Universal Uncertainty Relations

Shmuel Friedland,[1,*] Vlad Gheorghiu,[2,3,†] and Gilad Gour[2,‡]

[1]*Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, 851 S. Morgan Street, Chicago, Illinois 60607-7045, USA*

[2]*Institute for Quantum Science and Technology and Department of Mathematics and Statistics, University of Calgary, 2500 University Drive NW, Calgary, Alberta T2N 1N4, Canada*

[3]*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

(Received 10 May 2013; revised manuscript received 16 September 2013; published 3 December 2013)

Uncertainty relations are a distinctive characteristic of quantum theory that impose intrinsic limitations on the precision with which physical properties can be simultaneously determined. The modern work on uncertainty relations employs *entropic measures* to quantify the lack of knowledge associated with measuring noncommuting observables. However, there is no fundamental reason for using entropies as quantifiers; any functional relation that characterizes the uncertainty of the measurement outcomes defines an uncertainty relation. Starting from a very reasonable assumption of invariance under mere relabeling of the measurement outcomes, we show that Schur-concave functions are the most general uncertainty quantifiers. We then discover a fine-grained uncertainty relation that is given in terms of the majorization order between two probability vectors, significantly extending a majorization-based uncertainty relation first introduced in M. H. Partovi, Phys. Rev. A **84**, 052117 (2011). Such a vector-type uncertainty relation generates an infinite family of distinct scalar uncertainty relations via the application of arbitrary uncertainty quantifiers. Our relation is therefore universal and captures the essence of uncertainty in quantum theory.

DOI: 10.1103/PhysRevLett.111.230401 PACS numbers: 03.65.Ta, 03.67.Hk

Uncertainty relations lie at the core of quantum mechanics and are a direct manifestation of the noncommutative structure of the theory. In contrast to classical physics, where in principle any observable can be measured with arbitrary precision, quantum mechanics introduces severe restrictions on the allowed measurement results of two or more noncommuting observables. Uncertainty relations are not a manifestation of the experimentalists' (in)ability of performing precise measurements, but are inherently determined by the incompatibility of the measured observables.

The first formulation of the uncertainty principle was provided by Heisenberg [1], who noted that more knowledge about the position of a *single* quantum particle implies less certainty about its momentum and vice versa. He expressed the principle in terms of standard deviations of the momentum and position operators

$$\Delta X \cdot \Delta P \geq \frac{\hbar}{2}. \tag{1}$$

Robertson [2] generalized Heisenberg's uncertainty principle to any two arbitrary observables A and B as

$$\Delta A \cdot \Delta B \geq \frac{1}{2}|\langle\psi|[A, B]|\psi\rangle|. \tag{2}$$

A major drawback of Robertson's uncertainty principle is that it depends on the state $|\psi\rangle$ of the system. In particular, when $|\psi\rangle$ belongs to the null-space of the commutator $[A, B]$, the right upper bound becomes trivially zero. Deutsch [3] addressed this problem by providing an *entropic* uncertainty relation in terms of the Shannon entropies of any two nondegenerate observables, later improved by Maassen and Uffink [4] to

$$H(A) + H(B) \geq -2\log c(A, B). \tag{3}$$

Here, $H(A)$ is the Shannon entropy [5] of the probability distribution induced by measuring the state $|\psi\rangle$ of the system in the eigenbasis $\{|a_j\rangle\}$ of the oservable A (and similarly for B). The bound on the right-hand side $c(A, B) := \max_{m,n}|\langle a_m|b_n\rangle|$ represents the maximum overlap between the bases elements, and is independent of the state $|\psi\rangle$.

Recently, the study of uncertainty relations intensified [6,7] (see also [8,9] for recent surveys), and as a result various important applications have been discovered, ranging from security proofs for quantum cryptography [10–12], information locking [12], nonlocality [13], and the separability problem [14]. There were also recent attempts to generalize uncertainty relations to more than two observables. For this case relatively little is known [15–19], as the authors investigated only particular instances of the problem such as mutually unbiased bases.

