Entropic uncertainty from effective anticommutators

Jędrzej Kaniewski,* Marco Tomamichel, and Stephanie Wehner
Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543
(Received 3 March 2014; published 22 July 2014)

We investigate entropic uncertainty relations for two or more binary measurements, for example, spin-$\frac{1}{2}$ or polarization measurements. We argue that the effective anticommutators of these measurements, i.e., the anticommutators evaluated on the state prior to measuring, are an expedient measure of measurement incompatibility. Based on the knowledge of pairwise effective anticommutators we derive a class of entropic uncertainty relations in terms of conditional Rényi entropies. Our uncertainty relations are formulated in terms of effective measures of incompatibility, which can be certified in a device-independent fashion. Consequently, we discuss potential applications of our findings to device-independent quantum cryptography. Moreover, to investigate the tightness of our analysis we consider the simplest (and very well studied) scenario of two measurements on a qubit. We find that our results outperform the celebrated bound due to Maassen and Uffink [Phys. Rev. Lett. **60**, 1103 (1988)] and provide an analytical expression for the minimum uncertainty which also outperforms some recent bounds based on majorization.

DOI: 10.1103/PhysRevA.90.012332 PACS number(s): 03.67.Mn, 03.67.Dd

I. INTRODUCTION

Uncertainty relations tell us that quantum mechanics is inherently nondeterministic, i.e., that there exist experiments whose outcomes cannot be predicted with arbitrary precision. In the usual scenario we consider two distinct measurements, giving rise to random variables X and Y, respectively. The statement is of the form "if the two measurements are *incompatible* then it cannot be the case that both X and Y are close to being deterministic" and the statement must hold *regardless of the state of the system prior to measuring*. In other words, X or Y (or both) must be at least somewhat unpredictable, that is, random. To make this statement rigorous we need three ingredients: a measure of incompatibility, a measure of uncertainty, and a nontrivial relation between the two.

The study of uncertainty relations began when Heisenberg [1] and (more formally) Kennard [2] noticed that it is impossible to prepare a particle whose position and momentum are sharply peaked: The more localized a particle is, the more variable its momentum becomes, and vice versa. More generally, Robertson [3] showed that uncertainty might arise whenever two observables do not commute. Let ρ be the state of the system prior to the measurement. For an operator A, denote the expectation value of that operator by $\langle A\rangle = \mathrm{tr}(A\rho)$. For operators A and B, let $[A,B] = AB - BA$ be the *commutator* of A and B and let $\langle [A,B]\rangle$ be the *effective commutator*. Robertson's relation reads

$$\sigma_A \sigma_B \geqslant \tfrac{1}{2}|\langle [A,B]\rangle|,$$

where σ_X is the standard deviation of X, $\sigma_X^2 = \langle X^2\rangle - \langle X\rangle^2$. Note that this relation applies to both continuous-outcome (e.g., position or momentum) and discrete-outcome (e.g., spin or polarization) measurements.

In 1930 Schrödinger [4] proved a stronger relation:

$$\sigma_A^2 \sigma_B^2 \geqslant \left|\tfrac{1}{2}\langle \{A,B\}\rangle - \langle A\rangle\langle B\rangle\right|^2 + \left|\tfrac{1}{2}\langle [A,B]\rangle\right|^2,$$

where $\{A,B\} = AB + BA$ is the *anticommutator* of A and B and $\langle \{A,B\}\rangle$ is the *effective anticommutator*.

These early uncertainty relations are interesting from the foundational point of view but they suffer from two problems: (a) they are not tight in some important cases (e.g., for spin-$\frac{1}{2}$ particle with $A = \sigma_Z$, $B = \sigma_X$ and $\rho = \frac{\mathbb{1}}{2}$ the right-hand side is 0, despite both outcomes being maximally random, $\sigma_A = \sigma_B = 1$) and (b) their applications are limited because the standard deviation is not always a suitable measure of uncertainty.

To find uncertainty relations with applications in information theory and cryptography, entropies were employed as measures of uncertainty. Usually one considers a scenario in which we have a certain number of measurements and perform one of them uniformly at random. If we store the label of the measurement in K and the measurement outcome in X we obtain a joint probability distribution P_{XK}. Entropic uncertainty relations are simply lower bounds on a particular conditional entropy, $H(X|K)$, evaluated on the probability distribution P_{XK}.

The first entropic uncertainty relation was proved for position and momentum of an infinite-dimensional system in 1975 [5,6]. The first result for arbitrary rank-1 projective measurements performed on a finite-dimensional system was derived by Deutsch [7] and improved by Maassen and Uffink [8]. The latter state that for two projective rank-1 measurements on a d-dimensional system, described by measurement eigenvectors $\{|x_j\rangle\}_{j\in[d]}$ and $\{|y_j\rangle\}_{j\in[d]}$, we have

$$H(X|K) \geqslant -\tfrac{1}{2}\log_2 c,$$

where $H(X|K)$ is the conditional Shannon entropy and $c := \max_{j,k} |\langle x_j | y_k\rangle|^2$ is the *overlap* of the two measurements (note that this is independent of the state ρ prior to measurement). Entropic uncertainty relations became an active topic of research (see, for example, [9–12]) since entropies give operational meaning to the notion of uncertainty and thus find applications in many information processing and cryptographic tasks (see Ref. [13] for a recent review).

*j.kaniewski@nus.edu.sg

The authors of Ref. [14] considered a set of binary observables that pairwise anticommute (as operators) and they found that such measurements give rise to strong entropic uncertainty relations. While the case of perfect anticommutation is well understood, nothing is known about the case of partial (or approximate) anticommutation. This is a significant drawback since for most applications we need uncertainty relations which are robust against small perturbations. If we were interested in noneffective measures of anticommutation (e.g., the norm of the anticommutator) uncertainty statements would follow directly from standard, overlap-dependent uncertainty relations. However, for effective measures (e.g., effective anticommutator) the connection between the overlap and anticommutation is no longer valid (as explained in Appendix B) so new methods must be developed.

