Quantifying unsharpness of measurements via uncertainty

Yizhou Liu
Department of Engineering Mechanics, Tsinghua University, Beijing 100084, China

Shunlong Luo*
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

(Received 14 October 2021; revised 4 November 2021; accepted 16 November 2021; published 29 November 2021)

In the conventional axiomatic formalism of quantum mechanics, a physical measurement of a quantum observable is mathematically represented by the spectral decomposition of the Hermitian operator associated with the observable. This constitutes the scenario of a Lüders (projective) measurement, which includes a von Neumann (rank-one projective) measurement as a special and prominent instance. In this context, the measurement is called *sharp* in the sense that each measurement operator is an orthogonal projection (eigenprojection of the observable). In the modern operational formalism, a measurement is represented by a positive-operator valued measure (POVM), which consists of a family of non-negative definite operators (measurement operators, effects) summing to the identity. In this scenario, a measurement is called *unsharp* (fuzzy) if some measurement operators are not orthogonal projections. A natural question arises as to how to quantify unsharpness of a measurement. In this work, we address this issue in terms of uncertainty. For this purpose, we study a family of observables associated with a measurement and their uncertainty. By exploiting the difference between the (total) measurement uncertainty and the observable uncertainty, we are led to some information-theoretic quantifiers of unsharpness. We reveal their basic properties and illustrate them through some important measurements. In particular, we characterize Lüders measurements and equiangular POVMs as extreme measurements in terms of unsharpness.

DOI: 10.1103/PhysRevA.104.052227

I. INTRODUCTION

In the traditional formalism of quantum mechanics, a quantum system is mathematically described by a complex Hilbert space, with a quantum state represented by a non-negative definite operator with unit trace (density operator), and a physical observable represented by a Hermitian (self-adjoint) operator [1–3]. Measuring an observable K in a state ρ of a finite-dimensional system means performing the Lüders (projective) measurement $\mathcal{K} = \{\Pi_i : i = 1, 2, \ldots, n\}$ induced by the spectral decomposition $K = \sum_i \lambda_i \Pi_i$ [4]. It should be emphasized here that Π_i, as eigenprojectors of K, are mutually orthogonal projections. Such a measurement is called *sharp*, for the apparent reason that in such a setting, different measurement outcomes corresponding to different eigenvalues λ_i can be unambiguously discriminated due to orthogonality. In particular, when all Π_i are of rank one, the measurement reduces to a von Neumann measurement, the traditional one in quantum formalism.

When measuring an observable K (equivalently, the corresponding family of spectral projections $\mathcal{K}$) in the state ρ, one obtains the outcome λ_i with probability $p_i = \mathrm{tr}(\rho\Pi_i)$. Consequently, one gets a random variable (a function taking values in the measurement outcome space) Λ with probability distribution $P(\Lambda = \lambda_i) = p_i$. The average value (statistical expectation) $E(\Lambda) = \sum_i p_i\lambda_i$ of the random variable Λ coincides with the corresponding average value $E_\rho(K) = \mathrm{tr}(\rho K)$ of the quantum observable K (in the state ρ), that is,

$$E(\Lambda) = E_\rho(K), \tag{1}$$

which follows readily from $E(\Lambda) = \sum_i p_i\lambda_i = \sum_i \mathrm{tr}(\rho\Pi_i)\lambda_i = \mathrm{tr}[\rho(\sum_i \lambda_i E_i)] = \mathrm{tr}(\rho K) = E_\rho(K)$. This fundamental coincidence of average values, although rather trivial, leads one naturally to guess that the variance

$$\mathrm{Var}(\Lambda) = E[\Lambda - E(\Lambda)]^2 = \sum_i p_i\lambda_i^2 - \Big(\sum_i p_i\lambda_i\Big)^2$$

of the random variable Λ may also coincide with the corresponding variance

$$V_\rho(K) = \mathrm{tr}(\rho K^2) - (\mathrm{tr}\rho K)^2$$

of the quantum observable K (in the state ρ), that is,

$$\mathrm{Var}(\Lambda) = V_\rho(K). \tag{2}$$

This is indeed the case, as can be readily checked by noting that $K^2 = \sum_i \lambda_i^2 \Pi_i$ due to the spectral decomposition.

However, in the modern formalism of quantum mechanics, a measurement is represented by a positive-operator valued measure (POVM) $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$, with E_i non-negative definite operators (not necessarily projections)

*luosl@amt.ac.cn

summing to the identity operator [5–12] or, more generally, by an instrument [13,14]. For simplicity, we focus on measurements described by discrete POVMs. Such a measurement $\mathcal{E}$ is called *unsharp* (fuzzy) if at least one effect, say E_j, is not a projection. If we assign a real numerical value $\alpha_i \in \mathbb{R}$ to the outcome corresponding to the measurement operator (effect) E_i for each i, then similar to the above situation, we obtain a random variable Λ with probability distribution $p_i = P(\Lambda = \alpha_i) = \mathrm{tr}(\rho E_i)$ and a corresponding quantum observable $K = \sum_i \alpha_i E_i$ (not necessarily a spectral decomposition). In stark contrast to Eq. (1) for the average values, which is actually also true for any POVM, Eq. (2) for the variances may fail to be true for K induced by a general POVM due to unsharpness. In fact, the sharpness (i.e., the assumption of all measurement operators being orthogonal projections) of the measurement plays a crucial role in establishing Eq. (2) for Lüders measurements.

Motivated by the above well-known observation, we are naturally led to consider the failure of Eq. (2) as an indication of unsharpness of the measurement involved. This work is devoted to elaborating on this idea and quantifying unsharpness of measurements via uncertainty, which in turn is quantified by variance. More specifically, by exploring the difference between the variance of the random variable and that of the quantum observable, we introduce some quantifiers of unsharpness of measurements, reveal their basic properties, and provide some insight into the structure of measurements.

Due to the increasing interest and importance in both theoretical and practical considerations of general POVMs since the 1960s [15–37], issues related to quantitative aspects of unsharpness have been studied by many authors. For example, a noise operator (also called uncertainty operator) was introduced implicitly or explicitly in Refs. [30–33]. Applications of the noise operator and the associated decomposition of variance to uncertainty relations and other issues were investigated in Refs. [14,34,35]. The notion of resolution width of a measurement was used as a quantifier of unsharpness in Ref. [36], which has a clear interpretation in terms of actualizability. The problem of intrinsic unsharpness of a measuring device was addressed in [37], where an entropic measure of unsharpness based on the average Shannon entropy of probability distribution was explored. Inspired by these studies of unsharpness, we take a similar starting point to exploit the interplay between the variance of random variables and quantum observables. We introduce several matrices (different from the noise operator) and related quantities to characterize and quantify unsharpness. We note that although variance, as a simple and ubiquitous quantity, has been extensively and intensively studied and used ever since the inception of quantum mechanics, there are still novel aspects and applications awaiting exploration. Some recent applications of variance in characterizing coherence and nonclassicality were presented in Refs. [38,39].

This work is structured as follows. In Sec. II, as a preliminary discussion, we describe two different kinds of variance (uncertainty) associated with a POVM by introducing parameters as measurement values. In Sec. III, we quantify unsharpness of POVMs in terms of the uncertainty difference. We further reveal basic properties of the quantifiers of unsharpness. We illustrate unsharpness through some important POVMs. In Sec. IV, we characterize the Lüders measurements, the POVMs induced by equiangular tight frames, as well as the symmetric informationally complete POVMs (SIC-POVMs), as extremal measurements in terms of unsharpness. Finally, we summarize the results and discuss some perspectives in Sec. V. For simplicity, we focus on finite-dimensional quantum systems and discrete POVMs with finite elements (effects). Although the idea and methods may be adapted to the infinite-dimensional cases and some results can be carried over to the general cases, it is an important and highly nontrivial issue to address the general situation.

II. PRELIMINARY

To motivate and to facilitate later discussions, in this section we collect and derive some elementary properties concerning uncertainty of a POVM. These results are well known or can be straightforwardly verified. We will exploit uncertainty of a POVM in terms of two different versions of variance: variance of a random variable vs variance of a quantum observable (Hermitian operator). Both emerge naturally from a POVM coupled to a state.

