

State-Independent Error-Disturbance Tradeoff For Measurement Operators

S. S. Zhou,¹ Shengjun Wu,¹ and H. F. Chau²

¹*Kuang Yaming Honors School, Nanjing University, Nanjing, Jiangsu 210093, China*

²*Department of Physics, The University of Hong Kong, Pokfulam Road, Hong Kong*

In general, classical measurement statistics of a quantum measurement is disturbed by performing an additional incompatible quantum measurement beforehand. Using this observation, we introduce a state-independent definition of disturbance by relating it to the distinguishability problem between two classical statistical distributions — one resulting from a single quantum measurement and the other from a succession of two quantum measurements. Interestingly, we find an error-disturbance trade-off relation for any measurements in two-dimensional Hilbert space and for measurements with mutually unbiased bases in any finite-dimensional Hilbert space. This relation shows that error should be reduced to zero in order to minimize the sum of error and disturbance. We conjecture that a similar trade-off relation can be generalized to any measurements in an arbitrary finite-dimensional Hilbert space.

I. INTRODUCTION

The uncertainty principle has been regarded as a fundamental principle in quantum mechanics. It asserts that we cannot get the precise values of two physical observables in a quantum state, unless they are compatible. The well-known version of this principle was formulated by Heisenberg in 1927, namely, [1]

$$\Delta x \Delta p \geq \hbar/2. \quad (1)$$

A more general form of it can be written as

$$\varepsilon(A)\eta(B) \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}, \quad (2)$$

where $\varepsilon(A)$ is the error with which measurement of operator A is carried out, and $\eta(B)$ is the disturbance on the following measurement of operator B caused by measurement of A . In Eq. (1), Δx and Δp can be interpreted as error and disturbance when A and B are position and momentum operators. Mathematically, Eq. (2) comes from the Robertson's uncertainty relation [2]:

$$\sigma(A)\sigma(B) \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}, \quad (3)$$

where $\sigma(X) \equiv \langle \psi | X^2 | \psi \rangle - \langle \psi | X | \psi \rangle^2$ is the standard deviation of observable X in a quantum state $|\psi\rangle$. Note that while Eq. (3), usually regarded as a rigorous version of Heisenberg's uncertainty principle [3–5], can be proven mathematically, the justification of relation Eq. (2) is currently on hot debate because additional conditions have been used in its derivation [6]. More importantly, several experiments showed that Eq. (2) is violated [7–9]. Thus, the trade-off relation that the higher the precision of measuring A , the stronger the disturbance on B cannot be well-captured by Eq. (2).

Many important works in this area have been done, but the definitions of error and disturbance are still not settled [10]. Ozawa used noise-operator based error definition and proposed a “universally valid error-disturbance

relation” [11]:

$$\varepsilon(A)\eta(B) + \varepsilon(A)\sigma(B) + \sigma(A)\eta(B) \geq \frac{|\langle \psi | [A, B] | \psi \rangle|}{2}. \quad (4)$$

This uncertainty relation was later verified experimentally [7–9, 12–15] and inspired a lot of work on uncertainty relations [16–18], but some shortcomings were also pointed out [19, 20]. For example, it seems to violate the proposed operational constraint that the error and disturbance should be non-zero if the outcome distribution is deviated from what is expected according to Born's rule [19].

Using distance between distributions is another way to quantifying measurement errors [21–23], Busch *et al.* proved the original Heisenberg's error-disturbance relation in Eq. (2) by defining error and disturbance as figures of merit characteristic of the measuring devices [21, 24], which generated a debate over different approaches used in formalizing uncertainty relations [25, 26].

In this paper, we introduce a straightforward definition of error and disturbance. The key observation is that given an arbitrary quantum state, the measurement statistics of a measurement operation on that state is unchanged if and only if we perform an additional compatible measurement to the state beforehand. Thus, we may define the disturbance of \mathcal{B} (measurement of operator B) due to \mathcal{A} (measurement of operator A) as the distance between the two probability distributions of the measurement outcomes due to \mathcal{B} and $\mathcal{B} \circ \mathcal{A}$ maximized over all possible input quantum states. We introduce the definitions of error and disturbance in Sec. II and report a few basic properties of these quantities in Sec. III. Then, in Sec. IV, we prove the error-disturbance trade-off relation for the case of 2-dimensional Hilbert space. In particular, we derive a sharp lower bound of the sum of error and disturbance. We also give the trade-off relation in d -dimensional Hilbert space for a special but important case. Finally, we draw a few conclusions in Sec. V.

II. STATE-INDEPENDENT ERROR AND DISTURBANCE OF PROJECTIVE MEASUREMENTS

Suppose one is given a density matrix ρ in a d -dimensional Hilbert space with $d \geq 2$. Let \mathcal{A} be the projective measurements of operator A with rank-one projectors. (Unless otherwise stated, all measurements in this paper are associated with rank-one projectors. Note that our discussion can be easily extended to the case of a general positive operator-valued measurement. We restrain from doing so to avoid unnecessary notational and indexing complications.) The probability distribution obtained from applying \mathcal{A} to ρ is given by the vector

$$P_{\mathcal{A}}(\rho) = (p_i^{(\mathcal{A})}(\rho))_{i=1}^d \equiv (\langle a_i | \rho | a_i \rangle)_{i=1}^d, \quad (5)$$

where $|a_i\rangle\langle a_i|$ is the rank-one projector corresponding to the i^{th} measurement outcome. We now consider measuring ρ using another projective measurement \mathcal{A}' before feeding the resultant state to \mathcal{B} . We write the probability distribution of the measurement outcomes of \mathcal{A}' by $P_{\mathcal{A}'}(\rho)$. More importantly, the probability distribution of the final measurement outcomes of $\mathcal{B} \circ \mathcal{A}'$ is given by $P_{\mathcal{B} \circ \mathcal{A}'}(\rho) = P_{\mathcal{B}}(\rho')$ where $\rho' = \sum_i \langle a'_i | \rho | a'_i \rangle |a'_i\rangle\langle a'_i|$ with $|a'_i\rangle\langle a'_i|$ being the rank-one projector corresponding to the i^{th} measurement outcome of \mathcal{A}' .

