Improved entropic uncertainty relations and information exclusion relations

Patrick J. Coles[1] and Marco Piani[2]

[1]*Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, 117543 Singapore*
[2]*Institute for Quantum Computing and Department of Physics and Astronomy, University of Waterloo, N2L3G1 Waterloo, Ontario, Canada*

The uncertainty principle can be expressed in entropic terms, also taking into account the role of entanglement in reducing uncertainty. The information exclusion principle bounds instead the correlations that can exist between the outcomes of incompatible measurements on one physical system, and a second reference system. We provide a more stringent formulation of both the uncertainty principle and the information exclusion principle, with direct applications for, e.g., the security analysis of quantum key distribution, entanglement estimation, and quantum communication. We also highlight a fundamental distinction between the complementarity of observables in terms of uncertainty and in terms of information.

I. INTRODUCTION

A fundamental trait of quantum mechanics is the unavoidable uncertainty associated with measuring incompatible observables, i.e., the so-called uncertainty principle, which dates back to Heisenberg [1]. Kennard [2] formalised Heisenberg's original ideas in an uncertainty relation involving the products of standard deviations of the position y and momentum p_y observables, with the well-known inequality $\Delta y \Delta p_y \geqslant \hbar/2$. Robertson [3] generalised this to arbitrary Hermitian observables X and Z and found the uncertainty relation $\Delta X \Delta Z \geqslant \frac{1}{2}|\langle\psi|[X,Z]|\psi\rangle|$. From a conceptual point of view, though, standard deviation is an inadequate measure of uncertainty, when the latter is understood in terms of (lack of) knowledge of "which outcome" of a measurement, rather than in terms of the value of the outcome. Also, the right-hand side (r.h.s.) of Robertson's relation gives a trivial bound for states $|\psi\rangle$ that have zero expectation of the commutator, even if $|\psi\rangle$ is not a common eigenstate of X and Z. It has thus been proposed to use the *entropy* of the probability distribution of the outcomes as the measure of uncertainty [38].

The best known *entropic uncertainty relation* is probably the one by Maassen and Uffink [4]. They proved that, for any state ρ_A of a quantum system A with a finite dimension $d = \dim(\mathcal{H}_A)$, it holds

$$H(X) + H(Z) \geqslant q_{\mathrm{MU}}, \tag{1}$$

where $X = \{|x_j\rangle\}$ and $Z = \{|z_k\rangle\}$ indicate here orthonormal bases on $\mathcal{H}_A$, and $H(X) = -\sum_j p_j^x \log p_j^x$ is the Shannon entropy of the probability distribution $\{p_j^x = \langle x_j|\rho_A|x_j\rangle\}$ (similarly for $H(Z)$). The r.h.s. of (1) measures the strength of the knowledge tradeoff: the sum of the "ignorance" (as measured by entropy) about X and Z cannot be smaller than

$$q_{\mathrm{MU}} = \log(1/c_{\max}), \quad c_{\max} = \max_{j,k} c_{jk}, \quad c_{jk} = |\langle x_j|z_k\rangle|^2. \tag{2}$$

One has $q_{\mathrm{MU}} = 0$ (i.e., $c_{\max} = 1$) if and only if (iff) X and Z share a basis element, while q_{MU} is maximal, $q_{\mathrm{MU}} = \log d$ (i.e., $c_{\max} = 1/d$), iff X and Z are fully complementary, with $c_{jk} = 1/d$ for all j, k.