II. RESULTS AND OUTLINE

In this paper we prove uncertainty relations for an arbitrary set of binary observables (as usual we associate their outcomes with values ± 1). Given the knowledge of their pairwise effective anticommutators [cf. Eq. (1)] we derive lower bounds on conditional Rényi entropies [cf. Eqs. (8) and (9)] in two steps. In the first step we show that fixing the effective anticommutators imposes a simple geometric constraint on the expectation values of these observables (note that a probability distribution with two outcomes is fully characterized by its expectation value). In the second step we show that the constraint on expectation values implies a lower bound on entropic uncertainty.

Our relations have two desirable features. First, our measure of incompatibility is effective (state-dependent) and it can be certified experimentally based on ideas of Mayers and Yao [15,16], which leads to *device-independent uncertainty*. (Note that noneffective measures, like the overlap commonly used in entropic uncertainty relations, cannot be certified and so we can employ these relations only when the device is trusted.) Second, we can treat any (finite) number of observables. This is because we do not rely on a standard technique based on a reduction to qubits (Jordan's lemma), which only works for two observables, but instead use the full anticommutation structure of the set of observables.

We compare our results with existing bounds for the case of the Shannon entropy of two measurements. In particular, we improve on the celebrated Maassen-Uffink bound by providing an analytical bound that is strictly stronger for all nontrivial overlaps. We conclude the paper with a discussion of potential applications to device-independent quantum cryptography.

III. TECHNIQUES

A binary measurement consists of two positive semidefinite operators, $F_+, F_- \geqslant 0$, that add up to identity $F_+ + F_- = \mathbb{1}$. If we associate the outcomes with values ± 1 then the measurement can be written compactly as a binary *observable*, $A = F_+ - F_-$, which satisfies $-\mathbb{1} \leqslant A \leqslant \mathbb{1}$.

Suppose we are given a state, ρ, and a set of M binary observables, $\{A_j\}_{j\in[M]}$. Define the effective anticommutator between the jth and the kth observable as

$$\varepsilon_{jk} = \frac{\langle\{A_j, A_k\}\rangle}{2} = \frac{\mathrm{tr}(\{A_j, A_k\}\rho)}{2} \tag{1}$$

and note that ε_{jk} is real and $|\varepsilon_{jk}| \leqslant 1$. Let T be the *anticommutation matrix*, $[T]_{jk} = \varepsilon_{jk}$. For ease of presentation in the main paper we focus on projective observables, for which $[T]_{jj} = 1$ for all j. For a more general proof, which also covers generalized measurements, please refer to Appendix C. Let $g_j = \langle A_j\rangle$ be the expectation value of the jth observable. For binary observables the probability distribution of interest (as described in the introduction) can be written as

$$\Pr[X = x, K = k] = \frac{1}{M}\frac{1}{2}[1 + (-1)^x g_k]. \tag{2}$$

The conditional Rényi entropy [17] of order $\alpha > 1$ is defined as

$$H_\alpha(X|K) := \frac{\alpha}{1-\alpha}\log_2 \sum_k p_k \left(\sum_x p_{x|k}^\alpha\right)^{1/\alpha} \tag{3}$$

while the Shannon entropy equals $H(X|K) := \lim_{\alpha\to 1} H_\alpha(X|K)$. The goal is to prove lower bounds on $H_\alpha(X|K)$ and $H(X|K)$ (this is what we want) evaluated on the joint probability distribution (2) based on the knowledge of T (this is what we are given) and we do it in two steps. First, we show that T imposes a geometric condition on the expectation values of the observables. Then, we use this geometric condition to prove lower bounds on entropic uncertainty.

Let $g = (g_1, g_2, \ldots, g_M)$ be a (column) vector composed of expectation values. Clearly, g lies inside the (± 1)-hypercube, $g \in [-1,1]^M$, but we show that T imposes an extra geometric constraint on g. For this purpose, let a be an arbitrary real unit vector, $a \in [-1,1]^M$, and let $K = \sum_j a_j A_j$. Then

$$K^2 = \mathbb{1} + \frac{1}{2}\sum_{j\neq k} a_j a_k \{A_j, A_k\}.$$

For arbitrary operators the Cauchy-Schwarz inequality ensures that $[\mathrm{tr}(X^\dagger Y)]^2 \leqslant \mathrm{tr}(X^\dagger X)\cdot\mathrm{tr}(Y^\dagger Y)$. By setting $X = K\sqrt{\rho}$ and $Y = \sqrt{\rho}$ we find that

$$a^{\mathrm{T}} g g^{\mathrm{T}} a \leqslant a^{\mathrm{T}} T a.$$

Since this inequality holds for all choices of a, it is equivalent to the operator inequality

$$g g^{\mathrm{T}} \leqslant T. \tag{4}$$

This constraint admits an appealing geometrical interpretation: The matrix T defines an ellipsoid within the hypercube and the constraint restricts the vector g to lie inside that ellipsoid (see Fig. 1 for an example).

Moreover, an extension of the construction from Ref. [18] (Appendix C) shows that this characterization is tight: A vector of expectation values g and an anticommutation matrix T are compatible if and only if (iff) Eq. (4) holds.

To find lower bounds on a particular entropy ($H_\alpha(X|K)$ or $H(X|K)$) we just need to minimize it over the allowed set of expectation values. Note that for the probability distribution (2)

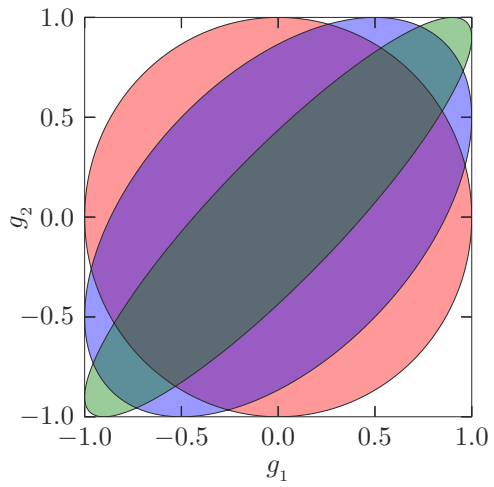

FIG. 1. (Color online) The allowed expectation values of two observables with a fixed effective anticommutator, $\varepsilon \in \{0, 0.5, 0.9\}$. For $\varepsilon = 0$ we get a circle, which becomes gradually elongated towards the corners as ε increases. Note that $\varepsilon > 0$ ($\varepsilon < 0$) forces the two expectation values to be correlated (anticorrelated), which results in an ellipse lying along the primary (secondary) diagonal. The deterministic points, corresponding to the corners, are only allowed for $|\varepsilon| = 1$.

