

**Majorization formulation of uncertainty in quantum mechanics**

M. Hossein Partovi\*

*Department of Physics and Astronomy, California State University, Sacramento, California 95819-6041, USA*

(Received 15 December 2010; published 23 November 2011)

Heisenberg's uncertainty principle is formulated for a set of generalized measurements within the framework of majorization theory, resulting in a partial uncertainty order on probability vectors that is stronger than those based on quasientropic measures. The theorem that emerges from this formulation guarantees that the uncertainty of the results of a set of generalized measurements without a common eigenstate has an inviolable lower bound which depends on the measurement set but not the state. A corollary to this theorem yields a parallel formulation of the uncertainty principle for generalized measurements corresponding to the entire class of quasientropic measures. Optimal majorization bounds for two and three mutually unbiased bases in two dimensions are calculated. Similarly, the leading term of the majorization bound for position and momentum measurements is calculated which provides a strong statement of Heisenberg's uncertainty principle in direct operational terms. Another theorem provides a majorization condition for the least-uncertain generalized measurement of a given state with interesting physical implications.

DOI: [10.1103/PhysRevA.84.052117](https://doi.org/10.1103/PhysRevA.84.052117)

PACS number(s): 03.65.Ta, 03.65.Ca, 03.67.Mn

**I. INTRODUCTION**

Heisenberg's uncertainty principle codifies certain inherent limitations on the simultaneous knowledge of observables in the microscopic realm, and as such constitutes one of the conceptual pillars of quantum theory [1]. The limitations implied by the uncertainty principle played an important role in the celebrated Bohr-Einstein debates and the formative years of quantum theory and have since served as a source of insight on the structure and behavior of microscopic systems. Having recognized that the noncommutativity of a pair of observables implies an irreducible indeterminacy in the simultaneous knowledge of their values, Heisenberg presented semiquantitative arguments to establish a universal lower bound of the order of Planck's constant for the product of their uncertainties. Heisenberg's arguments were subsequently formulated in a mathematically precise manner by Kennard [2] and extended by Robertson [3] and Schrödinger [4], all of whom adopted the square root of variance as the measure of uncertainty. The variance formulation of the uncertainty principle, which is the one familiar from textbook accounts, is often useful when applied to canonically conjugate observables for estimating spectral and structural properties of microsystems.

More than half a century later it was realized that entropy is a more effective measure for capturing the information theoretical aspects of uncertainty, especially when applied to noncanonical observables. The key idea was to quantify uncertainty as the information associated with the probabilities of measurement outcomes rather than the variance in the values of the measured observable. In a seminal paper [5], Deutsch argued that such a measure would be superior to the traditional one when dealing with observables of finite rank and developed an entropic formulation of the uncertainty principle based on an inviolable lower bound for the sum of Shannon entropies associated with projective measurements of noncommuting observables. He also emphasized that, in contrast to the variance formulation in its generalized form, the

irreducible lower bound in the entropic formulation depends on the observables but not the state.

In a subsequent work, this author showed that an extension of Deutsch's formulation to infinite-rank observables such as position and momentum, whether discrete or continuous, entails a nontrivial consideration of the resolution of the measuring device [6]. The resulting entropic measure was a more realistic description of uncertainty for canonical observables than the variance formulation, and in contrast to the latter, was mathematically well defined for any state of the system and an arbitrary set of observables, even if infinite rank. Consequently, while Deutsch's primary objective was to deal with the shortcomings of the variance formulation when applied to discrete spectra and finite-dimensional observables, the extension to continuous spectra such as position and momentum turned out to be quite potent as well.

The entropic formulation has since been refined by deriving sharper bounds [7,8] and extended by introducing mutually unbiased observables [8], generalized measurements [9], alternative entropy functions such as those of Tsallis [10] and Rényi [11], and other innovations [12]. Its applications include the quantum formulation of Jaynes' maximum entropy method and time-energy uncertainty relations [13] and, more recently, quantum cryptography, information locking, and entanglement detection [12,14]. In particular, the recent applications to quantum information and entanglement theory clearly underscore the importance of basing uncertainty considerations on probability vectors resulting from measurements rather than values of measured observables. It should also be noted here that the idea of entropy as a measure of uncertainty has an interesting earlier history [15].

In this paper we develop a majorization formulation of the uncertainty principle. Majorization provides a partial order on probability vectors which characterizes the degree of their disorder, or uncertainty, and is naturally suited for application to measurement results. It is based on the intuitive but surprisingly powerful notion that a probability vector which is a mixture of the permutations of another is more disordered. This simple condition gives rise to a measure

\*hpartovi@csus.edu

of uncertainty that is more discriminating than any measure based on a symmetric, concave function defined on probability vectors. We shall refer to this class of measures, which includes the Shannon and Tsallis (but not Rényi) entropies, as “quasientropic.” Thus the majorization order implies any order in the quasi-entropic class, but not vice versa [16]. As a result, the main theorem of the present formulation directly implies the existence of an entire class of scalar formulations of the uncertainty principle for generalized measurements based on quasientropic functions. This generalization serves, *inter alia*, to extend the standard (Shannon) entropic measure of uncertainty to a set of generalized measurements.

It is important to note here that while quasientropic order is total and based on a single inequality, the majorization order is partial (i.e., not every pair of probability vectors can be ordered) and rests on  $N - 1$  inequalities, where  $N$  is the number of nonzero entries in the less uncertain probability vector under comparison (see Sec. II B for details). Consequently, while the majorization order is equivalent to quasientropic ones for  $N \leq 2$ , it becomes progressively more stringent, and captures more information, with increasing  $N$ . It is this feature of majorization, namely the matching of its complexity to that of the object being probed, that makes it particularly effective as a comparator of disorder. Furthermore, when a pair of probability vectors is found to be incomparable according to the majorization order, it is a signal that there will be conflicting order determinations within the quasientropic class. This is because unanimity with respect to order determination within the quasientropic class guarantees comparability as well as agreement by the majorization order (the converse holds as well; see Sec. II B). Clearly then, when majorization finds a pair incomparable, it is because there is not sufficient reason to declare one or the other as more disordered, and, in fact, the order determined for the pair by one quasientropic measure, e.g., the standard measure based on Shannon entropy, will be contradicted by another. It should, therefore, come as no surprise that some of the most fundamental results of quantum information and entanglement theory are based on majorization relations [17]. We believe that its use in the present formulation likewise serves to extend the reach and power of the uncertainty principle.

The rest of this paper is organized as follows: In Sec. II A we describe measurement types and establish notation, in Sec. II B we introduce the relevant elements of majorization theory, and in Sec. II C we relate concepts of measurement uncertainty to majorization relations. In Secs. III A and III B we establish the central result of this paper, Theorem 1, parts A and B, and state the physical content of the uncertainty principle in majorization terms. In Sec. III C, we construct the class of quasientropic measures of uncertainty as a corollary to Theorem 1. We apply Theorem 1 to mutually unbiased observables in Secs. IV A and IV B and to position and momentum in Sec. V. In Sec. VI we establish the least-uncertain measurement of a quantum state in Theorem 2 and state its relation to the spectrum and von Neumann entropy of the state. We conclude the paper with a few remarks in Sec. VII.

## II. MEASUREMENT UNCERTAINTY AND MAJORIZATION

We begin our analysis by introducing the concepts and methods that underlie the majorization formulation of uncertainty in the following subsections.

### A. Measurements

To establish our notation and nomenclature, we find it convenient to start by defining types of measurement. A generalized measurement is defined by a set of measurement operators  $\{\hat{\mathcal{M}}_\alpha\}$  subject to the completeness condition  $\sum_\alpha \hat{E}_\alpha = \hat{1}$ , where  $\hat{E}_\alpha = \hat{\mathcal{M}}_\alpha^\dagger \hat{\mathcal{M}}_\alpha$  is called a measurement element and the index  $\alpha$  identifies the possible measurement outcomes [18]. Note that each measurement element is a bounded, positive, self-adjoint operator whose norm cannot exceed unity. The probability that outcome  $\alpha$  turns up in a measurement of the state  $\hat{\rho}$  is given by the Born rule  $\mathcal{P}_\alpha(\hat{\rho}) = \text{tr}[\hat{E}_\alpha \hat{\rho}]$ , with the postmeasurement state given by  $\hat{\rho}'_\alpha = \hat{\mathcal{M}}_\alpha \hat{\rho} \hat{\mathcal{M}}_\alpha^\dagger / \mathcal{P}_\alpha$  [19]. If  $\hat{\rho}$  is not a pure state, then  $\hat{\rho}'_\alpha$  will not, in general, be pure either unless  $\hat{\mathcal{M}}_\alpha$  is of rank 1 [20]. A measurement is rank 1 if every measurement operator is of rank 1. Rank-1 measurements are thus seen to have the highest resolution among generalized measurements in the sense that the range of their measurement operators consists of a single pure state.

