commute. It follows from the fact that $\tilde{\varepsilon}(\boldsymbol{M})$ is symmetric and from Lemma 2 that

$$\begin{aligned}\varepsilon(\hat{X}_1;\boldsymbol{M})\varepsilon(\hat{X}_2;\boldsymbol{M}) &\geqslant \det[\tilde{\varepsilon}(\boldsymbol{M})]\\ &= \det[K^{-1} - I]\det\big[\tilde{J}_S^{-1}\big] \geqslant \det\big[\tilde{J}_S^{-1}\big]\\ &= (\Delta X_1)^2(\Delta X_2)^2 - \mathcal{C}_s(\hat{X}_1,\hat{X}_2)^2.\end{aligned} \tag{73}$$

From the facts that

$$\Delta_Q X_\mu = \Delta X_\mu, \tag{74}$$

$$\mathcal{C}_Q(\hat{X}_1,\hat{X}_2) = \mathcal{C}_s(\hat{X}_1,\hat{X}_2), \tag{75}$$

(42) is proved. ■

V. NUMERICAL RESULTS OF FINDING ACHIEVABLE BOUND FOR ALL POVM MEASUREMENTS

Our trade-off relation (42) is rigorously proven for the measurements in $\mathcal{M}_{\text{all}}$ for $\dim\mathcal{H} = 2$ and $\mathcal{M}_{\text{noisy}}$ for $\dim\mathcal{H} \geqslant 3$. For higher-dimensional Hilbert spaces from $\dim\mathcal{H} = 3$ to 7, we numerically calculate the measurement errors of 10^9 randomly chosen POVMs in $\mathcal{M}_{\text{all}}$ for randomly chosen 10 pairs of quantum states and 2 observables $(\hat{\rho},\hat{X}_1,\hat{X}_2)$. We find that the calculated measurement errors satisfy (42). A typical example of the numerical calculation is shown in Fig. 2 (a). The area within the bound is blacked out by 10^9 data points with no point found outside of the bound. The range $\dim\mathcal{H} = 3$ to 7 includes prime numbers ($\dim\mathcal{H} = 3,5,7$), a power of prime ($\dim\mathcal{H} = 4$), and a composite number that is not a power of prime ($\dim\mathcal{H} = 6$). Therefore, we conjecture the following:

Conjecture 1. For all observables $\hat{X}_1$, $\hat{X}_2$ and quantum states $\hat{\rho}$, all POVMs $\boldsymbol{M} \in \mathcal{M}_{\text{all}}$ satisfy (42).

VI. CONCLUSION AND DISCUSSION

To summarize, we have formulated the complementarity of quantum measurement in a finite-dimensional Hilbert space by invoking quantum estimation theory. To quantify the information retrieved by the measurement, it is essential to take into account the estimation process. We have shown that the measurement errors of noncommuting observables satisfy the Heisenberg-type uncertainty relation, and find the stronger bound (42) that can be achieved for all quantum states and observables. The measurement schemes that achieve this bound can be implemented experimentally in cold-atom systems.

The bound set by (42) is proved for the measurement schemes that perform two projection measurements probabilistically. We numerically show that randomly generated POVM measurements satisfy the bound. Thus, we conjecture that (42) is satisfied for all quantum measurements. The rigorous proof of the conjecture remains a future problem.

ACKNOWLEDGMENTS

This work was supported by Grants-in-Aid for Scientific Research (KAKENHI 22340114 and 22103005), a Global COE Program “The Physical Sciences Frontier,” and the Photon Frontier Network Program, from MEXT of Japan. Y.W. and T.S. acknowledge support from JSPS (Grants No. 216681 and No. 208038, respectively).

APPENDIX: INVERSE OF FISHER INFORMATION MATRIX

In this section, we show how to calculate the inverses of classical and quantum Fisher information matrices.

First, we show $J(\boldsymbol{M})^{-1}$. By expanding each element of $\boldsymbol{M}$ as

$$\hat{M}_k = r_k + \boldsymbol{v}_k \cdot \hat{\boldsymbol{\lambda}}, \tag{A1}$$

the Fisher information matrix can be expressed as

$$J(\boldsymbol{M}) = \sum_{k=1}^{m} \frac{\boldsymbol{v}_k \boldsymbol{v}_k^{\mathrm{T}}}{p_k} = V P^{-1} V^{\mathrm{T}}, \tag{A2}$$

where V and P defined by

$$V = (\,\boldsymbol{v}_1 \cdots \boldsymbol{v}_m\,), \tag{A3}$$

$$P = \mathrm{diag}(p_1, \ldots, p_m) \tag{A4}$$

are $(d^2-1)\times m$ and $m \times m$ real matrices, respectively. The inverse $J(\boldsymbol{M})^{-1}$ can be obtained as

$$J(\boldsymbol{M})^{-1} = (V^{\mathrm{T}})^{-1}[P - PA(A^{-1}PA)^{-1}A^{-1}P]V^{-1}, \tag{A5}$$

where A is a real matrix whose column vectors $\boldsymbol{a}_i$ are linearly independent and satisfy $V\boldsymbol{a}_i = 0$. Then, A satisfies

$$VA = V(A^{\mathrm{T}})^{-1} = 0, \tag{A6}$$

$$V^{-1}V = V^{\mathrm{T}}(V^{\mathrm{T}})^{-1} = I - AA^{-1}, \tag{A7}$$

$$A^{-1}A = A^{\mathrm{T}}(A^{\mathrm{T}})^{-1} = I. \tag{A8}$$

Next, we show (22) and (23). By expanding $\hat{L}_i$ as

$$\hat{L}_i = a_i \hat{I} + \boldsymbol{b}_i \cdot \hat{\boldsymbol{\lambda}}, \tag{A9}$$

