Extremal properties of the variance and the quantum Fisher information

Géza Tóth[1,2,3,*] and Dénes Petz[4,5]

[1]*Department of Theoretical Physics, University of the Basque Country UPV/EHU, P.O. Box 644, E-48080 Bilbao, Spain*
[2]*IKERBASQUE, Basque Foundation for Science, E-48011 Bilbao, Spain*
[3]*Wigner Research Centre for Physics, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary*
[4]*Alfréd Rényi Institute of Mathematics, Reáltanoda utca 13-15, H-1051 Budapest, Hungary*
[5]*Department of Mathematical Analysis, Budapest University of Technology and Economics, H-1111 Budapest, Hungary*

We show that the variance is its own concave roof. For rank-2 density matrices and operators with zero diagonal elements in the eigenbasis of the density matrix, we prove analytically that the quantum Fisher information is four times the convex roof of the variance. Strong numerical evidence suggests that this statement is true even for operators with nonzero diagonal elements or density matrices with a rank larger than 2. We also find that within the different types of generalized quantum Fisher information considered in Petz [J. Phys. A **35**, 929 (2002)] and Gibilisco, Hiai, and Petz [IEEE Trans. Inf. Theory **55**, 439 (2009)], after appropriate normalization, the quantum Fisher information is the largest. Hence, we conjecture that the quantum Fisher information is four times the convex roof of the variance even for the general case.

I. INTRODUCTION

Metrology plays an important role in many areas of physics and engineering. Quantum metrology [1,2] is a novel development, which takes into account the quantum nature of matter when determining the achievable maximal accuracy in parameter estimation. For instance, quantum metrology shows that it is possible to reach a much higher accuracy in typical metrological tasks in a linear interferometer with highly entangled quantum systems than with nonentangled ones, in which no real quantum interaction took place between the particles [3–9]. In practical situations, the influence of noise can reduce considerably the gain in accuracy when applying entangled quantum states for metrology [10].

Quantum variance and quantum Fisher information are two fundamental notions of quantum metrology. When measuring a quantum-mechanical observable A, the uncertainty of the measurement outcome is described by the variance defined as

$$(\Delta A)^2_\varrho = \langle A^2 \rangle_\varrho - \langle A \rangle^2_\varrho, \tag{1}$$

where A is a Hermitian operator and ϱ is a density matrix. Connected to this uncertainty, when we would like to estimate a parameter of a quantum evolution based on measuring quantum observables, an uncertainty arises also in this parameter estimation. The most fundamental parameter estimation task is estimating the small parameter θ of the unitary dynamics $U = \exp(-iA\theta)$ via making measurements on the output state (see Fig. 1). The accuracy of estimating θ is bounded from below by the famous Cramér-Rao bound as

$$\Delta\theta \geqslant \frac{1}{\sqrt{F_Q^{\rm BC}[\varrho,A]}}, \tag{2}$$

where the quantum Fisher information is defined as [6–9,11–14]

$$F_Q^{\rm BC}[\varrho,A] = 2\sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j}|A_{ij}|^2. \tag{3}$$

Here λ_i are the eigenvalues of the density matrix and A_{ij} are the matrix elements of the operator A in the eigenbasis of the density matrix.

In Refs. [15–17], it has been shown that it is possible to define generalized variances and various types of generalized quantum Fisher information. In fact, the two notions are closely connected to each other as for every generalized variance there is a corresponding generalized quantum Fisher information. Thus, the question arises: Are the usual variance, Eq. (1), and the usual quantum Fisher information, Eq. (3), special among the generalized quantities? In this paper we will answer these questions affirmatively. Our findings show that fundamental quantities used in metrology have extremal properties.

In particular, in the first part of the paper we will define generalized variances. The definition presented is broader than that of Refs. [15–17]. Then, we will show that the usual variance is the smallest generalized variance and, connected to this fact, it fulfills the following important relation:

Theorem 1. Let us denote the decomposition of a density matrix into the mixture of pure states as

$$\varrho = \sum_k p_k |\Psi_k\rangle\langle\Psi_k|. \tag{4}$$

Then, the variance for a mixed state can be obtained through a concave roof construction as

$$(\Delta A)^2_\varrho = \sup_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k (\Delta A)^2_{\Psi_k}. \tag{5}$$

The proof of Theorem 1 will be given later.

In the second part of the paper we define the generalized quantum Fisher information. The definition presented is also broader than that of Refs. [15–17]. We will show that $F_Q^{\rm BC}$

_____

*toth@alumni.nd.edu; http://www.gtoth.eu
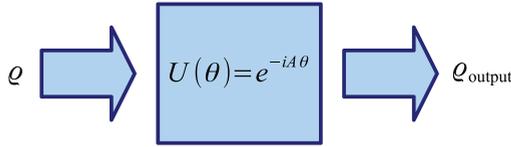
FIG. 1. (Color online) Basic parameter estimation task in quantum metrology. The small parameter θ must be estimated by making measurements on the output state ϱ_{output}.

defined in Eq. (3) is the largest among the types of generalized quantum Fisher information for the low rank case, and connected to this fact, it fulfills the following relation:

Theorem 2. For rank-2 density matrices and for an A with zero diagonal elements in the eigenbasis of the density matrix, the quantum Fisher information can be given as the convex roof of the variance as

$$F_Q^{\text{BC}}[\varrho,A] = 4 \inf_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k (\Delta A)^2_{\Psi_k}, \qquad (6)$$

where $\{p_k,|\Psi_k\rangle\}$ refers to a decomposition of ϱ of the type Eq. (4). The proof of Theorem 2 will also be given later.

We will also discuss that numerical calculations suggest that the left-hand side and the right-hand side of Eq. (6) are very close to each other, even when A has nonzero diagonal elements or the density matrix has a rank larger than 2.

Concerning the quantum Fisher information, we can also prove the following theorem.

Theorem 3. $F_Q^{\text{BC}}[\varrho,A]$ defined in Eq. (3) is maximal within the various types of quantum Fisher information defined by Refs. [15–17], if they are normalized such that for pure states they equal $4(\Delta A)^2$.

Concerning Theorem 3, it is important to note that all types of quantum Fisher information defined by Refs. [15–17] are convex. Based on Theorem 3 and the strong numerical evidence mentioned above, we formulate the following conjecture.

Conjecture 1. The quantum Fisher information $F_Q^{\text{BC}}[\varrho,A]$ defined in Eq. (3) is four times the convex roof of the variance for density matrices of any rank and any observable A.

The statements of Theorem 1 and Conjecture 1 can be concisely reformulated as follows: For any decomposition $\{p_k,|\Psi_k\rangle\}$ of the density matrix ϱ we have

$$\frac{1}{4}F_Q^{\text{BC}}[\varrho,A] \leqslant \sum_k p_k (\Delta A)^2_{\Psi_k} \leqslant (\Delta A)^2_\varrho, \qquad (7)$$

where the upper and the lower bounds are both tight.

The paper is organized as follows. In Sec. II, we define generalized variances and determine the minimal generalized variance. In Sec. III, we consider the generalized quantum Fisher information for the rank-2 case, and look for the maximal generalized quantum Fisher information. In Sec. IV, we present numerical calculations for density matrices with a higher rank. In Sec. V, we show that F_Q^{BC} is the largest among the types of quantum Fisher information defined by Refs. [15–17]. In the Appendix, we give some details of the semidefinite programs needed for Sec. IV.

II. GENERALIZED VARIANCES AND THE PROOF OF THEOREM 1

In this section, we will define the generalized variance and the minimal generalized variance. We will also show that the minimal generalized variance is the usual variance given in Eq. (1).

Definition 1. The generalized variance $\text{var}_\varrho(A)$ is defined by the following two requirements:

(1) For pure states, the generalized variance equals the usual variance

$$\text{var}_\Psi(A) = (\Delta A)^2_\Psi. \qquad (8)$$

(2) For mixed states, $\text{var}_\varrho(A)$ is concave in the state.

There are infinitely many generalized variances that fulfill the requirements of Definition 1. Next, we will show that it is possible to identify the smallest one.