A generalized measurement can always be considered a restriction of a more basic type of measurement, namely a *projective* measurement performed on an enlarged system, to the system under generalized measurement [18]. A projective measurement is commonly associated with an observable of the system, which would be represented by some self-adjoint operator  $\hat{M}$ . Such a measurement entails a partitioning of the spectrum of  $\hat{M}$  into a collection of subsets  $\{b_\alpha^M\}$  called *bins* [6]. This partition induces a corresponding partition of the Hilbert space into orthogonal subspaces with the corresponding projection operators being the measurement operators [21]. Thus, for projective measurements,  $\hat{\mathcal{M}}_\alpha^M = \hat{\Pi}_\alpha^M$ , where  $\hat{\Pi}_\alpha^M$  is the projection operator onto the subspace corresponding to  $b_\alpha^M$ . We call a projective measurement *maximal* if each bin consists of a single point of the spectrum of the measured observable. Note that a maximal measurement of an observable with a nondegenerate spectrum, the type usually described in textbook accounts, is rank 1. It is important to understand that physically realizable measurements are limited to a finite set of outcomes so any measurement of an observable with a continuous spectrum such as position or momentum, or even an infinite discrete spectrum such as the energy of a harmonic oscillator, must necessarily involve infinite-rank measurement operators and cannot be maximal [6]. Equivalently, only systems describable by finite-dimensional Hilbert spaces (such as spin systems) admit maximal measurements.

### B. Majorization relations

We now turn to a brief introduction to the basics of majorization relations, a topic that has found important applications in quantum information and entanglement theory in recent

years [16,17]. With every vector  $\lambda$  we associate another vector  $\lambda^\downarrow$  which is obtained from  $\lambda$  by arranging the components of the latter in a descending (i.e., nonincreasing) order. Then, given a pair of vectors  $\lambda^1$  and  $\lambda^2$ ,  $\lambda^1$  is said to be majorized by  $\lambda^2$  and written  $\lambda^1 < \lambda^2$  if  $\sum_{i=1}^j \lambda_i^{1\downarrow} \leq \sum_{i=1}^j \lambda_i^{2\downarrow}$  for  $j = 1, 2, \dots, d$ , where  $d$  is the larger of the two dimensions and trailing zeros are added where needed. An equivalent but intuitively more compelling definition is that  $\lambda^1 < \lambda^2$  if and only if  $\lambda^1$  equals a mixture of permutations of  $\lambda^2$ . As stated earlier, the majorization relation is a *partial* order, i.e., that not every two vectors are comparable under majorization. It is important to understand that this is not a shortcoming of majorization but rather a consequence of its more rigorous protocol for ordering uncertainty. Indeed, as mentioned in Sec. I, for any function  $F(\lambda)$  of the quasientropic type like Shannon and Tsallis entropies,  $\lambda^1 < \lambda^2$  implies  $F(\lambda^1) \geq F(\lambda^2)$  but not conversely [16]. On the other hand, if for *every* such function  $F(\lambda)$  we have  $F(\lambda^1) \geq F(\lambda^2)$ , then  $\lambda^1 < \lambda^2$ . Clearly, the majorization relation as a comparator of disorder is stronger than any single quasientropic measure and, in a sense, is equivalent to all such measures taken collectively.

Another concept needed in the following is that of the *infimum* of a set of  $N$  vectors, defined as the vector that is majorized by every element of the set and, in turn, majorizes any vector with that property [22]. The *supremum* is similarly defined as the vector that majorizes every element of the set and is, in turn, majorized by any vector with that property [23]. To construct these extremal vectors, we consider a vector  $\mu^{\text{inf}}$  with components

$$\mu_0^{\text{inf}} = 0, \quad \mu_j^{\text{inf}} = \min \left( \sum_{i=1}^j \lambda_i^{1\downarrow}, \sum_{i=1}^j \lambda_i^{2\downarrow}, \dots, \sum_{i=1}^j \lambda_i^{N\downarrow} \right), \quad 1 \leq j \leq d_{\text{max}}, \quad (1)$$

where  $d_{\text{max}}$  is the largest dimension found in the set [24]. The desired infimum is then given by

$$\lambda_i^{\text{inf}} = [\text{inf}(\lambda^1, \lambda^2, \dots, \lambda^N)]_i = \mu_i^{\text{inf}} - \mu_{i-1}^{\text{inf}}, \quad (2)$$

where  $1 \leq i \leq d_{\text{max}}$ . One can show that  $\lambda^{\text{inf}}$  as given by Eq. (2) is a descending sequence [25]. While the construction given in Eq. (2) guarantees that  $\lambda^{\text{inf}}$  is majorized by every element of the set, the descending property guarantees that any other vector with that property is in turn majorized by  $\lambda^{\text{inf}}$ .

A parallel construction to the above with “min” in Eq. (1) replaced with “max” yields a sequence that majorizes every element of the set, but one that does not necessarily emerge in a descending order and may, therefore, fail to be majorized by any other sequence that has the same property as required. In such a case, the sequence so obtained,  $\tilde{\lambda}$ , can be modified by a “flattening” process that, while maintaining the property of majorizing every element of the set, culminates in a sequence that is descending as well. The flattening process starts with  $\tilde{\lambda}$  and, for every pair of components violating the descending property, i.e.,  $\tilde{\lambda}_{i+1} > \tilde{\lambda}_i$ , replaces the pair by their mean such that the updated elements are  $\tilde{\lambda}'_i = \tilde{\lambda}'_{i+1} = (\tilde{\lambda}_i + \tilde{\lambda}_{i+1})/2$ . This process of flattening is then continued until a descending sequence corresponding to the supremum  $\lambda^{\text{sup}}$  is obtained [26].

### C. Uncertainty

We are now in a position to characterize uncertainty by means of majorization relations. The probability vector  $\mathcal{P}^X(\hat{\rho})$  resulting from a measurement  $X$  on a state  $\hat{\rho}$  is said to be *uncertain* if it is majorized by  $\mathcal{I} = (1, 0, \dots, 0)$  but not equal to it. As such,  $\mathcal{P}^X(\hat{\rho})$  is said to be strictly majorized by  $\mathcal{I}$  and written  $\mathcal{P}^X(\hat{\rho}) < \mathcal{I}$ . Similarly,  $\hat{\rho}$  is more uncertain with respect to measurement  $X$  than with respect to  $Y$  if  $\mathcal{P}^X(\hat{\rho}) < \mathcal{P}^Y(\hat{\rho})$ . Furthermore, we define the joint uncertainty of a pair of measurements  $X$  and  $Y$  by means of the outer product  $\mathcal{P}^X \otimes \mathcal{P}^Y$ , i.e.,  $\mathcal{P}_{\alpha\beta}^{X \oplus Y} = \mathcal{P}_\alpha^X \mathcal{P}_\beta^Y$ . Since  $H(\mathcal{P}^X \otimes \mathcal{P}^Y) = H(\mathcal{P}^X) + H(\mathcal{P}^Y)$ , where  $H(\cdot)$  is the Shannon entropy function, this definition is seen to be consistent with its entropic counterpart. As stated earlier,  $\mathcal{P}^X < \mathcal{P}^Y$  implies  $H(\mathcal{P}^X) \geq H(\mathcal{P}^Y)$  but not conversely. Note that the foregoing definitions naturally extend to an arbitrary number of states and measurements.

It is worth repeating here that the partial nature of the majorization order implies that not all measurement results are comparable under our uncertainty order. A concomitant of partial order is that the infimum or supremum of a set of probability vectors will in general not be a member of the set. In other words, a set of probability vectors does not in general have a least-uncertain element, although it may have multiple elements that are not more uncertain than any other probability vector [27]. While one may wish to do away with these unfamiliar features, they are nevertheless necessary and meaningful features of a high-resolution comparator of uncertainty, as discussed in Sec. I.

## III. MAJORIZATION FORMULATION OF THE UNCERTAINTY PRINCIPLE

The information theoretical expression of the uncertainty principle may be stated as the requirement that the information available from a set of measurements performed on a system is subject to an irreducible level of uncertainty unless the measurements have a common eigenstate. In the latter case, the outcome of every measurement would be certain and unique, and the information so obtained would be complete for the given set of measurements. An important additional requirement is that the said irreducible level of uncertainty be a property of the measurement set and not depend on the state of the system. The entropic formulation implements these requirements by identifying the sum of the entropies associated with the measurements in the set as the measure of their joint uncertainty [5]. The majorization formulation, by contrast, relies on the outer product of probability vectors resulting from the measurements as representative of their joint uncertainty. An example of this definition for a set of two measurements was stated in Sec. II C. The corresponding irreducible level of uncertainty is then defined to be the supremum of such outer products as all possible states are considered. This supremum is then guaranteed to be less uncertain than the probability vector resulting from the measurement of any possible state of the system. As also noted in Sec. II C, the Shannon entropy of the outer product of a set of probability vectors equals the sum of their entropies, which guarantees that the two formulations are consistent. It also implies that a formulation

of the uncertainty principle in terms of majorization implies a parallel entropic formulation, as will be demonstrated in Sec. III C.