and from (20), we obtain

$$\begin{aligned} a_i + \boldsymbol{\theta}\cdot\boldsymbol{b}_i &= 0,\\ a_i\boldsymbol{\theta} + G_{\boldsymbol{\theta}}\boldsymbol{b}_i &= \boldsymbol{e}_i, \end{aligned} \tag{A10}$$

$$\Rightarrow \boldsymbol{b}_i = (G_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}})^{-1}\boldsymbol{e}_i, \tag{A11}$$

where $\boldsymbol{e}_i$ is a unit vector whose ith element is 1, and $G_{\boldsymbol{\theta}}$ is a symmetric matrix whose ij element is defined as

$$[G_{\boldsymbol{\theta}}]_{ij} := \tfrac{1}{2}\langle\{\hat{\lambda}_i, \hat{\lambda}_j\}\rangle. \tag{A12}$$

Therefore, the symmetric logarithmic derivative Fisher information can be written as

$$J_S = (G_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}})^{-1}, \tag{A13}$$

and its inverse can be obtained as (22).

To derive (23), we expand $\hat{L}_i'$ as

$$\hat{L}_i' = c_i \hat{I} + \boldsymbol{d}_i \cdot \hat{\boldsymbol{\lambda}}. \tag{A14}$$

Since $\hat{L}_i'$ is non-Hermitian, c_i is complex and $\boldsymbol{d}_i$ is a complex vector. From (21), these coefficients satisfy

$$\begin{aligned} c_i + \boldsymbol{\theta}\cdot\boldsymbol{d}_i &= 0,\\ c_i\boldsymbol{\theta} + (G_{\boldsymbol{\theta}} - F_{\boldsymbol{\theta}})\boldsymbol{d}_i &= \boldsymbol{e}_i, \end{aligned} \tag{A15}$$

$$\Rightarrow \boldsymbol{d}_i = (G_{\boldsymbol{\theta}} - F_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}})^{-1}\boldsymbol{e}_i, \tag{A16}$$

where $F_{\boldsymbol{\theta}}$ is a Hermitian matrix whose ij element is defined as

$$[F_{\boldsymbol{\theta}}]_{ij} := \tfrac{1}{2}\langle[\hat{\lambda}_i, \hat{\lambda}_j]\rangle. \tag{A17}$$

Therefore, the right logarithmic derivative Fisher information can be written as

$$J_R = (G_{\boldsymbol{\theta}} + F_{\boldsymbol{\theta}} - \boldsymbol{\theta}\boldsymbol{\theta}^{\mathrm{T}})^{-1}, \tag{A18}$$

and (23) is derived.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927). English translation: J. A. Wheeler and H. Zurek, *Quantum Theory and Measurement* (Princeton University Press, Princeton, New Jersey, 1983), p. 62.
[2] E. H. Kennard, Z. Phys. **44**, 326 (1927).
[3] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[4] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983); H. Maassen and J. B. M. Uffink, *ibid.* **60**, 1103 (1988).
[5] N. Bohr, Phys. Rev. **48**, 696 (1935).
[6] C. W. Helstrom, Phys. Lett. A **25**, 101 (1967).
[7] C. W. Helstrom, J. Stat. Phys. **1**, 231 (1969).
[8] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[9] E. Arthurs and J. L. Kelly, Bell Syst. Tech. J. **44**, 725 (1965); E. Arthurs and M. S. Goodman, Phys. Rev. Lett. **60**, 2447 (1988).
[10] M. Ozawa, Phys. Lett. A **320**, 367 (2004).
[11] H. P. Yuen and M. Lax, IEEE Trans. Inf. Theory **19**, 740 (1973).
[12] H. Nagaoka, Trans. Jpn. Soc. Indust. Appl. Math. **1**, 305 (1991) (in Japanese). English translation: M. Hayashi, *Asymptotic Theory of Quantum Statistical Inference: Selected Papers* (World Scientific, Singapore, 2005), p. 133.
[13] T. Sagawa and M. Ueda, Phys. Rev. A **77**, 012313 (2008).
[14] B.-G. Englert, Phys. Rev. Lett. **77**, 2154 (1996).
[15] P. Busch and C. Shilladay, Phys. Rep. **435**, 1 (2006).
[16] A. Kuzmich *et al.*, Phys. Rev. A **60**, 2346 (1999).
[17] Y. Takahashi *et al.*, Phys. Rev. A **60**, 4974 (1999).
[18] A. Kuzmich *et al.*, Phys. Rev. Lett. **85**, 1594 (2000).
[19] T. Takano *et al.*, Phys. Rev. Lett. **102**, 033601 (2009).
[20] Y. Watanabe, T. Sagawa, and M. Ueda, Phys. Rev. Lett. **104**, 020401 (2010).
[21] E. L. Lehmann and G. Casella, *Theory of Point Estimation* (Springer, Berlin, 1998)
[22] H. Cramér, *Mathematical Methods of Statistics* (Princeton University, Princeton, New Jersey, 1946).
[23] E. H. Moore, Bull. Amer. Math. Soc. **26**, 394 (1920).
[24] R. Penrose, Proc. Cambridge Philos. Soc. **51**, 406 (1950).
[25] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[26] C. W. Helstrom, Phys. Lett. A **25**, 101 (1967).
[27] D. Petz, Linear Algebra Appl. **244**, 81 (1996).
[28] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).
[29] R. D. Gill and S. Massar, Phys. Rev. A **61**, 042312 (2000).