the expression (3) simplifies to

$$H_\alpha(X|K) = \frac{\alpha}{1-\alpha} \log_2 \frac{\sum_k w_\alpha(g_k)}{M},$$

where $w_\alpha(g) = [(\frac{1+g}{2})^\alpha + (\frac{1-g}{2})^\alpha]^{1/\alpha}$. Now, the task is to minimize $H_\alpha(X|K)$ over the ellipsoid, or, equivalently, to solve

$$\max: \sum_k w_\alpha(g_k) \quad \text{subject to} \quad g g^{\mathrm{T}} \leqslant T.$$

Unfortunately, this seemingly natural task turns out to be rather difficult even in the simplest cases. Therefore, we consider a relaxation of the problem, in which we optimize over a sphere whose radius is determined by the largest semiaxis of the ellipsoid, denoted by r (see Fig. 2 for an example). Note that

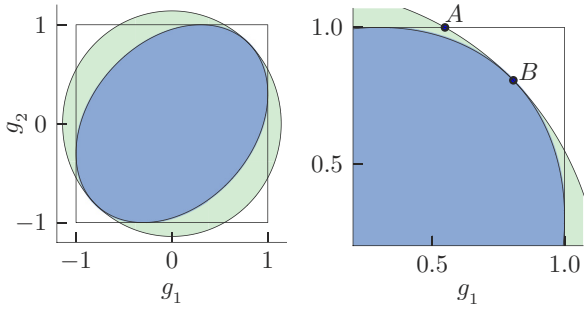

FIG. 2. (Color online) The spherical relaxation for two measurements with $\varepsilon = 0.3$. Optimization is performed over a circle (light color) rather than an ellipse (dark color). Points A and B are the optimal solutions to the relaxed optimisation problem (6) for convex ($\alpha \in (1, \frac{3}{2}]$) and concave ($\alpha \in [2, \infty)$) functions, respectively.

$r = \|T\|$, the spectral norm of T.

$$\max: \sum_k w_\alpha(g_k) \quad \text{subject to} \quad g \in [-1,1]^M, \quad \sum_k g_k^2 \leqslant r. \tag{5}$$

Note that we added the hypercube constraints explicitly since they are not implied by the relaxed, spherical constraint. This approach has the advantage that it compresses the whole anticommutation matrix into just one number — its norm. More important, the relaxed problem can be solved analytically for most values of α as explained below.

Since neither the objective function nor the constraints of the optimization problem (5) depend on the sign of g_k we can restrict ourselves to non-negative expectation values. This allows us to define $t_k = g_k^2$ and the problem becomes

$$\max: \sum_y w_\alpha(\sqrt{t_k}) \quad \text{subject to} \quad t \in [0,1]^M, \ \sum_k t_k \leqslant r. \tag{6}$$

Since the objective function is monotone we can assert that the optimal solution satisfies $\sum_k t_k = r$.

For $\alpha \in (1, \frac{3}{2}]$ the function $w_\alpha(\sqrt{t})$ is convex in t (Appendix D) and since the maximum of a convex function over a convex set is achieved at an extremal point, the optimal value must be achieved at an assignment of the form

$$t_k = \begin{cases} 1 & \text{for} \quad 1 \leqslant k \leqslant \lfloor r \rfloor, \\ r - \lfloor r \rfloor & \text{for} \quad k = \lfloor r \rfloor + 1, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

Hence, for $\alpha \in (1, \frac{3}{2}]$ we arrive at the following bound, which constitutes our main result:

$$H_\alpha(X|K) \geqslant H_\alpha(Y|K), \quad \text{where}$$
$$\Pr[Y = y, K = k] = \frac{1}{M}\frac{1}{2}[1 + (-1)^y \sqrt{t_k}\,] \tag{8}$$

and t_k refers to the optimal assignment [Eq. (7)]. This can be extended to the Shannon entropy by taking the limit of $\alpha \to 1$ yielding $H(X|K) \geqslant H(Y|K)$.

For $\alpha \in [2, \infty)$ the function $w_\alpha(\sqrt{t})$ is concave and since it is also symmetric the minimum is achieved for $t_k = \frac{r}{M}$ for all k. Therefore, we have

$$H_\alpha(X|K) \geqslant H_\alpha(Y), \quad \text{where}$$
$$\Pr[Y = y] = \frac{1}{2}\left(1 + (-1)^y \sqrt{\frac{r}{M}}\right). \tag{9}$$

Note that in both cases these bounds are functions of M and r only and, hence, can be computed easily.

IV. COMPARISON WITH EXISTING BOUNDS

Although effective anticommutators play a central role in our work, it is more common to state uncertainty relations in terms of the overlap. Let us consider two projective rank-1 measurements on a qubit and the conditional Shannon entropy that arises. We look for bounds of the form $H(X|K) \geqslant q(c)$ and, as stated in the introduction, the celebrated result of Maassen and Uffink [8] reads

$$q_{\mathrm{MU}}(c) = -\tfrac{1}{2} \log_2 c.$$

While this is known to be tight for the extreme values of the overlap, $c \in \{\frac{1}{2}, 1\}$, it is not tight in the interior (see,