In the following we will establish a theorem that embodies the above statements in a mathematically precise manner. We will treat the cases where measurement elements have discrete spectra and behave similarly to Hermitian matrices acting on finite-dimensional spaces, namely *compact* measurement elements, separately from other cases such as canonical observables and continuous spectra [28]. The reason for this separate treatment is simply that the very notion of an eigenstate may lose proper mathematical meaning for noncompact measurement elements [29]. Needless to say, the physical content of the uncertainty principle is unaffected by these mathematical details.

### A. Compact measurement elements

Here we consider a measurement set whose elements are compact. Since measurement elements are bounded, positive, self-adjoint operators, the condition of compactness will guarantee that they only have a discrete spectrum consisting of positive eigenvalues, each with a finite multiplicity, except possibly for a clustering of eigenvalues near zero [28]. Since we will primarily be concerned with eigenvalues away from zero, compact elements essentially behave as positive Hermitian matrices of finite dimension for our purposes. Note, however, that while finite-rank measurement elements are necessarily compact, the converse is not true. Note also that most discussions of uncertainty in the literature deal with finite-rank, hence, compact, elements and often just rank-1 projection operators as in the case of maximal projective measurements.

*Theorem 1A.* Let  $\{\mathcal{P}^X(\hat{\rho}), \mathcal{P}^Y(\hat{\rho}), \dots, \mathcal{P}^Z(\hat{\rho})\}$  be the set of probability vectors resulting from a set of generalized measurements  $\{X, Y, \dots, Z\}$  with compact elements on the state  $\hat{\rho}$ . Then

$$\mathcal{P}^X(\hat{\rho}) \otimes \mathcal{P}^Y(\hat{\rho}) \otimes \dots \otimes \mathcal{P}^Z(\hat{\rho}) \prec \mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} \prec \mathcal{I}, \quad (3)$$

where

$$\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \sup_{\hat{\rho}} [\mathcal{P}^X(\hat{\rho}) \otimes \mathcal{P}^Y(\hat{\rho}) \otimes \dots \otimes \mathcal{P}^Z(\hat{\rho})], \quad (4)$$

unless the measurement elements  $\{\hat{E}^X, \hat{E}^Y, \dots, \hat{E}^Z\}$  have a common eigenstate in which case  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$ .

It is important to note that  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  depends on the measurement set but is independent of the state  $\hat{\rho}$ . Note also that as the supremum of all possible measurement outcomes for the measurement set, it is the probability vector that sets the irreducible lower bound to uncertainty for the set. As such, it is analogous to variance or entropic lower bounds for existing formulations of the uncertainty principle. Unlike the scalar bounds of the variance and entropic formulations, however,  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  is, in general, a vector quantity whose dimension is variable and grows with the complexity of the measurement set. In addition,  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  is in general not realizable on any state of the system [23] except under special conditions. In other words, there is generally no such thing as a “minimum uncertainty state” within the majorization framework, as alluded to in Sec. II C. An important exception

to this statement is the special case of zero uncertainty for which  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$ , signaling the existence of a common eigenstate for the measurement elements as asserted by Theorem 1A. We will illustrate these and other properties of majorization bounds for three archetypal cases in Secs. IV and V.

To establish Theorem 1A, we need to show that  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  is strictly majorized by  $\mathcal{I}$ , equivalently, that its largest component is strictly less than unity, if the measurement elements do not have a common eigenstate. Suppose, on the contrary, that the largest component of  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  does equal unity while the measurement elements in the set do not possess a common eigenstate. There then must exist a set of indices  $(\alpha^*, \beta^*, \dots, \gamma^*)$  such that  $\sup_{\hat{\rho}} [\mathcal{P}_{\alpha^*}^X(\hat{\rho}) \mathcal{P}_{\beta^*}^Y(\hat{\rho}) \dots \mathcal{P}_{\gamma^*}^Z(\hat{\rho})] = 1$  or, in terms of measurement elements,

$$\sup_{\hat{\rho}} [\text{tr}(\hat{E}_{\alpha^*}^X \hat{\rho}) \text{tr}(\hat{E}_{\beta^*}^Y \hat{\rho}) \dots \text{tr}(\hat{E}_{\gamma^*}^Z \hat{\rho})] = 1. \quad (5)$$

Since the elements of each measurement are positive, Hermitian operators whose sum equals the identity operator, e.g.  $\sum_{\alpha} \hat{E}_{\alpha}^X = \hat{1}$ , they must all be bounded operators with norms not exceeding unity, i.e.,  $\|\hat{E}_{\alpha}^X\| \leq 1$  for every  $\alpha$ , with similar conditions for all other measurement elements in the set. But this implies that  $\text{tr}(\hat{E}_{\alpha^*}^X \hat{\rho}) \leq 1$ , with similar conditions for all measurement elements in the set. Consequently, the only way Eq. (5) can be satisfied under the stated constraint on the norms is that (a) each measurement element is of unit norm and (b) each trace term in Eq. (5) equals unity, i.e.,

$$\sup_{\hat{\rho}} \text{tr}(\hat{E}_{\alpha^*}^X \hat{\rho}, \hat{E}_{\beta^*}^Y \hat{\rho}, \dots, \hat{E}_{\gamma^*}^Z \hat{\rho}) = (1, 1, \dots, 1). \quad (6)$$

Note that by probability conservation, if a subscript on an entry in the left-hand side of Eq. (6) is changed, the corresponding entry on the right-hand side must vanish. Physically, Eq. (6) implies the existence of states for which the outcome of every measurement is essentially determinate, and the joint results are basically without any uncertainty, since the outcome of the measurement set  $\{X, Y, \dots, Z\}$  will very nearly all be events in those measurement “bins” that correspond to the subscript set  $(\alpha^*, \beta^*, \dots, \gamma^*)$ .

Mathematically, on the other hand, we note that the measurement elements in Eq. (6) are positive, compact Hermitian operators of unit norm, with discrete eigenvalues and corresponding eigenfunctions that are complete. In addition, all nonzero eigenvalues have finite multiplicity. We may, therefore, conclude that there exists a state  $\hat{\rho}^*$  that realizes the equalities of Eq. (6), i.e., that  $\text{tr}(\hat{E}_{\alpha^*}^X \hat{\rho}^*) = \text{tr}(\hat{E}_{\beta^*}^Y \hat{\rho}^*) = \dots = \text{tr}(\hat{E}_{\gamma^*}^Z \hat{\rho}^*) = 1$ , and that  $\hat{\rho}^*$  is a common eigenstate of the measurement elements  $\{\hat{E}_{\alpha^*}^X, \hat{E}_{\beta^*}^Y, \dots, \hat{E}_{\gamma^*}^Z\}$  with eigenvalues unity (and of all other elements with eigenvalue zero) [30]. However, this conclusion contradicts our starting assumption, thereby completing the proof of Theorem 1A.

### B. Noncompact measurement elements

In the foregoing paragraph we used the compactness property of the measurement elements to deduce the existence of a common eigenstate for them from Eq. (6). In the general case where noncompact elements may be present, the measurement elements may not even have properly defined eigenstates or eigenvalues, common or otherwise [29]. Of



course, the physical content of the uncertainty principle is still captured by Eq. (6). We will, therefore, use the physically equivalent notion of an *approximate* eigenstate in this case, defined as follows: If for some number  $a$  and any  $\epsilon > 0$  there exists a state  $\psi(\epsilon)$  such that  $\|(\hat{A} - a)\psi(\epsilon)\| < \epsilon$ , then  $\psi(\epsilon)$  is said to be an approximate eigenstate of  $\hat{A}$ .

Clearly, Eq. (6) implies the existence of an approximate common eigenstate for the measurement elements therein, thus yielding the desired result. Nevertheless, it is useful to develop a separate formulation and an alternative proof for this general case, especially with a view to deriving majorization bounds for position and momentum measurements in Sec. V.

*Theorem 1B.* Let  $\{\mathcal{P}^X(\hat{\rho}), \mathcal{P}^Y(\hat{\rho}), \dots, \mathcal{P}^Z(\hat{\rho})\}$  be the set of probability vectors resulting from a set of generalized measurements  $\{X, Y, \dots, Z\}$  on the state  $\hat{\rho}$ . Equation (3) of Theorem 1A then holds unless the measurement elements  $\{\hat{E}^X, \hat{E}^Y, \dots, \hat{E}^Z\}$  have an approximate common eigenstate in which case  $\mathcal{P}_{\sup}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$ .