In general, $P_{\mathcal{B} \circ \mathcal{A}'}(\rho)$ is different from $P_{\mathcal{B}}(\rho)$ as measurement changes the state of a quantum system. We would like to know how a change in the intermediate measurement \mathcal{A}' affects the change of $P_{\mathcal{B}}(\rho)$ through their classical statistics of their measurement outcomes only. With this motivation in mind, for any given metric $D(\cdot, \cdot)$ of an Euclidean space, we define the **state-dependent error** between $P_{\mathcal{A}}(\rho)$ and $P_{\mathcal{A}'}(\rho)$, and the **state-dependent disturbance** between $P_{\mathcal{B}}(\rho)$ and $P_{\mathcal{B}}(\rho')$ as

$$\varepsilon_{\rho}(\mathcal{A}, \mathcal{A}') = D(P_{\mathcal{A}}(\rho), P_{\mathcal{A}'}(\rho)) \quad (6)$$

and

$$\eta_{\rho}(\mathcal{A}', \mathcal{B}) = D(P_{\mathcal{B}}(\rho), P_{\mathcal{B}}(\rho')), \quad (7)$$

respectively. Since our goal is to study the maximum pointwise deviation in the distribution of the measurement outcomes, we use the metric based on the infinity norm, namely,

$$D(x, y) = \max_i |x_i - y_i|. \quad (8)$$

We now define the **state-independent error** and **state-independent disturbance** by

$$\varepsilon(\mathcal{A}, \mathcal{A}') = \max_{\rho} \varepsilon_{\rho}(\mathcal{A}, \mathcal{A}') \quad (9)$$

and

$$\eta(\mathcal{A}', \mathcal{B}) = \max_{\rho} \eta_{\rho}(\mathcal{A}', \mathcal{B}). \quad (10)$$

From now on, the terms ‘‘error’’ and ‘‘disturbance’’ refer to the state-independent versions unless otherwise stated. Note that these definitions meet the proposed operational constraint [19] for $\varepsilon(\mathcal{A}, \mathcal{A}') = 0$ if and only if $\mathcal{A} = \mathcal{A}'$ and $\eta(\mathcal{A}', \mathcal{B}) = 0$ if and only if $\mathcal{A}' = \mathcal{B}$.

Finally, to obtain a trade-off relation between error and disturbance in one measurement, that is, to find out how much we need to sacrifice on one to lower the other, just as what Heisenberg did, we introduce the **state-independent overall error**

$$\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) = \max_{\rho} (\varepsilon_{\rho}(\mathcal{A}, \mathcal{A}') + \eta_{\rho}(\mathcal{A}', \mathcal{B})). \quad (11)$$

Clearly, $\varepsilon + \eta \geq \Delta$.

III. BASIC PROPERTIES OF STATE-INDEPENDENT ERROR AND DISTURBANCE

Since

$$\begin{aligned} \varepsilon_{\rho}(\mathcal{A}, \mathcal{A}') &= \max_i |\text{tr}(\rho(|a_i\rangle\langle a_i| - |a'_i\rangle\langle a'_i|))| \\ \eta_{\rho}(\mathcal{A}', \mathcal{B}) &= \max_i |\text{tr}(\rho(|b_i\rangle\langle b_i| - \sum_j |a'_j\rangle\langle a'_j|)|^2 \langle a'_j|))|, \end{aligned} \quad (12)$$

we have

$$\varepsilon(\mathcal{A}, \mathcal{A}') = \max_i R(|a_i\rangle\langle a_i| - |a'_i\rangle\langle a'_i|) \quad (13)$$

and

$$\eta(\mathcal{A}', \mathcal{B}) = \max_i R(|b_i\rangle\langle b_i| - \sum_k |\langle b_i | a'_k \rangle|^2 |a'_k\rangle\langle a'_k|). \quad (14)$$

Here, $R(\cdot)$ is the spectral radius of a matrix (the largest of absolute values of the eigenvalues). Similarly, we have

$$\begin{aligned} \Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) &= \max_{i,j,\pm} R(|a_i\rangle\langle a_i| - |a'_i\rangle\langle a'_i| \pm \\ &|b_j\rangle\langle b_j| \mp \sum_k |\langle b_j | a'_k \rangle|^2 |a'_k\rangle\langle a'_k|). \end{aligned} \quad (15)$$

Note that the maximum of ε_{ρ} can be attained by a pure state ρ ; and similarly for η_{ρ} and $\varepsilon_{\rho} + \eta_{\rho}$.

Property 1 (Maximum Error and Disturbance). *The error satisfies*

$$\varepsilon(\mathcal{A}, \mathcal{A}') \leq 1 \quad (16)$$

for any $\mathcal{A}, \mathcal{A}'$, with equality if and only if $\langle a'_i | a_i \rangle = 0$ for some i . In addition, the disturbance obeys

$$\eta(\mathcal{A}', \mathcal{B}) \leq 1 - 1/d; \quad (17)$$

for any $\mathcal{A}', \mathcal{B}$, with equality if and only if there is an i such that $|b_i\rangle$ is unbiased in $\{|a'_j\rangle\}_{j=1}^d$. That is to say, $|\langle a'_j | b_i \rangle|^2 = 1/d$ for any j .

Proof. The rank of the matrix $|a_i\rangle\langle a_i| - |a'_i\rangle\langle a'_i|$ is at most 2. Thus, the spectral radius of this matrix can be calculated easily as $|a_i\rangle$ and $|a'_i\rangle - \langle a_i|a'_i\rangle|a_i\rangle$ are orthogonal. Hence, Eq. (13) becomes

$$\varepsilon(\mathcal{A}, \mathcal{A}') = \max_i \sqrt{1 - |\langle a'_i|a_i\rangle|^2}. \quad (18)$$

Consequently, $\varepsilon(\mathcal{A}, \mathcal{A}') \leq 1$ with equality holds when there exists i such that $\langle a'_i|a_i\rangle = 0$.