Consider a measurement in a d-dimensional quantum system represented by a POVM,

$$\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\},$$

which consists of non-negative definite operators E_i summing to the identity, that is, $E_i \geqslant 0$ and $\sum_i E_i = \mathbf{1}$ (resolution of identity) [5–12]. Each measurement operator E_i, which is called an effect, corresponds to an experimental outcome. If all $\mathrm{tr}(E_i)$ are equal, then the POVM is called homogenous, and if all effects are of rank one, then the POVM is called rank one. When all E_i are orthogonal projections, the POVM reduces to a Lüders measurement and one recovers the conventional measurement formalism in terms of the spectral decomposition of a Hermitian operator representing the quantum observable. The probability of obtaining the ith outcome of the measurement (when the system is in a state ρ) is stipulated as

$$p_i = \mathrm{tr}(\rho E_i)$$

via the celebrated Born rule.

In the above setup, the measurement outcomes are only distinguished by the label i, and no physical values are assigned to them. In practice, and for our purpose of quantitative study of uncertainty and unsharpness in an operational way, it is helpful to introduce parameters to represent the measurement outcomes in order to employ the full power of quantitative analysis. Thus we argument the POVM $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ with a vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$, and call the combination $(\mathcal{E}, \boldsymbol{\alpha})$ a measurement scenario, with α_i interpreted as the measurement value corresponding to the effect E_i. Associated with a measurement scenario $(\mathcal{E}, \boldsymbol{\alpha})$, there are two natural objects, which are of special significance in our approach to unsharpness, as follows:

(1) The quantum observable

$$K_{\boldsymbol{\alpha}} = \sum_i \alpha_i E_i, \tag{3}$$

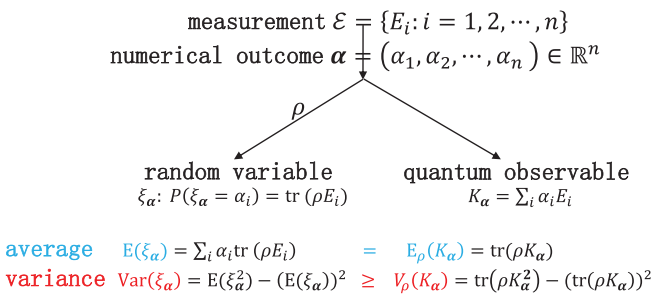

FIG. 1. A measurement scenario $(\mathcal{E}, \boldsymbol{\alpha})$ induces two natural objects: random variable $\xi_{\boldsymbol{\alpha}}$ (via coupling with state ρ) vs quantum observable $K_{\boldsymbol{\alpha}}$. Their averages agree, but their variances in general disagree. This difference may be exploited to quantify the unsharpness (deviation from the Lüders measurements) of the measurement $\mathcal{E}$.

which is a Hermitian operator summarizing the weighted effects. Here, $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$. In the conventional formalism of measurement, if $\mathcal{K} = \{\Pi_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement and $\boldsymbol{\alpha} = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ (the vector consisting of the spectrum of a quantum observable $K = \sum_i \lambda_i \Pi_i$), then $K_{\boldsymbol{\alpha}} = K$ is exactly the original observable inducing the measurement $\mathcal{K}$.

(2) The random variable $\xi_{\boldsymbol{\alpha}}$ with probability distribution $p_i = P(\xi_{\boldsymbol{\alpha}} = \alpha_i) = \mathrm{tr}(\rho E_i), \quad i = 1, 2, \ldots, n,$ which records the measurement outcomes of performing the measurement $\mathcal{E}$ in a state ρ and the corresponding probabilities as a function (rather than an operator). This should be compared with the quantum observable $K_{\boldsymbol{\alpha}}$ defined by Eq. (3). The essential distinction lies in that the quantum observable is a Hermitian operator (independent of the state ρ) constructed from the POVM via a linear combination of effects, while the random variable (depending on the state ρ) is a function derived from measuring the POVM (in the state ρ) with the probabilities determined via the Born rule. Thus it is more rigorous to denote the random variable $\xi_{\boldsymbol{\alpha}}$ as $\xi_{\boldsymbol{\alpha},\rho}$. But, for notational simplicity, we just use the simpler one. We illustrate the comparison between them in Fig. 1.

For the quantum observable $K_{\boldsymbol{\alpha}}$, we have the average value

$$E_\rho(K_{\boldsymbol{\alpha}}) = \mathrm{tr}(\rho K_{\boldsymbol{\alpha}})$$

and the conventional variance

$$V_\rho(K_{\boldsymbol{\alpha}}) = \mathrm{tr}\big(\rho K_{\boldsymbol{\alpha}}^2\big) - (\mathrm{tr}\rho K_{\boldsymbol{\alpha}})^2. \tag{4}$$

For the random variable $\xi_{\boldsymbol{\alpha}}$, we have the average value

$$E(\xi_{\boldsymbol{\alpha}}) = \sum_i p_i \alpha_i,$$

which coincides with the average value $E_\rho(K_{\boldsymbol{\alpha}}) = \mathrm{tr}(\rho K_{\boldsymbol{\alpha}})$ of the quantum observable $K_{\boldsymbol{\alpha}}$ (in the state ρ), as can be readily checked from

$$E(\xi_{\boldsymbol{\alpha}}) = \sum_i \mathrm{tr}(\rho E_i)\alpha_i = \mathrm{tr}\Big(\rho \sum_i E_i \alpha_i\Big) = \mathrm{tr}(\rho K_{\boldsymbol{\alpha}}).$$

This coincidence indicates that the quantum observable $K_{\boldsymbol{\alpha}}$ is indistinguishable from the random variable $\xi_{\boldsymbol{\alpha}}$ in terms of their average values. For the variance of the random variable $\xi_{\boldsymbol{\alpha}}$, we have

$$\mathrm{Var}(\xi_\alpha) = E\big(\xi_{\boldsymbol{\alpha}}^2\big) - (E\xi_{\boldsymbol{\alpha}})^2 = \sum_i p_i \alpha_i^2 - \Big(\sum_i p_i \alpha_i\Big)^2. \tag{5}$$

For notational convenience and to explicitly indicate the dependence on $\mathcal{E}$, ρ, and $\boldsymbol{\alpha}$, we denote the above variance as

$$V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \mathrm{Var}(\xi_{\boldsymbol{\alpha}}). \tag{6}$$

If we introduce an $n \times n$ matrix $\boldsymbol{V}_\rho(\mathcal{E})$ with matrix entries

$$[\boldsymbol{V}_\rho(\mathcal{E})]_{ij} = \delta_{ij}\mathrm{tr}(\rho E_i) - \mathrm{tr}(\rho E_i)\mathrm{tr}(\rho E_j),$$

then

$$V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \boldsymbol{\alpha}\boldsymbol{V}_\rho(\mathcal{E})\boldsymbol{\alpha}^T,$$

which is a quadratic form of $\boldsymbol{\alpha} \in \mathbb{R}^n$. Here, T denotes transpose of the vectors.

The variance $V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \mathrm{Var}(\xi_{\boldsymbol{\alpha}})$ of the random variable $\xi_{\boldsymbol{\alpha}}$ summarizes the uncertainty of the measurement outcomes and has the following properties.

Lemma 1. It holds that

(a) $V_\rho(\mathcal{E}, \boldsymbol{\alpha}) \geqslant 0, \forall\ \boldsymbol{\alpha}$ or, equivalently, the real matrix $\boldsymbol{V}_\rho(\mathcal{E})$ is non-negative definite. $V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = 0$ if and only if all α_i with $P(\xi_{\boldsymbol{\alpha}} = \alpha_i) \neq 0$ are equal, i.e., the random variable $\xi_{\boldsymbol{\alpha}}$ is a constant with probability 1.

(b) $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$ possesses unitary covariance in the sense that

$$V_{U\rho U^\dagger}(U\mathcal{E}U^\dagger, \boldsymbol{\alpha}) = V_\rho(\mathcal{E}, \boldsymbol{\alpha}),$$

where U is any unitary operator on the system Hilbert space and $U\mathcal{E}U^\dagger = \{UE_iU^\dagger : i = 1, 2, \ldots, n\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$.

(c) $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$ is concave in ρ and convex in $\boldsymbol{\alpha}$.