In most of the recent work on uncertainty relations, entropy functions like the Shannon and Renyi entropies are used to quantify uncertainty. However, in the context of the uncertainty principle, these entropies are by no reason the most adequate to use. Indeed, as we show here, other

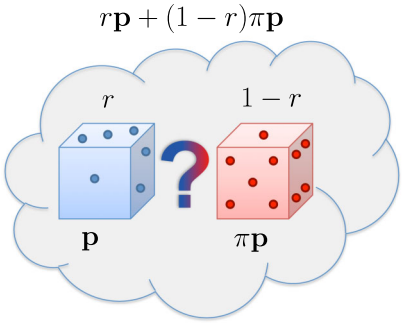

FIG. 1 (color online). With probability r Alice samples from a random variable (blue dice), and with probability $1-r$, Alice samples from its relabeling (red dice), but at the end of the protocol she "forgets" where she sampled from. The resulting probability distribution $r\boldsymbol{p} + (1-r)\pi\boldsymbol{p}$ is more uncertain than the initial one associated with the blue (red) dice $\boldsymbol{p}$ ($\pi\boldsymbol{p}$).

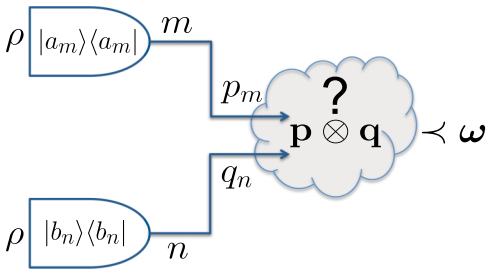

FIG. 2 (color online). A quantum state is measured using two orthonormal bases. We collect the induced joint probability distribution in a vector $\boldsymbol{p} \otimes \boldsymbol{q}$ and quantify its uncertainty in terms of a majorization relation, independently of the state ρ.

functions can be more suitable in providing a quantitative description for the uncertainty principle.

Our approach is based on using majorization [20] to quantify uncertainty. The idea of using majorization to study uncertainty relations was first introduced in [21], and here we build on these ideas and provide explicit closed formulas.

Uncertainty is related to the "spread" of a probability distribution, or, equivalently, to the ability of learning that probability distribution. Intuitively a less spread distribution is more certain than a more widely spread. For example, in a d-dimensional sample space, the probability distribution $\boldsymbol{p} = (1, 0, \ldots, 0)$ is the most certain, whereas the distribution $\boldsymbol{q} = (1/d, 1/d, \ldots, 1/d)$ is the most uncertain. What are then the minimum requirements that a good measure of uncertainty has to satisfy?

In his seminal paper [3] on *entropic* uncertainty relations, Deutsch pointed out that the standard deviation Δ can be increased by mere relabelling of the random variables associated with the measurements. He therefore concluded that the relation in (1) can not be used as a quantitative description of the uncertainty principle.

Following Deutsch observation, we assume here that the uncertainty about a random variable cannot decrease under a relabelling of its alphabet; i.e., the uncertainty associated with a probability vector $\boldsymbol{p}$ cannot be larger than the uncertainty associated with a relabelled version of it, $\pi\boldsymbol{p}$, where π is some permutation matrix. In fact, both uncertainties are the same as permutations acting on a probability space are reversible. Next, we make the reasonable assumption that uncertainty cannot decrease by forgetting information (discarding); see Fig. 1. We call this very reasonable presumption monotonicity under random relabeling. This will be our *only* requirement for a measure of uncertainty. We therefore conclude that any reasonable measure of uncertainty is a function only of the probability vector, is invariant under permutations of its elements, and must be nondecreasing under a random relabeling of its argument.