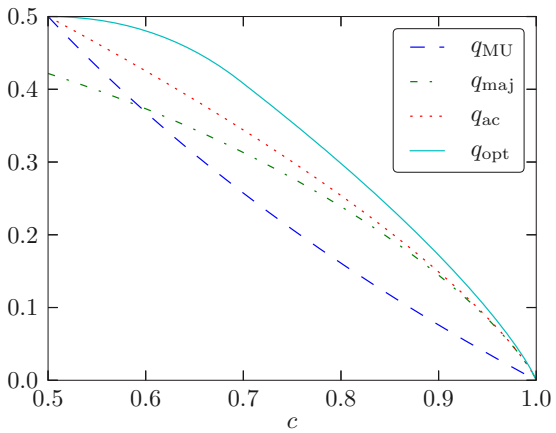

FIG. 3. (Color online) Comparison of various lower bounds on $H(X|K)$ as a function of the overlap c. The quantity $q_{\mathrm{ac}}(c)$ is defined in Eq. (10).

e.g., [19,20]). It turns out that our results might be applied to this case to give an improvement for all intermediate values of c. We take advantage of the fact that for projective measurements on a qubit there is a one-to-one mapping between the effective anticommutator and the overlap, $c = (1 + |\varepsilon|)/2$. Therefore, we can write our bound (8) as a function of the overlap

$$q_{\mathrm{ac}}(c) = \frac{1}{2}h\left(\frac{1+\sqrt{|\varepsilon|}}{2}\right) = \frac{1}{2}h\left(\frac{1+\sqrt{2c-1}}{2}\right), \qquad (10)$$

where $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy. In Fig. 3 we compare these bounds with a bound recently developed using a majorization technique [21,22] (and very recently [23]), denoted $q_{\mathrm{maj}}(c)$, and the largest state-independent lower bound, denoted $q_{\mathrm{opt}}(c)$. (For $c \gtrsim 0.7$ there is an analytic expression for q_{opt} due to Ghirardi *et al.* [19], while for $c \lesssim 0.7$ one needs to resort to numerics.)

V. APPLICATIONS TO QUANTUM CRYPTOGRAPHY

Recently, in the context of quantum cryptography, there has been a lot of interest in self-testing [15,16,24] and device-independent security [25]. In self-testing the task is to characterize the internal working of a device by analyzing observed correlations alone. This characterization then allows proof of security of a cryptographic protocol executed using that device. (The term *device-independent* comes from the fact that we did not assume how the device works but we deduced it from the statistics.)

Uncertainty relations constitute an important ingredient of many device-independent security proofs (see Ref. [26] for an example in quantum key distribution and Ref. [27] for a very recent example in randomness expansion). An interesting development would be to prove device-independent security for two-party cryptography, for example, in the bounded [28,29] or noisy [30,31] storage models. (In the case of trusted devices, security based on uncertainty relations was proved in the bounded storage model [32] and for relativistic bit commitment [33].)

Our results fit into this framework since we derive uncertainty from effective anticommutators, which can be certified experimentally. To certify effective anticommutation between two observables it is enough to observe Clauser-Horne-Shimony-Holt (CHSH) violation (see, e.g., Ref. [34]). To extend this result to multiple observables we resort to a game proposed by Slofstra [35], which can be seen as a combination of multiple CHSH games in which one of the parties is not told which particular subgame they are playing (see Appendix E for details). This testing procedure produces bounds on the effective anticommutator of every pair of observables, which implies an upper bound on the norm of the anticommutation matrix, r. Then, we use Eqs. (8) and (9) to obtain explicit entropic bounds, hence leading us to *device-independent uncertainty*.

VI. CONCLUSION

Drawing from early uncertainty relations we have shown that it is possible to derive entropic uncertainty relations for binary observables from effective anticommutation. The effective anticommutators seem to be natural objects to study and give rise to strong uncertainty relations. Moreover, since they can be certified (self-tested) our uncertainty relations are expected to have applications in device-independent cryptography. Investigating these potential applications is the most interesting open question arising from our research. Another more foundational line of research could investigate whether our approach can be extended to allow for quantum side information.

ACKNOWLEDGMENTS

This work is funded by the Ministry of Education (MOE) and National Research Foundation Singapore, as well as MOE Tier 3 Grant "Random numbers from quantum processes" (MOE2012-T3-1-009). We thank Thomas Vidick and Matthew McKague for useful discussions.

APPENDIX A: PRELIMINARIES

For an integer n, let $[n] = \{1,2,\ldots,n\}$. Let $\mathcal{H}$ denote a finite-dimensional Hilbert space of dimension $d = \dim \mathcal{H}$ and let $\mathcal{H}(\mathcal{H})$ denote the set of Hermitian operators acting on $\mathcal{H}$. Let $\mathcal{S}(\mathcal{H})$ denote the set of quantum states on $\mathcal{H}$: $\rho \in \mathcal{S}(\mathcal{H}) \iff \rho \in \mathcal{H}(\mathcal{H}), \rho \geqslant 0, \operatorname{tr}\rho = 1$. A binary observable, $\Gamma \in \mathcal{H}(\mathcal{H})$, is a Hermitian operator which satisfies $-\mathbb{1}_d \leqslant \Gamma \leqslant \mathbb{1}_d$, where $\mathbb{1}_d$ denotes the identity matrix of dimension d.

Let $\{\Gamma_j\}$ be a set of Hermitian, traceless, anticommuting observables acting on a d-dimensional Hilbert space:

$$\Gamma_j = \Gamma_j^\dagger, \quad \operatorname{tr}\Gamma_j = 0 \text{ and } \{\Gamma_j, \Gamma_k\} = 2\delta_{jk}\mathbb{1}_d.$$

Note that such a set can always be found, regardless of the number of observables required, as long as the dimension is high enough (e.g., by Jordan-Wigner transformation; see Ref. [14] for details). Let us first show that if we build a quantum state out of these operators then a simple Bloch-sphere-type condition holds.

Lemma A.1. Let x be a real vector. The operator

$$\rho = \frac{1}{d}\left(\mathbb{1}_d + \sum_j x_j \Gamma_j\right)$$

corresponds to a valid quantum state iff $\sum_j x_j^2 \leqslant 1$.