Definition 2. The minimal generalized variance $\text{var}_\varrho^{\min}(A)$ is defined by the following two requirements:

(1) For pure states, it equals the usual variance

$$\text{var}_\Psi^{\min}(A) = (\Delta A)^2_\Psi, \qquad (9)$$

(2) For mixed states, it is defined through a *concave roof* construction [18],

$$\text{var}_\varrho^{\min}(A) = \sup_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k (\Delta A)^2_{\Psi_k}, \qquad (10)$$

where $\{p_k,|\Psi_k\rangle\}$ refers to a decomposition of ϱ of the type Eq. (4).

Note the counterintuitive fact that while a supremum is used to define Eq. (10), it is the minimal generalized variance and not the maximal one. Let us show that it is indeed the minimal generalized variance. It is clear that for any generalized variance, we have

$$\text{var}_\varrho(A) \geqslant \text{var}_\varrho^{\min}(A). \qquad (11)$$

If this were not true then it would be possible that for some decomposition of the type Eq. (4) we have $\text{var}_\varrho(A) < \sum_k p_k \text{var}_{\Psi_k}(A)$, which would contradict the concavity of the variance. On the other hand, Eq. (10) is concave in the state, since it is defined by a concave roof construction, thus it fulfills Definition 1. Hence, Eq. (10) is the minimal generalized variance.

Next we will prove two lemmas that we need later in the proof of the central theorem of the section.

Lemma 1. For rank-2 states, the minimal generalized variance is the usual variance given in Eq. (1). In particular, this statement is true for all qubit states.

Proof. For all decompositions of the form Eq. (4), the minimal variance is bounded from above and below as

$$(\Delta A)^2_\varrho \geqslant \text{var}_\varrho^{\min}(A) \geqslant \sum_k p_k (\Delta A)^2_{\Psi_k}. \qquad (12)$$

The first inequality in Eq. (12) is due to Eq. (11), the second one is due to the concavity of the variance.

In order to proceed to prove Lemma 1, we need to know an important property of the usual variance $(\Delta A)^2$. It can be written in the following way:

$$(\Delta A)^2_\varrho = \sum_k p_k \big[(\Delta A)^2_{\Psi_k} + (\langle A \rangle_{\Psi_k} - \langle A \rangle_\varrho)^2 \big]. \qquad (13)$$

Here the summation is over the pure subensembles. $(\Delta A)^2_{\Psi_k}$ is the variance in the kth subensemble, while the $(\langle A \rangle_{\Psi_k} - \langle A \rangle_\varrho)^2$ term is the square of the difference between the expectation value for the subensemble and the expectation value for the entire ensemble. One can call the two terms the "quantum" and the "classical" part of the variance, respectively.

Based on Eq. (13) we can state the following. If for a quantum state $\tilde{\varrho}$ there is a decomposition $\{\tilde{p}_k, |\tilde{\Psi}_k\rangle\}$ such that the subensemble expectation values equal the expectation value for the entire ensemble (i.e., $\langle A \rangle_{\tilde{\Psi}_k} = \langle A \rangle_{\tilde{\varrho}}$ for all k) then for such a decomposition we have

$$(\Delta A)^2_{\tilde{\varrho}} = \sum_k \tilde{p}_k (\Delta A)^2_{\tilde{\Psi}_k}. \quad (14)$$

Based on Eq. (12), we can now state that for $\tilde{\varrho}$ the minimum generalized variance of A and the usual variance are equal,

$$(\Delta A)^2_{\tilde{\varrho}} = \mathrm{var}^{\min}_{\tilde{\varrho}}(A). \quad (15)$$

We will now show that such a decomposition, in which the subensemble expectation values are all equal, always exists for rank-2 density matrices. Hence, we will show that for all such states the usual quantum variance is the minimal generalized variance.

Let us take an eigendecomposition of the state ϱ,

$$\varrho = p|\Psi_1\rangle\langle\Psi_1| + (1-p)|\Psi_2\rangle\langle\Psi_2|. \quad (16)$$

Thus, in the $|\Psi_1\rangle$ and $|\Psi_2\rangle$ basis we can write the density matrix ϱ as

$$\varrho = \begin{bmatrix} p & 0 \\ 0 & 1-p \end{bmatrix}. \quad (17)$$

Let us define now the family of states

$$|\Psi_\phi\rangle = \sqrt{p}|\Psi_1\rangle + \sqrt{1-p}|\Psi_2\rangle e^{i\phi}. \quad (18)$$

For the states (18), the expectation value of the operator A can be written as

$$\langle \Psi_\phi | A | \Psi_\phi \rangle = \langle A \rangle_\varrho + 2\sqrt{p(1-p)}\mathrm{Re}(\langle\Psi_1|A|\Psi_2\rangle e^{i\phi}). \quad (19)$$

Clearly, there is an angle ϕ_1 such that

$$\mathrm{Re}(\langle\Psi_1|A|\Psi_2\rangle e^{i\phi_1}) = 0. \quad (20)$$

For this angle, the expectation value of A in the state $|\Psi_{\phi_1}\rangle$ is the same as its expectation value in ϱ, since the second term in Eq. (19) is zero. In the basis of the states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, we can write the projection operators onto $|\Psi_{\phi_1}\rangle$ as

$$|\Psi_{\phi_1}\rangle\langle\Psi_{\phi_1}| = \begin{bmatrix} p & \sqrt{p(1-p)}e^{-i\phi_1} \\ \sqrt{p(1-p)}e^{+i\phi_1} & 1-p \end{bmatrix}. \quad (21)$$

The same is true for $\phi_1 + \pi$. In the basis of the states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, we can write the projection operator onto $|\Psi_{\phi_1+\pi}\rangle$ as

$$|\Psi_{\phi_1+\pi}\rangle\langle\Psi_{\phi_1+\pi}|$$
$$= \begin{bmatrix} p & -\sqrt{p(1-p)}e^{-i\phi_1} \\ -\sqrt{p(1-p)}e^{+i\phi_1} & 1-p \end{bmatrix}. \quad (22)$$

Based on Eqs. (21) and (22), the state ϱ can be decomposed as

$$\varrho = \tfrac{1}{2}\left(|\Psi_{\phi_1}\rangle\langle\Psi_{\phi_1}| + |\Psi_{\phi_1+\pi}\rangle\langle\Psi_{\phi_1+\pi}|\right), \quad (23)$$

and the two component states have $\langle\Psi_{\phi_1}|A|\Psi_{\phi_1}\rangle = \langle A \rangle_\varrho$ and $\langle\Psi_{\phi_1+\pi}|A|\Psi_{\phi_1+\pi}\rangle = \langle A \rangle_\varrho$, as expected. ∎
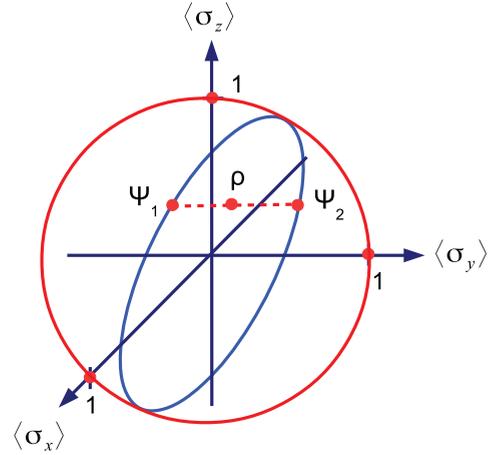
FIG. 2. (Color online) A single-qubit mixed state ϱ is decomposed as Eq. (23) into the mixture of pure states $|\Psi_{\phi_1}\rangle$ and $|\Psi_{\phi_1+\pi}\rangle$ defined in the text. Quantum states are represented by points in the $(\langle\sigma_x\rangle, \langle\sigma_y\rangle, \langle\sigma_z\rangle)$ space. The sphere with radius 1 is the Bloch sphere. Density matrices ϱ' for which the expectation value of A is the same as for ϱ fulfill the linear condition $\mathrm{Tr}[A(\varrho - \varrho')] = 0$ and are on a plane. The elliptic curve is the cross section between this plane and the surface of the Bloch sphere. Points corresponding to ϱ, $|\Psi_{\phi_1}\rangle$ and $|\Psi_{\phi_1+\pi}\rangle$ are all on this plane, thus the operator A has the same expectation value in these three states.