To establish this result, we first note that, by definition,

$$\|\hat{E}_\alpha^X + \hat{E}_\beta^Y + \dots + \hat{E}_\gamma^Z\| \geq \text{tr}[(\hat{E}_\alpha^X + \hat{E}_\beta^Y + \dots + \hat{E}_\gamma^Z)\hat{\rho}]$$

for any (normalized) density operator  $\hat{\rho}$ . The right-hand side of this inequality is, by definition, equal to  $[\mathcal{P}_\alpha^X(\hat{\rho}) + \mathcal{P}_\beta^Y(\hat{\rho}) + \dots + \mathcal{P}_\gamma^Z(\hat{\rho})]$ , which is a sum of non-negative numbers. Since the arithmetic mean of a set of non-negative numbers is never exceeded by their geometric mean, we have the inequality

$$[\mathcal{P}_\alpha^X(\hat{\rho}) + \mathcal{P}_\beta^Y(\hat{\rho}) + \dots + \mathcal{P}_\gamma^Z(\hat{\rho})]/n \geq [\mathcal{P}_\alpha^X(\hat{\rho})\mathcal{P}_\beta^Y(\hat{\rho}) \dots \mathcal{P}_\gamma^Z(\hat{\rho})]^{1/n}$$

which governs the set of probability vectors.

Combining the above pair of inequalities, we arrive at the important conclusion that

$$\mathcal{P}_\alpha^X(\hat{\rho})\mathcal{P}_\beta^Y(\hat{\rho}) \dots \mathcal{P}_\gamma^Z(\hat{\rho}) \leq (\|\hat{E}_\alpha^X + \hat{E}_\beta^Y + \dots + \hat{E}_\gamma^Z\|/n)^n, \quad (7)$$

where  $n$  is the number of measurements in the set.

At this point we follow the proof of Theorem 1A by assuming, contrary to Theorem 1B, that  $\mathcal{P}_{\sup}^{X \oplus Y \oplus \dots \oplus Z} = \mathcal{I}$  while the measurement elements do not possess an approximate common eigenstate. But then there must exist a set of indices  $(\alpha^*, \beta^*, \dots, \gamma^*)$  such that  $\sup_{\hat{\rho}} [\mathcal{P}_{\alpha^*}^X(\hat{\rho})\mathcal{P}_{\beta^*}^Y(\hat{\rho}) \dots \mathcal{P}_{\gamma^*}^Z(\hat{\rho})] = 1$ . However, this equality together with the inequality in Eq. (7) imply that

$$\|\hat{E}_{\alpha^*}^X + \hat{E}_{\beta^*}^Y + \dots + \hat{E}_{\gamma^*}^Z\| = n. \quad (8)$$

Since the measurement elements appearing in Eq. (8) are positive operators with norms not exceeding unity, we must conclude that they are all in fact of unit norm. Equation (8) further implies that for any  $\epsilon > 0$ , there must exist a state  $\hat{\rho}^*$  such that  $|1 - \text{tr}(\hat{E}\hat{\rho}^*)| < \epsilon$ , where  $\hat{E}$  stands for every measurement element in Eq. (8). But this implies the existence of an approximate common eigenstate contrary to our assumption, thus completing the proof of Theorem 1B.

We are now in a position to summarize the physical content of the uncertainty principle in the framework of majorization theory.

*The uncertainty principle.* The joint results of a set of generalized measurements of a given state are no less uncertain than a probability vector that depends on the measurement set

but not the state and is itself uncertain unless the measurement elements have a common eigenstate.

In the above statement, we have dropped the qualification ‘‘approximate’’ in referring to eigenfunctions since this is the common practice in the physics literature as well as the fact that it makes little difference for the physical content of the uncertainty principle.

### C. Quasientropic formulations of uncertainty

Theorems 1A and 1B immediately imply a parallel formulation of the uncertainty principle for generalized measurements based on quasientropic measures. We recall that a quasientropic measure of uncertainty is any symmetric, concave function of the components of  $\mathcal{P}^{X \oplus Y \oplus \dots \oplus Z}(\hat{\rho})$  resulting from generalized measurements  $\{X, Y, \dots, Z\}$  on the state  $(\hat{\rho})$  [31]. Specifically, for every quasientropic function  $F$ , we define a scalar uncertainty measure according to

$$\mathcal{U}^{X \oplus Y \oplus \dots \oplus Z}(F, \hat{\rho}) = F[\mathcal{P}^{X \oplus Y \oplus \dots \oplus Z}(\hat{\rho})] - F(\mathcal{I}), \quad (9)$$

where we have normalized the measure such that it vanishes when measurement results have zero uncertainty and is positive otherwise.

A special class of quasientropic measures is obtained if we choose

$$F[\mathcal{P}^{X \oplus Y \oplus \dots \oplus Z}(\hat{\rho})] = \sum_{\alpha, \beta, \dots, \gamma} f[\mathcal{P}_\alpha^X(\hat{\rho})\mathcal{P}_\beta^Y(\hat{\rho}) \dots \mathcal{P}_\gamma^Z(\hat{\rho})], \quad (10)$$

where  $f$  is a concave function of a single variable. Note that  $F$  as constructed in Eq. (10) is manifestly symmetric and, as a sum of concave functions, it is also concave.

The standard entropic measure of uncertainty [5] corresponds to the choice  $f(x) = H(x) = -x \ln(x)$  in Eq. (10), in which case the sum on the right-hand side simplifies to the sum of Shannon entropies for each measurement. This example serves to demonstrate that the class of measures introduced in Eq. (9) is a vast generalization of the standard entropic measure of uncertainty, not only in the functional form of the uncertainty measure but also in the fact that it allows for any number of generalized measurements. Furthermore, as stated in the following formulation of the uncertainty principle based on quasientropic functions, the corresponding lower bounds to uncertainty are given by the associated majorization bound.

*Corollary 1.* Let  $\{\mathcal{P}^X(\hat{\rho}), \mathcal{P}^Y(\hat{\rho}), \dots, \mathcal{P}^Z(\hat{\rho})\}$  be the set of probability vectors resulting from a set of generalized measurements  $\{X, Y, \dots, Z\}$  on the state  $\hat{\rho}$ . Then, for any uncertainty measure  $\mathcal{U}^{X \oplus Y \oplus \dots \oplus Z}(F, \hat{\rho})$  as defined in Eq. (9), and with  $\mathcal{P}_{\sup}^{X \oplus Y \oplus \dots \oplus Z}$  as defined in Eq. (4), we have

$$\mathcal{U}^{X \oplus Y \oplus \dots \oplus Z}(F, \hat{\rho}) \geq \mathcal{U}_{\min}^{X \oplus Y \oplus \dots \oplus Z} = F[\mathcal{P}_{\sup}^{X \oplus Y \oplus \dots \oplus Z}] > 0, \quad (11)$$

unless the measurement elements  $\{\hat{E}^X, \hat{E}^Y, \dots, \hat{E}^Z\}$  have an approximate common eigenstate or a common eigenstate if the elements are compact. In either of these cases,  $\mathcal{U}^{X \oplus Y \oplus \dots \oplus Z}(F, \hat{\rho})$  vanishes.

This general result is an immediate consequence of Theorems 1A and 1B and the quasientropic nature of the underlying measures. It is important to understand that the uncertainty bound given in Eq. (11) is valid, and, in fact,

optimal, for the entire class of quasientropic measures. The optimality property of  $F[\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}]$  is a consequence of the optimality of  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$ , which is, in fact, the defining characteristic of the latter. As such,  $F[\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}]$  cannot be expected to be optimal for individual members of the quasientropic class such as the Shannon or Tsallis measures.

As an example, we will calculate the Shannon-entropic uncertainty bound given by Eq. (11) for maximal projective measurements of the three components of a spin- $\frac{1}{2}$  system. The majorization bound for this measurement,  $\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}$ , is given in Eq. (16) of Sec. IV B. Therefore, the desired entropic bound is given by  $H(\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z})$ , so we can write

$$\mathcal{U}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}(H, \hat{\rho}) \geq H(\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}) = 1.23. \quad (12)$$

The fact that a single majorization bound  $\mathcal{P}_{\text{sup}}^{X \oplus Y \oplus \dots \oplus Z}$  generates an uncertainty bound for the entire quasientropic class of measures is, of course, a consequence of its power as a comparator of disorder, as discussed in Sec. I.

In the remainder of this paper we will explore certain consequences of the majorization formulation of the uncertainty principle. Our objective will be to illustrate the power and reach of the majorization formulation, primarily in applications that are familiar from the traditional variance and entropic formulations of the uncertainty principle.

#### IV. MUTUALLY UNBIASED OBSERVABLES

Our first application is to projective measurements of mutually unbiased observables in two-dimensional Hilbert spaces.