Proof. Clearly, ρ is Hermitian and of unit trace, hence, we just need to verify that it is also positive semi-definite. Let $F = \sum_j x_j \Gamma_j$ and note that F^2 is proportional to $\mathbb{1}_d$. Therefore, F can be written as

$$F = \left(\sum_j x_j^2\right)^{1/2} (2P - \mathbb{1}_d),$$

where P is a $d/2$-dimensional projector and $\operatorname{tr} P = d/2$ (note that this implies that d must be even). Clearly, $\rho \geqslant 0$ is equivalent to $\mathbb{1}_d + F \geqslant 0$, which is satisfied iff $\sum_j x_j^2 \leqslant 1$. ∎

APPENDIX B: ANTICOMMUTATION VS THE OVERLAP

In this section we show that the overlap is related to the norm of the anticommutator. We also show that if we make the measures effective (i.e., state dependent) this connection is no longer valid and the two quantities can be vastly different.

For two projective, binary measurements $\{P_0, P_1\}$ and $\{Q_0, Q_1\}$ their overlap is defined as

$$c = \max_{b,b'} \|P_b Q_{b'}\|,$$

where $\|\cdot\|$ denotes the Schatten ∞ norm. By Jordan's lemma the measurement operators can be simultaneously block diagonalized

$$P_b = \bigoplus_k P_b^k \quad \text{and} \quad Q_b = \bigoplus_k Q_b^k$$

and the blocks are of dimension at most 2. If the kth block is trivial (dimension 1) then one of the projectors is rank 1, while the other is zero: e.g., $P_0^k = |0\rangle\langle 0|, P_1^k = 0$ and similarly for Q_0, Q_1. If the block is of dimension 2 then each of the four projectors is rank 1:

$$P_0 = |p_0\rangle\langle p_0|,\ P_1 = |p_1\rangle\langle p_1|,$$
$$Q_0 = |q_0\rangle\langle q_0|,\ Q_1 = |q_1\rangle\langle q_1|,$$

where $\langle p_0|p_1\rangle = \langle q_0|q_1\rangle = 0$ and

$$|p_0\rangle\langle p_0| + |p_1\rangle\langle p_1| = |q_0\rangle\langle q_0| + |q_1\rangle\langle q_1| = R_k$$

and R_k is the projector on the two-dimensional support of the kth block. The corresponding observables $A = P_0 - P_1$, $B = Q_0 - Q_1$ are block diagonalized in the same way to give

$$A = \bigoplus_k A^k \quad \text{and} \quad B = \bigoplus_k B^k$$

and it is easy to verify that for both types of blocks we have

$$\max_{b,b'} \left\|P_b^k Q_{b'}^k\right\| = \tfrac{1}{2} + \tfrac{1}{4}\|\{A^k, B^k\}\|.$$

Taking maximum over the blocks leads to a relationship between the overlap and the norm of the anticommutator:

$$c = \tfrac{1}{2} + \tfrac{1}{4}\|\{A, B\}\|.$$

Without any knowledge of the state no further improvement is possible. However, if we know the state we can refine our measures by making them effective (or state-dependent). This is easy in case of the anticommutator by defining

$$\varepsilon = \tfrac{1}{2}\operatorname{tr}(\{A, B\}\rho).$$

In the case of the overlap the extension is less obvious. Let us consider a simplified version of the quantity proposed in Ref. [34]:

$$c^* = \sum_k \operatorname{tr}(R_k \rho) \max_{b,b'} \left\|P_b^k Q_{b'}^k\right\|,$$

which is a weighted average of the blockwise overlaps. Now, we want to show that these two quantities can be very different. Consider a system of dimension 4, where the measurements are

$$P_0 = |0\rangle\langle 0| + |2\rangle\langle 2|, \quad P_1 = |1\rangle\langle 1| + |3\rangle\langle 3|,$$
$$Q_0 = |0\rangle\langle 0| + |3\rangle\langle 3|, \quad Q_1 = |1\rangle\langle 1| + |2\rangle\langle 2|,$$

and the state is $\rho = \frac{1}{2}(|0\rangle\langle 0| + |2\rangle\langle 2|)$. It is easy to verify that $c^* = 1$, since in each of the two blocks the measurements are actually the same (up to relabelling). In fact, any definition of effective overlap based on the block-diagonal form should give us full overlap. On the other hand, the effective anticommutator equals $\varepsilon = 0$. Hence, the overlap approach shows no incompatibility, while according to the effective anticommutator the observables are maximally incompatible. The reason for these seemingly contradictory conclusions is that the definition of the effective overlap implicitly takes into account the subspace information. The effective anticommutator does not have access to the subspace information and this ignorance results in uncertainty. This also demonstrates why the uncertainty guarantees based on the effective anticommutator might not hold if one conditions on additional classical information (in this case the subspace information).

APPENDIX C: THE ELLIPSOID CONDITION

Suppose we are given a state, ρ, and a set of M binary observables, $\{A_j\}_{j\in[M]}$. Let g be the (column) vector of expectation values, $g_j = \langle A_j\rangle$, and let T be the anticommutation matrix

$$T_{jk} = \begin{cases} \langle A_j^2\rangle & \text{if} \quad j = k, \\ \langle\{A_j, A_k\}\rangle/2 & \text{otherwise.} \end{cases}$$

Lemma C.1. Any valid combination of g and T satisfies $gg^{\mathrm{T}} \leqslant T$.

Proof. Let a be an arbitrary real unit vector, $a \in [-1,1]^M$, and let $K = \sum_j a_j A_j$. Then

$$K^2 = \sum_j a_j^2 A_j^2 + \frac{1}{2}\sum_{j\neq k} a_j a_k \{A_j, A_k\}.$$

Consider Hermitian operators $X = K\sqrt{\rho}$ and $Y = \sqrt{\rho}$. Note that $\operatorname{tr}(X^\dagger Y) = \operatorname{tr}(K\rho) = \sum_j a_j g_j = a^{\mathrm{T}} g$, $\operatorname{tr}(X^\dagger X) = \operatorname{tr}(K^2\rho) = a^{\mathrm{T}} T a$ and $\operatorname{tr}(Y^\dagger Y) = \operatorname{tr}(\rho) = 1$. Therefore, the

Cauchy-Schwarz inequality, $[\mathrm{tr}(X^\dagger Y)]^2 \leqslant \mathrm{tr}(X^\dagger X)\,\mathrm{tr}(Y^\dagger Y)$, implies that

$$a^{\mathrm{T}} g g^{\mathrm{T}} a \leqslant a^{\mathrm{T}} T a.$$

Since this inequality holds for all choices of a, it is equivalent to the operator inequality

$$g g^{\mathrm{T}} \leqslant T. \qquad \blacksquare$$

Moreover, the following lemma shows that this characterization is tight.