(d) If $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement, then $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$ reduces to the conventional variance $V_\rho(K_{\boldsymbol{\alpha}})$ of the quantum observable $K_{\boldsymbol{\alpha}} = \sum_i \alpha_i E_i$ (in the state ρ) as defined by Eq. (4), i.e.,

$$V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = V_\rho(K_{\boldsymbol{\alpha}}).$$

Although the above properties are simple or well known, they are important and, for completeness, we sketch the proof. Item (a) is clear since $V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \mathrm{Var}(\xi_{\boldsymbol{\alpha}})$ is the variance of a real-valued random variable $\xi_{\boldsymbol{\alpha}}$, and thus is always a non-negative number. The variance of a random variable vanishes if and only if it is a constant with probability 1. Consequently, $V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \mathrm{Var}(\xi_{\boldsymbol{\alpha}}) = 0$ holds if and only if ξ_α is a constant with probability 1, that is, all α_i with $P(\xi_{\boldsymbol{\alpha}} = \alpha_i) \neq 0$ are equal.

Items (b) and (c) can be checked straightforwardly.

Item (d) follows from

$$\begin{aligned} V_\rho(\mathcal{E}, \boldsymbol{\alpha}) &= \sum_i p_i \alpha_i^2 - \Big(\sum_i p_i \alpha_i\Big)^2 \\ &= \sum_i \mathrm{tr}(\rho E_i)\alpha_i^2 - \Big[\sum_i \mathrm{tr}(\rho E_i)\alpha_i\Big]^2 \\ &= \mathrm{tr}\Big(\rho \sum_i E_i \alpha_i^2\Big) - \Big[\mathrm{tr}\Big(\rho \sum_i E_i \alpha_i\Big)\Big]^2 \\ &= \mathrm{tr}(\rho K_{\boldsymbol{\alpha}}^2) - (\mathrm{tr}\rho K_{\boldsymbol{\alpha}})^2 \\ &= V_\rho(K_{\boldsymbol{\alpha}}). \end{aligned}$$

In the above derivation, the assumption of Lüders measurement plays a crucial role since, under this condition, E_i are orthogonal projections and $K_{\boldsymbol{\alpha}} = \sum_i \alpha_i E_i$ is the spectral decomposition of $K_{\boldsymbol{\alpha}}$, which implies that $K_{\boldsymbol{\alpha}}^2 = \sum_i \alpha_i^2 E_i$. This is not true in general if $K_{\boldsymbol{\alpha}} = \sum_i \alpha_i E_i$ is not a spectral decomposition.

Now a natural question arises as to the relation between $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$ (variance of the random variable $\xi_{\boldsymbol{\alpha}}$) and $V_\rho(K_{\boldsymbol{\alpha}})$ (variance of the quantum observable $K_{\boldsymbol{\alpha}}$). To address this issue, we introduce

$$F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = V_\rho(\mathcal{E}, \boldsymbol{\alpha}) - V_\rho(K_{\boldsymbol{\alpha}}), \tag{7}$$

which serves as a starting point for our approach to unsharpness. The symbol F refers to fuzziness (unsharpness). To gain a clearer picture of the structure of the above quantity, let us define an $n \times n$ matrix $\boldsymbol{F}_\rho(\mathcal{E})$ with matrix entries

$$[\boldsymbol{F}_\rho(\mathcal{E})]_{ij} = \delta_{ij}\mathrm{tr}(\rho E_i) - \mathrm{tr}(\rho\{E_i, E_j\}), \tag{8}$$

where $\{E_i, E_j\} = (E_iE_j + E_jE_i)/2$ denotes the anticommutator (Jordan product, symmetric product) of the operators. It can be directly checked that

$$F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \boldsymbol{\alpha}\boldsymbol{F}_\rho(\mathcal{E})\boldsymbol{\alpha}^T, \tag{9}$$

which shows that $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$, like $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$, is a quadratic form of $\boldsymbol{\alpha}$.

In this context, we recall some early studies of unsharpness. For $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$, let $\boldsymbol{\alpha}^2 = (\alpha_1^2, \alpha_2^2, \ldots, \alpha_n^2)$ (Hadamard product, entrywise product of vectors). The noise operator

$$N_{\boldsymbol{\alpha}} = K_{\boldsymbol{\alpha}^2} - K_{\boldsymbol{\alpha}}^2 \tag{10}$$

was introduced and studied in Refs. [14,30–36]. From Eqs. (5)–(7), we have

$$\begin{aligned} F_\rho(\mathcal{E}, \boldsymbol{\alpha}) &= V_\rho(\mathcal{E}, \boldsymbol{\alpha}) - V_\rho(K_{\boldsymbol{\alpha}}) \\ &= \mathrm{Var}(\xi_{\boldsymbol{\alpha}}) - V_\rho(K_{\boldsymbol{\alpha}}) \\ &= \sum_i \alpha_i^2 \mathrm{tr}(\rho E_i) - \Big[\sum_i \alpha_i \mathrm{tr}(\rho E_i)\Big]^2 \\ &\quad - \mathrm{tr}\big(\rho K_{\boldsymbol{\alpha}}^2\big) + \big[\mathrm{tr}(\rho K_{\boldsymbol{\alpha}})\big]^2 \\ &= \sum_i \alpha_i^2 \mathrm{tr}(\rho E_i) - \mathrm{tr}\big(\rho K_{\boldsymbol{\alpha}}^2\big) \\ &= \mathrm{tr}(\rho N_{\boldsymbol{\alpha}}), \end{aligned}$$

which shows that the average value of the noise operator $N_{\boldsymbol{\alpha}}$ is the fuzziness $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$, i.e.,

$$F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \mathrm{tr}(\rho N_{\boldsymbol{\alpha}}). \tag{11}$$

We stress that $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ is expressed as a quadratic form of $\boldsymbol{\alpha}$ with the corresponding matrix $\boldsymbol{F}_\rho(\mathcal{E})$ being an operator on $\mathbb{C}^n$, while the noise operator $N_{\boldsymbol{\alpha}}$ acts on the d-dimensional system Hilbert space. These two quite different operators are related as

$$\boldsymbol{\alpha}\boldsymbol{F}_\rho(\mathcal{E})\boldsymbol{\alpha}^T = \mathrm{tr}(\rho N_{\boldsymbol{\alpha}}). \tag{12}$$

The quantity $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ defined by Eq. (9) has the following basic properties.

Lemma 2. It holds that

(a) $F_\rho(\mathcal{E}, \boldsymbol{\alpha}) \geqslant 0, \forall\ \boldsymbol{\alpha}$ or, equivalently, the real matrix $\boldsymbol{F}_\rho(\mathcal{E})$ is non-negative definite. If $F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = 0$ for all ρ and $\boldsymbol{\alpha} \in \mathbb{R}^n$, then E_i are orthogonal projections, that is, $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement.

(b) $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ possesses unitary covariance in the sense that

$$F_{U\rho U^\dagger}(U\mathcal{E}U^\dagger, \boldsymbol{\alpha}) = F_\rho(\mathcal{E}, \boldsymbol{\alpha}),$$

where U is any unitary operator and $U\mathcal{E}U^\dagger = \{UE_iU^\dagger : i = 1, 2, \ldots, n\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$.

(c) $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ is affine in ρ and convex in $\boldsymbol{\alpha}$.

We proceed to establish the above results.

To prove item (a), we first define

$$X = K_{\boldsymbol{\alpha}}\sqrt{\rho} = \sum_i \alpha_i E_i \sqrt{\rho}, \quad \boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n,$$

then $X^\dagger = \sqrt{\rho}K_{\boldsymbol{\alpha}} = \sum_i \alpha_i \sqrt{\rho}E_i$. By $\sum_i E_i = \mathbf{1}$ and twice use of the Cauchy-Schwarz inequality, we have

$$\begin{aligned} |\mathrm{tr}(X^\dagger X)|^2 &= \Big|\mathrm{tr}\Big(X^\dagger \sum_i \alpha_i E_i \sqrt{\rho}\Big)\Big|^2 \\ &= \Big|\sum_i \mathrm{tr}\Big[(X^\dagger \sqrt{E_i})(\alpha_i \sqrt{E_i}\sqrt{\rho})\Big]\Big|^2 \\ &\leqslant \Big|\sum_i \Big[\mathrm{tr}(X^\dagger \sqrt{E_i})(X^\dagger \sqrt{E_i})^\dagger\Big]^{\frac{1}{2}} \\ &\quad \times \Big[\mathrm{tr}(\alpha_i \sqrt{E_i}\sqrt{\rho})(\alpha_i \sqrt{E_i}\sqrt{\rho})^\dagger\Big]^{\frac{1}{2}}\Big|^2 \\ &= \Big|\sum_i \Big[\mathrm{tr}(X^\dagger E_i X)\Big]^{\frac{1}{2}}\Big[\mathrm{tr}(\alpha_i^2 E_i \rho)\Big]^{\frac{1}{2}}\Big|^2 \\ &\leqslant \sum_i \mathrm{tr}(X^\dagger E_i X) \sum_i \mathrm{tr}\big(\alpha_i^2 E_i \rho\big) \\ &= \mathrm{tr}(X^\dagger X) \sum_i \mathrm{tr}\big(\alpha_i^2 E_i \rho\big), \end{aligned}$$

which implies that

$$\mathrm{tr}(X^\dagger X) \leqslant \sum_i \mathrm{tr}\big(\alpha_i^2 E_i \rho\big).$$