We formulate the above requirements quantitatively using Birkhoff's theorem [22,23], which states that the convex hull of permutation matrices is the class of doubly stochastic matrices (their components are non-negative real numbers, and each row and column sums to 1). The Birkhoff theorem thus implies that a probability vector $\boldsymbol{q}$ obtained from $\boldsymbol{p}$ by a random relabeling is more uncertain than the latter if and only if the two are related by a doubly stochastic matrix, $\boldsymbol{q} = D\boldsymbol{p}$, which is equivalent to $\boldsymbol{q} \prec \boldsymbol{p}$. The last equation is known as a majorization relation [20] and consists of a system of d inequalities. (A vector $\boldsymbol{x} \in \mathbb{R}^d$ *is majorized by* a vector $\boldsymbol{y} \in \mathbb{R}^d$, and write $\boldsymbol{x} \prec \boldsymbol{y}$, whenever $\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}$ for all $1 \leq k \leq d-1$, with $\sum_{j=1}^{d} x_j^{\downarrow} = \sum_{j=1}^{d} y_j^{\downarrow}$. The down-arrow notation denotes that the components of the corresponding vector are ordered in decreasing order, $x_1^{\downarrow} \geq x_2^{\downarrow} \geq \cdots \geq x_d^{\downarrow}$.) The above discussion implies that any measure of uncertainty has to preserve the partial order induced by majorization. The class of functions that preserve this order are the Schur-concave functions. These are functions Φ on a d-dimensional probability space, Φ: $\mathbb{R}^d \rightarrow \mathbb{R}$, for which $\Phi(\boldsymbol{x}) \geq \Phi(\boldsymbol{y})$, whenever $\boldsymbol{x} \prec \boldsymbol{y}$, $\forall\, \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$. We therefore define a measure of uncertainty as being any non-negative Schur-concave function that takes the value zero on the vector $\boldsymbol{x} = (1, 0, \ldots, 0)$. The last requirement is not essential but is convenient as it ensures that the measure is non-negative.

Our definition for a measure of uncertainty is very general and resulted solely from requiring monotonicity under random relabeling; it also encompasses the most common entropy functions used in information theory, but it is not restricted to them. As we are not concerned with asymptotic regimes, we use in the following the most general Φ to quantify uncertainty, without making any assumptions about its functional form.

Having defined what a measure of uncertainty is, we now use it to study uncertainty relations. Let ρ be a mixed state on a d-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^d$. For simplicity of the exposition, we first consider two basis

(projective) measurements. We denote the two orthonormal bases of $\mathcal{H}$ by $\{|a_m\rangle\}_{m=1}^d$ and $\{|b_n\rangle\}_{n=1}^d$. We also denote by $p_m(\rho) = \langle a_m|\rho|a_m\rangle$ and $q_n(\rho) = \langle b_n|\rho|b_n\rangle$ the two probability distributions obtained by measuring ρ with respect to these bases. We collect the numbers $p_m(\rho)$ and $q_n(\rho)$ into two probability vectors $\boldsymbol{p}(\rho)$ and $\boldsymbol{q}(\rho)$, respectively. The goal of our work is to bound the uncertainty about $\boldsymbol{p}(\rho)$ and $\boldsymbol{q}(\rho)$ by a quantity that depends only on the bases elements but not on the state ρ. The object of our investigation is therefore the joint probability distribution $\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho)$.

The main result of our article is an uncertainty relation of the form, as schematically depicted in Fig. 2

$$\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho) \prec \boldsymbol{\omega}, \quad \forall\ \rho, \tag{4}$$

where $\boldsymbol{\omega}$ is some vector independent of ρ that we explicitly calculate. A majorization uncertainty relation of a similar form was first introduced by Partovi in [21]; however, his right-hand side of the majorization relation is not explicit but written in terms of supremum over all density matrices, which makes it difficult to calculate. We call (4) a universal uncertainty relation (UUR) as, for any measure of uncertainty Φ,

$$\Phi[\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho)] \geq \Phi(\boldsymbol{\omega}), \quad \forall\ \rho. \tag{5}$$