The uncertainty principle inspired the original proposal for quantum cryptography [5]. However, the uncertainty relations known at the time did not take into account the possibility for an eavesdropper to have quantum correlations, i.e., entanglement [6], with the system being measured. Hence, those relations could not be directly used to prove cryptographic security. Berta et al. [7] filled such a gap, generalizing the uncertainty relation (1) to take into account the possible use of a quantum memory. The latter would allow Bob, who is supposed to have access to a quantum system B that may be entangled to Alice's system A, to violate (1) [7]. Berta et al. showed that nonetheless, for any bipartite state ρ_{AB}, Bob's uncertainty about the result of measurements in the X and Z bases on Alice's system is bounded by

$$H(X|B) + H(Z|B) \geqslant q_{\mathrm{MU}} + H(A|B), \tag{3}$$

where $H(A|B) = H(\rho_{AB}) - H(\rho_B)$ is the *conditional von Neumann entropy*, with $H(\sigma) = -\mathrm{Tr}(\sigma \log \sigma)$ the von Neumann entropy, and ρ_B the reduced state of ρ_{AB} on B. $H(X|B)$ can be interpreted as Bob's ignorance about the result of Alice's measurement of X on A, given that Bob has access to the system B (similarly for $H(Z|B)$) [39]. The two terms $H(X|B)$ and $H(Z|B)$ are non-negative since they represent classical uncertainties, but $H(A|B)$ can be negative if ρ_{AB} is entangled [6], so that the effect of entanglement is to weaken the knowledge tradeoff. While equation (3) reduces to (1) when B is a trivial system, if AB are maximally entangled, $\rho_{AB} = |\phi\rangle\langle\phi|$, $|\phi\rangle = (1/\sqrt{d})\sum_i |i\rangle|i\rangle$, we have $H(A|B) = -\log d \leqslant -q_{\mathrm{MU}}$ independently of X and Z, and the r.h.s. of (3) gives a *trivial* bound on Bob's uncertainty. The generality of (3) opens up a range of applications, e.g., in entanglement witnessing [7–9] and in the security analysis of quantum key distribution [10, 11].

A. Summary of results

One main result of this article is to improve the bound in (3) by replacing q_{MU} with a larger parameter almost always strictly greater than q_{MU}. Another result is the improvement of Hall's "information exclusion principle" [12], which regards the mutual information between the outcomes of measurements on one physical system, and a second system correlated with the first system. Mutual information is a measure of correlations, and is the central quantity in, e.g., communication theory [13]. It quantifies the number of bits of information gained—equivalently, the reduction of ignorance—about X when given access to Y, and can indeed be defined as $I(X:Y) = H(X) - H(X|Y)$. Hall's idea was essentially to reformulate the uncertainty principle in terms of mutual information, as follows. Let X and Z be two orthonormal bases on system A, and let Y be a classical register that may be correlated to A. Then

$$I(X:Y) + I(Z:Y) \leqslant r_{\mathrm{H}}, \quad r_{\mathrm{H}} = \log(d^2 \cdot c_{\max}). \tag{4}$$

Hall's bound says that one cannot probe the register Y in order to obtain complete information about both the X and Z observables, if these two observables have a small value of $c_{\max}$ (defined in (2)). Bounds on the sum of complementary information terms have been called *information exclusion relations* [12, 14, 15]. They have not been studied as much as uncertainty relations [40], and the best known information exclusion relation, Eq. (4), is actually not a very strong bound, as pointed out by Grudka et al. [16]. Grudka et al. have attempted to remedy this by conjecturing a stronger information exclusion relation. They found numerical evidence, and proved analytically in some special cases, that

$$I(X:Y) + I(Z:Y) \leqslant r_{\mathrm{G}}, \quad r_{\mathrm{G}} = \log\left(d \cdot \sum_{\text{d largest}} c_{jk}\right), \tag{5}$$

with the sum over the largest d terms of the matrix $[c_{jk}]$ (again, see (2)). Since $\sum_{\text{d largest}} c_{jk} \leqslant d \cdot c_{\max}$, we have $r_{\mathrm{G}} \leqslant r_{\mathrm{H}}$ (potentially with strict inequality) and, if true, (5) would be an improvement over Hall's bound. In what follows we shall actually prove a *stronger* version of Grudka et al.'s conjecture. Furthermore we will extend it to the much more general case of quantum memory, where Y is replaced by a general quantum system.