In Fig. 2, we show an example to demonstrate how Lemma 1 works for single-qubit states. The three coordinate axes are the expectation values of the three Pauli spin matrices. Points on the surface of the Bloch ball correspond to pure states, points within the ball correspond to rank-2 mixed states. The operator A has the same expectation value in the states ϱ, $|\Psi_{\phi_1}\rangle$ and $|\Psi_{\phi_1+\pi}\rangle$. Note that σ_z has also the same expectation value in these states.

Next we will consider the decomposition of states with a rank larger than 2.

Lemma 2. Let us consider an eigendecomposition of a density matrix

$$\varrho_0 = \sum_{k=1}^{r_0} \lambda_k |\Psi_k\rangle\langle\Psi_k| \quad (24)$$

with all $\lambda_k > 0$. Let us denote the rank of the density matrix as $r(\varrho_0) = r_0$. We consider density matrices for which $r_0 \geqslant 3$. Let us denote by A_0 the expectation value of an operator A in the state ϱ_0,

$$\mathrm{Tr}(A\varrho_0) = A_0. \quad (25)$$

We claim that for any A, ϱ_0 can always be decomposed as

$$\varrho_0 = p\varrho_- + (1-p)\varrho_+, \quad (26)$$

such that $r(\varrho_-) < r_0$, $r(\varrho_+) < r_0$, and

$$\mathrm{Tr}(A\varrho_+) = \mathrm{Tr}(A\varrho_-) = A_0. \quad (27)$$

Proof. Let us consider the eigenvalues λ_k in the eigendecomposition of ϱ, Eq. (24). First, note that from the conditions

$\text{Tr}(\varrho) = 1$ and Eq. (25), it follows for the λ_k's that

$$\sum_{k=1}^{r_0} \lambda_k = 1, \tag{28a}$$

$$\sum_{k=1}^{r_0} \lambda_k a_k = A_0, \tag{28b}$$

where we introduced for the subensemble expectation values the notation

$$a_k = \langle \Psi_k | A | \Psi_k \rangle. \tag{29}$$

Let us now define a family of states as

$$\varrho(c, \Delta\vec{\lambda}) = \sum_{k=1}^{r_0} (\lambda_k + c\Delta\lambda_k)|\Psi_k\rangle\langle\Psi_k|, \tag{30}$$

where c is a real parameter and $\Delta\lambda_k$ fulfill

$$\sum_{k=1}^{r_0} \Delta\lambda_k = 0, \tag{31a}$$

$$\sum_{k=1}^{r_0} a_k \Delta\lambda_k = 0. \tag{31b}$$

Equation (31a) ensures that the state $\varrho(c, \Delta\vec{\lambda})$ has a unit trace, while due to Eq. (31b) for all c

$$\text{Tr}[A\varrho(c, \Delta\vec{\lambda})] = A_0. \tag{32}$$

We can rewrite Eqs. (31a) and (31b) as

$$\begin{bmatrix} a_1 & a_2 & \dots & a_{r_0-1} & a_{r_0} \\ 1 & 1 & \dots & 1 & 1 \end{bmatrix} \Delta\vec{\lambda} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{33}$$

For $r_0 \geqslant 3$, there is always a $\Delta\vec{\lambda} \neq 0$ fulfilling Eq. (33). It is clear that, if $\varrho(c, \Delta\vec{\lambda})$ has non-negative eigenvalues, then $\varrho(c, \Delta\vec{\lambda})$ is a physical state. In the following, for compactness, we will omit $\Delta\vec{\lambda}$ from the argument of ϱ.

Let us now examine the properties of the $\varrho(c)$ family. It is clear that

$$\varrho(0) = \varrho_0. \tag{34}$$

Moreover, it is also clear that around $c = 0$ there is some range of values for c, such that $\varrho(c)$ is physical. So first, let us start to increase c gradually from zero until one of the eigenvalues becomes zero. Let us denote this value by c_+. Thus, $\varrho(c)$ is physical for

$$0 \leqslant c \leqslant c_+. \tag{35}$$

We have $r(\varrho(c_+)) < r_0$. Moreover, $\varrho(c)$ is nonphysical for $c > c_+$. A similar thing happens, if we start from $c = 0$ by decreasing c gradually until one of the eigenvalues becomes zero. Let us denote this value by c_-. Thus, $\varrho(c)$ is physical for

$$c_- \leqslant c \leqslant 0. \tag{36}$$

Again, we have $r(\varrho(c_-)) < r_0$. Moreover, $\varrho(c)$ is nonphysical for $c < c_-$. In summary, for any choice of ϱ and $\Delta\lambda_k$ fulfilling Eqs. (31b) and (31a), there is a family of states $\varrho(c)$ defined in Eq. (30) such that $\varrho(c)$ is physical for

$$c_- \leqslant c \leqslant c_+. \tag{37}$$
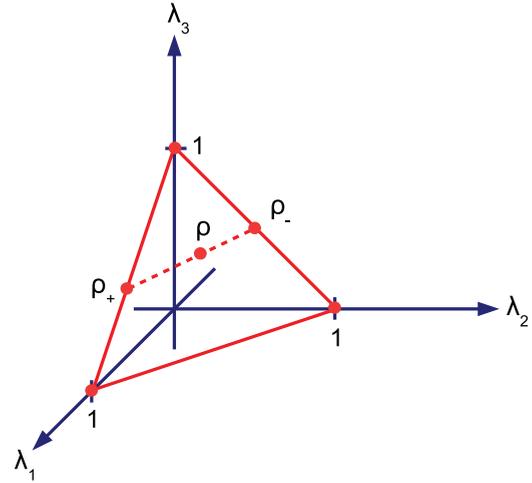
FIG. 3. (Color online) The rank-3 mixed state ϱ_0 is decomposed as Eq. (26) into the mixture of two rank-2 states, ϱ_- and ϱ_+. The coordinate axes are the eigenvalues of the density matrix, while we assume that all density matrices have the same eigenstates as ϱ. Points corresponding to states ϱ' for which $\text{Tr}(\varrho'A) = \text{Tr}(\varrho_0 A)$ are on the dashed line. For ϱ_+ we have $\lambda_2 = 0$, while for ϱ_- we have $\lambda_1 = 0$.

and $r(\varrho(c_+)) < r_0$, and $r(\varrho(c_-)) < r_0$. One can explicitly reconstruct these bounds as

$$c_+ = \min_k \frac{\lambda_k}{\Theta(-\Delta\lambda_k)} \tag{38}$$

and

$$c_- = -\min_k \frac{\lambda_k}{\Theta(\Delta\lambda_k)}, \tag{39}$$

where

$$\Theta(x) = \begin{cases} x & \text{if } x \geqslant 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{40}$$

Let us now construct the decomposition Eq. (26). Direct calculations can show that the following decomposition is of the desired form:

$$\varrho_- = \varrho(c_-), \tag{41}$$

and

$$\varrho_+ = \varrho(c_+), \tag{42}$$

while the mixing probability is

$$p = \frac{c_+}{c_+ - c_-}. \tag{43}$$

∎

In Fig. 3, we presented an example showing how Lemma 2 works for a rank-3 state. A point $(\lambda_1, \lambda_2, \lambda_3)$ in the coordinate system corresponds to $\varrho = \sum_{k=1}^{3} \lambda_k |\Psi_k\rangle\langle\Psi_k|$. All points have to fulfill Eq. (28a) and $\lambda_k \geqslant 0$ for $k = 1, 2, 3$. Thus, all points corresponding to physical states are in the triangle. Interior points of the triangle correspond to rank-3 states, while points on the boundary correspond to rank-2 and rank-1 states.