Eq. (14) can be rewritten as

$$\eta(\mathcal{A}', \mathcal{B}) = \max_{\psi, i} \left| |\langle \psi|b_i\rangle|^2 - \sum_k |\langle \psi|a'_k\rangle|^2 |\langle b_i|a'_k\rangle|^2 \right|. \quad (19)$$

Note that the spectral radius of the matrix $|b_i\rangle\langle b_i| - \sum_k |\langle b_i|a'_k\rangle|^2 |a'_k\rangle\langle a'_k|$ is unchanged by multiplying a phase $e^{i\delta}$ to $|a'_k\rangle$. Hence, we may always set each $\langle b_i|a'_k\rangle$ to be real and non-negative. Then by the Perron-Frobenius theorem [27], there is always a positive eigenvalue corresponding to the spectral radius. In other words,

$$\eta(\mathcal{A}', \mathcal{B}) = \max_{\psi, i} |\langle \psi|b_i\rangle|^2 - \sum_k |\langle \psi|a'_k\rangle|^2 |\langle b_i|a'_k\rangle|^2. \quad (20)$$

Using the Cauchy-Schwarz inequality,

$$\begin{aligned} |\langle \psi|b_i\rangle|^2 &\leq \left(\sum_k |\langle \psi|a'_k\rangle| |\langle a'_k|b_i\rangle| \right)^2 \\ &\leq \sum_k |\langle \psi|a'_k\rangle|^2 |\langle b_i|a'_k\rangle|^2 \sum_k 1^2. \end{aligned} \quad (21)$$

So $|\langle \psi|a'_k\rangle|^2 |\langle b_i|a'_k\rangle|^2 \geq |\langle \psi|b_i\rangle|^2 / d$. Hence, Eq. (17) is true. Moreover, the equality holds if and only if there exists an i such that $|\langle \psi|b_i\rangle| = 1$ and $|\langle b_i|a'_1\rangle|^2 = |\langle b_i|a'_2\rangle|^2 = \dots = |\langle b_i|a'_d\rangle|^2$. In other words, $|b_i\rangle$ is unbiased in $\{|a'_k\rangle\}_{k=1}^d$. \square

Property 1 can be understood as follows. Suppose we take $\rho = |a_i\rangle\langle a_i|$, the measurement operation \mathcal{A} will give a post-measurement state $|a_i\rangle$ for sure. In contrast, when we perform the measurement \mathcal{A}' , we have no chance of getting the state $|a'_i\rangle$ if $|a_i\rangle$ and $|a'_i\rangle$ are orthogonal. This is a situation in which the error between the two measurement outcomes of \mathcal{A} and \mathcal{A}' on the same state ρ is maximized. In addition, suppose there is an i such that $|b_i\rangle$ is unbiased in $\{|a'_j\rangle\}_{j=1}^d$. Then by choosing $\rho = |b_i\rangle\langle b_i|$, the measurement operation \mathcal{A}' will completely erase the information on the distribution of ρ in $\{|b_j\rangle\}_{j=1}^d$. Hence, the disturbance $\eta(\mathcal{A}', \mathcal{B})$ is maximized.

Property 2 (Reducibility). *Suppose A, A' and B are operators on a Hilbert space $\mathbf{H} = \mathbf{H}_1 \oplus \mathbf{H}_2 \oplus \dots \oplus \mathbf{H}_l$ such that $|a_i\rangle$'s and $|a'_i\rangle$'s are in $\mathbf{H}_1 \cup \mathbf{H}_2 \cup \dots \cup \mathbf{H}_l$, then*

$$\varepsilon(\mathcal{A}, \mathcal{A}') = \max_{1 \leq k \leq l} \varepsilon(\mathcal{A}|_{\mathbf{H}_k}, \mathcal{A}'|_{\mathbf{H}_k}). \quad (22)$$

If we substitute $|b_i\rangle$'s for $|a_i\rangle$'s in above conditions, we have similarly

$$\eta(\mathcal{A}', \mathcal{B}) = \max_{1 \leq k \leq l} \eta(\mathcal{A}'|_{\mathbf{H}_k}, \mathcal{B}|_{\mathbf{H}_k}). \quad (23)$$

Proof. We only prove Eq. (23) as the proof of Eq. (22) is similar. Let $\mathbf{B}_k = \text{Span}\{|b_i\rangle : |b_i\rangle \in \mathbf{H}_k\}$. Obviously, $\mathbf{B}_k \subseteq \mathbf{H}_k$. And since $\mathbf{H} = \mathbf{B}_1 \oplus \mathbf{B}_2 \oplus \dots \oplus \mathbf{B}_l$, we know that $\mathbf{H}_k = \mathbf{B}_k$. Therefore, for any $|b_i\rangle, |a'_j\rangle$ not in same \mathbf{H}_k , $\langle b_i|a'_j\rangle = 0$. Suppose $|b_i\rangle \in \mathbf{B}_{k_i} = \mathbf{H}_{k_i}$, Eq. (14) becomes

$$\begin{aligned} &\eta(\mathcal{A}', \mathcal{B}) \\ &= \max_i R(|b_i\rangle\langle b_i| - \sum_{|a'_j\rangle \in \mathbf{H}_{k_i}} |\langle b_i|a'_j\rangle|^2 |a'_j\rangle\langle a'_j|) \\ &= \max_{k, i \in \{i' | k_{i'}=k\}} R(|b_i\rangle\langle b_i| - \sum_{|a'_j\rangle \in \mathbf{H}_{k_i}} |\langle b_i|a'_j\rangle|^2 |a'_j\rangle\langle a'_j|) \\ &= \max_k \eta(\mathcal{A}'|_{\mathbf{H}_k}, \mathcal{B}|_{\mathbf{H}_k}). \end{aligned} \quad (24)$$

Thus, Eq. (23) holds. \square

Property 3 (Subsystems). *Suppose A, A' and B are operators on the Hilbert space $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2 \otimes \dots \otimes \mathbf{H}_l$, then*

$$\varepsilon(\mathcal{A}, \mathcal{A}') \geq \max_{1 \leq k \leq l} \varepsilon(\mathcal{A}|_{\mathbf{H}_k}, \mathcal{A}'|_{\mathbf{H}_k}) \quad (25)$$

and

$$\eta(\mathcal{A}', \mathcal{B}) \geq \max_{1 \leq k \leq l} \eta(\mathcal{A}'|_{\mathbf{H}_k}, \mathcal{B}|_{\mathbf{H}_k}). \quad (26)$$