This is equivalent to

$$\mathrm{tr}\big(\rho K_{\boldsymbol{\alpha}}^2\big) \leqslant \sum_i \alpha_i^2 p_i$$

since $\mathrm{tr}(X^\dagger X) = \mathrm{tr}(\sqrt{\rho}K_{\boldsymbol{\alpha}}K_{\boldsymbol{\alpha}}\sqrt{\rho}) = \mathrm{tr}(\rho K_{\boldsymbol{\alpha}}^2)$ and $\mathrm{tr}(E_i\rho) = p_i$. Now, from Eq. (3), we have $\mathrm{tr}(\rho K_{\boldsymbol{\alpha}}) = \sum_i \alpha_i p_i$, and from Eqs. (4) and (6), we obtain

$$V_\rho(K_{\boldsymbol{\alpha}}) \leqslant V_\rho(\mathcal{E}, \boldsymbol{\alpha}),$$

which is equivalent to $F_\rho(\mathcal{E}, \boldsymbol{\alpha}) \geqslant 0$, the desired result. We remark that the non-negativity of the difference of the variance already appeared explicitly or implicitly in Refs. [14,30–36], and our alternative proof here is only for completeness and for illuminating further features of the variance difference.

If $F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = 0$ holds for all ρ and $\boldsymbol{\alpha}$, then by Eq. (9), we know that $\boldsymbol{F}_\rho(\mathcal{E})$ must be the zero matrix. Thus, we have

$$[\boldsymbol{F}_\rho(\mathcal{E})]_{ii} = \mathrm{tr}\big[\rho\big(E_i - E_i^2\big)\big] = 0, \quad \forall\ \rho,$$

which implies that $E_i - E_i^2 = 0$ (i.e., E_i are projectors). Also, when $i \neq j$,

$$[\boldsymbol{F}_\rho(\mathcal{E})]_{ij} = -\mathrm{tr}(\rho\{E_i, E_j\}) = 0, \quad \forall\, \rho,$$

which implies $\mathrm{tr}(\{E_i, E_j\}) = 0$. Because E_i and E_j are both non-negative definite operators, it follows that $E_iE_j = 0$ for $i \neq j$. Consequently, E_i are orthogonal projections, i.e., $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement. Conversely, if $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement, it can be readily verified that $F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = 0$ for all ρ and $\boldsymbol{\alpha}$. In this case, $\boldsymbol{F}_\rho(\mathcal{E})$ is the zero matrix.

Item (b) can be checked straightforwardly.

For item (c), since $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ can be expressed as

$$F_\rho(\mathcal{E}, \boldsymbol{\alpha}) = \sum_i \mathrm{tr}(\rho E_i)\alpha_i^2 - \mathrm{tr}\big(\rho K_{\boldsymbol{\alpha}}^2\big),$$

we see clearly that it is affine in ρ. Because $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ is a quadratic form of the vector $\boldsymbol{\alpha}$, it follows that it is convex in $\boldsymbol{\alpha}$.

In view of Lemma 2, we call $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$, as defined by Eq. (7), the *extra variance* of the measurement scenario $(\mathcal{E}, \boldsymbol{\alpha})$ (in the state ρ). In contrast, we call $V_\rho(\mathcal{E}, \boldsymbol{\alpha})$, as defined by Eq. (6), the *total variance*, and $V_\rho(K_{\boldsymbol{\alpha}})$, as defined by Eq. (4), the *irreducible variance*, of the measurement scenario $(\mathcal{E}, \boldsymbol{\alpha})$ (in the state ρ). Consequently, we have classified the variance of a measurement scenario into three categories, which satisfy the decomposition relation:

Total variance = Irreducible variance + Extra variance.

Symbolically,

$$V_\rho(\mathcal{E}, \boldsymbol{\alpha}) = V_\rho(K_{\boldsymbol{\alpha}}) + F_\rho(\mathcal{E}, \boldsymbol{\alpha}).$$

Such a kind of decomposition of variance was studied by several authors [14,33,35,36]. This is a motivation for our investigations of the unsharpness of measurements in the next section. The above decomposition is reminiscent of several information-theoretical decompositions in the literature: (1) the decomposition of total correlations into quantum correlations and classical correlations [40–42], (2) the decomposition of a quantum observable into a quantum fluctuation part and a classical part [43], and (3) the decomposition of total uncertainty into quantum uncertainty and classical uncertainty [44].

We emphasize that although we will only regard the numerical vector $\boldsymbol{\alpha}$ associated with the measurement effects as an ancillary tool (extrinsic quantity) for quantifying intrinsic unsharpness, in fact, these numerical values are essential in studying uncertainties and are of basic physical significance in practical experiments.

III. QUANTIFYING UNSHARPNESS OF MEASUREMENTS

After the above preparation, now we proceed to study the unsharpness of measurements in terms of uncertainty. The parameter vector $\boldsymbol{\alpha}$ plays a significant role in the study of measurement uncertainty and actual recording of outcomes. In our approach, we have separated the effects of $\boldsymbol{\alpha}$ from those of the original measurement $\mathcal{E}$ since we have expressed $F_\rho(\mathcal{E}, \boldsymbol{\alpha})$ as a quadratic form of $\boldsymbol{\alpha}$ in Eq. (9). Thus, in order to introduce some intrinsic quantities for the measurement $\mathcal{E}$, we only need to focus on the matrix $\boldsymbol{F}_\rho(\mathcal{E})$ defined by Eq. (8), which summarizes certain features of extra uncertainty of the measurement $\mathcal{E}$ (in the state ρ) via Eq. (9), i.e., unsharpness of $\mathcal{E}$. This quantity still depends on the state ρ. In order to get some quantities intrinsic only to the measurement, it is desirable to get rid of the state dependence. Two intuitive and simple approaches manifest themselves immediately.

(1) We may take ρ simply to be the maximally mixed state $\mathbf{1}/d$; then we come to the matrix

$$\boldsymbol{F}_{\mathbf{1}/d}(\mathcal{E}) = (r_{ij}),$$

with $r_{ij} = [\delta_{ij}\mathrm{tr}(E_i) - \mathrm{tr}(E_iE_j)]/d$ and $i, j = 1, 2, \ldots, n$.

(2) We may take the average of $\boldsymbol{F}_\rho(\mathcal{E})$ over the state ρ. Since $\boldsymbol{F}_\rho(\mathcal{E})$ is affine in ρ, this will yield essentially the same result as the above.

For notational simplicity, we omit the subscript in $\boldsymbol{F}_{\mathbf{1}/d}(\mathcal{E})$ and denote

$$\boldsymbol{F}(\mathcal{E}) = \boldsymbol{F}_{\mathbf{1}/d}(\mathcal{E}),$$

which may be called the unsharpness matrix of $\mathcal{E}$. The basic properties of the $n \times n$ matrix $\boldsymbol{F}(\mathcal{E})$ are as follows.

Proposition 1. It holds that

(a) $\boldsymbol{F}(\mathcal{E}) \geqslant \mathbf{0}$ (zero matrix), that is, $\boldsymbol{F}(\mathcal{E})$ is a non-negative definite matrix. Moreover, the equality holds if and only if $\mathcal{E} = \{\Pi_i : i = 1, 2, \ldots, n\}$ is a Lüders measurement.