After proving Lemma 1 and Lemma 2, we will prove the main theorem of the section.

Proof of Theorem 1. Theorem 1 is equivalent to the statement that for systems of any dimension, the minimal generalized variance is the usual variance,

$$\text{var}_\varrho^{\min}(A) = (\Delta A)^2_\varrho. \tag{44}$$

For rank-2 states, this is true because of Lemma 1. According to Lemma 2, any state with a rank larger than 2 can be decomposed into the mixture of lower rank states that have the same expectation value for A as the original state has. The lower rank states can then be decomposed into the mixture of states with an even lower rank, until we reach rank-2 states. Thus, any state ϱ can be decomposed into the mixture of the form Eq. (4) such that $\langle\Psi_k|A|\Psi_k\rangle = \text{Tr}(A\varrho)$. Hence, similarly ro the proof of Lemma 1, the statement of Theorem 1 follows [19]. ∎

We will present two examples in order to clarify certain issues concerning generalized variances.

Example 1. Let us consider the state

$$\varrho = \tfrac{1}{2}(|+1\rangle\langle+1|+|-1\rangle\langle-1|), \tag{45}$$

where $|\pm1\rangle$ denote the eigenstates of the Pauli spin matrix σ_z. For this state, we have $(\Delta\sigma_z)^2 = +1$. Let us look at the decomposition

(1) $p_1 = \frac{1}{2}, |\Psi_1\rangle = \frac{1}{\sqrt{2}}(|+1\rangle+|-1\rangle)$,

(2) $p_2 = \frac{1}{2}, |\Psi_2\rangle = \frac{1}{\sqrt{2}}(|+1\rangle-|-1\rangle)$.

For this decomposition $(\Delta\sigma_z)^2_{\Psi_k} = +1$ and $\langle\Psi_k|\sigma_z|\Psi_k\rangle = \text{Tr}(\sigma_z\varrho) = 0$. Thus, we presented a decomposition such that the expectation value of A is the same for both subensembles $|\Psi_k\rangle$.

Example 2. Let us consider the generalized variance

$$\text{var}_\varrho^{\text{quadratic}}(A) = \text{Tr}(A^2\varrho) - \text{Tr}(A\varrho)^2 + 1 - \text{Tr}(\varrho^2). \tag{46}$$

For pure states, it equals the usual variance. For mixed states, it is larger than that and it is concave in the state. It is in fact a sum of the usual variance and the linear entropy, which is zero for pure states and is concave in the state.

III. GENERALIZED QUANTUM FISHER INFORMATION AND THE PROOF OF THEOREM 2

In this section, we will define the generalized quantum Fisher information and the maximal generalized quantum Fisher information. We will show that for rank-2 density matrices, the maximal quantum Fisher information is the usual quantum Fisher information.

Definition 3. The generalized quantum Fisher information $F_Q[\varrho, A]$ is defined by the following two requirements:

(1) For pure states, we have

$$F_Q[|\Psi\rangle\langle\Psi|, A] = 4(\Delta A_\Psi)^2. \tag{47}$$

The factor 4 appears to keep the consistency with the existing literature [6].

(2) For mixed states, $F_Q[\varrho, A]$ is convex in the state.

The usual quantum Fisher information, Eq. (3), is one of the types of generalized quantum Fisher information. Next, we will present another well-known generalized quantum Fisher information.

Example 3. Let us consider the Wigner-Yanase skew information defined as [20]

$$I[\varrho, A] = \text{Tr}(A^2\varrho) - \text{Tr}(A\varrho^{1/2}A\varrho^{1/2}). \tag{48}$$

For pure states it equals $(\Delta A)^2_\varrho$ and it is convex in the state. Thus $4I[\varrho, A]$ fulfills Definition 3.

There are infinitely many types of generalized quantum Fisher information that fulfill the requirements of Definition 3. Next, we will show that it is possible to identify the largest one.

Definition 4. The maximal generalized quantum Fisher information, $F_Q^{\max}[\varrho, A]$, is defined by the following two requirements:

(1) For pure states, it equals four times the usual variance,

$$F_Q^{\max}[|\Psi\rangle\langle\Psi|, A] = 4(\Delta A)^2_\Psi. \tag{49}$$

(2) For mixed states, it is defined through a *convex roof* construction [18],

$$F_Q^{\max}[\varrho, A] = 4 \inf_{\{p_k, |\Psi_k\rangle\}} \sum_k p_k(\Delta A)^2_{\Psi_k}. \tag{50}$$

It is clear that for any generalized quantum Fisher information, we have

$$F_Q[\varrho, A] \leqslant F_Q^{\max}[\varrho, A]. \tag{51}$$

This can be proven similarly as it has been shown that the minimal generalized variance is smaller than or equal to all other generalized variances.

Next we will present the main theorem of this section.

Proof of Theorem 2. Let us consider a density matrix ϱ of rank 2 given as

$$\varrho = q|\phi_1\rangle\langle\phi_1| + (1-q)|\phi_2\rangle\langle\phi_2|, \tag{52}$$

and a self-adjoint observable A. We will show that for any ϱ and A exists a decomposition into a mixture of projectors of the form Eq. (4) such that

$$F_Q^{\text{BC}}[\varrho, A] = 4\sum_k p_k(\Delta A)^2_{\Psi_k}. \tag{53}$$

Knowing that

$$\sum_k p_k 4(\Delta A)^2_{\Psi_k} \geqslant F_Q^{\max}[\varrho, A] \geqslant F_Q^{\text{BC}}[\varrho, A], \tag{54}$$

if a decomposition of the type Eq. (53) exists then

$$F_Q^{\text{BC}}[\varrho, A] = F_Q^{\max}[\varrho, A]. \tag{55}$$

Note that the first inequality in Eq. (53) is due to the convexity of the quantum Fisher information, the second one is due to Definition 4.

We consider the case where the diagonal elements of A are zero, hence A can explicitly be obtained as

$$A = |A_{12}|(|\phi_1\rangle\langle\phi_2|e^{+i\alpha} + |\phi_2\rangle\langle\phi_1|e^{-i\alpha}) \tag{56}$$

and $|A_{12}|$ and α are real. In fact, the the phase α is irrelevant. Thus, in our computation only $|A_{12}|$ is essential, so we can reduce the problem to the two-dimensional space generated by $|\phi_1\rangle$ and $|\phi_2\rangle$, and instead of A, it is sufficient to take

into account $|A_{12}|$. For the usual quantum Fisher information, Eq. (3), we have

$$F_Q^{BC}[\varrho,A] = 4(2q-1)^2|A_{12}|^2. \qquad (57)$$

For the right-hand side of Eq. (53) we have

$$4\sum_k p_k(\Delta A)^2_{\Psi_k} = 4[\text{Tr}(\varrho A^2) - \sum_k p_k|\langle\Psi_k|A|\Psi_k\rangle|^2]. \quad (58)$$

Based on Eqs. (57) and (58), Eq. (53) can be rewritten as

$$4(2q-1)^2|A_{12}|^2 = 4[\text{Tr}(\varrho A^2) - \sum_k p_k|\langle\Psi_k|A|\Psi_k\rangle|^2]. \quad (59)$$

This is the equality to be solved. Using that

$$A^2 = |A_{12}|^2(|\phi_1\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|), \qquad (60)$$

we obtain $\text{Tr}(\varrho A^2) = |A_{12}|^2$. Substituting this into Eq. (59), we get

$$4q(1-q)|A_{12}|^2 = \sum_k p_k|\langle\Psi_k|A|\Psi_k\rangle|^2. \qquad (61)$$

Finally, we can present a desired decomposition of the form Eq. (4) for ϱ fulfilling Eq. (61). It is given by two subensembles as

$$p_1 = \tfrac{1}{2}, \quad |\Psi_1\rangle = \sqrt{q}|\phi_1\rangle + \sqrt{1-q}|\phi_2\rangle, \qquad (62)$$

and

$$p_2 = \tfrac{1}{2}, \quad |\Psi_2\rangle = \sqrt{q}|\phi_1\rangle - \sqrt{1-q}|\phi_2\rangle. \qquad (63)$$

∎

IV. NUMERICAL EVIDENCE FOR THE VALIDITY OF THEOREM 2 FOR MATRICES WITH A RANK LARGER THAN 2

After presenting a proof for Eq. (6) for the rank-2 case, in this section we examine numerical evidence for larger systems. Thus, our approach is similar to the one followed in Ref. [21], where the concurrence formula has been derived for rank-2 mixed states and for mixed states with a higher rank it has been supported by numerical evidence.