The UUR (4) generates in fact an infinite family of uncertainty relations of the form (5), one for each Φ. The right hand side of (5) provides a single-number lower bound on the uncertainty of the joint measurement results. Whenever Φ is additive under tensor products (e.g., Renyi entropies, minus the logarithm of the G concurrence [24] or minus the logarithm of the minimum nonzero component of the probability distribution), (5) splits as

$$\Phi[\boldsymbol{p}(\rho)] + \Phi[\boldsymbol{q}(\rho)] \geq \Phi(\boldsymbol{\omega}). \tag{6}$$

We now construct the d^2-dimensional vector $\boldsymbol{\omega}$ appearing on the right-hand side of our UUR (4). Let $\mathcal{I}_k \subset [d] \times [d]$ be a subset of k distinct pair of indices (m, n), where $[d]$ is the set of natural numbers ranging from 1 to d. Let

$$\Omega_k := \max_{\mathcal{I}_k} \max_{\rho} \sum_{(m,n)\in\mathcal{I}_k} p_m(\rho) q_n(\rho), \tag{7}$$

where the outer maximum is over all subsets $\mathcal{I}_k$ with cardinality k and the inner maximum is taken over all density matrices. Then the vector $\boldsymbol{\omega}$ in the UUR (4) is given by

$$\boldsymbol{\omega} = (\Omega_1, \Omega_2 - \Omega_1, \ldots, \Omega_d - \Omega_{d-1}, 0, \ldots, 0). \tag{8}$$

Moreover, we show in the Appendix that $\Omega_k = 1$, for all $d \leq k \leq d^2$.

The quantities Ω_k in (7) can be in general difficult to calculate explicitly, as they involve an optimization problem. However, in the Appendix we show that the first two elements can be computed explicitly as

$$\Omega_1 = \frac{1}{4}[1 + c]^2, \qquad \Omega_2 = \frac{1}{4}[1 + c']^2, \tag{9}$$

where $c := \max_{m,n}|\langle a_m|b_n\rangle|$, and

$$c' := \max\sqrt{|\langle a_m|b_n\rangle|^2 + |\langle a_{m'}|b_{n'}\rangle|^2},$$

where the maximum is taken over all indices $m = m'$ and $n \neq n'$, and over all indexes $n = n'$ and $m \neq m'$.

For $k > 2$, we upper bound each Ω_k in (7) by

$$\tilde{\Omega}_k := \max_{\substack{\mathcal{R},\mathcal{S} \\ |\mathcal{R}|+|\mathcal{S}|=k+1}} \max_{\rho} \left(\sum_{m\in\mathcal{R}} p_m(\rho)\right)\left(\sum_{n\in\mathcal{S}} q_n(\rho)\right) \tag{10}$$

$$= \frac{1}{4} \max_{\substack{\mathcal{R},\mathcal{S} \\ |\mathcal{R}|+|\mathcal{S}|=k+1}} \left\| \sum_{m\in\mathcal{R}} |a_m\rangle\langle a_m| + \sum_{n\in\mathcal{S}} |b_n\rangle\langle b_n| \right\|_\infty^2 \leq 1, \tag{11}$$

where $\mathcal{R}$ ($\mathcal{S}$) are subsets of distinct indices from $[d]$, $|\mathcal{R}|$ ($|\mathcal{S}|$) denotes the size (number of elements) of $\mathcal{R}$ ($|\mathcal{S}|$), and $\|\cdot\|_\infty$ denotes the infinity operator norm—which, for positive operators (as it is in our case), coincides with the maximum eigenvalue of its argument. Moreover, $\tilde{\Omega}_d = 1$. Note that $\Omega_k = \tilde{\Omega}_k$ for $k = 1, 2$, and otherwise $\Omega_k \leq \tilde{\Omega}_k$ since for a fixed ρ, all the terms in the sum of (7) are strictly contained in the expression (10). The equality in (11) is nontrivial and follows from the main technical Theorem of this article (See Theorem 1 in the Appendix): $\max_\rho \mathrm{Tr}(\rho A)\,\mathrm{Tr}(\rho B) = (1/4)\|A + B\|_\infty^2$, for two projections A and B.