Proof. Suppose $l = 2$. Then, $|a'_j\rangle$ can be represented by $|a'^1_{m'}\rangle \otimes |a'^2_{n'}\rangle$, where $|a'^1_{m'}\rangle \in \mathbf{H}_1$ and $|a'^2_{n'}\rangle \in \mathbf{H}_2$. Similarly, we write $|b_i\rangle = |b^1_m\rangle \otimes |b^2_n\rangle$. Then,

$$\begin{aligned} &R(|b_i\rangle\langle b_i| - \sum_j |\langle b_i|a'_j\rangle|^2 |a'_j\rangle\langle a'_j|) \\ &\geq \max_{|\psi_1\rangle, |\psi_2\rangle} \left| |\langle \psi_1|b^1_m\rangle|^2 |\langle \psi_2|b^2_n\rangle|^2 \right. \\ &\quad \left. - \sum_{m', n'} \left| |\langle \psi_1|a'^1_{m'}\rangle|^2 |\langle \psi_2|a'^2_{n'}\rangle|^2 |\langle b^1_m|a'^1_{m'}\rangle|^2 |\langle b^2_n|a'^2_{n'}\rangle|^2 \right| \right|^2 \\ &\geq \max_{|\psi_1\rangle} \left| |\langle \psi_1|b^1_m\rangle|^2 - \sum_{m', n'} \left| |\langle \psi_1|a'^1_{m'}\rangle|^2 |\langle b^2_n|a'^2_{n'}\rangle|^4 |\langle b^1_m|a'^1_{m'}\rangle|^2 \right| \right|^2 \\ &\geq \max_{|\psi_1\rangle} \left| |\langle \psi_1|b^1_m\rangle|^2 - \sum_{m'} \left| |\langle \psi_1|a'^1_{m'}\rangle|^2 |\langle b^1_m|a'^1_{m'}\rangle|^2 \right| \right|^2 \\ &\geq R(|b^1_m\rangle\langle b^1_m| - \sum_{m'} |\langle b^1_m|a'^1_{m'}\rangle|^2 |a'^1_{m'}\rangle\langle a'^1_{m'}|). \end{aligned} \quad (27)$$

We conclude that $\eta(\mathcal{A}, \mathcal{B}) \geq \eta(\mathcal{A}|_{\mathbf{H}_1}, \mathcal{B}|_{\mathbf{H}_1})$. By induction of l , we show the validity of Eq. (26). Eq. (25) can be proven in a similar way. \square

Property 4 (Upper bound Estimation).

$$\eta(\mathcal{A}', \mathcal{B}) \leq \max_i \left[\left(1 - \frac{1}{d} \right) \left(1 - \sum_k |\langle a'_k|b_i\rangle|^4 \right) \right]^{1/2}. \quad (28)$$

The equality holds when $d = 2$. When $d > 2$, the equality holds if and only if there exists an i such that $|b_i\rangle$ is unbiased in $\{|a'_j\rangle\}_{j=1}^d$, or $\mathcal{A}' = \mathcal{B}$. Furthermore,

$$\eta(\mathcal{A}', \mathcal{B}) \leq \max_{i,j} |\langle a'_j | b_i \rangle| \sum_{k \neq j} |\langle a'_k | b_i \rangle|. \quad (29)$$

The equality holds when $d = 2$. When $d > 2$, the equality holds if and only if for an i which maximizes $\max_j |\langle a'_j | b_i \rangle| \sum_{k \neq j} |\langle a'_k | b_i \rangle|$, for all j, k such that $\langle a'_{j,k} | b_i \rangle \neq 0$, $|\langle a'_j | b_i \rangle| = |\langle a'_k | b_i \rangle|$.

Proof. We first prove Eq. (28). Suppose the eigenvalues of the matrix $M = |b_i\rangle \langle b_i| - \sum_{j=1}^d |\langle b_i | a'_j \rangle|^2 |a'_j\rangle \langle a'_j|$ be λ_j 's. Arrange these eigenvalues so that $|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_d| = \eta(\mathcal{A}', \mathcal{B})$. Since $\text{tr}(M) = 0$, we have $\sum_{i=1}^d \lambda_i = 0$. Moreover,

$$\begin{aligned} \text{tr}(M^2) &= \lambda_d^2 + \sum_{i=1}^{d-1} \lambda_i^2 \geq \lambda_d^2 + \frac{1}{d-1} \left(\sum_{i=1}^{d-1} \lambda_i \right)^2 \\ &= \frac{d\lambda_d^2}{d-1}, \end{aligned} \quad (30)$$

and $\text{tr}(M^2) = 1 - \sum_k |\langle a'_k | b_i \rangle|^4$. Hence, Eq. (28) is true. Moreover, the equality holds if and only if $\lambda_1 = \lambda_2 = \dots = \lambda_{d-1} = -\lambda_d/(d-1)$. A straightforward calculation shows that if $|b_i\rangle$ is unbiased in $\{|a'_j\rangle\}_{j=1}^d$ or there exists a j such that $|\langle a'_j | b_i \rangle| = 1$, the above equations hold.

Next, we prove Eq. (29). According to the Frobenius theorem [28, 29],

$$|\lambda_d| \leq \max_j \sum_{k=1}^d |M_{kj}| = \max_j |\langle a'_j | b_i \rangle| \sum_{k \neq j} |\langle a'_k | b_i \rangle|. \quad (31)$$

And this equality holds if and only if $d = 2$, or for all j, k such that $\langle a'_{j,k} | b_i \rangle \neq 0$, $|\langle a'_j | b_i \rangle| = |\langle a'_k | b_i \rangle|$. \square