We show a method to compute the infimum over convex decompositions in Eq. (50) by mapping this problem to an optimization over symmetric separable states [22]. While the optimization for symmetric separable states can be computed only for $d = 2$, we get a lower bound on the infimum by optimizing for symmetric quantum states with a positive partial transpose (PPT), since states with a positive partial transpose are a superset of separable states [23]. We will also consider a smaller superset, the set of symmetric quantum states with a PPT symmetric extension [24]. Note that the optimization for density matrices and operators of dimension d is equivalent to optimization over density matrices of rank d allowing for density matrices and operators of any size.

The optimization problem for $F_Q^{\max}[\varrho,A]$ defined in Eq. (50) can be rewritten as follows:

$$F_Q^{\max}[\varrho,A]$$
$$= 4\left(\langle A^2\rangle_\varrho - \sup_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k \langle A\rangle^2_{\Psi_k}\right), \qquad (64)$$

where $\{p_k,|\Psi_k\rangle\}$ refers to a decomposition of ϱ of the type Eq. (4).

Next, we will rewrite the term quadratic in expectation values as a term linear in expectation values of an operator acting on a bipartite system as

$$F_Q^{\max}[\varrho,A]$$
$$= 4\left(\langle A^2\rangle_\varrho - \sup_{\{p_k,|\Psi_k\rangle\}} \sum_k p_k \langle A\otimes A\rangle_{\Psi_k\otimes\Psi_k}\right). \quad (65)$$

Further transformations lead to a form in which instead of a sum of expectation values we have an expectation value of a mixture as

$$F_Q^{\max}[\varrho,A]$$
$$= 4\left(\langle A^2\rangle_\varrho - \sup_{\{p_k,|\Psi_k\rangle\}} \langle A\otimes A\rangle_{\sum_k p_k|\Psi_k\rangle\langle\Psi_k|^{\otimes 2}}\right). \quad (66)$$

Hence, we arrive at a form based on an optimization over symmetric separable states:

$$F_Q^{\max}[\varrho,A] = 4\left(\langle A^2\rangle_\varrho - \sup_{\substack{\varrho_{ss}\in S_s,\\ \text{Tr}_1(\varrho_{ss})=\varrho}} \langle A\otimes A\rangle_{\varrho_{ss}}\right)$$
$$= 2\inf_{\substack{\varrho_{ss}\in S_s,\\ \text{Tr}_1(\varrho_{ss})=\varrho}} \langle(A\otimes 1 - 1\otimes A)^2\rangle_{\varrho_{ss}}, \qquad (67)$$

where S_s is the set of symmetric separable states. States in S_s are mixtures of symmetric product states [22,25],

$$\sum_k p_k|\Psi_k\rangle\langle\Psi_k|^{\otimes 2}. \qquad (68)$$

It is not possible to compute the right-hand side of Eq. (67) in general. However, it is possible to obtain a series of lower bounds on it based on the theory of PPT symmetric extensions [24]. Let us introduce the notation for such bounds:

$$\mathcal{B}_{SE_n}(A,\varrho) := 2\inf_{\substack{\varrho_{SE_n}\in S_{SE_n}\\ \text{Tr}_1(\varrho_{SE_n})=\varrho}} \langle(A\otimes 1 - 1\otimes A)^2\rangle_{\varrho_{SE_n}}, \qquad (69)$$

where S_{SE_n} are the symmetric bipartite states with a PPT symmetric extension of size n. We note that S_{SE_2} is identical to the set of symmetric PPT states,

$$S_{SE_2} \equiv S_{SPPT}, \qquad (70)$$

hence we will use the two notations interchangeably in the rest of the paper. With the definition Eq. (69), for $n > 3$ we can write

$$F_Q^{\max}[\varrho,A] \geqslant \cdots \geqslant \mathcal{B}_{SE_{(n+1)}}(A,\varrho) \geqslant \mathcal{B}_{SE_n}(A,\varrho)$$
$$\geqslant \cdots \geqslant \mathcal{B}_{SE_3}(A,\varrho) \geqslant \mathcal{B}_{SPPT}(A,\varrho). \qquad (71)$$

All these bounds can be obtained using semidefinite programming [24].

The nonincreasing series of bounds in Eq. (71) exists since, in general, states having a PPT symmetric extension for n qudits are a subset of states having a PPT symmetric extension for $m < n$ qudits. In entanglement theory PPT states are typically separable and entangled states with a PPT property are very special [23]. Thus, the bound $\mathcal{B}_{SPPT}$ based

on an optimization over PPT states is expected to be very close to the value obtained from optimization over separable states for most A's and ϱ's. The bound $\mathcal{B}_{SE_3}$ computed based on an optimization over S_{SE_3} is expected to be even closer to the one obtained from an optimization over separable states.

We carried out numerical tests using random A and ϱ matrices. We considered 10 000 random trials for dimensions $d = 2,3,4$ for computing the bound $\mathcal{B}_{SPPT}$ defined in Eq. (69). The matrix A was generated as follows. First a matrix M was obtained such that all the real and imaginary parts of its elements were independent random numbers with a normal distribution centered around 0 and with a variance $\sigma = 1$. Then, the Hermitian matrix A was obtained from M as $\frac{1}{2}(M + M^\dagger)$. The density matrix ϱ was generated based on the method described in Ref. [26]. For each randomly chosen A and ϱ we determined the relative difference between $\mathcal{B}_{SPPT}(\varrho, A)$ given in Eq. (69) and $F_Q^{BC}[\varrho, A]$. We used the program packages SEDUMI and YALMIP for semidefinite programming, while QUBIT4MATLAB v4.0 was used for calculations connected to quantum physics [27–29]. MATLAB uses double precision arithmetic.

To test our method, we considered first the $d = 2$ case for A matrices with zero diagonal elements. According to Theorem 2 and knowing that the set of PPT states and the set of separable states are the same for two qubits, $F_Q^{BC}[\varrho, A]$ is equal to the bound $\mathcal{B}_{SPPT}(\varrho, A)$. The results are in the first row of Table I. We found that the largest relative difference is below 2×10^{-6}, while the average relative difference and the standard deviation of the relative difference are much smaller. The nonzero value must be the result of the finite precision of the numerical calculations.

After testing the method for a case that have been proven analytically in Theorem 2, we also used it to explore the cases that have not been proven yet. We carried out calculations for A matrices with nonzero diagonal elements for $d = 2,3,4$. The results of the numerical tests can be seen in the second, third, and fourth rows of Table I. We also carried out similar numerical tests for dimensions $d = 3,4$ for computing the bound $\mathcal{B}_{SE_3}$ based on an optimization for symmetric states with an $N = 3$ PPT symmetric extension given in Eq. (69) and compare it to the quantum Fisher information. The results of the numerical tests can be seen in the bottom part of Table I.

We can see that for all these trials, the relative difference between the two bounds and the value of F_Q^{BC} is smaller than 10^{-5}, while the average relative difference is smaller than 10^{-8}, and the standard deviation of the relative difference is smaller than 10^{-7}. Thus, the numerical tests suggest that it should be examined carefully through analytical calculations whether

$$F_Q^{max}[\varrho, A] = F_Q^{BC}[\varrho, A] \tag{72}$$

is true in general. Based on this strong evidence, we conjecture that this is the case. The semidefinite programs necessary to calculate $\mathcal{B}_{SPPT}(\varrho, A)$ and $\mathcal{B}_{SE_3}(\varrho, A)$ are outlined in the Appendix.