Besides improving both the uncertainty relation (3) and the information exclusion relation (4), this article provides the insight that the complementarity of uncertainty (i.e., a limit on the knowledge about the outcomes of complementary observables) and the complementarity of information (i.e., a limit on the correlations between the outcomes of complementary observables and some external system) differ both conceptually and practically. Hence, from the quantitative point of view, we should not expect to have the same complementarity factor appearing in uncertainty relations and information exclusion relations. What makes Hall's bound weak is the use in (4) of the same parameter $c_{\max}$ as in (1).

In what follows, we first give a simplified presentation of our results in Secs. II and III and then discuss their implications in Secs. IV and V. We then give a more detailed presentation, discussing the generalisation of our results for arbitrary positive operator valued measures (POVMs) in Sec. VI, and giving more details on our state-independent approach in Sec. VII. The main technical proofs are given in the Appendix.

II. IMPROVED UNCERTAINTY RELATION

Our main technical result is an entropic uncertainty relation that, much like (3), accounts for the possible reduction of Bob's uncertainty about Alice's system thanks to the entanglement between systems A and B. Before presenting our strongest result, we focus on a simple corollary that gives intuition about the nature of our improvement (see Appendix A 3 for the proof).

Corollary 1. *For any bipartite state ρ_{AB}, and any orthonormal bases $X = \{|x_j\rangle\}$ and $Z = \{|z_k\rangle\}$ on $\mathcal{H}_A$,*

$$H(X|B) + H(Z|B) \geqslant q' + H(A|B), \tag{6}$$

where $H(X|B) = H(\rho_{XB}) - H(\rho_B)$, with $\rho_{XB} = (\mathcal{X} \otimes \mathcal{I})(\rho_{AB})$ and $\mathcal{X}(\cdot) = \sum_j |x_j\rangle\langle x_j|(\cdot)|x_j\rangle\langle x_j|$ (similarly for $H(Z|B)$), and

$$q' = q_{\mathrm{MU}} + \frac{1}{2}(1 - \sqrt{c_{\max}}) \log \frac{c_{\max}}{c_2}, \tag{7}$$

where c_2 is the second largest entry of the matrix $[c_{jk}]$.

Notice that, for small $c_{\max}$, like in the case of almost complementary X and Z, one has $q' \approx \log(1/\sqrt{c_{\max} c_2})$—to be compared with $q_{\mathrm{MU}} = \log(1/c_{\max})$. So our bound nicely captures the importance of both $c_{\max}$ and c_2, i.e., takes into account more information about the relation between the two bases. Clearly $q' \geqslant q_{\mathrm{MU}}$ in general. Furthermore $q' > q_{\mathrm{MU}}$ iff there is exactly one pair $(\hat{j}, \hat{k})$ such that $c_{\max} = c_{\hat{j}\hat{k}}$, with $c_{\max} < 1$. In the special case where the system A is a qubit, it is immediate to check that necessarily $c_{\max} = c_2$, hence $q' = q_{\mathrm{MU}}$. However, for $d \geqslant 3$, we have $q' > q_{\mathrm{MU}}$ for *almost all* pairs of bases (X, Z). Indeed, in $d \geqslant 3$ a typical unitary—seen here as the unitary that connects the two bases, i.e., $X = \{|x_j\rangle\} = \{U|z_j\rangle\}$—has $c_2 < c_{\max} < 1$, see Sec. VII B. We remark that, even for the simple improvement provided by Corollary 1, the gap between q' and q_{MU} can become arbitrarily large. In Sec. VII C, we give an example where the gap $q' - q_{\mathrm{MU}}$ diverges as the logarithm of the dimension d of A [41].

We now state our main technical result, from which all of our other relations follow (see Appendix A 1 for the proof). We first replace the bound q_{MU} in (3) with a *state-dependent* bound $q(\rho_A)$, and then define a new state-independent bound.