Similar to the definition of the vector $\boldsymbol{\omega}$ in (8), we construct the vector $\tilde{\boldsymbol{\omega}}$ as in (8) by replacing Ω_k with $\tilde{\Omega}_k$. A simple calculation (see Appendix) shows that

$$\boldsymbol{p}(\rho) \otimes \boldsymbol{q}(\rho) \prec \tilde{\boldsymbol{\omega}}, \quad \forall\ \rho. \tag{12}$$

Therefore $\tilde{\boldsymbol{\omega}}$ provides a (weaker) lower-bound for the UUR (4), but which is now explicitly computable.

To appreciate the generality of our UUR (4), we compare in Fig. 3 the best known lower bounds for the uncertainty of the measurement in two bases with our induced uncertainty relation (5), in which we take Φ to be the Shannon entropy H. We consider the region in which $c > 0.83$, for which the best known bound [25] has an explicit analytical form. We note that our bound overperforms [25] in a large number of instances (around 90% of the time). For $c < 0.83$, our bound tend to be slightly worse than [25], but this is expected since our uncertainty relation is valid for *all* measures of uncertainty and is not optimized for a specific one such as Shannon's. Next we take $\Phi = H_\infty$ in (5) and note that we recover Maassen's and Uffink bound [4] for the minimum entropy, which is tight. Finally, choosing $\Phi = H_\alpha$ (Renyi-α entropy) in (5) provides yet a novel entropic uncertainty relation valid for *all* values of the parameter α.

We now extend our results to the most general case of $L \geq 2$ positive operator valued measures (POVMs). Denote by $\{\Pi_{\alpha_\ell}^{(\ell)}\}_{\alpha_\ell=1}^{N_\ell}$ the ℓth POVM, with $1 \leq \ell \leq L$.

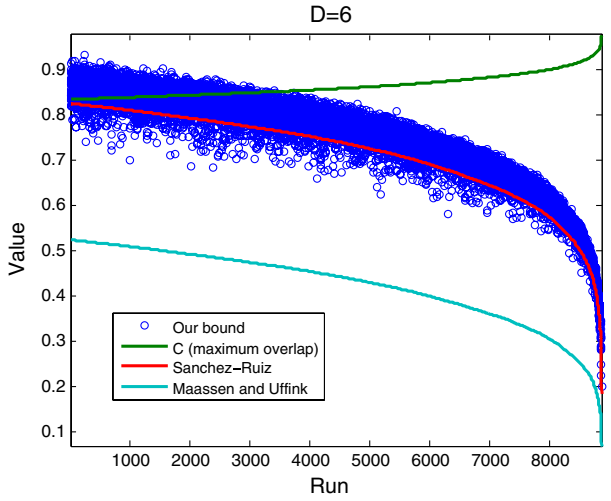

FIG. 3 (color online). For each run we randomly generate two orthonormal bases and a random state $|\psi\rangle$ in a 6-dimensional Hilbert space, then compute our lower bound $H(\boldsymbol{\omega})$ (right hand side of (5), using the Shannon entropy H as a measure of uncertainty); $c = \max_{m,n}|\langle a_m|b_n\rangle|$ denotes the maximum overlap between the two bases.