From Eq. (6)–Eq. (8), we have

$$\eta^c(\mathcal{A}', \mathcal{B}) \equiv \max_{\rho=|b_i\rangle\langle b_i|} \eta_\rho(\mathcal{A}', \mathcal{B}) = \max_i \left(1 - \sum_k |\langle a'_k | b_i \rangle|^4 \right), \quad (32)$$

where $\eta^c(\mathcal{A}', \mathcal{B})$ is the so-called calibration disturbance [10], which measures the disturbance with respect to eigenvectors of B as original states. Moreover, it can also be interpreted as the maximum among probabilities that an original eigenstate $|b_i\rangle$ evolves into another eigenstate $|b_j\rangle$ ($j \neq i$) after $\mathcal{B} \circ \mathcal{A}'$. Calibration disturbance is another way to define a state-independent disturbance [21]. We already know from the definitions of worst-case disturbance and calibration disturbance obeys $\eta^c(\mathcal{A}', \mathcal{B}) \leq \eta(\mathcal{A}', \mathcal{B})$. What Eq. (28) shows is that

$$\eta(\mathcal{A}', \mathcal{B}) \leq \left[\left(1 - \frac{1}{d} \right) \cdot \eta^c(\mathcal{A}', \mathcal{B}) \right]^{1/2}. \quad (33)$$

This relation provides an upper bound of worst-case disturbance $\eta(\mathcal{A}', \mathcal{B})$ – the geometric mean of $\eta^c(\mathcal{A}', \mathcal{B})$ and

the maximum possible disturbance for measurements in d -dimensional Hilbert space, namely, $1 - 1/d$. Incidentally, for $\varepsilon(\mathcal{A}, \mathcal{A}')$, we have

$$\varepsilon(\mathcal{A}, \mathcal{A}') = [\varepsilon^c(\mathcal{A}, \mathcal{A}')]^{1/2}, \quad (34)$$

where $\varepsilon^c(\mathcal{A}', \mathcal{A})$ is defined analogously.

In Fig. 1, we show performance of these two estimations for the case of $d = 3$. We find that the two upper bounds have their own advantages in different situations.

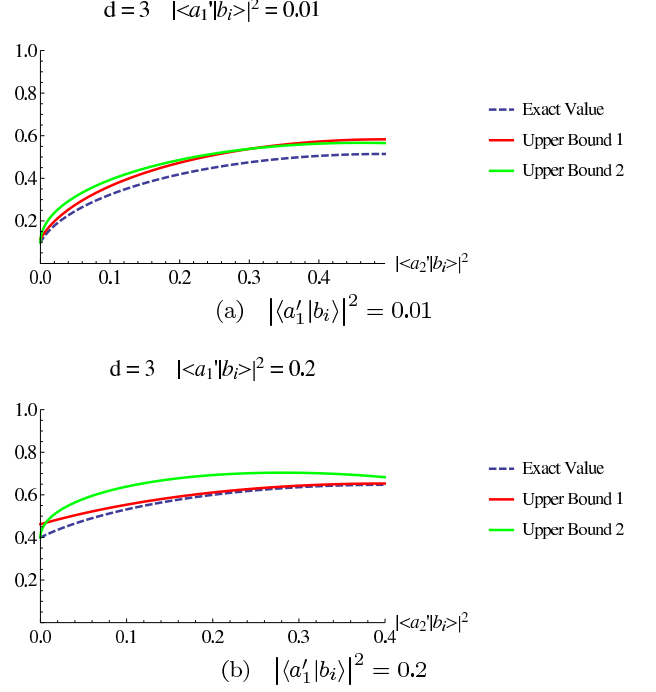


FIG. 1. (color online) Performance of the upper bound estimation in $d = 3$ case: We fix the value of $|\langle a_1' | b_i \rangle|^2$, and plot the value (before choosing maximum among different i 's) of $\eta(\mathcal{A}', \mathcal{B})$ and the two upper bounds (“Upper Bound 1” corresponds to Eq. (28) and “Upper Bound 2” corresponds to Eq. (29)), choosing $|\langle a_2' | b_i \rangle|^2$ as independent variable. In this figure, we assume $|\langle a_1' | b_i \rangle|^2 \leq |\langle a_2' | b_i \rangle|^2 \leq |\langle a_3' | b_i \rangle|^2$.

IV. TRADE-OFF RELATION

We now report the trade-off relation between error and disturbance.

Theorem 1 (Trade-off in Two-Dimensional Hilbert space). *When $d = 2$, the following trade-off between error and disturbance holds:*

$$\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B}) \geq \eta(\mathcal{A}, \mathcal{B}), \quad (35)$$

$$\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) \geq \eta(\mathcal{A}, \mathcal{B}), \quad (36)$$

where $\eta(\mathcal{A}, \mathcal{B}) = \frac{1}{2} \sin \theta$. Moreover, the equalities are both attained when $|a'_i\rangle = |a_i\rangle$ for all i . Here $\theta = \arccos |\mathbf{a} \cdot \mathbf{b}|$

with \mathbf{a}, \mathbf{b} being the Bloch vectors of eigenstates of A and B .

Proof. We use Bloch vectors to represent density matrix and different bases. That is,

$$\begin{aligned} \rho &= \frac{I + \mathbf{n} \cdot \boldsymbol{\sigma}}{2}, \quad |a_i\rangle \langle a_i|_{i=1,2} = \frac{I \pm \mathbf{a} \cdot \boldsymbol{\sigma}}{2}, \\ |b_i\rangle \langle b_i|_{i=1,2} &= \frac{I \pm \mathbf{b} \cdot \boldsymbol{\sigma}}{2}, \quad |a'_i\rangle \langle a'_i|_{i=1,2} = \frac{I \pm \mathbf{a}' \cdot \boldsymbol{\sigma}}{2}, \end{aligned} \quad (37)$$

where $\mathbf{a}, \mathbf{a}', \mathbf{b}$ are unit vectors on Bloch sphere and \mathbf{n} is in/on Bloch sphere ($|\mathbf{n}| \leq 1$). Straightforward calculations gives

$$\begin{aligned} \varepsilon_\rho(\mathcal{A}, \mathcal{A}') &= \frac{1}{2} |(\mathbf{a} - \mathbf{a}') \cdot \mathbf{n}| \\ &\leq \frac{1}{2} |\mathbf{a} - \mathbf{a}'| = \varepsilon(\mathcal{A}, \mathcal{A}') \end{aligned} \quad (38)$$

and

$$\begin{aligned} \eta_\rho(\mathcal{A}', \mathcal{B}) &= \frac{1}{2} |(\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}') \cdot \mathbf{n}| \\ &\leq \frac{1}{2} |\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}'| = \eta(\mathcal{A}', \mathcal{B}). \end{aligned} \quad (39)$$