Theorem 2. *For any bipartite state ρ_{AB}, and any orthonormal bases $X = \{|x_j\rangle\}$ and $Z = \{|z_k\rangle\}$ on $\mathcal{H}_A$,*

$$H(X|B) + H(Z|B) \geqslant q(\rho_A) + H(A|B), \tag{8}$$

where, from c_{jk} in (2), we define

$$q(\rho_A) = \max\{q(\rho_A, X, Z), q(\rho_A, Z, X)\}, \tag{9a}$$

$$q(\rho_A, X, Z) = \sum_j p_j^x \log(1/\max_k c_{jk}), \tag{9b}$$

$$q(\rho_A, Z, X) = \sum_k p_k^z \log(1/\max_j c_{jk}). \tag{9c}$$

Hence, the following state-independent bound holds:

$$H(X|B) + H(Z|B) \geqslant q + H(A|B), \quad q = \min_{\rho_A} q(\rho_A). \tag{10}$$

It is clear that $q(\rho_A) \geqslant q_{\mathrm{MU}}$, since averaging over j or k gives a larger value than minimising. For A a qubit ($d = 2$), we have that $\max_k c_{jk}$ is independent of j and hence $q(\rho_A) = q_{\mathrm{MU}}$. But $q(\rho_A) \geqslant q'$ (see the proof of Cor. 1), so that, for $d \geqslant 3$, $q(\rho_A) \geqslant q' > q_{\mathrm{MU}}$ for all states, for almost all choices of X and Z. Hence the lower bound q of (10) is an improvement over q_{MU}. By using the minimax theorem, see Sec. VII A, we obtain

$$q = \max_{0 \leqslant p \leqslant 1} \lambda_{\min}[\Delta(p)], \tag{11}$$

where $\lambda_{\min}[\cdot]$ denotes the minimum eigenvalue and $\Delta(p) = p\Delta_{XZ} + (1-p)\Delta_{ZX}$, with $\Delta_{XZ} = \sum_j \log(1/\max_k c_{jk})|x_j\rangle\langle x_j|$ and $\Delta_{ZX} = \sum_k \log(1/\max_j c_{jk})|z_k\rangle\langle z_k|$. Thus, computing q can be done by finding the minimum eigenvalue of particular matrices, a straightforward numerical calculation. Furthermore, by setting $p = 1/2$ in (11) one can get a bound *still certified to be at least as large as* $q' \geqslant q_{\mathrm{MU}}$. In general, we have

$$q \geqslant \lambda_{\min}[\Delta(1/2)] \geqslant q' \geqslant q_{\mathrm{MU}}.$$

Example 1. Let $d = 3$, $Z = \{|0\rangle, |1\rangle, |2\rangle\}$, and $X = \{U|0\rangle, U|1\rangle, U|2\rangle\}$, with

$$U = \begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/\sqrt{6} & -\sqrt{2/3} & 1/\sqrt{6} \end{pmatrix}.$$

We have $q_{\mathrm{MU}} = \log(3/2) \approx 0.58$, $q' \approx 0.62$, $\lambda_{\min}[\Delta(1/2)] \approx 0.64$, and $q \approx 0.64$. Furthermore, our state-dependent bound is often *much* better than q_{MU}: if the reduced state is maximally mixed then $q(\mathbb{1}/3) = (2/3)\log 3 \approx 1.06$, while numerically averaging over all pure states gives $\langle q(|\psi\rangle)\rangle_{|\psi\rangle} \approx 1.07$.

Other attempts have been made to strengthen Eq. (3) [17–19], or the less general relation Eq. (1) [20–22]. Refs. [21, 22] took a majorisation approach; however, their bounds can be weaker than (1) when X and Z have a large q_{MU} value. Ref. [17] added a term to the r.h.s. of (3) that depends on the quantum discord [23] of the state ρ_{AB}; that same term (see [17]) can be added to the r.h.s. of our result (8) if one wishes. Ref. [18] (Ch. 7) replaced q_{MU} in (3) with a state-dependent bound $\hat{q}(\rho_A)$, like we did in (8); however in their case they have $\min_{\rho_A} \hat{q}(\rho_A) = q_{\mathrm{MU}}$, so unlike our result it does not lead to a strengthened state-independent bound.