(b) $\boldsymbol{F}(\mathcal{E})$ is concave for the inner mixture of $\mathcal{E}$ in the sense that

$$\boldsymbol{F}[c\mathcal{E} + (1-c)\mathcal{G}] \geqslant c\boldsymbol{F}(\mathcal{E}) + (1-c)\boldsymbol{F}(\mathcal{G}),$$

where $c \in [0, 1]$, and $c\mathcal{E} + (1-c)\mathcal{G} = \{cE_i + (1-c)G_i : i = 1, 2, \ldots, n\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ and $\mathcal{G} = \{G_i : i = 1, 2, \ldots, n\}$.

(c) $\boldsymbol{F}(\mathcal{E})$ is concave for the outer mixture of $\mathcal{E}$ in the sense that

$$\boldsymbol{F}[c\mathcal{E} \cup (1-c)\mathcal{G}] \geqslant c\boldsymbol{F}(\mathcal{E}) \oplus (1-c)\boldsymbol{F}(\mathcal{G}),$$

where $c \in [0, 1]$, and $c\mathcal{E} \cup (1-c)\mathcal{G} = \{cE_i : i = 1, 2, \ldots, n\} \cup \{(1-c)G_j : j = 1, 2, \ldots, m\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ and $\mathcal{G} = \{G_j : j = 1, 2, \ldots, m\}$.

Before proving the above statements, we first illustrate the operational meaning of the above two types of mixture of measurements, which are different procedures for combining measurements into new ones. The inner mixture concerns two measurements $\mathcal{E}$ and $\mathcal{G}$ with the same number of (ordered) effects, and arises from performing a measurement whose ith effect is a probabilistic combination of the ith effect of $\mathcal{E}$ and that of $\mathcal{G}$. This leads to the measurement $\{cE_i + (1-c)G_i : i = 1, 2, \ldots, n\}$, which is precisely the inner mixture $c\mathcal{E} + (1-c)\mathcal{G}$ of $\mathcal{E}$ and $\mathcal{G}$ (convex combination). In contrast, the outer mixture concerns two measurements $\mathcal{E}$ and $\mathcal{G}$ whose numbers of effects can be arbitrary, and arises from performing the measurement $\mathcal{E}$ with probability c and the measurement $\mathcal{G}$ with probability $1-c$. This combined measurement leads to the measurement $\{cE_1, cE_2, \ldots, cE_n, (1-c)G_1, (1-c)G_2, \ldots, (1-c)G_m\}$, which is precisely the outer mixture $c\mathcal{E} \cup (1-c)\mathcal{G}$ (randomization). For further elaboration on the operational mixtures of measurements, see Sec. 3.1.5 in Ref. [13].

Now we proceed to prove Proposition 1.

Item (a) follows readily from Lemma 2 [item (a)]. From item (a), we see that the unsharpness of any

Lüders measurement vanishes. Phrased alternatively, a Lüders measurement is sharp, as it should be.

For item (b), note that the Gram matrix $\boldsymbol{Z} = (z_{ij})$ with matrix entries $z_{ij} = \mathrm{tr}[(E_i - G_i)(E_j - G_j)]$ is clearly non-negative definite. Now direct manipulation leads to

$$\boldsymbol{F}[c\mathcal{E} + (1-c)\mathcal{G}] - c\boldsymbol{F}(\mathcal{E}) - (1-c)\boldsymbol{F}(\mathcal{G}) = c(1-c)\boldsymbol{Z},$$

which implies the desired result.

In order to establish item (c), for any real vectors $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_m) \in \mathbb{R}^m$, denote $\boldsymbol{\alpha\beta} = (\alpha_1, \alpha_2, \ldots, \alpha_n, \beta_1, \beta_2, \ldots, \beta_m) \in \mathbb{R}^{n+m}$, and consider the three measurement scenarios $(\mathcal{E}, \boldsymbol{\alpha})$, $(\mathcal{G}, \boldsymbol{\beta})$, and $[c\mathcal{E} \cup (1-c)\mathcal{G}, \boldsymbol{\alpha\beta}]$ with the associated quantum observables $K_{\boldsymbol{\alpha}} = \sum_i \alpha_i E_i$, $J_{\boldsymbol{\beta}} = \sum_j \beta_j G_j$, and

$$S_{\boldsymbol{\alpha\beta}} = \sum_i \alpha_i c E_i + \sum_j \beta_j (1-c) G_j = cK_{\boldsymbol{\alpha}} + (1-c)J_{\boldsymbol{\beta}}.$$

The corresponding random variables are $\xi_{\boldsymbol{\alpha}}, \eta_{\boldsymbol{\beta}}$, and $\zeta_{\boldsymbol{\alpha\beta}}$ with probability distributions $P(\xi_{\boldsymbol{\alpha}} = \alpha_i) = \mathrm{tr}(\rho E_i)$, $P(\eta_{\boldsymbol{\beta}} = \beta_j) = \mathrm{tr}(\rho G_j)$, and

$$P(\zeta_{\boldsymbol{\alpha\beta}} = \alpha_i) = \mathrm{tr}(\rho c E_i), \qquad i = 1, 2, \ldots, n,$$
$$P(\zeta_{\boldsymbol{\alpha\beta}} = \beta_j) = \mathrm{tr}[\rho(1-c)G_j], \qquad j = 1, 2, \ldots, m.$$

Direct calculation shows that

$$\mathrm{Var}(\zeta_{\boldsymbol{\alpha\beta}}) - c\mathrm{Var}(\xi_{\boldsymbol{\alpha}}) - (1-c)\mathrm{Var}(\eta_{\boldsymbol{\beta}}) = c(1-c)[E(\xi_{\boldsymbol{\alpha}}) - E(\eta_{\boldsymbol{\beta}})]^2 \geqslant 0,$$

which is equivalent to

$$V_\rho[c\mathcal{E} \cup (1-c)\mathcal{G}, \boldsymbol{\alpha\beta}] \geqslant cV_\rho(\mathcal{E}, \boldsymbol{\alpha}) + (1-c)V_\rho(\mathcal{G}, \boldsymbol{\beta}).$$

Noting that the variance $V_\rho(K)$ is convex in the observable K, we have

$$V_\rho(S_{\boldsymbol{\alpha\beta}}) \leqslant cV_\rho(K_{\boldsymbol{\alpha}}) + (1-c)V_\rho(J_{\boldsymbol{\beta}}).$$

Combining the above two inequalities yields

$$F_\rho[c\mathcal{E} \cup (1-c)\mathcal{G}, \boldsymbol{\alpha\beta}] \geqslant cF_\rho(\mathcal{E}, \boldsymbol{\alpha}) + (1-c)F_\rho(\mathcal{G}, \boldsymbol{\beta}).$$

In particular, taking $\rho = \mathbf{1}/d$ to be the maximally mixed state and recalling that $F_{\mathbf{1}/d}(\cdot) = F(\cdot)$, we obtain

$$\begin{aligned}
&\boldsymbol{\alpha\beta}\boldsymbol{F}[c\mathcal{E} \cup (1-c)\mathcal{G}](\boldsymbol{\alpha\beta})^T \\
&= F[c\mathcal{E} \cup (1-c)\mathcal{G}, (\boldsymbol{\alpha}, \boldsymbol{\beta})] \\
&\geqslant cF(\mathcal{E}, \boldsymbol{\alpha}) + (1-c)F(\mathcal{G}, \boldsymbol{\beta}) \\
&= c\boldsymbol{\alpha}\boldsymbol{F}(\mathcal{E})\boldsymbol{\alpha}^T + (1-c)\boldsymbol{\beta}\boldsymbol{F}(\mathcal{G})\boldsymbol{\beta}^T \\
&= \boldsymbol{\alpha\beta}[c\boldsymbol{F}(\mathcal{E}) \oplus (1-c)\boldsymbol{F}(\mathcal{G})](\boldsymbol{\alpha\beta})^T,
\end{aligned}$$

which implies the desired result,

$$\boldsymbol{F}[c\mathcal{E} \cup (1-c)\mathcal{G}] \geqslant c\boldsymbol{F}(\mathcal{E}) \oplus (1-c)\boldsymbol{F}(\mathcal{G}).$$

In view of Proposition 1, we call $\boldsymbol{F}(\mathcal{E})$ the unsharpness matrix of $\mathcal{E}$. In order to obtain a measure of unsharpness as a single numerical quantity (rather than a matrix), we may take a suitable functional of the matrix. Recall that a unitarily invariant norm $\|\cdot\|$ of matrices satisfies $\|UXV\| = \|X\|$ for any unitary operators U and V [45,46]. Typical unitarily invariant norms include the Frobenius norm, the l^p norm with $p \geqslant 1$, as well as the Ky Fan norm. Actually, any unitarily invariant norm has the form $||X|| = g(s_1, s_2, \ldots, s_n)$ with s_i the singular values of the matrix X, and g is any non-negative, subadditive, symmetric function on $\mathbb{R}^n_+$ satisfying $g(s_1, s_2, \ldots, s_n) = 0$ if and only if $s_1 = s_2 = \cdots = s_n = 0$ [47]. For any unitarily invariant norm $\|\cdot\|$ of matrices, one may employ $\|\boldsymbol{F}(\mathcal{E})\|$ as a numerical quantifier of unsharpness. For simplicity, we define the quantity

$$f(\mathcal{E}) = \|\boldsymbol{F}(\mathcal{E})\|_1$$

as a measure of unsharpness (fuzziness) of the measurement $\mathcal{E}$, with the smaller norm corresponding to the smaller unsharpness (larger sharpness) of the measurement. Here, $\|X\|_1 = \sum_{ij} |x_{ij}|$ is the l^1 norm of the matrix $X = (x_{ij})$. This quantity has the following properties.