The quantity N_ℓ denotes the number of elements in the lth POVM, and the index α_ℓ labels its elements, with $\alpha_\ell = 1, 2, \ldots, N_\ell$. A measurement of ρ with the ℓth POVM $\Pi^{(\ell)}$ induces a probability distribution vector $\boldsymbol{p}^{(\ell)}(\rho) = (p_1^{(\ell)}(\rho), p_2^{(\ell)}(\rho), \ldots, p_{N_\ell}^{(\ell)}(\rho))$. We discover a UUR of the form

$$\bigotimes_{\ell=1}^{L} \boldsymbol{p}^{(\ell)}(\rho) \prec \boldsymbol{\omega}, \quad \forall\ \rho, \tag{13}$$

where the quantity on the left-hand side represents the joint probability distribution induced by measuring ρ with each POVM $\Pi^{(\ell)}$. Here,

$$\boldsymbol{\omega} = (\Omega_1, \Omega_2 - \Omega_1, \Omega_3 - \Omega_2, \ldots, \Omega_N - \Omega_{N-1}), \tag{14}$$

where $N \equiv N_1 N_2 \ldots N_L$, and for $k = 1, 2, \ldots, N$

$$\Omega_k := \max_{\mathcal{I}_k} \max_{\rho} \sum_{(\alpha_1,\ldots,\alpha_L)\in\mathcal{I}_k} p_{\alpha_1}^{(1)}(\rho) p_{\alpha_2}^{(2)}(\rho) \cdots p_{\alpha_L}^{(L)}(\rho), \tag{15}$$

where $\mathcal{I}_k \subset [N_1] \times [N_2] \times \cdots \times [N_L]$ is a subset of k distinct string of indices $(\alpha_1, \ldots, \alpha_L)$ (here $[N_j]$ is the set of natural numbers ranging from 1 to N_j).

Since the above quantities Ω_k can in general be difficult to calculate explicitly, we have found tight upper bounds $\tilde{\Omega}_k$ that do not involve an optimization over all states ρ. Our upper bound $\Omega_k \le \tilde{\Omega}_k$ is given by

$$\tilde{\Omega}_k := \max_{\substack{\mathcal{S}_1,\ldots,\mathcal{S}_L \\ \sum_{\ell=1}^{L} |\mathcal{S}_\ell| = L+k-1}} \left\| \frac{1}{L} \sum_{\ell=1}^{L} \left(\sum_{\alpha_\ell \in \mathcal{S}_\ell} \Pi_{\alpha_j}^{(\ell)} \right) \right\|_\infty^L \le 1, \tag{16}$$

where $\mathcal{S}_\ell$ denotes a subset of distinct indices from $[N_\ell]$, $|\mathcal{S}_\ell|$ denotes the size (number of elements) of $\mathcal{S}_\ell$. Note that by definition $\Omega_N = 1$. We define the vector $\tilde{\boldsymbol{\omega}}$ as in (14) by replacing Ω_k with $\tilde{\Omega}_k$, and then show in the Appendix that the UUR in Eq. (13) holds with $\boldsymbol{\omega}$ replaced by $\tilde{\boldsymbol{\omega}}$.

Note that an $L > 2$ measurement uncertainty relation can be trivially generated by summing pairwise two-measurement uncertainty relations, one for each pair of observables. Our UUR (13) is more powerful and is not of this form. This fact can be seen most clearly in a set of measurement operators in which any two observables share a common eigenvector. In this case, a two-measurement uncertainty relation will provide a trivial lower bound of zero; hence, the pairwise sum must also be zero. However, the vector $\boldsymbol{\omega}$ in (14) is in general different from $(1, 0, \ldots, 0)$, (see example 1 in Sec. B of the Supplemental Material [26]), thus providing a nontrivial bound for the UUR (13) (or the induced family obtained by applying various uncertainty measures Φ on it). Finally, the UUR (13) is not restricted to MUBs or particular values of L, but is valid for any number of arbitrary bases.