On one hand,

$$\begin{aligned} &\eta(\mathcal{A}', \mathcal{B}) + \varepsilon(\mathcal{A}, \mathcal{A}') \\ &= \frac{1}{2} |\mathbf{a} - \mathbf{a}'| + \frac{1}{2} |\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}'| \\ &\geq \frac{1}{2} |\mathbf{b} \cdot \mathbf{a}' \mathbf{a} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}'| + \frac{1}{2} |\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}'| \\ &\geq \min_x \frac{1}{2} |x \mathbf{a} - \mathbf{b}| = \frac{1}{2} |(1 - \mathbf{a} \mathbf{a}) \cdot \mathbf{b}| = \frac{1}{2} |\mathbf{a} \times \mathbf{b}| \end{aligned} \quad (40)$$

and

$$\begin{aligned} &\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) \\ &= \max_{\mathbf{n}} \left(\frac{1}{2} |(\mathbf{a} - \mathbf{a}') \cdot \mathbf{n}| + \frac{1}{2} |(\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}') \cdot \mathbf{n}| \right) \\ &\geq \max_{\mathbf{n}} \frac{1}{2} |\mathbf{b} \cdot \mathbf{a}'| \frac{1}{2} |(\mathbf{a} - \mathbf{a}') \cdot \mathbf{n}| + \frac{1}{2} |(\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}') \cdot \mathbf{n}| \\ &\geq \frac{1}{2} |\mathbf{b} - \mathbf{b} \cdot \mathbf{a}' \mathbf{a}'| \geq \frac{1}{2} |(1 - \mathbf{a} \mathbf{a}) \cdot \mathbf{b}| = \frac{1}{2} |\mathbf{a} \times \mathbf{b}|. \end{aligned} \quad (41)$$

On the other hand, when $\mathbf{a}' = \mathbf{a}$, $\eta(\mathcal{A}', \mathcal{B}) + \varepsilon(\mathcal{A}, \mathcal{A}') = \Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) = \frac{1}{2} |\mathbf{b} - \mathbf{b} \cdot \mathbf{a} \mathbf{a}|$. Therefore, $\min_{\mathcal{A}'} \eta(\mathcal{A}', \mathcal{B}) + \varepsilon(\mathcal{A}, \mathcal{A}') = \min_{\mathcal{A}'} \Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) = \frac{1}{2} |\mathbf{a} \times \mathbf{b}|$. \square

This result gives a lower bound of the overall error and the sum of error and disturbance, and therefore provide a version of uncertainty principle in the case of two-dimensional Hilbert space. We could make $\sin \theta = 1$ by choosing $\mathbf{a} \cdot \mathbf{b} = 0$. This means that the overall error is the largest when $\{|a_i\rangle\}_{i=1}^d$ and $\{|b_i\rangle\}_{i=1}^d$ are mutually unbiased.

Fig. 2 shows the performance of the version of uncertainty principle in Theorem 1. It shows that the

values of $\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B})$ and $\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B})$ do not change smoothly because we have absolute values, maxima and minima in the definitions of error and disturbance. In particular, we find that $\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B})$ and $\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B})$ are not differentiable near $\mathcal{A}' = \mathcal{A}$.

On the other hand, it turns out that the minima of the overall error Δ and the sum of error and disturbance are both attained when $\{|a'_i\rangle\}_{i=1,2}$ are exactly $\{|a_i\rangle\}_{i=1,2}$, regardless of the choice of $\{|b_i\rangle\}_{i=1,2}$. This is somewhat unexpected for one would guess that \mathcal{A}' should be somewhere “between” \mathcal{A} and \mathcal{B} in order to minimize the overall error. However, according to Theorem 1, the best approach to minimize the overall error in 2-dimensional Hilbert space during the consecutive measurement process is simply to leave the first measurement device unchanged.

Another noteworthy observation is that when $d = 2$

$$\min_{\mathcal{A}'} (\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B})) = \min_{\mathcal{A}'} \Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}), \quad (42)$$

Because when $|a_i\rangle = |a'_i\rangle$ for all i , $\Delta(\mathcal{A}, \mathcal{A}, \mathcal{B}) = \eta(\mathcal{A}, \mathcal{B})$ and $\varepsilon(\mathcal{A}, \mathcal{A}) = 0$, revealing that it is no easier to reduce the overall error than the sum of error and disturbance.

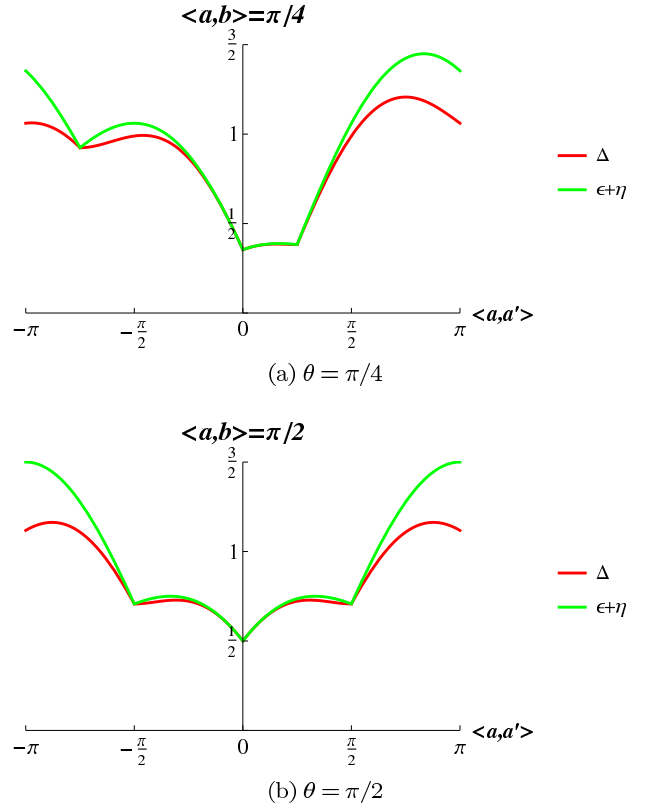


FIG. 2. (color online) We plot the value of $\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B})$ and $\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B})$ respectively when \mathbf{a}' is in $\text{Span}\{\mathbf{a}, \mathbf{b}\}$. Independent variable $\langle \mathbf{a}, \mathbf{a}' \rangle$ denotes the angle between \mathbf{a} and \mathbf{a}' . This figure shows clearly that when $\mathbf{a}' = \mathbf{a}$, $\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B})$ and $\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B})$ are the smallest.