Proposition 2. It holds that

(a) $f(\mathcal{E}) \geqslant 0$, and the equality holds if and only if $\mathcal{E}$ is a Lüders measurement.

(b) $f(\mathcal{E})$ is concave for the inner mixture of $\mathcal{E}$ in the sense that

$$f[c\mathcal{E} + (1-c)\mathcal{G}] \geqslant cf(\mathcal{E}) + (1-c)f(\mathcal{G}),$$

where $c \in [0, 1]$, and $c\mathcal{E} + (1-c)\mathcal{G} = \{cE_i + (1-c)G_i : i = 1, 2, \ldots, n\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ and $\mathcal{G} = \{G_i : i = 1, 2, \ldots, n\}$.

(c) $f(\mathcal{E})$ is concave for the outer mixture of $\mathcal{E}$ in the sense that

$$f[c\mathcal{E} \cup (1-c)\mathcal{G}] \geqslant cf(\mathcal{E}) + (1-c)f(\mathcal{G}),$$

where $c \in [0, 1]$, and $c\mathcal{E} \cup (1-c)\mathcal{G} = \{cE_i : i = 1, 2, \ldots, n\} \cup \{(1-c)G_j : j = 1, 2, \ldots, m\}$ for $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ and $\mathcal{G} = \{G_j : j = 1, 2, \ldots, m\}$.

We sketch the proof. Clearly, item (a) follows readily from Proposition 1 [item (a)].

Item (b) follows from Proposition 1 [item (b)] as

$$\begin{aligned}
\|\boldsymbol{F}(c\mathcal{E} + (1-c)\mathcal{G})\|_1 &\geqslant \|c\boldsymbol{F}(\mathcal{E}) + (1-c)\boldsymbol{F}(\mathcal{G})\|_1 \\
&= c\|\boldsymbol{F}(\mathcal{E})\|_1 + (1-c)\|\boldsymbol{F}(\mathcal{G})\|_1.
\end{aligned}$$

The last equation is due to the fact that each entry of the matrix $\boldsymbol{F}(\mathcal{E})$ is of the same sign with the corresponding entry of the matrix $\boldsymbol{F}(\mathcal{G})$. In fact, all diagonal entries of both matrices $\boldsymbol{F}(\mathcal{E})$ and $\boldsymbol{F}(\mathcal{G})$ are non-negative, and all off-diagonal entries of both matrices are nonpositive.

To prove item (c), note that from Proposition 1 [item (c)], we have

$$\|\boldsymbol{F}(c\mathcal{E} \cup (1-c)\mathcal{G})\|_1 \geqslant \|c\boldsymbol{F}(\mathcal{E}) \oplus (1-c)\boldsymbol{F}(\mathcal{G})\|_1.$$

But for the l^1 norm, we have $\|X \oplus Y\|_1 = \|X\|_1 + \|Y\|_1$ and, consequently,

$$\|\boldsymbol{F}[c\mathcal{E} \cup (1-c)\mathcal{G}]\|_1 \geqslant c\|\boldsymbol{F}(\mathcal{E})\|_1 + (1-c)\|\boldsymbol{F}(\mathcal{G})\|_1,$$

which is the desired result.

In Ref. [37], the interesting entropic quantity (called device uncertainty)

$$D_\rho(\mathcal{E}) = \sum_{i=1}^n \sum_{k=1}^{m_i} \mathrm{tr}(\rho\Pi_{ik})h(\lambda_{ik})$$

was introduced to characterize the unsharpness of the measurement $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ via the spectral

decompositions

$$E_i = \sum_{k=1}^{m_i} \lambda_{ik} \Pi_{ik}$$

of all constituent effects E_i. Here, $h(\lambda_{ik}) = -\lambda_{ik} \ln \lambda_{ik}$ is the entropy function. This quantity has several nice properties. Due to the need for spectral decompositions of all the effects, it may be difficult to calculate. Our quantifiers of unsharpness, apart from sharing similar properties, are easier to calculate and directly accessible to experiments since only variancelike quantities are involved.

Let us illustrate unsharpness through some important measurements.

(1) *Weak measurement.* For the weak measurement $\mathcal{E}(t) = \{E_1(t), E_2(t)\}$ on a qubit system with $t \in (0, 1]$ and [48]

$$E_1(t) = \frac{1-t}{2}|0\rangle\langle 0| + \frac{1+t}{2}|1\rangle\langle 1|,$$
$$E_2(t) = \frac{1+t}{2}|0\rangle\langle 0| + \frac{1-t}{2}|1\rangle\langle 1|,$$

straightforward calculation yields

$$\boldsymbol{F}(\mathcal{E}(t)) = \frac{1-t^2}{4}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

We see that the above matrix is a decreasing (in the sense of matrix order) function of the parameter $t \in (0, 1]$, which captures the strength of the measurement. This is consistent with our physical intuition: The smaller t is, the farther away $\mathcal{E}(t)$ is from a Lüders measurement, and the larger the unsharpness is. In particular, for $t = 1$, the measurement is a Lüders measurement and the unsharpness vanishes. Moreover,

$$f(\mathcal{E}(t)) = \|\boldsymbol{F}(\mathcal{E}(t))\|_1 = 1 - t^2.$$

(2) *Equiangular POVM.* An equiangular tight frame with n elements is a set $\{|\psi_i\rangle : i = 1, 2, \ldots, n\}$ of n pure states (unit vectors) in a d-dimensional Hilbert space satisfying $|\langle\psi_i|\psi_j\rangle|^2 = a$ for $i \neq j$ and $\sum_i |\psi_i\rangle\langle\psi_i| = b\mathbf{1}$ for some constants $a \geqslant 0, b > 0$ [49–53]. Such a structure can be considered as equidistant points in a complex projective space. It is known that if such a frame exists, then

$$d \leqslant n \leqslant d^2, \quad a = \frac{n-d}{d(n-1)}, \quad b = \frac{n}{d}.$$

However, it may happen that equiangular tight frames with n elements fail to exist for certain $n \in [d, d^2]$. For instance, if $d = 3$, then there is no equiangular tight frame for $n = 5, 8$ [52].

Any equiangular tight frame $\{|\psi_i\rangle : i = 1, 2, \ldots, n\}$ induces naturally a measurement $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$, with

$$E_i = \frac{d}{n}|\psi_i\rangle\langle\psi_i|.$$

In this case, $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ is actually a homogenous rank-one POVM. We will call it an equiangular POVM (EA-POVM). In particular, when $n = d^2$, an EA-POVM is a symmetric, informationally complete measurement (SIC-POVM) [54–59]. In some sense, EA-POVMs are natural generalizations of von Neumann measurements.

From the definition of an equiangular tight frame, we readily obtain that

$$\boldsymbol{F}(\mathcal{E}) = \frac{n-d}{n^2}\mathbf{1} - \frac{n-d}{n^2(n-1)}\begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & 1 \\ 1 & 1 & \cdots & 0 \end{pmatrix},$$

where $\mathbf{1}$ is the identity matrix of order n. Consequently,

$$f(\mathcal{E}) = \|\boldsymbol{F}(\mathcal{E})\|_1 = 2\left(1 - \frac{d}{n}\right).$$

In particular, for $n = d$, we have $f(\mathcal{E}) = 0$, which corresponds to von Neumann measurements, and for $n = d^2$, we have

$$f(\mathcal{E}) = 2\left(1 - \frac{1}{d}\right),$$

which is the unsharpness of SIC-POVMs.