Lemma C.2. Let T be a $M \times M$ real, positive semidefinite matrix and let $g \in [-1,1]^M$ be a real vector such that $g g^{\mathrm{T}} \leqslant T$. Then, there exists a quantum state and measurements that give g as the vector of expectation values and T as the anticommutation matrix.

Proof. Since $T \geqslant 0$ there exists a $M \times r$ real matrix R, such that $RR^{\mathrm{T}} = T$ and $r = \mathrm{rk}(T)$. Let the jth observable be

$$A_j = \sum_{i=1}^{r} R_{ji} \Gamma_i,$$

which implies that $\{A_j, A_k\} = 2T_{jk}\mathbb{1}_d$. Therefore, the anticommutation matrix is reproduced correctly independent of the state.

Consider an operator defined as

$$\rho = \frac{1}{d}\left(\mathbb{1}_d + \sum_{j=1}^{r} x_j \Gamma_j\right).$$

It is easy to verify that if ρ corresponds to a valid state then the resulting vector of expectation values equals $g = Rx$. Since $\mathrm{rk}(R) = r$, R has a left inverse, namely a $r \times M$ matrix Q such that $QR = \mathbb{1}_r$, and x can be calculated as $x = Qg$. To verify that ρ corresponds to a valid state we must check that $x^{\mathrm{T}} x \leqslant 1$ which follows directly from the fact that

$$x x^{\mathrm{T}} = Q g g^{\mathrm{T}} Q^{\mathrm{T}} \leqslant Q T Q^{\mathrm{T}} = Q R R^{\mathrm{T}} Q^{\mathrm{T}} = \mathbb{1}_r,$$

where we used the assumption $g g^{\mathrm{T}} \leqslant T$. $\blacksquare$

As a corollary we obtain a lower bound on the dimension of the system necessary to reproduce a particular choice of g and T.

Corollary C.1. To reproduce correctly g and T it is sufficient to use $r = \mathrm{rk}(T)$ anticommuting observables which can be realized in dimension $d = 2^{\lceil \frac{r-1}{2} \rceil}$.

APPENDIX D: CONVEXITY AND CONCAVITY OF $w_\alpha(\sqrt{t})$

For completeness recall the definition of $w_\alpha(x)$ for $x \in [-1,1]$:

$$w_\alpha(x) = \left[\left(\frac{1+x}{2}\right)^\alpha + \left(\frac{1-x}{2}\right)^\alpha\right]^{1/\alpha}.$$

Lemma D.1. The function $w_\alpha(\sqrt{t})$ for $t \in [0,1]$ is convex for $\alpha \in (1, \frac{3}{2}]$ and concave for $\alpha \in [2, \infty)$.

Proof. Let us write $w_\alpha(\sqrt{t})$ as

$$w_\alpha(\sqrt{t}) = \tfrac{1}{2}[g_\alpha(t)]^{1/\alpha},$$

$$\text{where} \quad g_\alpha(t) = (1+\sqrt{t})^\alpha + (1-\sqrt{t})^\alpha.$$

Calculating the derivatives gives

$$\frac{d}{dt} w_\alpha(\sqrt{t}) = \frac{1}{2\alpha} g_\alpha(t)^{(1-\alpha)/\alpha} g'_\alpha(t),$$

$$\begin{aligned}\frac{d^2}{dt^2} w_\alpha(\sqrt{t}) &= \frac{1-\alpha}{2\alpha^2} g_\alpha(t)^{(1-2\alpha)/\alpha} [g'_\alpha(t)]^2 \\ &\quad + \frac{1}{2\alpha} g_\alpha(t)^{(1-\alpha)/\alpha} g''_\alpha(t) \\ &= \frac{g_\alpha(t)^{(1-2\alpha)/\alpha}}{2\alpha^2} [(1-\alpha)[g'_\alpha(t)]^2 + \alpha g_\alpha(t) g''_\alpha(t)].\end{aligned}$$

Therefore, what we are interested in is the sign of

$$h_\alpha(t) = \frac{1-\alpha}{\alpha^2}[g'_\alpha(t)]^2 + \frac{1}{\alpha} g_\alpha(t) g''_\alpha(t). \tag{D1}$$

It is easy to verify that

$$g'_\alpha(t) = \frac{\alpha}{2\sqrt{t}}[(1+\sqrt{t})^{\alpha-1} - (1-\sqrt{t})^{\alpha-1}],$$

$$g''_\alpha(t) = \frac{\alpha(\alpha-1)}{4t} g_{\alpha-2}(t) - \frac{g'_\alpha(t)}{2t}.$$

Expanding the terms gives

$$\frac{1-\alpha}{\alpha^2}[g'_\alpha(t)]^2 = \frac{1-\alpha}{4t}[(1+\sqrt{t})^{2(\alpha-1)} + (1-\sqrt{t})^{2(\alpha-1)} - 2(1-t)^{\alpha-1}],$$

$$\begin{aligned}\frac{1}{\alpha} g_\alpha(t) g''_\alpha(t) &= \frac{\alpha-1}{4t} g_\alpha(t) g_{\alpha-2}(t) - \frac{1}{2\alpha t} g_\alpha(t) g'_\alpha(t) \\ &= \frac{\alpha-1}{4t}[(1+\sqrt{t})^{2(\alpha-1)} + (1-\sqrt{t})^{2(\alpha-1)} + 2(1+t)(1-t)^{\alpha-2}] \\ &\quad - \frac{1}{4t\sqrt{t}}[(1+\sqrt{t})^{2\alpha-1} - (1-\sqrt{t})^{2\alpha-1} - 2\sqrt{t}(1-t)^{\alpha-1}].\end{aligned}$$