V. MAXIMAL QUANTUM FISHER INFORMATION WITHIN THE SET DEFINED BY REFS. [15,16]

In this section, first we will review how Refs. [15,16] define the variances and types of quantum Fisher information. We will denote them by vâr and $\hat{F}$, in order to distinguish them from the generalized variances and types of quantum Fisher information considered in this paper. Then, we will show that these definitions, apart from a constant factor, fulfill Definition 1 and Definition 3. This is a surprise since these quantities in Refs. [15,16] were defined based on ideas very different from the ones presented in this paper. We also show that the set of variances and types of quantum Fisher information satisfying Definition 1 and Definition 3 are a broader set than the corresponding quantities presented in Refs. [15,16]. Finally, we will show that the usual quantum Fisher information given in Eq. (3), after appropriate normalization, is the largest even within the different types of quantum Fisher information considered in Refs. [15,16].

A. Summary of the basic results of Refs. [15,16]

The basic idea of Refs. [15,16] is that for each standard matrix monotone function $f : \mathbb{R}^+ \to \mathbb{R}^+$, a generalized variance and a corresponding quantum Fisher information are defined. The notion standard means that $f(1) = 1$ and $f(t) = tf(t^{-1})$. First, let us consider the generalized variances.

Definition 5. The variance according to Refs. [15,16] is defined as

$$\text{vâr}_\varrho^f(A) = \langle A, \mathbb{J}_\varrho^f(A) \rangle - (\text{Tr}\varrho A)^2, \tag{73}$$

where A is Hermitian, the scalar product is defined as $\langle A, B \rangle = \text{Tr}(AB)$, and

$$\mathbb{J}_\varrho^f(A) = f(\mathbb{L}_\varrho \mathbb{R}_\varrho^{-1}) \mathbb{R}_\varrho, \tag{74}$$

where

$$\mathbb{L}_\varrho(A) = \varrho A, \quad \mathbb{R}_\varrho(A) = A\varrho.$$

Computing $\mathbb{J}_\varrho(A)$ can be simplified knowing that for $f(x) = \sum_k c_k x^k$ we have [15,16]

$$\mathbb{J}_\varrho^f(A) = \sum_k c_k \varrho^k A \varrho^{(1-k)}. \tag{75}$$

Thus, Eq. (75) can be used to calculate $\mathbb{J}_\varrho(A)$ for any polynomial of x.

Moreover, it is also useful to define the mean based on f as

$$m_f(a,b) = af\left(\frac{b}{a}\right) \tag{76}$$

and use it instead of f. The $f(1) = 1$ normalization condition corresponds to the condition $m_f(a,a) = a$ for the means. The $f(t) = tf(t^{-1})$ requirement corresponds to $m_f(a,b) = m_f(b,a)$. A list of quantum Fisher informations generated by various well-known means $m_f(a,b)$ can be found in Refs. [15–17].

After we discussed the generalized variances, we will turn our attention to the quantum Fisher information.

Definition 6. The quantum Fisher information according to Refs. [15,16] is defined as

$$\hat{F}^f(\varrho; A) = \text{Tr}(A\mathbb{J}_f^{-1}(\varrho)A). \tag{77}$$

TABLE I. Top row: Statistics of the relative difference between the quantum Fisher information and $\mathcal{B}_{\text{SPPT}}(\varrho,A)$ for 10 000 random tests. For each test, a random A with zero diagonal elements and a random ϱ is generated. The largest relative difference, the average relative difference and the standard deviation of the relative difference (i.e., the square root of the variance) is shown for two qudits for $d = 2$. SPPT indicates that the optimization is taken over PPT symmetric states. Second, third, and fourth rows: the same for random A with nonzero diagonal elements $d = 2,3,4$. Bottom two rows: the same for $d = 3,4$ for the relative difference between the quantum Fisher information and $\mathcal{B}_{\text{SE}_3}(\varrho,A)$. SE_3 indicates that the optimization is taken over symmetric states with a PPT symmetric extension for $N = 3$ qudits.

Task	Largest rel. diff.	Average rel. diff.	Standard dev. of rel. diff.
2SPPT, zeros in the diagonal	1.4793×10^{-6}	3.6194×10^{-10}	2.2321×10^{-8}
2SPPT	9.2505×10^{-6}	1.3313×10^{-9}	9.8533×10^{-8}
3SPPT	2.3552×10^{-8}	8.8486×10^{-10}	9.9692×10^{-10}
4SPPT	9.0345×10^{-9}	1.3448×10^{-9}	9.8216×10^{-10}
3SE$_3$	7.7749×10^{-8}	4.2694×10^{-10}	3.5161×10^{-9}
4SE$_3$	9.5506×10^{-9}	1.0429×10^{-9}	1.0404×10^{-9}

Again, the most important types of generalized quantum Fisher information correspond to famous means appearing often in mathematics.

For the arithmetic mean $m_f(a,b) = \frac{a+b}{2}$, the quantum Fisher information is defined as

$$\hat{F}^{\text{BC}}(\varrho; A) = \int_0^\infty \text{Tr}[\exp(-t\varrho/2)A \exp(-t\varrho/2)A]dt. \quad (78)$$

Other means generate other types of quantum Fisher information. Details can be found in Refs. [15–17].

B. The usual quantum Fisher information is the maximal one within a set of quantum Fisher informations defined in Refs. [15,16]

After reviewing the results of Refs. [15,16], we will connect the results of the present paper to that of Refs. [15,16]. In particular, we will show that the usual quantum Fisher information, Eq. (3), is the largest within the set of quantum Fisher informations defined by Refs. [15,16], if they are normalized such that for pure states they equal four times the usual variance given in Eq. (1).

In order to proceed, we will now prove the following lemma concerning the behavior of the family of variances defined in Refs. [15,16] for pure states.

Lemma 3. For pure states we have

$$\hat{\text{var}}_\varrho^f(A) = 2m_f(1,0)(\Delta A)^2. \quad (79)$$

Thus, for pure states the variance of Refs. [15,16] equals the usual variance times a constant.

Proof. It can be shown that Eq. (73) can be rewritten as [16]

$$\hat{\text{var}}_\varrho^f(A) = \sum_{ij} m_f(\lambda_i,\lambda_j)|A_{ij}|^2 - \left|\sum \lambda_i A_{ii}\right|^2, \quad (80)$$

where λ_j are the eigenvalues of ϱ, and A is given in the basis of the eigenvectors of the density matrix. Here we took advantage of the fact that $m_f(a,a) = 1$. One can see that for pure states ($\lambda_1 = 1$, $\lambda_k = 0$ for $k \geqslant 2$) we have

$$\hat{\text{var}}_\varrho^f(A) = 2\sum_{j>1} m_f(1,0)|A_{1j}|^2. \quad (81)$$

Hence, Eq. (79) follows. Note that there is a factor of 2 in Eq. (81) since $m_f(\lambda_i,\lambda_j)$ in Eq. (80) is nonzero for the pure state considered for $i = 1$, $j > 1$ and for $j = 1$, $i > 1$. ∎

Next, we will turn our attention to the quantum Fisher information. We define the following notation:

$$\hat{F}_Q^f[\varrho,A] = \hat{F}^f(\varrho; i[\varrho,A]). \quad (82)$$

With this definition, for the arithmetic mean $m_f(a,b) = \frac{a+b}{2}$, $\hat{F}_Q^f[\varrho,A]$ equals the usual quantum Fisher information given in Eq. (3) denoted in the physics literature as $F_Q[\varrho,A]$ [6].