IV. CHARACTERIZING EA-POVM VIA MINIMAL UNSHARPNESS

Since, for general n, the existence of an EA-POVM with n effects is still an open issue, in this section, we only consider the case when an EA-POVM $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ with n effects E_i exists. Such a measurement is highly symmetric in the sense that both $\mathrm{tr}(E_i)$ and $\mathrm{tr}(E_i E_j)$ (for $i \neq j$) are constants. In fact,

$$\mathrm{tr}(E_i) = \frac{d}{n}, \quad \mathrm{tr}(E_i E_j) = \frac{n-d}{d(n-1)}, \quad i \neq j.$$

Due to their symmetry, one may expect that EA-POVMs have certain extreme properties. Indeed, we have the following result.

Proposition 3. EA-POVMs achieve the minimal value of unsharpness in the sense that for any homogenous POVM $\mathcal{E}$, it holds that

$$\|\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})\| \leqslant \|\boldsymbol{F}(\mathcal{E})\|, \tag{13}$$

where $\mathcal{E}^{\mathrm{EA}}$ denotes any EA-POVM with the same number of effects as that in $\mathcal{E}$, and $\|\cdot\|$ is any unitarily invariant norm of matrices.

To establish the above result, we invoke the theory of majorization [45,46,60]. Let $\mathcal{E}^{\mathrm{EA}} = \{E_i^{\mathrm{EA}} : i = 1, 2, \ldots, n\}$ have n effects, and for any real vector $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, let $x^{\downarrow} = (x_1^{\downarrow}, x_2^{\downarrow}, \ldots, x_n^{\downarrow})$ be the vector obtained by rearranging $\boldsymbol{x}$ in a nonincreasing order. Recall that the vector $\boldsymbol{x}$ weakly majorizes another vector $\boldsymbol{y}$, denoted by $\boldsymbol{x} \succ_{\mathrm{w}} \boldsymbol{y}$, means that [60]

$$\sum_{i=1}^{j} x_i^{\downarrow} \geqslant \sum_{i=1}^{j} y_i^{\downarrow}, \quad j = 1, 2, \ldots, n.$$

If, furthermore, $\sum_{i=1}^{n} x_i^{\downarrow} = \sum_{i=1}^{n} y_i^{\downarrow}$, then one says that $\boldsymbol{x}$ majorizes $\boldsymbol{y}$, denoted by $\boldsymbol{x} \succ \boldsymbol{y}$. Heuristically, if $\boldsymbol{x} \succ \boldsymbol{y}$, then the vector $\boldsymbol{y}$ is "flatter" (more "chaotic," more "spread out," and more "uniform") than $\boldsymbol{x}$.

Let $\boldsymbol{\lambda}(X)$ and $\boldsymbol{s}(X)$ denote the eigenvalue spectrum and singular value spectrum of a matrix X, respectively, arranged

as vectors in $\mathbb{R}^n$. By the properties of the EA-POVM $\mathcal{E}^{\mathrm{EA}} = \{E_i^{\mathrm{EA}} : i = 1, 2, \ldots, n\}$, we get the matrix entries

$$[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})]_{ij} = \begin{cases} \frac{n-d}{n^2}, & i = j \\ -\frac{n-d}{n^2(n-1)}, & i \neq j. \end{cases}$$

Consequently, the matrix $\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})$ has the spectrum

$$\boldsymbol{\lambda}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})] = \frac{n-d}{n(n-1)}(1, 1, \cdots, 1, 0) \in \mathbb{R}^n.$$

In particular,

$$\mathrm{tr}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})] = (n-1)\frac{n-d}{n(n-1)} = 1 - \frac{d}{n}.$$

For any homogenous POVM $\mathcal{E} = \{E_i : i = 1, 2, \ldots, n\}$ with n effects, let $\rho_i = nE_i/d$; then $\rho_i \geqslant 0$ and, from $\mathrm{tr}(E_i) = d/n$ (homogeneity), we know that $\mathrm{tr}(\rho_i) = 1$, which implies that $\mathrm{tr}(\rho_i^2) \leqslant 1$. Consequently,

$$\begin{aligned} \mathrm{tr}[\boldsymbol{F}(\mathcal{E})] &= 1 - \frac{1}{d}\sum_{i=1}^{n}\mathrm{tr}(E_i^2) = 1 - \frac{1}{d}\left(\frac{d}{n}\right)^2\sum_{i=1}^{n}\mathrm{tr}(\rho_i^2) \\ &\geqslant 1 - \frac{1}{d}\left(\frac{d}{n}\right)^2 n = 1 - \frac{d}{n} = \mathrm{tr}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})]. \end{aligned} \tag{14}$$

Therefore,

$$\begin{aligned} \boldsymbol{\lambda}[\boldsymbol{F}(\mathcal{E})] &\succ \frac{\mathrm{tr}[\boldsymbol{F}(\mathcal{E})]}{n-1}(1, 1, \cdots, 1, 0) \\ &\succ_{\mathrm{w}} \frac{\mathrm{tr}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})]}{n-1}(1, 1, \cdots, 1, 0) \\ &= \boldsymbol{\lambda}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})]. \end{aligned}$$

Here, the above "$\succ$" follows from the fact that $\boldsymbol{F}(\mathcal{E})$ always has a zero eigenvalue with eigenvector $(1, 1, \ldots, 1) \in \mathbb{R}^n$ since

$$\sum_j [\boldsymbol{F}(\mathcal{E})]_{ij} = \frac{1}{d}\left[\mathrm{tr}(E_i) - \sum_j \mathrm{tr}(E_i E_j)\right] = 0, \quad \forall\ j.$$

The above "$\succ_{\mathrm{w}}$" follows from inequality (14). Because the real matrix $\boldsymbol{F}(\cdot)$ is non-negative definite, we have $\boldsymbol{\lambda}[\boldsymbol{F}(\cdot)] = \boldsymbol{s}[\boldsymbol{F}(\cdot)]$. Consequently, the above weak majorization relation is just $\boldsymbol{s}[\boldsymbol{F}(\mathcal{E})] \succ_{\mathrm{w}} \boldsymbol{s}[\boldsymbol{F}(\mathcal{E}^{\mathrm{EA}})]$. Since any unitarily invariant norm is monotone with respect to the partial order induced by weak majorization of the vectors of singular values of the matrices [45], the desired Eq. (13) follows.

V. DISCUSSION

Given the increasing interest and usage of general measurements described by POVMs in both foundational and experimental aspects of quantum mechanics, in particular in quantum information science, it is desirable to characterize how a POVM deviates from a traditional Lüders measurement. While the latter is usually referred to as sharp, the former is termed unsharp (fuzzy) if it is not a Lüders measurement. Thus a natural question arises as to how to quantify the degree of unsharpness of a POVM. Following previous studies in Refs. [14,30–37], we have explored this issue in terms of the uncertainty of random variables and quantum observables induced by a POVM. We have quantified the unsharpness of a discrete POVM (measurement) at two levels: matrix level and numerical level. Unitarily invariant norms and majorization are exploited to produce various quantities of unsharpness. Our quantifiers of unsharpness are intrinsic quantities depending only on the original POVM in a rather simple way. They are easy to compute since they involve only pairwise overlaps between the effects in the POVM. Furthermore, a characterization of equiangular POVMs as extreme measurements in terms of unsharpness is naturally obtained.

It remains to investigate the theoretical implications and experimental applications of unsharpness in characterizing the structure of POVMs. In particular, classifying and ordering POVMs according to certain measures of unsharpness may be useful in revealing information-theoretic aspects of quantum measurements.

We have not addressed the issue of the operational meaning of the quantifiers of unsharpness, which is an important problem under further investigation. Since these quantifiers are all expressed as the difference of the variance of the observables derived from POVMs, some hints of operational significance may be read from that of the variance. Of course, it is more relevant to reveal their operational meaning in quantum foundations and quantum information tasks.

It is well known that by the celebrated Naimark dilation [61–63], any POVM can be dilated to a Lüder measurement in a higher-dimensional ambient system. Consequently, the relation between dilation and unsharpness is of basic importance for further investigations. Some studies of applications of Naimark dilation in characterizing and quantifying entanglement are presented in Refs. [64,65].