##### A. Two spin- $\frac{1}{2}$ components

The simplest example of Eq. (3) is a maximal projective measurement of a pair of mutually unbiased observables in a two-dimensional Hilbert space, e.g., a measurement of  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$  on a spin-1/2 system. The state of a spin-1/2 system can, in general, be represented as  $\hat{\rho} = (1 + \hat{\sigma} \cdot \mathbf{p})/2$ , corresponding to a polarization vector  $\mathbf{p}$ . The measurement elements, on the other hand, are  $\hat{E}_{1,2}^{\sigma_x} = (1 \pm \hat{\sigma}_x)/2$  and  $\hat{E}_{1,2}^{\sigma_y} = (1 \pm \hat{\sigma}_y)/2$ . A calculation using these quantities gives  $\mathcal{P}^{\sigma_x}(\hat{\rho}) = [(1 + p_x)/2, (1 - p_x)/2]$  and  $\mathcal{P}^{\sigma_y}(\hat{\rho}) = [(1 + p_y)/2, (1 - p_y)/2]$ , whereby we find

$$\mathcal{P}^{\sigma_x \oplus \sigma_y}(\hat{\rho}) = [(1 + p_x)(1 + p_y)/4, (1 + p_x)(1 - p_y)/4, (1 - p_x)(1 + p_y)/4, (1 - p_x)(1 - p_y)/4]. \quad (13)$$

The next step is to find the supremum of  $\mathcal{P}^{\sigma_x \oplus \sigma_y}(\hat{\rho})$  as  $\hat{\rho}$ , or, equivalently,  $\mathbf{p}$ , is varied. Following the construction of Eq. (2) *et seq.*, we first determine  $\mu_1^{\text{sup}}$  by finding the maximum value of a single component of  $\mathcal{P}^{\sigma_x \oplus \sigma_y}(\hat{\rho})$  in Eq. (13), then  $\mu_2^{\text{sup}}$  by finding the maximum value of the sum of a pair of components of  $\mathcal{P}^{\sigma_x \oplus \sigma_y}(\hat{\rho})$ , and so on. Implementing this process, we find

$$\mu^{\text{sup}} = [0, (1 + 1/\sqrt{2})^2/4, 1, 1, 1], \quad (14)$$

where  $\mu_1^{\text{sup}}$  obtains for  $p_z = 0$  and  $\|\mathbf{p}_x\| = |p_y| = 1/\sqrt{2}$ , and the next three components for  $\mathbf{p}$  equal to a unit vector along

either the  $x$  or  $y$  axis. Using Eq. (14) and the counterpart of Eq. (2) for the supremum, we arrive at

$$\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y} = [(1.5 + \sqrt{2})/4, (2.5 - \sqrt{2})/4, 0, 0]. \quad (15)$$

Thus, a measurement of  $\hat{\sigma}_x$  and  $\hat{\sigma}_y$  on any spin-1/2 system will yield results no more certain than this supremum. Notice that while the supremum in Eq. (15) is less uncertain than is possible for any state of the system, it is not itself a possible probability vector in any actual measurement [23]. This is, therefore, an instance of a measurement where there is no ‘‘minimum uncertainty’’ state in the majorization sense, as discussed in Sec. III A.

We note in passing here that the infimum for the above measurement,  $\mathcal{P}_{\text{inf}}^{\sigma_x \oplus \sigma_y}$ , is trivially realized on an unpolarized state (i.e., for  $\mathbf{p} = 0$ ) with all four components equal 1/4. As such, it represents the state of maximum uncertainty for the measurement. If  $\hat{\rho}$  is restricted to pure states, on the other hand, the supremum is still given by Eq. (15) but the infimum is found to be  $(1/2, 1/2, 0, 0)$ , which is realized when  $\hat{\rho}$  is an eigenstate of one or the other of the two observables.

##### B. Three spin- $\frac{1}{2}$ components

The extension of the above analysis to the case of three mutually unbiased observables, e.g., all three components of  $\hat{\boldsymbol{\sigma}}$  in the foregoing example, is analogous but requires the full machinery of the calculation of the supremum. Here the measurement set is  $\sigma_x \oplus \sigma_y \oplus \sigma_z$ , with the six measurement elements given by  $\hat{E}_{1,2}^{\sigma_x} = (1 \pm \hat{\sigma}_x)/2$ ,  $\hat{E}_{1,2}^{\sigma_y} = (1 \pm \hat{\sigma}_y)/2$ , and  $\hat{E}_{1,2}^{\sigma_z} = (1 \pm \hat{\sigma}_z)/2$ . The state is parametrized as above, so  $\hat{\rho} = (1 + \hat{\sigma} \cdot \mathbf{p})/2$ . A straightforward calculation of the eight components of  $\mathcal{P}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}(\hat{\rho})$  now gives  $(1 \pm p_x)(1 \pm p_y)(1 \pm p_z)/8$ , which extends the result given in Eq. (13) to three observables.

The procedure for finding  $\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}$  is the same as above, i.e., maximizing a single element of  $\mathcal{P}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}(\hat{\rho})$ , then the sum of a pair of components, and so on, followed by the ‘‘flattening’’ process described in Sec. II B to obtain a descending sequence. While doable analytically, this calculation is more conveniently done numerically [32]. Using either method, one finds for the first two components of  $\mu^{\text{sup}}$  the values  $\mu_1^{\text{sup}} = (1 + 1/\sqrt{3})^3/8$  and  $\mu_2^{\text{sup}} = (1 + 1/\sqrt{2})^2/4$ . The next two components emerge in an ascending order and must, therefore, be flattened, i.e., replaced by their mean. It turns out that these four components add up to unity, thus implying that the next four components vanish. The desired supremum is then found from  $\mu^{\text{sup}}$  and is given by

$$\mathcal{P}_{\text{sup}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z} = (0.491, 0.238, 0.136, 0.136, 0, 0, 0, 0). \quad (16)$$

Thus, any measurement of the three components of a spin- $\frac{1}{2}$  system will yield results more uncertain than the supremum given in Eq. (16), while the latter itself cannot be reached in any physically realizable measurement.

As in the case of two spin components,  $\mathcal{P}_{\text{inf}}^{\sigma_x \oplus \sigma_y \oplus \sigma_z}$  is trivially realized on an unpolarized state with all eight components equal. With  $\hat{\rho}$  restricted to pure states, the supremum is unchanged while the infimum is found to be  $(0.250, 0.250, 0.250, 0.104, 0.062, 0.040, 0.034, 0.011)$ .

Needless to say, the results given in Eqs. (15) and (16) conform to the requirements of Theorem 1A.

### V. CANONICALLY CONJUGATE OBSERVABLES

Here we consider position and momentum, the archetypal example of the uncertainty principle for canonically conjugate observables. Our objective is to calculate the leading component of  $\mathcal{P}_{\text{sup}}^{x\oplus p}$  for a projective measurement of position and momentum in one dimension since the knowledge of this component is sufficient to determine whether  $\mathcal{P}_{\text{sup}}^{x\oplus p} \prec \mathcal{I}$  as required by Theorem 1B. As expected, we will find that this condition is fulfilled in this case as well.

A projective measurement of position in one dimension,  $\hat{x}$ , entails a set of measurement bins corresponding to intervals of the  $x$  axis (where the detectors are positioned) [6]. Let  $[x_{1,\alpha}, x_{2,\alpha}]$  be the  $\alpha$ th bin, with  $\hat{\Pi}_\alpha^x$  the corresponding Hilbert space projection operator, and, similarly,  $\hat{\Pi}_\beta^p$  and  $[p_{1,\beta}, p_{2,\beta}]$  for the momentum ( $\hat{p}$ ) measurement [33]. Note that these projection operators are the measurement elements for this measurement set, i.e.,  $\hat{E}_\alpha^x = \hat{\Pi}_\alpha^x$  and  $\hat{E}_\beta^p = \hat{\Pi}_\beta^p$ . The explicit representation of these projection operators in coordinate space are as follows:

$$\begin{aligned} \langle x | \hat{\Pi}_\alpha^x | x' \rangle &= \delta(x - x') \Theta(x_{1,\alpha} - x) \Theta(x - x_{2,\alpha}), \\ \langle x | \hat{\Pi}_\beta^p | x' \rangle &= \frac{1}{2\pi} \exp[i\bar{p}_\beta(x - x')] \frac{\sin[\Delta p_\beta(x - x')/2]}{(x - x')/2}, \end{aligned} \quad (17)$$

where  $\bar{p}_\beta = (p_{2,\beta} + p_{1,\beta})/2$  and  $\Delta p_\beta = p_{2,\beta} - p_{1,\beta}$ . We have also set  $\hbar = 1$  in Eq. (17) to simplify the writing.

Our objective is to find the maximum value of  $\mathcal{P}_{\alpha\beta}^{x\oplus p}$ , which, by definition, equals  $\mathcal{P}_\alpha^x(\hat{\rho})\mathcal{P}_\beta^p(\hat{\rho}) = \text{tr}(\hat{\Pi}_\alpha^x \hat{\rho}) \text{tr}(\hat{\Pi}_\beta^p \hat{\rho})$ , as  $\hat{\rho}$  is varied. Since  $\hat{\rho}$  is a convex mixture of pure states, the desired maximum will be realized on pure states for which  $\hat{\rho} \rightarrow |\psi\rangle\langle\psi|$ . We are thus looking to maximize  $\langle\psi|\hat{\Pi}_\alpha^x|\psi\rangle\langle\psi|\hat{\Pi}_\beta^p|\psi\rangle/\langle\psi|\psi\rangle^2$  by varying  $|\psi\rangle$  or, equivalently,  $\langle\psi|$ .