It is instructive to analyze the differences between the quantum information $\hat{F}^f(\varrho,A)$ appearing in Refs. [15,16] and, in general, in the mathematics literature and $F_Q^f[\varrho,A]$, which appears in the physics literature. The difference between the two notions is that $1/\hat{F}^f(\varrho,A)$ bounds the squared uncertainty when the small parameter t of

$$\varrho_{\text{output}}^{\text{Math.}}(t) = \varrho + At \quad (83)$$

is estimated [15]. Here A is a matrix, and hence in this case the output density matrix is a linear function of t. In contrast, $1/F_Q^f[\varrho,A]$ bounds the squared uncertainty when estimating θ in

$$\varrho_{\text{output}}(t) = \exp(-iA\theta)\varrho \exp(+iA\theta), \quad (84)$$

as mentioned in the Introduction. The definition of $\hat{F}^f(\varrho,A)$ is not clear for pure states, as for such states $\varrho + At$ is aphysical for $t > 0$ or $t < 0$, while Eq. (84) describes a physical dynamics that leads to a valid density matrix for all t.

Next, we will prove a connection between the various types of generalized quantum Fisher information and the usual variance for pure states.

Lemma 4. For pure states, the generalized quantum Fisher information $\hat{F}_Q^f[\varrho,A]$ is proportional to the usual variance given in Eq. (1),

$$\hat{F}_Q^f[\varrho,A] = \frac{2(\Delta A)^2}{m_f(1,0)}. \quad (85)$$

Proof. Equation (77) leads to Ref. [16]

$$\hat{F}^f(\varrho; A) = \sum_{i,j} \frac{1}{m_f(\lambda_i, \lambda_j)} |A_{ij}|^2, \tag{86}$$

where λ_j are the eigenvalues of ϱ, and A is given in the basis of the eigenvectors of the density matrix. Substituting Eq. (86) into Eq. (82), we obtain

$$\hat{F}^f_Q[\varrho, A] = \sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{m_f(\lambda_i, \lambda_j)} |A_{ij}|^2. \tag{87}$$

For $m_f(a,b) = \frac{a+b}{2}$, we obtain the usual quantum Fisher information given in Eq. (3). The lemma can be proven substituting $\lambda_1 = 1, \lambda_k = 0$ for $k \geqslant 2$ into Eq. (87). ∎

Next, we will present the central theorems of this section.

Theorem 4. After appropriate normalization, the generalized variances defined in Refs. [15,16] fulfill Definition 1. However, the opposite is not true. Not all generalized variances that fulfill Definition 1 belong to the generalized variances defined in Refs. [15,16].

Proof. It has been proven that the generalized variances of Refs. [15,16] are concave [30]. Based on this fact and Lemma 3, we see that

$$\text{var}^f_\varrho(A) = \frac{\hat{\text{var}}^f_\varrho(A)}{2m_f(1,0)} \tag{88}$$

is a generalized variance in the sense of Definition 1. At this point the question arises: Are the two definitions equivalent? We find that this is not the case. Equation (46) is a generalized variance according to Definition 1, but cannot be written in the form Eq. (73) remembering that $\mathbb{J}_\varrho(A)$ is defined in Eq. (75). This is essentially due to the fact that ϱ^2 appears in the definition Eq. (46). On the other hand, Refs. [15,16] define a family of variances based on a single-variable function f, while Definition 1 does not provide an explicit formula for obtaining such quantities. ∎

Theorem 5. After appropriate normalization, the types of quantum Fisher information defined in Refs. [15,16] fulfill Definition 3.

Proof. It has been proven that the types of generalized quantum Fisher information of Refs. [15,16] are convex [30]. Based on Lemma 4, we can also see that

$$F^f_Q[\varrho, A] = 2m_f(1,0)\hat{F}^f_Q[\varrho, A] \tag{89}$$

is a generalized quantum Fisher information in the sense of Definition 3. ∎

Next, we discuss, which quantum Fisher information is the largest from the family considered in Refs. [15,16]. It has been proven that $\hat{F}^f(\varrho; A)$ is the smallest for $f(x) = \frac{1+x}{2}$, that is, for the usual quantum Fisher information [15,16]. Moreover, it has been also found that the quantum Fisher information for $f(x) = \frac{2x}{1+x}$, i.e., for the harmonic mean, is the smallest [15,16]. This can be seen as follows. In the expression for the quantum Fisher information, Eq. (86), appears the $m_f(a,b)$ mean. The arithmetic mean $m_f(a,b) = \frac{a+b}{2}$ is the largest among the means, while the harmonic mean is the smallest.

With a similar argument, we can prove that $\hat{F}^f[\varrho, A]$ defined in Eq. (87) is the largest for $f(x) = \frac{1+x}{2}$, i.e., for the arithmetic mean. It can also be proven that it is the smallest for the harmonic mean.

According to Theorem 4, the variance fitting Definition 1 is of the form Eq. (88). Moreover, according to Theorem 5, the types of quantum Fisher information fitting the Definition 3 is of the form Eq. (89). Both differ from the definitions of Refs. [15,16] by a normalization factor. For which $f(x)$ is the quantum Fisher information $F^f_Q[\varrho, A]$ the largest, if it is normalized such that for pure states it equals $4(\Delta A)^2$? In order to answer this question, based on the derivation above, let us write down explicitly the variances and types of quantum Fisher information of Refs. [15,16], together with the appropriate normalization factors such that for pure states they equal to the usual variance, Eq. (1), and to four times the usual variance, respectively. Based on Eqs. (79), (80), (85), and (87), the two definitions are as follows:

$$\text{var}^f_\varrho(A) = \frac{1}{2} \sum_{i,j} \frac{m_f(\lambda_i, \lambda_j)}{m_f(1,0)} |A_{ij}|^2 - \left| \sum_i \lambda_i A_{ii} \right|^2,$$
$$\tag{90}$$
$$F^f_Q[\varrho, A] = 2 \sum_{i,j} \frac{m_f(1,0)}{m_f(\lambda_i, \lambda_j)} (\lambda_i - \lambda_j)^2 |A_{ij}|^2.$$

For each mean $m_f(a,b)$, Equation (90) provides a generalized variance fulfilling Definition 1 and a generalized quantum Fisher information fulfilling Definition 3. It can be seen that the same $m_f(a,b)$ corresponds to the minimal variance and to the maximal quantum Fisher information. However, from Theorem 1, we already know that the minimal variance is the usual variance given in Eq. (1), corresponding to $f(x) = \frac{1+x}{2}$, i.e., to the arithmetic mean $m_f(a,b) = \frac{a+b}{2}$. Hence, Theorem 3 follows.

VI. CONCLUSIONS

We have defined generalized variances and types of quantum Fisher information. We have shown that the smallest variance is the usual variance since it is its own concave roof. We have also shown that the largest quantum Fisher information is the usual quantum Fisher information for rank-2 density matrices and for operators that have zero diagonal elements in the eigenbasis of the density matrix. In other words, the quantum Fisher information is four times the convex roof of the variance for such matrices. We found strong numerical evidence that the quantum Fisher information is very close to the convex roof even for matrices with nonzero diagonal elements and for matrices with a larger rank. We related our findings to the generalized variances and types of quantum Fisher information defined in Refs. [15,17]. We showed that the usual quantum Fisher information is the largest within the set of types of quantum Fisher information of Refs. [15,16], if appropriate normalization is applied. Hence we conjectured that the usual quantum Fisher information is the convex roof of the variance for density matrices and Hermitian operators of any size.

In the future, it would be important to prove Conjecture 1. Since convex roofs appear in entanglement theory very often, our findings help to establish further connections between the quantum Fisher information and entanglement theory [6–9,31–34]. For instance, it would be important to examine the consequences of our findings concerning the detection of entanglement with variances [35] or the quantum Fisher information [6–9], or the estimation of the quantum Fisher information based on measurements [36,37]. It would also be important to point out connections to the recent findings in Refs. [38,39], which, in another context, also relate the quantum Fisher information to convex roof constructions.

APPENDIX: SEMIDEFINITE PROGRAMS USED FOR OBTAINING THE GLOBAL OPTIMUM

Here we show how to compute $\mathcal{B}_{\mathrm{SPPT}}(\varrho,A) \equiv \mathcal{B}_{\mathrm{SE}_2}(\varrho,A)$ and $\mathcal{B}_{\mathrm{SE}_3}(\varrho,A)$ defined in Eq. (69) using semidefinite programming.