We have worked only in finite-dimensional systems and discrete POVMs. The most general setup of measurements in quantum theory is formulated in terms of operations and instruments [13,14,18], which is playing an increasingly important role in the foundational studies of quantum theory. It is desirable to extend the results to infinite-dimensional quantum systems, such as continuous systems in quantum optics. It seems this is highly nontrivial. We note that a fundamental difference between finite dimensions and infinite dimensions is that the maximally mixed states (proportional to the identity operators) only exist in finite dimensions, while our present approach to get a quantifier independent of the states relies on the existence of the maximally mixed states. This shows that a naive and direct generalization from finite dimensions to infinite dimensions is impossible, but new methods and ideas are required.

ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China, Grant No. 2020YFA0712700, the National Natural Science Foundation of China, Grants No. 11875317 and No. 61833010, and the ORIC project from TEEP.

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1958).

[2] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).

[3] W. Pauli, *General Principles of Quantum Mechanics* (Springer, Berlin, 1980).

[4] G. Lüders, Über die zustandstiderung durch den messprozess, Ann. Phys. (Leipzig) **8**, 322 (1951).

[5] E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).

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

[7] S. P. Gudder, *Stochastic Methods in Quantum Mechanics* (North Holland, Amsterdam, 1978).

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

[9] K. Kraus, *States, Effects, and Operations* (Springer, Berlin, 1983).

[10] P. Busch, M. Grabowski, and P. Lahti, *Operational Quantum Physics* (Springer, Berlin, 1995).

[11] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, 1995).

[12] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[13] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory: From Uncertainty to Entanglement* (Cambridge University Press, Cambridge, 2012).

[14] P. Busch, P. Lahti, J.-P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer, Berlin, 2016).

[15] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories, Z. Phys. **181**, 233 (1964).

[16] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories, II, Commun. Math. Phys. **4**, 331 (1967).

[17] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories, III, Commun. Math. Phys. **9**, 12 (1968).

[18] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).

[19] B. Mielnik, Generalized quantum mechanics, Commun. Math. Phys. **37**, 221 (1974).

[20] S. T. Ali and G. G. Emch, Fuzzy observables in quantum mechanics, J. Math. Phys. **15**, 176 (1974).

[21] G. Ludwig, *Foundations of Quantum Mechanics*, Vol. I (Springer, New York, 1983); *ibid.,* Vol. II (Springer, New York, 1985).

[22] P. Busch, P. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, 2nd ed. (Springer, Berlin, 1996).

[23] S. Gudder, Sharp and unsharp quantum effects, Adv. Appl. Math. **20**, 169 (1998).

[24] P. Busch and S. Gudder, Effects as functions on projective Hilbert space, Lett. Math. Phys. **47**, 329 (1999).

[25] T. Heinonen, P. Lahti, and K. Ylinen, Covariant fuzzy observables and coarse-graining, Rep. Math. Phys. **53**, 425 (2004).

[26] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[27] Y. G. Liang, R. W. Spekkens, and H. M. Wiseman, Specker's parable of the over-protective seer: A road to contextuality, nonlocality and complementarity, Phys. Rep. **506**, 1 (2011).

[28] S. T. Ali, C. Carmeli, T. Heinosaari, and A. Toigo, Commutative POVMs and fuzzy observables, Found. Phys. **39**, 593 (2009).

[29] T. Heinosaari and J.-P. Pellonpää, Extreme commutative quantum observables are sharp, J. Phys. A **44**, 315303 (2011).

[30] M. J. W. Hall, Algebra for generalized quantum observables, arXiv:quant-ph/0302007.

[31] P. Busch, T. Heinonen, and P. Lahti, Noise and disturbance in quantum measurement, Phys. Lett. A **320**, 261 (2004).

[32] M. Ozawa, Uncertainty relations for joint measurements of noncommuting observables, Phys. Lett. A **320**, 367 (2004).

[33] S. Massar, Uncertainty relations for positive-operator-valued measures, Phys. Rev. A **76**, 042114 (2007).

[34] P. Busch, T. Heinonen, P. Lahti, Heisenberg's uncertainty principle, Phys. Rep. **452**, 155 (2007).

[35] P. Busch, P. Lahti, and R. F. Werner, Quantum root-mean-square error and measurement uncertainty relations, Rev. Mod. Phys. **86**, 1261 (2014).

[36] C. Carmeli, T. Heinonen, and A. Toigo, Intrinsic unsharpness and approximate repeatability of quantum measurements, J. Phys. A **40**, 1303 (2007).

[37] K. Baek and W. Son, Unsharpness of generalized measurement and its effects in entropic uncertainty relations, Sci. Rep. **6**, 30228 (2016).

[38] Y. Sun and S. Luo, Coherence as uncertainty, Phys. Rev. A **103**, 042423 (2021).

[39] Y. Zhang and S. Luo, Quantum states as observables: Their variance and nonclassicality, Phys. Rev. A **102**, 062211 (2020).

[40] L. Henderson and V. Vedral, Classical, quantum and total correlations, J. Phys. A **34**, 6899 (2001).

[41] H. Ollivier and W. H. Zurek, Quantum Discord: A Measure of the Quantumness of Correlations, Phys. Rev. Lett. **88**, 017901 (2001).

[42] S. Luo, Quantum discord for two-qubit systems, Phys. Rev. A **77**, 042303 (2008).

[43] S. Luo, Statistics of local value in quantum mechanics, Intl. J. Theor. Phys. **41**, 1713 (2002).

[44] S. Luo, Quantum versus classical uncertainty, Theor. Math. Phys. **143**, 681 (2005).

[45] R. A. Horn, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, 1991).

[46] R. Bhatia, *Matrix Analysis* (Springer, Berlin, 2013).

[47] G. W. Stewart and J. Sun, *Matrix Perturbation Theory* (Academic, Boston, 1990).

[48] O. Oreshkov and T. A. Brun, Weak Measurements are Universal, Phys. Rev. Lett. **95**, 110409 (2005).

[49] P. W. H. Lemmens and J. J. Seidel, Equiangular lines, J. Algebra **24**, 494 (1973).

[50] G. Zauner, Quantum designs: Foundations of a noncommutative theory of designs, Ph.D. thesis, University of Vienna, 1999.

[51] J. M. Renes, Equiangular tight frames from Paley tournaments, Linear Algebra Appl. **426**, 497 (2007).

[52] F. Szöllösi, All complex equiangular tight frames in dimension 3, arXiv:1402.6429v1.

[53] S. F. Waldron, *An Introduction to Finite Tight Frames* (Springer, Berlin, 2018).

[54] A. J. Scott, Tight informationally complete quantum measurements, J. Phys. A **39**, 13507 (2006).

[55] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, Symmetric informationally complete quantum measurements, J. Math. Phys. **45**, 2171 (2004).

[56] D. M. Appleby, C. A. Fuchs, and H. Zhu, Group theoretic, Lie algebraic and Jordan algebraic formulations of the SIC existence problem, Quantum Inf. Comput. **15**, 0061 (2015).

[57] C. A. Fuchs, M. C. Hoang, and B. C. Stacey, The SIC question: History and state of play, Axioms **6**, 21 (2017).

[58] J. B. DeBrota and B. C. Stacey, Lüders channels and the existence of symmetric-informationally complete measurements, Phys. Rev. A **100**, 062327 (2019).

[59] J. B. DeBrota, C. A. Fuchs, and B. C. Stacey, Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems, Phys. Rev. Res. **2**, 013074 (2020).

[60] A. W. Marshall, *Inequalities: Theory of Majorization and Its Applications* (Springer, New York, 2011).

[61] M. A. Naimark, Spectral functions of a symmetric operator, Izv. Akad. Nauk SSSR Ser. Mater. **4**, 277 (1940).

[62] M. A. Naimark, On a representation of additive operator set functions, Dokl. Akad. Nauk SSSR **41**, 359 (1943).

[63] N. I. Akhiezer and I. M. Glazman, *Theory of Linear Operators in Hilbert Space* (Dover, New York, 1963).

[64] N. Li and S. Luo, Classical states versus separable states, Phys. Rev. A **78**, 024303 (2008).

[65] S. Luo, Entanglement as minimal discord over state extensions, Phys. Rev. A **94**, 032129 (2016).