A variation with respect to  $\langle\psi|$  gives the eigenvalue equation

$$\hat{\mathcal{L}}_+ |\psi^*\rangle = |\psi^*\rangle, \quad (18)$$

where

$$\hat{\mathcal{L}}_\pm = (\hat{\Pi}_\alpha^x / \mathcal{P}_\alpha^{x*} \pm \hat{\Pi}_\beta^p / \mathcal{P}_\beta^{p*}) / 2 \quad (19)$$

and where we have used a star to signify an optimized quantity so  $\mathcal{P}_\alpha^{x*} = \langle\psi^*|\hat{\Pi}_\alpha^x|\psi^*\rangle/\langle\psi^*|\psi^*\rangle$  and  $\mathcal{P}_\beta^{p*} = \langle\psi^*|\hat{\Pi}_\beta^p|\psi^*\rangle/\langle\psi^*|\psi^*\rangle$ . The two operators  $\hat{\mathcal{L}}_\pm$  defined above are clearly bounded and self-adjoint, with  $\hat{\mathcal{L}}_+$  positive as well. Furthermore,  $\hat{\mathcal{L}}_-$  has a vanishing expectation value in the optimized state  $|\psi^*\rangle$  by definition.

Using Eq. (18) and the fact that  $\langle\psi^*|\hat{\mathcal{L}}_-|\psi^*\rangle = 0$ , we find that  $\langle\psi^*|\hat{\mathcal{L}}_+\hat{\mathcal{L}}_- + \hat{\mathcal{L}}_-\hat{\mathcal{L}}_+|\psi^*\rangle = 0$ . If we then apply the definitions of  $\mathcal{P}_\alpha^{x*}$  and  $\mathcal{P}_\beta^{p*}$  to this equation, we arrive at the equality  $\mathcal{P}_\alpha^{x*} = \mathcal{P}_\beta^{p*}$ . This equality is a consequence of the symmetry with respect to the  $\hat{x} \rightleftharpoons \hat{p}$  exchange in the above optimization problem.

At this point we appeal to Eq. (7) of Sec. III B, which implies that  $\mathcal{P}_\alpha^x(\hat{\rho})\mathcal{P}_\beta^p(\hat{\rho}) \leq \frac{1}{4} \|\hat{\Pi}_\alpha^x + \hat{\Pi}_\beta^p\|^2$ . This inequality, in view of

$\mathcal{P}_\alpha^{x*} = \mathcal{P}_\beta^{p*}$ , in turn, implies that  $\mathcal{P}_\alpha^{x*} = \mathcal{P}_\beta^{p*} \leq \frac{1}{2} \|\hat{\Pi}_\alpha^x + \hat{\Pi}_\beta^p\|$ . Comparing this to Eq. (18), we conclude that

$$\mathcal{P}_\alpha^{x*} = \mathcal{P}_\beta^{p*} = \frac{1}{2} \|\hat{\Pi}_\alpha^x + \hat{\Pi}_\beta^p\|. \quad (20)$$

Thus, the eigenvalue of unity in Eq. (18) is in fact the maximum for the operator  $\hat{\mathcal{L}}_+$ . Our task then is to find  $\|\hat{\Pi}_\alpha^x + \hat{\Pi}_\beta^p\|$ .

To that end, we left-multiply Eq. (18) by  $\hat{\mathcal{L}}_+$ , use Eq. (18) again together with the idempotent property of the projection operators to eliminate all operators except  $\hat{\Pi}_\beta^p \hat{\Pi}_\alpha^x$ , and rewrite the resulting equation in terms of  $|\phi^*\rangle = \hat{\Pi}_\alpha^x |\psi^*\rangle$ . The result is the transformed equation

$$\hat{\Pi}_\alpha^x \hat{\Pi}_\beta^p \hat{\Pi}_\alpha^x |\phi^*\rangle = (2\mathcal{P}_\alpha^{x*} - 1)^2 |\phi^*\rangle. \quad (21)$$

Thus,  $(2\mathcal{P}_\alpha^{x*} - 1)^2$  equals the largest eigenvalue of the positive operator  $\hat{\Pi}_\alpha^x \hat{\Pi}_\beta^p \hat{\Pi}_\alpha^x$ , which we denote by  $\mu_{\text{max}}^2$ . Consequently, the desired maximum,  $\mathcal{P}_{\alpha\beta}^{x\oplus p} = (\mathcal{P}_\alpha^{x*})^2$ , equals  $\frac{1}{4}(1 + \mu_{\text{max}})^2$ . At this point, we observe that  $\|\hat{\Pi}_\alpha^x \hat{\Pi}_\beta^p \hat{\Pi}_\alpha^x\| \leq 1$  since it involves a product of projection operators. This, in turn, implies that  $\mu_{\text{max}} \leq 1$  and, consequently,  $\mathcal{P}_{\alpha\beta}^{x\oplus p} \leq 1$  as well. Furthermore,  $\mathcal{P}_{\alpha\beta}^{x\oplus p} = 1$  is excluded since it would imply  $\|\hat{\Pi}_\alpha^x + \hat{\Pi}_\beta^p\| = 2$  by Eq. (19) and an approximate common eigenstate for  $\|\hat{\Pi}_\alpha^x\|$  and  $\|\hat{\Pi}_\beta^p\|$  which is not possible for finite  $x$  and  $p$  bins.

To simplify the notation, we can without a loss of generality translate the  $x$  and  $p$  axes such that the two bins  $\alpha$  and  $\beta$  for which the optimal values above are reached are symmetrically centered at  $x = 0$  and  $p = 0$  with  $x_{2,\alpha} = -x_{1,\alpha} = \Delta x/2$  and  $p_{2,\beta} = -p_{1,\beta} = \Delta p/2$ . Using the representations of the projection operators given in Eq. (17), we can write Eq. (21) as the following integral equation:

$$\frac{1}{2\pi} \int_{-\Delta x/2}^{+\Delta x/2} dx' \frac{\sin[\Delta p_\beta(x - x')/2]}{(x - x')/2} \phi^*(x') = \mu_{\text{max}}^2 \phi^*(x), \quad (22)$$

where  $\phi^*(x) = \langle x | \phi^*\rangle$  is the wave function corresponding to the state  $|\phi^*\rangle = \hat{\Pi}_\alpha^x |\psi^*\rangle$ , with  $x$  and  $x'$  restricted to the interval  $[-\Delta x/2, +\Delta x/2]$ . It is convenient to rescale Eq. (22) by measuring  $x$  and  $x'$  in units of  $\Delta x$ . The resulting equation can then be written as

$$\frac{1}{\pi} \int_{-1/2}^{+1/2} d\xi' \frac{\sin[s\pi(\xi - \xi')]}{(\xi - \xi')} f(\xi') = \mu^2 f(\xi), \quad (23)$$

where  $s = (\Delta x)(\Delta p)/2\pi\hbar$  and  $\xi$  and  $\xi'$  are restricted to the interval  $[-1/2, +1/2]$ . Note that  $\hbar$  has been restored to the expression for  $s$  here.

The largest eigenvalue of the integral equation (23) is thus equal to  $\mu_{\text{max}}^2$ . Its kernel, on the other hand, is a positive operator bounded by unity according to Eq. (19) *et seq.* In addition, the square of this operator has a finite trace, which implies that the kernel belongs to the Hilbert-Schmidt class of operators and is, therefore, compact as well [28]. This confirms that the spectrum of Eq. (23) is discrete and of finite multiplicity (except possibly for zero), confined to the interval from zero to 1, and can cluster only around zero. Furthermore, the sum of the eigenvalues of the operator, which may be found by calculating its trace, equals  $s$ .

The spectrum of Eq. (23) can be intuitively captured by considering  $(\Delta x \Delta p)$  as the ‘‘volume of phase space’’ and of

$s = (\Delta x)(\Delta p)/2\pi\hbar$  as the “number of states” as well as the number of (nonzero) eigenvalues. For a large phase-space volume, there are  $s$  eigenstates with nonzero eigenvalues nearly equal to unity so their sum should be of the order of  $s$ , which agrees with the trace of the kernel found above. Indeed, a perturbative treatment of Eq. (23) for large values of  $s$  confirms this interpretation [34]. Indeed for  $s \rightarrow \infty$ , we find, using the stationary phase approximation, that the kernel of Eq. (23) effectively approaches  $\delta(\xi - \xi')$  and  $\mu^2 = \mu_{\max}^2 \rightarrow 1$ . This is the limit of low-precision measurements with position or momentum bins approaching the entire span of the  $x$  or  $p$  space and  $\mathcal{P}_{\alpha\beta}^{x\oplus p^*} \rightarrow 1$ . Therefore, the measurement results approach (but do not reach) the limit of zero uncertainty for large  $s$ , albeit for measurements that yield correspondingly little information. The physically interesting limit is, of course, the opposite extreme of small bins and high resolution where the effects of the uncertainty principle are most strongly manifested. We will next consider that limit.