Semidefinite programming can be used to look for the infimum over symmetric PPT states, i.e., finding $\mathcal{B}_{\mathrm{SE}_2}(\varrho,A)$ given in Eq. (69). In this way we obtain an upper bound on the infimum for separable states as has also been used in Ref. [40]. The corresponding task can be formulated as a standard semidefinite program as

minimize $\quad 2\langle (A \otimes 1 - 1 \otimes A)^2 \rangle_{\varrho_0},$

subject to $\quad \varrho_0 = \sum_{m,n} |s_m(2,d)\rangle\langle s_n(2,d)|(\varrho_s)_{mn},$ (A1)

$$\varrho_0 \geqslant 0, \quad \mathrm{Tr}(\varrho_0) = 1, \quad \varrho_0^{T_1} \geqslant 0, \quad \mathrm{Tr}_2(\varrho_0) = \varrho.$$

Here $|s_n(2,d)\rangle$ denotes the basis states of the symmetric subspace for 2 qudits of dimension d. In order to ensure that the optimization is only over symmetric states, we define the density matrix ϱ_s, which is a two-qudit symmetric state given in the basis of the symmetric states. ϱ_0 is a density matrix given in the product basis, and T_1 means partial transpose according to the first subsystem.

Semidefinite programming can also be used to look for the infimum over symmetric states that have a PPT symmetric extension over N qudits appearing in $\mathcal{B}_{\mathrm{SE}_N}(\varrho,A)$ defined in Eq. (69). The corresponding task can be formulated as a semidefinite program as

minimize $\quad 2\langle (A \otimes 1 - 1 \otimes A)^2 \rangle_{\varrho_{12}},$

subject to $\varrho_0 = \sum_{m,n} |s_m(N,d)\rangle\langle s_n(N,d)|(\varrho_s)_{mn}, \quad \varrho_0 \geqslant 0,$
$$\mathrm{Tr}(\varrho_0) = 1, \quad \varrho_0^{T_{1,2,\ldots M}} \geqslant 0 \quad \text{for} \quad M = 1,2,\ldots,\lfloor \tfrac{N}{2} \rfloor,$$
$$\varrho_{12} = \mathrm{Tr}_{3,4,\ldots N}(\varrho_0), \quad \mathrm{Tr}_2(\varrho_{12}) = \varrho, \quad \text{(A2)}$$

where ϱ_{12} is the two-qudit reduced state of ϱ_0 and $\lfloor x \rfloor$ is the integer part of x. Now ϱ_s and ϱ_0 are N-qudit states. Here $|s_n(N,d)\rangle$ denotes the basis states of the symmetric subspace for N qudits of dimension d. For N qubits, these are

$$|s_n(N,d)\rangle = \binom{n}{N}^{-1/2} \sum_k \mathcal{P}_k(|1\rangle^{\otimes n}|0\rangle^{\otimes(N-n)}), \quad \text{(A3)}$$

for $n = 0,1,\ldots,N$, where the summation is over all different permutations. Now T_I means partial transposition according to the group of qudits given in I. Such an optimization has already been used in Ref. [41] for studying the entanglement properties of symmetric quantum states.

[1] V. Giovanetti, S. Lloyd, and L. Maccone, Science **306**, 1330 (2004).

[2] V. Giovannetti, S. Lloyd, and L. Maccone, Phys. Rev. Lett. **96**, 010401 (2006).

[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009); O. Gühne and G. Tóth, Phys. Rep. **474**, 1 (2009).

[4] A. S. Sørensen, L.-M. Duan, J. I. Cirac, and P. Zoller, Nature (London) **409**, 63 (2001).

[5] A. S. Sørensen and K. Mølmer, Phys. Rev. Lett. **86**, 4431 (2001).

[6] L. Pezzé and A. Smerzi, Phys. Rev. Lett. **102**, 100401 (2009).

[7] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezzé, and A. Smerzi, Phys. Rev. A **85**, 022321 (2012).

[8] G. Tóth, Phys. Rev. A **85**, 022322 (2012).

[9] R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, P. Hyllus, L. Pezzé, and A. Smerzi, Phys. Rev. Lett. **107**, 080504 (2011).

[10] R. Demkowicz-Dobrzański, J. Kołodyński, and M. Guţă, Nat. Commun. **3**, 1063 (2012).

[11] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).

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

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

[14] D. Petz, *Quantum Information Theory and Quantum Statistics* (Springer-Verlag, Heidelberg, 2008).

[15] D. Petz, J. Phys. A **35**, 929 (2002).

[16] P. Gibilisco, F. Hiai, and D. Petz, IEEE Trans. Inf. Theory **55**, 439 (2009).

[17] For a review on generalized variances, types of Fisher information, and covariances, see D. Petz and C. Ghinea, in *Introduction to Quantum Fisher Information*, QP-PQ: Quantum Probability and White Noise Analysis. Vol. 27, edited by R. Rebolledo and M. Orszag (World Scientific, Singapore, 2011), pp. 261–281.

[18] For the application of convex roofs in quantum information science see Ref. [21], and W. K. Wootters, Phys. Rev. Lett. **80**, 2245 (1998); C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996); A. Uhlmann, Entropy **12**, 1799 (2010).

[19] For an extension of the problem to covariance matrices, see Z. Léka and D. Petz, Some Decompositions of Matrix Variances (unpublished).

[20] E. P. Wigner and M. M. Yanase, Proc. Natl. Acad. Sci. USA **49**, 910 (1963).

[21] S. Hill and W. K. Wootters, Phys. Rev. Lett. **78**, 5022 (1997).

[22] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).

[23] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996); M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).

[24] A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri, Phys. Rev. A **69**, 022308 (2004); **71**, 032333 (2005).

[25] Every symmetric separable state can be written in the form Eq. (68); see J. K. Korbicz, J. I. Cirac, and M. Lewenstein, Phys. Rev. Lett. **95**, 120502 (2005).

[26] The random matrix is uniformly distributed according to the Hilbert-Schmidt norm. H.-J. Sommers and K. Życzkowski, J. Phys. A **37**, 8457 (2004).

[27] J. Sturm, SEDUMI, a MATLAB Toolbox for Optimization Over Symmetric Cones, http://sedumi.ie.lehigh.edu/.

[28] J. Löfberg, Yalmip: A Toolbox for Modeling and Optimization in MATLAB, in Proceedings of the CACSD Conference, Taipei, Taiwan, 2004; MATLAB, http://users.isy.liu.se/johanl/yalmip/.

[29] G. Tóth, Comput. Phys. Commun. **179**, 430 (2008).

[30] The generalized variances of Refs. [15,16] are concave and the various types of quantum Fisher information $\hat{F}^f(\varrho; A)$ are convex in both variables. Hence, $\hat{F}^f_Q[\varrho, A]$ is also convex in ϱ. See F. Hiai and D. Petz, Publ. RIMS Kyoto University **48**, 525 (2012).

[31] P. Hyllus, O. Gühne, and A. Smerzi, Phys. Rev. A **82**, 012337 (2010).

[32] F. Fröwis and W. Dür, Phys. Rev. Lett. **106**, 110402 (2011).

[33] F. Fröwis and W. Dür, Phys. Rev. Lett. **109**, 170401 (2012).

[34] F. Fröwis and W. Dür, New J. Phys. **14**, 093039 (2012).

[35] For entanglement detection based on local uncertainty relations, see O. Gühne, Phys. Rev. Lett. **92**, 117903 (2004).

[36] P. Hyllus (private communication).

[37] G. Tóth, D. Petz, O. Gühne, and P. Hyllus (unpublished).

[38] B. M. Escher, R. L. de Matos Filho, and L. Davidovich, Nat. Phys. **7**, 406 (2011).

[39] B. M. Escher, L. Davidovich, N. Zagury, and R. L. de Matos Filho, Phys. Rev. Lett. **109**, 190404 (2012).

[40] G. Tóth, W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, and H. Weinfurter, New J. Phys. **11**, 083002 (2009).

[41] G. Tóth and O. Gühne, Phys. Rev. Lett. **102**, 170503 (2009).