Unfortunately, the absence of Bloch representation means that the proof techniques in Theorem 1 cannot

be used to tackle the trade-off between error and disturbance when $d > 2$. However, the same conclusion in [Theorem 1](#) still works in some special but interesting cases.

Theorem 2 (Trade-off for Mutually Unbiased Bases in d -dimensional Hilbert space). *For all $d \geq 2$, if eigenstates of A and B form a pair of mutually unbiased bases (MUB), then we have*

$$\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B}) \geq \eta(\mathcal{A}, \mathcal{B}) = 1 - 1/d. \quad (43)$$

When $\{|a'_i\rangle\}_{i=1}^d = \{|a_i\rangle\}_{i=1}^d$, the equality holds.

Proof. According to [Eq. \(18\)](#), we need to prove the following inequality:

$$\begin{aligned} & \max_j \sqrt{1 - |\langle a_j | a'_j \rangle|^2} \\ & + \max_i R(|b_i\rangle \langle b_i| - \sum_k |\langle b_i | a'_k \rangle|^2 |a'_k\rangle \langle a'_k|) \\ & \geq \max_i R(|b_i\rangle \langle b_i| - \sum_k |\langle b_i | a_k \rangle|^2 |a_k\rangle \langle a_k|). \end{aligned} \quad (44)$$

Using the triangle inequality for spectral radius [\[30\]](#), the above inequality is true if we can prove that for any $|b_i\rangle$,

$$\begin{aligned} & \max_j \sqrt{1 - |\langle a_j | a'_j \rangle|^2} \\ & \geq R\left(\sum_k |\langle b_i | a_k \rangle|^2 |a_k\rangle \langle a_k| - \sum_k |\langle b_i | a'_k \rangle|^2 |a'_k\rangle \langle a'_k|\right). \end{aligned} \quad (45)$$

Since $|\langle b_i | a_k \rangle|^2 = 1/d$, the LHS of the above inequality equals to $\max_j |1/d - |\langle b_i | a'_j \rangle|^2|$. It is then sufficient to prove that

$$\left|1/d - |\langle b_i | a'_j \rangle|^2\right| \leq \sqrt{1 - |\langle a_j | a'_j \rangle|^2} \quad (46)$$

for any j . Using $\sum_k |\langle a_k | a'_j \rangle|^2 = 1$ and $|\langle b_i | a_k \rangle|^2 = 1/d$, it is clear that the value of $|\langle b_i | a'_j \rangle|^2 = |\sum_k \langle b_i | a_k \rangle \langle a_k | a'_j \rangle|^2$ is between

$$\frac{1}{d} \left(1 \pm \sum_{k \neq l} |\langle a_k | a'_j \rangle| \cdot |\langle a_l | a'_j \rangle|\right).$$

Hence, all we need to prove is that

$$\frac{1}{d} \sum_{k \neq l} |\langle a_k | a'_j \rangle| \cdot |\langle a_l | a'_j \rangle| \leq \sqrt{1 - |\langle a_j | a'_j \rangle|^2}. \quad (47)$$

Set $x = 1 - |\langle a_j | a'_j \rangle|^2$. If $x = 0$, [Eq. \(47\)](#) is obviously

true. Otherwise,

$$\begin{aligned} & \sum_{k \neq l} |\langle a_k | a'_j \rangle| \cdot |\langle a_l | a'_j \rangle| \\ & = 2\sqrt{1-x} \sum_{m \neq j} |\langle a_m | a'_j \rangle| + \sum_{\substack{k \neq l \\ k, l \neq j}} |\langle a_k | a'_j \rangle| \cdot |\langle a_l | a'_j \rangle| \\ & \leq 2\sqrt{(1-x)(d-1)} \sqrt{\sum_{m \neq j} |\langle a_m | a'_j \rangle|^2} \\ & \quad + \sum_{\substack{k \neq l \\ k, l \neq j}} \frac{|\langle a_k | a'_j \rangle|^2 + |\langle a_l | a'_j \rangle|^2}{2} \\ & \leq 2\sqrt{(1-x)(d-1)}\sqrt{x} + x(d-2). \end{aligned} \quad (48)$$

Because

$$\begin{aligned} d\sqrt{x} & \geq 2\sqrt{(1-x)(d-1)}\sqrt{x} + x(d-2) \\ \iff d & \geq 2\sqrt{(1-x)(d-1)} + \sqrt{x}(d-2) \\ \iff d & \geq 2\sqrt{d-1} \sin \phi + (d-2) \cos \phi \\ \iff d & \geq \sqrt{(2\sqrt{d-1})^2 + (d-2)^2} = d, \end{aligned} \quad (49)$$

[Eq. \(47\)](#) is proved, and so is [Theorem 2](#). \square

The above result is of great importance because the case that the rank-one projectors of two measurement operators form a pair of MUBs is believed to maximize the sum of error and disturbance. [Theorem 2](#) confirms this belief for projective measurements in any finite-dimensional Hilbert space. Moreover, [Theorem 2](#) tells us again when $\mathcal{A}' = \mathcal{A}$, the sum of error and disturbance is minimized.

It is natural to ask if [Theorem 1](#) is true for general measurement operations when the Hilbert space dimension $d > 2$. Our numerical simulations find that while [Theorem 1](#) does not hold when $d > 2$, the theorem seems to be correct by adopting the following slightly relaxed definition of $\varepsilon(\mathcal{A}, \mathcal{B})$,

$$\varepsilon(\mathcal{A}, \mathcal{B}) = \min_{\tilde{\mathcal{B}}} \varepsilon(\mathcal{A}, \tilde{\mathcal{B}}), \quad (50)$$

where the minimum is over all possible projective measurements with rank one projectors $\tilde{\mathcal{B}}$ corresponding to the measurement operation \mathcal{B} . This relaxed definition of ε means that we only distinguish between the classical probability distributions resulting from measurements \mathcal{A} and \mathcal{B} with no regard to the labeling of different measurement outcomes. With this relaxed definition for ε , we make the following conjecture.