For small bins and high-resolution measurements, we expect no more than one small nonzero eigenvalue for Eq. (23). To verify this expectation, we observe that as  $s \rightarrow 0$ ,  $\sin[s\pi(\xi - \xi')]/(\xi - \xi') \rightarrow s\pi$ ,  $u(\xi) \rightarrow \theta(1 - 4\xi^2)$  (arbitrary normalization), and  $\mu^2 = \mu_{\max}^2 \rightarrow s$ , where  $\theta$  is the usual step function. Thus the largest eigenvalue approaches the sum of all eigenvalues, confirming that there is just one nonzero eigenvalue in this limit. Using its value, we find  $(\mathcal{P}_{\alpha}^{x^*}) \rightarrow \frac{1}{2}(1 + \sqrt{s})$  as  $s \rightarrow 0$  and, therefore,

$$\mathcal{P}_{\alpha\beta}^{x\oplus p^*} \xrightarrow{s \rightarrow 0} \frac{1}{4}(1 + 2\sqrt{s}). \quad (24)$$

The corresponding wave function,  $\psi^*(x) = \langle x|\psi^* \rangle$ , can be constructed by reference to the two projections  $\phi^*(x) = \langle x|\hat{\Pi}_{\alpha}^x|\psi^* \rangle$  and  $\chi^*(x) = \langle x|\hat{\Pi}_{\beta}^p|\psi^* \rangle$ . The first of these was defined above and is directly related to  $u(\xi) \rightarrow \theta(1 - 4\xi^2)$  as found in the foregoing paragraph. The second,  $\chi^*(x)$ , can be constructed from the first using the  $\hat{x} \rightleftharpoons \hat{p}$  symmetry mentioned above. Omitting the details of these steps, we can state the result as follows:

$$\psi^*(x) \xrightarrow{s \rightarrow 0} \frac{1}{\sqrt{2\Delta x}}\theta(\Delta x^2 - 4x^2) + \sqrt{\frac{\Delta p}{4\pi\hbar}} \frac{\sin(x\Delta p/2\hbar)}{(x\Delta p/2\hbar)}. \quad (25)$$

In view of the  $\hat{x} \rightleftharpoons \hat{p}$  symmetry mentioned above, it is instructive to consider the momentum-space representation of this equation,  $\tilde{\psi}^*(p) = \langle p|\psi^* \rangle$ , which can be found by Fourier transformation:

$$\tilde{\psi}^*(p) \xrightarrow{s \rightarrow 0} \frac{1}{\sqrt{2\Delta p}}\theta(\Delta p^2 - 4p^2) + \sqrt{\frac{\Delta x}{4\pi\hbar}} \frac{\sin(p\Delta x/2\hbar)}{(p\Delta x/2\hbar)}. \quad (26)$$

The expected symmetry is clearly in evidence between Eqs. (25) and (26). It is important to remember here that  $\Delta x$  and  $\Delta p$  are bin sizes and in effect represent the resolution of the measuring devices. They should not be confused with variances.

The above results, i.e.,  $(\mathcal{P}_{\alpha}^{x^*}) = (\mathcal{P}_{\beta}^{p^*}) \rightarrow \frac{1}{2}(1 + \sqrt{s})$  as  $s \rightarrow 0$ , show that the state with the sharpest simultaneous values of position and momentum will turn up the two values only 25% of the time. Note also that as a characterization

of the limitations on what is knowable in a measurement of position and momentum, this is a more poignant statement of the uncertainty principle than the Heisenberg inequality or the optimal entropic bound [11]. This is so even though we have only used the leading component of  $\mathcal{P}^{x\oplus p^*}$  for its derivation. As pointed out earlier,  $|\psi^* \rangle$  is not a “minimum uncertainty” state in the usual sense as it maximizes only the leading component of the probability vector. The calculation of the remaining terms of  $\mathcal{P}^{x\oplus p^*}$  in the high-precision limit, which will impose limitations on the readouts of more than one position or momentum bin at a time, is an interesting unsolved problem.

## VI. LEAST-UNCERTAIN MEASUREMENT OF A STATE

Given a state  $\hat{\rho}$ , different measurements performed on it will give rise to probability vectors of varying uncertainty. Is there a measurement  $X^*$  which results in a minimally uncertain probability vector such that  $\mathcal{P}^X(\hat{\rho}) < \mathcal{P}^{X^*}(\hat{\rho})$  for any measurement  $X$ ? This question is, of course, meaningful only if restricted to measurements of comparable precision and interesting only if applied to high-precision measurements such as those that are rank 1, since low-precision measurements can yield probability vectors of arbitrarily low uncertainty. We will, therefore, look for  $X^*$  among rank-1 measurements. To that end, we will first develop a sharpened version of Eq. (29) of Ref. [35].

*Lemma.* Suppose a rank-1 measurement  $X$  is performed on a state  $\hat{\rho}$  resulting in states  $\{\hat{\rho}'_{\alpha}\}$  with probabilities  $\{\mathcal{P}_{\alpha}^X\}$ . We then have (i)  $\oplus_{\alpha} \mathcal{P}_{\alpha}^X \lambda(\hat{\rho}'_{\alpha}) < \lambda(\hat{\rho})$ , where

$$\sup_X [\oplus_{\alpha} \mathcal{P}_{\alpha}^X \lambda(\hat{\rho}'_{\alpha})] = \lambda^{\downarrow}(\hat{\rho}), \quad (27)$$

and (ii) there exists a rank-1 projective measurement  $X^*$  that realizes the above supremum.

Above,  $\lambda^{\downarrow}(\cdot)$  is the spectrum of a density matrix in a descending order and “ $\oplus$ ” denotes a direct sum of spectra as defined in Ref. [35]. The majorization condition in part (i) of this lemma was proved in Ref. [35]; cf. Eq. (29) therein.

For the equality in part (i), as well as in part (ii), we need to demonstrate only the existence of a rank-1 measurement that matches the right-hand side of Eq. (27), thus realizing the supremum. One can readily verify that the projective measurement defined by  $\hat{\Pi}_{\alpha}^* = |\alpha\rangle\langle\alpha|$ , where  $\{|\alpha\rangle\langle\alpha|\}$  are the eigenstates of  $\hat{\rho}$  arranged according to descending eigenvalues, is, in fact, the desired measurement  $X^*$ . This is because with  $X$  restricted to rank-1 measurements, generalized or projective, every  $\hat{\rho}'_{\alpha}$  is pure so  $\lambda^{\downarrow}(\hat{\rho}'_{\alpha})$  equals  $(1, 0, \dots, 0)$  and  $\oplus_{\alpha} \mathcal{P}_{\alpha}^X \lambda(\hat{\rho}'_{\alpha}) = \mathcal{P}^X(\hat{\rho})$ . For  $X^*$ , on the other hand,  $\mathcal{P}_{\alpha}^{X^*}(\hat{\rho}) = \text{tr}(|\alpha\rangle\langle\alpha|\hat{\rho}) = \lambda_{\alpha}^{\downarrow}(\hat{\rho})$ . Putting the last two equalities together, we arrive at

$$\oplus_{\alpha} \mathcal{P}_{\alpha}^{X^*} \lambda(\hat{\rho}'_{\alpha}) = \lambda^{\downarrow}(\hat{\rho}), \quad (28)$$

showing that measurement  $X^*$  realizes the supremum in Eq. (27). This proves the lemma, which allows us to state the following theorem.

*Theorem 2.* The probability vector resulting from a rank-1 generalized measurement of a state is majorized by the spectrum of that state,  $\mathcal{P}^X(\hat{\rho}) < \lambda(\hat{\rho})$ . Furthermore, for any quasientropic measure  $F$ ,  $F[\lambda(\hat{\rho})] \leq F[\mathcal{P}^X(\hat{\rho})]$ . In particular,



the von Neumann entropy of a state is the infimum of the Shannon entropy of all rank-1 measurements of that state, i.e.,  $S(\hat{\rho}) \leq H[\mathcal{P}^X(\hat{\rho})]$ . In addition, there is a projective measurement  $X^*$  which satisfies  $\mathcal{P}^{X^*}(\hat{\rho}) = \lambda(\hat{\rho})$  and saturates the stated inequalities.

The second conclusion of Theorem 2 follows from the fact that  $\lambda^1 \prec \lambda^2$  implies  $F(\lambda^1) \geq F(\lambda^2)$  if  $F$  is a quasientropic measure (i.e., if it is a symmetric, concave function). The choice of the Shannon entropy function  $H$  for  $F$  then yields the next part of Theorem 2. This part was also established in the first paper of Ref. [13] for projective measurements.

Theorem 2 provides a fundamental characterization of the von Neumann entropy as the greatest lower bound of all rank-1 measurement (Shannon) entropies. The majorization statement of the theorem, on the other hand, provides an operational meaning for the spectrum of a quantum state (which is the set of eigenvalues of its density operator) as the least-uncertain probability vector resulting from rank-1 measurements of that state. Not surprisingly, the corresponding least-uncertain measurement  $X^*$ , which saturates all majorization and quasientropic relations above, reproduces the premeasurement state on average, i.e., in the absence of postselection.