Therefore,

$$h_\alpha(t) = \frac{1}{t}\left((1-t)^{\alpha-2}\left[\alpha - \frac{1+t}{2}\right] - \frac{1}{4\sqrt{t}}[(1+\sqrt{t})^{2\alpha-1} - (1-\sqrt{t})^{2\alpha-1}]\right).$$

Since we are only interested in the sign of Eq. (D1), we consider

$$2\alpha - 1 - t - \frac{(1-t)^{3/2}}{2\sqrt{t}}\left[\left(\frac{1+\sqrt{t}}{1-\sqrt{t}}\right)^{\alpha-1/2} - \left(\frac{1-\sqrt{t}}{1+\sqrt{t}}\right)^{\alpha-1/2}\right]. \tag{D2}$$

Here, it is convenient to introduce hyperbolic functions. Let $e^{2x} = (1+\sqrt{t})/(1-\sqrt{t})$, which means that $t \in [0,1]$ is mapped onto $x \in [0,\infty)$. Then, we have

$$x = \operatorname{arctanh}\sqrt{t}, \quad t = \tanh^2 x, \quad \text{and} \quad 1 - t = \frac{1}{\cosh^2 x}$$

and expression (D2) becomes

$$\begin{aligned} &2\alpha - 1 - \tanh^2 x - \frac{\sinh[x(2\alpha-1)]}{\sinh x \cosh^2 x} \\ &\quad = 2(\alpha-1) + \frac{\sinh x - \sinh[x(2\alpha-1)]}{\sinh x \cosh^2 x}. \end{aligned}$$

Note that $2\sinh x \cosh^2 x = \sinh 2x \cosh x = (\sinh 3x + \sinh x)/2$. The sign is the same as the sign of

$$\frac{\alpha-1}{2}\sinh 3x + \frac{1+\alpha}{2}\sinh x - \sinh[x(2\alpha-1)],$$

which we can Taylor expand. Note that this is an odd function and the coefficients are

$$c_k(\alpha) = \frac{1}{2k!}[(\alpha-1)3^k + 1 + \alpha - 2(2\alpha-1)^k].$$

To show convexity (concavity) it suffices to show that all the coefficients are positive (negative). Since $c_k(\alpha)$ is a polynomial and it vanishes at $\alpha = 1$ it must be divisible by $(\alpha - 1)$.

$$\begin{aligned} (2\alpha-1)^k &= \sum_{j=0}^{k}\binom{k}{j}(\alpha-1)^j\alpha^{k-j} \\ &= \alpha^k + (\alpha-1)\sum_{j=0}^{k-1}\binom{k}{j+1}(\alpha-1)^j\alpha^{k-j-1}, \\ 1+\alpha-2\alpha^k &= (1-\alpha) + 2\alpha(1-\alpha^{k-1}) \\ &= (1-\alpha)\left(1 + 2\sum_{j=1}^{k-1}\alpha^j\right). \end{aligned}$$

Putting everything together gives

$$\begin{aligned} c_k(\alpha) &= \frac{\alpha-1}{2k!}p_k(\alpha), \\ \text{where} \quad p_k(\alpha) &= 3^k - 1 - 2\sum_{j=1}^{k-1}\alpha^j \\ &\quad - 2\sum_{j=0}^{k-1}\binom{k}{j+1}(\alpha-1)^j\alpha^{k-j-1}. \end{aligned}$$

Note that for $\alpha \geqslant 1$, $p_k(\alpha)$ is monotonically decreasing in α, so it has at most one zero. Therefore, $c_k(\alpha)$ has at most two zeros (the first one at $\alpha = 1$). By checking

$$\begin{aligned} c_k\left(\frac{3}{2}\right) &= \frac{1}{2k!}\left(\frac{3^k+5}{2} - 2^{k+1}\right) \geqslant 0, \\ c_k(2) &= \frac{1}{2k!}(3 - 3^k) \leqslant 0, \end{aligned}$$

we conclude that the other zero is always there and is contained within $\alpha \in (\frac{3}{2},2)$. Hence for $\alpha \in (1,\frac{3}{2}] \cup [2,\infty)$ all the coefficients have the same sign, which proves convexity or concavity of the original function. ■

APPENDIX E: THE CERTIFICATION PROCEDURE

This certification procedure assumes that both devices are memoryless, i.e., every round is identical and independent of each other.

Suppose we are given a measurement device (Alice) with M different settings, which correspond to different binary observables, $\{A_j\}_{j\in[M]}$. The goal of the certification procedure is to characterize the anticommutation matrix T, or more specifically the effective pairwise commutators

$$\varepsilon_{jk} = \tfrac{1}{2}\langle\{A_j, A_k\}\rangle = \tfrac{1}{2}\operatorname{tr}(\{A_j, A_k\}\rho).$$

Ideally, since we are interested in large uncertainty, we would like our measurements to exactly anticommute, i.e., $\varepsilon_{jk} = 0$ for $j \neq k$.

To perform device-independent certification we need an auxiliary device (Bob), which in our case is a measurement device with $2\binom{M}{2}$ settings denoted by $B_{jk,t}$, where $j,k \in [M], j \neq k$ and $t \in \{0,1\}$ that shares entanglement with the first device. Following the procedure proposed by Slofstra [35] we estimate the following quantity for all pairs (j,k), $j \neq k$:

$$\beta_{jk} := \langle A_j \otimes (B_{jk,0} + B_{jk,1}) + A_k \otimes (B_{jk,0} - B_{jk,1})\rangle.$$

Since this is clearly equivalent to the CHSH game, we can see the entire procedure as a combination of multiple CHSH subgames in which Alice is not told which subgame she is playing. Therefore, we can apply a standard result from Ref. [34], which establishes a trade-off between the observed violation and the effective anticommutator of the observables used by Alice (in fact, the same trade-off applies on Bob's side but since we do not want to certify the auxiliary device we do not need it). More specifically, we have

$$|\varepsilon_{jk}| \leqslant \frac{\beta_{jk}}{4}\sqrt{8 - \beta_{jk}^2} := c_{jk}.$$

While this does not allow us to find the anticommutation matrix explicitly, we can place an upper bound on its norm. It is easy to see that $\|T\| \leqslant \|T'\|$, where

$$T'_{jk} = \begin{cases} 1 & \text{if} \quad j = k, \\ c_{jk} & \text{otherwise.} \end{cases}$$

Therefore, the observed statistics allows us to bound $\|T\|$, which turns out to be sufficient for our applications.