Conjecture 3 (Trade-off in d -dimensional Hilbert space). *For all $d \geq 2$, the following trade-off between error and disturbance holds:*

$$\varepsilon(\mathcal{A}, \mathcal{A}') + \eta(\mathcal{A}', \mathcal{B}) \geq f(\mathcal{A}, \mathcal{B}) \quad (51)$$

and

$$\Delta(\mathcal{A}, \mathcal{A}', \mathcal{B}) \geq f(\mathcal{A}, \mathcal{B}), \quad (52)$$

where

$$f(\mathcal{A}, \mathcal{B}) = \min\{\varepsilon(\mathcal{A}, \mathcal{B}), \eta(\mathcal{A}, \mathcal{B})\}. \quad (53)$$

Moreover, the equalities are both attained either when for all i , $|a'_i\rangle = |a_i\rangle$, or when for all i , $|a'_i\rangle = |b_i\rangle$.

If our conjecture can be proved, the validity of our trade-off relation will be extended to any operator in Hilbert space. Our sharp lower bound $f(\mathcal{A}, \mathcal{B})$ will then serve as a importance physical quantity characterizing the trade-off between error and disturbance in consecutive measurement process.

V. CONCLUSION

To summarize, we have introduced new definitions of error and disturbance in quantum measurement that

are both state- and eigenvalue-independent. An error-disturbance trade-off for a single qubit as well as for a special case of finite-dimensional Hilbert space have been proven. In particular, it implies that the lower bound of the sum of error and disturbance can be obtained when $\mathcal{A}' = \mathcal{A}$. Physically, it means it is not worth sacrificing error to lower the disturbance. Finally, we introduce [Conjecture 3](#) which is a non-trivial extension of [Theorem 1](#) to d -dimensional Hilbert space with $d > 2$. Our numerical simulations supports this conjecture; and if it can be proved, an error-disturbance trade-off just like Heisenberg's uncertainty principle can be derived.

ACKNOWLEDGMENTS

S. S. Zhou and S. Wu are supported by the National Natural Science Foundation of China under Grants No. 11275181 and No. 11475084, and the Fundamental Research Funds for the Central Universities under Grant No. 20620140531. H. F. Chau is supported in part by the RGC Grant HKU 700712P of the Hong Kong SAR Government.

-
- [1] W. Heisenberg, *Zeitschrift für Physik* **43**, 172 (1927).
 - [2] H. P. Robertson, *Physical Review* **34**, 163 (1929).
 - [3] J. V. Neumann, *Mathematical foundations of quantum mechanics*, 2 (Princeton university press, 1955).
 - [4] D. Bohm, *Quantum Theory*, Dover Books on Physics Series (Dover Publications, 1951).
 - [5] L. Schiff, *Quantum Mechanics*, International series in pure and applied physics (McGraw-Hill, 1955).
 - [6] M. Ozawa, in *Journal of Physics: Conference Series*, Vol. 504 (IOP Publishing, 2014) p. 012024.
 - [7] G. Sulyok, S. Sponar, J. Erhart, G. Badurek, M. Ozawa, and Y. Hasegawa, *Physical Review A* **88**, 022110 (2013).
 - [8] S.-Y. Baek, F. Kaneda, M. Ozawa, and K. Edamatsu, *Scientific reports* **3** (2013).
 - [9] F. Kaneda, S.-Y. Baek, M. Ozawa, and K. Edamatsu, *Physical review letters* **112**, 020402 (2014).
 - [10] P. Busch, P. Lahti, and R. F. Werner, *Reviews of Modern Physics* **86**, 1261 (2014).
 - [11] M. Ozawa, *Physical Review A* **67**, 042105 (2003).
 - [12] A. Lund and H. M. Wiseman, *New Journal of Physics* **12**, 093011 (2010).
 - [13] L. A. Rozema, A. Darabi, D. H. Mahler, A. Hayat, Y. Soudagar, and A. M. Steinberg, *Physical Review Letters* **109**, 100404 (2012).
 - [14] J. Erhart, S. Sponar, G. Sulyok, G. Badurek, M. Ozawa, and Y. Hasegawa, *Nature Physics* **8**, 185 (2012).
 - [15] M. Ringbauer, D. N. Biggerstaff, M. A. Broome, A. Fedrizzi, C. Branciard, and A. G. White, *Physical review letters* **112**, 020401 (2014).
 - [16] M. J. Hall, *Physical Review A* **69**, 052113 (2004).
 - [17] C. Branciard, *Proceedings of the National Academy of Sciences* **110**, 6742 (2013).
 - [18] C. Branciard, *Physical Review A* **89**, 022124 (2014).
 - [19] K. Korzekwa, D. Jennings, and T. Rudolph, *Physical Review A* **89**, 052108 (2014).
 - [20] J. Dressel and F. Nori, *Physical Review A* **89**, 022106 (2014).
 - [21] P. Busch, P. Lahti, and R. F. Werner, *Physical review letters* **111**, 160405 (2013).
 - [22] F. Buscemi, M. J. Hall, M. Ozawa, and M. M. Wilde, *Physical review letters* **112**, 050401 (2014).
 - [23] P. J. Coles and F. Furrer, *Physics Letters A* **379**, 105 (2015).
 - [24] P. Busch, P. Lahti, and R. F. Werner, *Journal of Mathematical Physics* **55**, 042111 (2014).
 - [25] M. Ozawa, arXiv preprint [arXiv:1308.3540](#) (2013).
 - [26] P. Busch, P. Lahti, and R. F. Werner, arXiv preprint [arXiv:1402.3102](#) (2014).
 - [27] R. B. Bapat and T. E. Raghavan, *Nonnegative matrices and applications*, Vol. 64 (Cambridge University Press, 1997).
 - [28] B. K. Butler and P. H. Siegel, *Linear Algebra and its Applications* **439**, 1468 (2013).
 - [29] N. Derzko and A. Pfeffer, *Mathematics of Computation* **19**, 62 (1965).
 - [30] A. W. Marshall, I. Olkin, and B. Arnold, *Inequalities: theory of majorization and its applications* (Springer Science & Business Media, 2010).