Note that in the trivial case of a pure state,  $\hat{\rho} = |\psi\rangle\langle\psi|$ , we find  $\lambda(|\psi\rangle) = (1, 0, \dots, 0)$  and  $\mathcal{P}^{X^*}(|\psi\rangle) = \mathcal{I}$ , corresponding to zero uncertainty. In this case, the optimal measurement  $X^*$  is given by the orthogonal projections  $\hat{\Pi}_1^* = |\psi\rangle\langle\psi|$  and  $\hat{\Pi}_i^* = |i\rangle\langle i|$ ,  $i = 2, 3, \dots$ , where the states  $\{|i\rangle\}$  together with  $|\psi\rangle$  form a complete orthonormal set. In case of a pure state, then, the least-uncertain rank-1 measurement reproduces the premeasurement state with certainty. This amounts to a full identification of the state of the quantum system, which embodies all available information about the system and zero uncertainty. In the case of a mixed state, on the other hand, the optimal measurement  $X^*$  turns up the uncertainties that result from the impurity of the quantum state. In either case,  $X^*$  embodies all available information about the system, as would be expected of a least-uncertain measurement of highest resolution.

## VII. CONCLUDING REMARKS

It is worth recalling here that majorization is a more refined comparator of uncertainty than those based on a scalar condition and that the corresponding formulation of the uncertainty principle is more stringent than the quasientropic or variance formulations. We have explained the reasons that set the majorization order apart from the others and have demonstrated the consequences in several instances. The price of this generality is a mathematically more complex scheme, as one should expect in going from a scalar to a vector formulation. Generally speaking, one may expect the majorization formulation to be most suitable when dealing with overarching information-theoretical aspects of quantum systems, such as those considered in Corollary 1, Secs. V, and VI. As such, it complements the entropic and variance formulations at a fundamental level.

In developing majorization bounds, we resorted to the concept of the supremum of a set of probability vectors. It is important to distinguish the supremum from a *maximal element* for such a set, which would be defined as an element that is not majorized by any other element in the set. A set of probability vectors may include many maximal elements but no supremum, i.e., while the supremum does exist as a probability vector, it may not be a member of the set. This is the reason why there is in general no “minimum uncertainty” state within the majorization formulation.

In this paper we have presented a few basic applications of the majorization formulation of uncertainty, with many others remaining to be worked out. A challenging case is the calculation of the nonleading components of  $\mathcal{P}_{\alpha\beta}^{x\oplus p^*}$  for position and momentum measurements. Another interesting application currently under development is the use of majorization uncertainty bounds for entanglement detection.

## ACKNOWLEDGMENTS

I would like to thank Y. Huang for useful comments. This work was in part supported by a grant from California State University, Sacramento.

- 
- [1] W. Heisenberg, *Zeit. Phys.* **43**, 172 (1927).  
 [2] E. H. Kennard, *Zeit. Phys.* **44**, 326 (1927).  
 [3] H. P. Robertson, *Phys. Rev.* **34**, 163 (1929).  
 [4] E. Schrödinger, *Sitzungsber. Preuss. Akad. Wiss. Berlin, Math. Phys. Kl.* **19**, 296 (1930).  
 [5] D. Deutsch, *Phys. Rev. Lett.* **50**, 631 (1983).  
 [6] M. H. Partovi, *Phys. Rev. Lett.* **50**, 1883 (1983).  
 [7] I. Bialynicki-Birula, *Phys. Lett. A* **103**, 253 (1984).  
 [8] K. Kraus, *Phys. Rev. D* **35**, 3070 (1987); H. Maassen and J. B. M. Uffink, *Phys. Rev. Lett.* **60**, 1103 (1988).  
 [9] F. E. Schroeck, *J. Math. Phys.* **30**, 2078 (1989); M. J. W. Hall, *Phys. Rev. A* **55**, 100 (1997); S. Massar, *ibid.* **76**, 042114 (2007); A. E. Rastegin, e-print arXiv:0810.0083v2 (2009).  
 [10] A. Rajagopal, *Phys. Lett. A* **205**, 32 (1995).  
 [11] I. Bialynicki-Birula, *Phys. Rev. A* **74**, 052101 (2006).  
 [12] See S. Wehner and A. Winter, *New J. Phys.* **12**, 025009 (2010), for a survey and references to the literature.  
 [13] R. Blankenbecler and M. H. Partovi, *Phys. Rev. Lett.* **54**, 373 (1985); M. H. Partovi and R. Blankenbecler, *ibid.* **57**, 2887 (1986).  
 [14] The use of uncertainty relations for detecting entanglement is reviewed in O. Gühne and M. Lewenstein, *Phys. Rev. A* **70**, 022316 (2004).  
 [15] The early contributions of H. Everett, I. I. Hirschmann Jr., W. Beckner, I. Bialynicki-Birula, and J. Mycielski are reviewed in M. D. Srinivas, *Pramana J. Phys.* **60**, 1137 (2003).  
 [16] A. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, 1979), Chap. 1 and pp. 64 and 108.  
 [17] M. A. Nielsen, *Phys. Rev. Lett.* **83**, 436 (1999); M. A. Nielsen and J. Kempe, *ibid.* **86**, 5184 (2001); M. A. Nielsen and G. Vidal, *Quantum Inf. Comp.* **1**, 76 (2001); M. H. Partovi, *Phys. Rev. Lett.* **103**, 230502 (2009).

- [18] M. A. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [19] In this paper we will use a caret to distinguish operators and assume all density operators such as  $\hat{\rho}$  to have unit trace.
- [20] The rank of an operator is the dimension of its range, so  $\hat{\rho}$  is rank 1 if it has the structure  $c|\phi\rangle\langle\psi|$ , where  $c$  is a complex number. Clearly,  $\hat{\rho}\hat{\rho}^\dagger$  will then equal a multiple of the pure state  $|\phi\rangle\langle\phi|$  for any  $\hat{\rho}$ , pure or mixed.
- [21] Properly speaking, these are *orthogonal* projection operators, although we will continue to use the abbreviated nomenclature in referring to them.
- [22] M. H. Partovi, *Phys. Rev. Lett.* **103**, 230502 (2009).
- [23] The infimum or supremum of a (finite or infinite) set of vectors always exists but need not be a member of the set [24].
- [24] Here we have stated the definition of the infimum for a set of  $N$  vectors. In case the number of vectors is infinite, the function “min” must be replaced with “glb” (greatest lower bound). Similarly, “max” must be replaced with “lub” (least upper bound) in the definition of the supremum. In such cases, we would be dealing with the greatest lower, or the least upper, bound of a bounded set of real numbers which always exists.
- [25] Each term of the form  $\sum_{i=1}^j \lambda_i^a$ ,  $1 \leq a \leq N$ , is a partial sum of a descending sequence of positive numbers. As the minimum of such terms,  $\mu_j^{\text{inf}}$  inherits that property, whence the descending character of  $\lambda_i^{\text{inf}}$ . By contrast,  $\mu_j^{\text{sup}}$  is the *maximum* of such terms and does not necessarily inherit the descending property in question.
- [26] A MATHEMATICA program for computing the infimum and supremum of a set of probability vectors using the routines outlined in the text is posted at [<http://www.csus.edu/indiv/p/partovimh/inf-sup.nb>]. As a check on this program, we have compared the infimum and supremum as given by this program to those found by a direct search over a random sample of 100 000 5-dimensional (and a sample of 10 000 for 25-dimensional) probability vectors. The results are in agreement within the expected numerical error, which decreases with sample size.
- [27] Such vectors are referred to as *maximal* elements in the mathematics literature.
- [28] For a definition and properties of compact operators, see M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis* (Academic Press, 1972), Chap. VI and p. 203.
- [29] Specifically, the spectrum of a noncompact Hermitian operator such as a projection operator onto an interval of position or momentum space (treated in Sec. V) may include a continuous component which would preclude the existence of mathematically well-defined eigenstates or eigenvalues; see Ref. [28]. A familiar example of this mathematical difficulty is the position operator which does not have proper eigenvalues or eigenstates, even if the system is confined to a finite region of space.
- [30] The largest eigenvalue of a compact operator equals its norm and has finite multiplicity, exactly as in the case of a finite-dimensional Hermitian matrix. Therefore, an eigenstate always exists that has the norm as its eigenvalue.
- [31] Quasientropic functions belong to a larger class known as *Schur-concave* functions. A Schur-concave function  $G$  is defined by the property that  $\lambda^1 < \lambda^2$  implies  $G(\lambda^1) \geq G(\lambda^2)$ . The Schur-convex class of functions is defined analogously; see Ref. [16].
- [32] Numerical calculations were done by using the code given in Ref. [26] and searching over the magnitude and direction of the polarization vector  $\mathbf{p}$ .
- [33] Note that, our notation notwithstanding, at least one measurement bin must be infinite in size, e.g.,  $|x| \geq L$ , to account for position outcomes outside the apparatus boundaries, where  $L$  is of the order of the apparatus size.
- [34] The properties of the kernel of the integral equation as stated in the foregoing paragraph justify the perturbation calculation.
- [35] M. A. Nielsen, *Phys. Rev. A* **63**, 022114 (2001).