To summarize, we derived two explicit closed form uncertainty relations, Eq. (4), which is valid for measurements in two orthonormal bases, and (13), which is the generalization of the former to the most general setting of L POVMs. Our relations are "fine-grained"; they do not depend on a single number (such as the maximum overlap between base elements), but on *all* components of the vector $\tilde{\boldsymbol{\omega}}$, which we compute explicitly, via a majorization relation. Our uncertainty relations are *universal* and capture the essence of uncertainty in quantum mechanics, as they are not quantified by particular measures of uncertainty such as Shannon or Renyi entropies.

We did not explore here which bases provide the most uncertain measurement results for the UURs. One may conjecture that MUBs are the suitable candidates. Indeed, this seems to be the case, and we conjecture that Ω_k in (7) is given by $\Omega_k = 1/4(1 + \sqrt{k/d})^2$, which can then be used to construct [see (8)] the vector $\boldsymbol{\omega}^{\mathrm{MUB}}$ for the UUR (4). The conjecture is strongly supported by numerical simulations. Moreover, we observed that for bases that are not MUBs, the best $\boldsymbol{\omega}$ we were able to find (numerically) always majorizes $\boldsymbol{\omega}^{\mathrm{MUB}}$, i.e., $\boldsymbol{\omega}^{\mathrm{MUB}} \prec \boldsymbol{\omega}$. This provides strong support for the initial assumption that MUBs provide the most uncertain measurement outcomes.

Another important direction of investigation is the extension of the results presented here to uncertainty relations with quantum memory [27]. These are particularly useful in the context of quantum cryptography. However, such an extension is nontrivial and is left for future work.

The authors thank Robert Spekkens and Marco Piani for useful comments and discussions. V. Gheorghiu and G. Gour acknowledge support from the Natural Sciences and Engineering Research Council (NSERC) of Canada and from the Pacific Institute for the Mathematical Sciences (PIMS). Shmuel Friedland was partially supported by NSF Grant No. DMS-1216393.

*friedlan@uic.edu
†vgheorgh@gmail.com
‡gour@ucalgary.ca

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927).
[2] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[3] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[4] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 2005), 2nd ed.
[6] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011).
[7] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, Phys. Rev. Lett. **108**, 210405 (2012).
[8] S. Wehner and A. Winter, New J. Phys. **12**, 025009 (2010).
[9] I. Bialynicki-Birula and L. Rudnicki, in *Statistical Complexity*, edited by K. Sen (Springer, Netherlands, 2011), p. 1.
[10] I. B. Damgaard, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, in *Advances in Cryptology—CRYPTO '07, Vol. 4622 of Lecture Notes in Computer Science* (Springer, New York, 2007), p. 360.
[11] R. Konig, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theory **58**, 1962 (2012).
[12] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008).
[13] J. Oppenheim and S. Wehner, Science **330**, 1072 (2010).
[14] O. Gühne, Phys. Rev. Lett. **92**, 117903 (2004).
[15] I. D. Ivanovic, J. Phys. A **25**, L363 (1992).
[16] J. Sanchez-Ruiz, Phys. Lett. A **173**, 233 (1993).
[17] M. A. Ballester and S. Wehner, Phys. Rev. A **75**, 022319 (2007).
[18] S. Wu, S. Yu, and K. Mølmer, Phys. Rev. A **79**, 022104 (2009).
[19] S. Wehner and A. Winter, J. Math. Phys. (N.Y.) **49**, 062105 (2008).
[20] M. Albert W., O. Ingram, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications (2nd Edition)*, Springer Series in Statistics (Springer, New York, 2011).
[21] M. H. Partovi, Phys. Rev. A **84**, 052117 (2011).
[22] G. Birkhoff, Univ. Nac. Tucumán. Rev. Ser. A **5**, 147 (1946).
[23] R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).
[24] G. Gour, Phys. Rev. A **71**, 012318 (2005).
[25] J. I. de Vicente and J. Sánchez-Ruiz, Phys. Rev. A **77**, 042110 (2008).
[26] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.230401 for detailed mathematical proofs.
[27] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nat. Phys. **6**, 659 (2010).