For completeness, we also provide an explicit description of devices that achieve the maximum violation for all subgames. Suppose Alice and Bob share a maximally entangled state of

dimension $d = 2^{\lceil \frac{M-1}{2} \rceil}$

$$|\Psi\rangle_{AB} = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} |k\rangle_A |k\rangle_B$$

and that their measurements are

$$A_j = \Gamma_j \quad \text{and} \quad B_{jk,t} = \frac{\Gamma_j^{\mathrm{T}} + (-1)^t \Gamma_k^{\mathrm{T}}}{\sqrt{2}},$$

where $\{\Gamma_j\}$ is a set of anticommuting observables acting on d-dimensional Hilbert space as defined in Appendix A. It is easy to check that for every pair $(j,k), j \neq k$, we obtain

$$\langle \Psi | A_j \otimes (B_{jk,0} + B_{jk,1}) + A_k \otimes (B_{jk,0} - B_{jk,1}) | \Psi \rangle = 2\sqrt{2},$$

which implies $\varepsilon_{jk} = 0$. Hence, we have certified a device that performs M exactly anticommuting measurements.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927).
[2] E. H. Kennard, Z. Phys. **44**, 326 (1927).
[3] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[4] E. Schrödinger, Sitzungsberichte der Preußischen Akademie der Wissenschaften, Physikalisch-mathematische Klasse **14**, 296 (1930).
[5] W. Beckner, Ann. Math. **102**, 159 (1975).
[6] I. Białynicki-Birula and J. Mycielski, Comm. Math. Phys. **44**, 129 (1975).
[7] D. Deutsch, Phys. Rev. Lett. **50**, 631 (1983).
[8] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[9] J. I. de Vicente and J. Sánchez-Ruiz, Phys. Rev. A **77**, 042110 (2008).
[10] G. M. Bosyk, M. Portesi, A. Plastino, and S. Zozor, Phys. Rev. A **84**, 056101 (2011).
[11] Ł. Rudnicki, S. P. Walborn, and F. Toscano, Phys. Rev. A **85**, 042115 (2012).
[12] S. Zozor, G. M. Bosyk, and M. Portesi, J. Phys. A **46**, 465301 (2013).
[13] S. Wehner and A. Winter, New J. Phys. **12**, 025009 (2010).
[14] S. Wehner and A. Winter, J. Math. Phys. **49**, 062105 (2008).
[15] D. Mayers and A. Yao, in *Proceedings of the 39th Annual Symposium on Foundations of Computer Science, FOCS'98* (IEEE Computer Society, Washington, DC, 1998), p. 503.
[16] D. Mayers and A. Yao, Quant. Inf. Comp. **4**, 273 (2004).
[17] S. Arimoto, in *Topics in Information Theory*, Series Colloq. Math. Soc. J. Bolyai, Vol. 16, edited by I. Csiszar and P. Elias (North Holland, Amsterdam, 1977), pp. 41–52.
[18] B. S. Tsirelson, Lett. Math. Phys. **4**, 93 (1980).
[19] G. Ghirardi, L. Marinatto, and R. Romano, Phys. Lett. A **317**, 32 (2003).
[20] P. J. Coles and M. Piani, Phys. Rev. A **89**, 022112 (2014).
[21] S. Friedland, V. Gheorghiu, and G. Gour, Phys. Rev. Lett. **111**, 230401 (2013).
[22] Z. Puchała, Ł. Rudnicki, and K. Życzkowski, J. Phys. A **46**, 272002 (2013).
[23] Ł. Rudnicki, Z. Puchała, and K. Życzkowski, Phys. Rev. A **89**, 052115 (2014).
[24] M. McKague and M. Mosca, in *Theory of Quantum Computation, Communication and Cryptography*, Lecture Notes in Computer Science, Vol. 6519, edited by W. van Dam, V. M. Kendon, and S. Severini (Springer, Berlin, Heidelberg, 2011), pp. 113–130.
[25] A. Acín, N. Gisin, and L. Masanes, Phys. Rev. Lett. **97**, 120405 (2006).
[26] C. C. W. Lim, C. Portmann, M. Tomamichel, R. Renner, and N. Gisin, Phys. Rev. X **3**, 031006 (2013).
[27] C. A. Miller and Y. Shi, in *Proceedings of the 46th Annual ACM Symposium on Theory of Computing, STOC'14* (ACM, New York, 2014), pp. 417–426.
[28] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner, in *46th Annual IEEE Symposium on Foundations of Computer Science, 2005, FOCS 2005* (IEEE, Piscataway, NJ, 2005), pp. 449–458.
[29] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, in *Advances in Cryptology-CRYPTO 2007*, Lecture Notes in Computer Science, Vol. 4622, edited by A. Menezes (Springer, Berlin, Heidelberg, 2007), pp. 360–378.
[30] S. Wehner, C. Schaffner, and B. M. Terhal, Phys. Rev. Lett. **100**, 220502 (2008).
[31] R. König, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theory **58**, 1962 (2012).
[32] N. H. Y. Ng, M. Berta, and S. Wehner, Phys. Rev. A **86**, 042315 (2012).
[33] J. Kaniewski, M. Tomamichel, E. Hänggi, and S. Wehner, IEEE Trans. Inf. Theory **59**, 4687 (2013).
[34] M. Tomamichel and E. Hänggi, J. Phys. A **46**, 055301 (2013).
[35] W. Slofstra, J. Math. Phys. **52**, 102202 (2011).