III. IMPROVED INFORMATION EXCLUSION RELATION

As a corollary of (8), we prove Grudka et al.'s conjectured information exclusion relation [16]. Furthermore, we actually strengthen their bound and extend it to the case of quantum memory. In order to fully appreciate this, let us first consider the extension of Hall's result to the case of quantum memory, i.e., we replace the classical system Y with a general quantum system B. A corollary of (3) is:

$$I(X:B) + I(Z:B) \leqslant r_{\mathrm{H}} - H(A|B). \tag{12}$$

Improving (4), this result allows for entanglement between A and B. It says that the trade-off in correlations is weakened if $H(A|B)$ is negative, i.e. if ρ_{AB} is strongly entangled. After all, in the maximally entangled case, $I(X:B) = I(Z:B) = \log d$, so in such a case the bound on the r.h.s. must be no smaller than $2\log d$.

Now consider the following information exclusion relation, a corollary of our uncertainty relation (8).

Corollary 3. *For any bipartite state ρ_{AB},*

$$I(X:B) + I(Z:B) \leqslant r - H(A|B), \tag{13}$$

with

$$r = \min\{r(X, Z), r(Z, X)\}, \tag{14a}$$

$$r(X, Z) = \log(d \sum_j \max_k c_{jk}), \tag{14b}$$

$$r(Z, X) = \log(d \sum_k \max_j c_{jk}). \tag{14c}$$

Proof. Write $H(X|B) = H(X) - I(X:B)$ (similarly for $H(Z|B)$), rearrange (8), and use $H(Z) \leqslant \log d$ to get

$$I(X:B) + I(Z:B) \leqslant \log d + H(X) - q(\rho_A, X, Z) - H(A|B).$$

Now, $H(X) - q(\rho_A, X, Z) = \sum_j p_j^x \log(\max_k c_{jk}/p_j^x) \leqslant \log(\sum_j \max_k c_{jk})$, where we used the concavity of the log. Bringing d inside the log completes the proof. A similar bound holds when interchanging X and Z. □

This allows us to conclude

Corollary 4. *Grudka et al.'s conjecture,* (5)*, is true.*

Proof. Consider the d different terms $\{\max_k c_{jk}\}_j$ appearing in $r(X,Z)$; these may not be the d largest terms of the matrix $[c_{jk}]$, hence summing over them is smaller than computing $\sum_{\text{d largest}} c_{jk}$. So $r(X,Z) \leqslant r_{\text{G}}$ (see (5)), thus $r \leqslant r_{\text{G}}$. Also, if we set $B = Y$, where Y is classical, then we have $H(A|Y) \geqslant 0$. Combining this with $r \leqslant r_{\text{G}}$ and (13) proves (5). □

We emphasise that Eq. (13) goes well beyond Grudka et al.'s conjecture: it strengthens (5) by replacing r_{G} by r, and it generalises the result to the case of quantum memory, allowing for arbitrary (possibly non-classical) correlations between A and B. In general, we have

$$r \leqslant r_{\text{G}} \leqslant r_{\text{H}}.$$

In the qubit case ($d = 2$), we have equality $r = r_{\text{G}} = r_{\text{H}}$. To see a case where all three are different, consider the qutrit example given in Ex. 1. In this case we have $r_{\text{H}} = \log 6$, $r_{\text{G}} = \log 5$, and $r = \log(9/2)$. Note that r can be calculated analytically given the coefficients c_{jk} of (2).

IV. UNCERTAINTY VERSUS INFORMATION

One key conceptual insight of our work is to draw a distinction between the complementarity of uncertainty and the complementarity of information. The factor $c_{\max}$ naturally appears—via $q_{\text{MU}} = \log(1/c_{\max})$—in uncertainty relations like (1) and (3). But we should not expect it to be the right factor to capture the complementarity of information. While our work shows that uncertainty relations can be improved by replacing q_{MU} with q as in (10), a much more dramatic improvement is given by replacing $r_{\text{H}} = \log(d^2 \cdot c_{\max})$ with r, i.e., going from the information exclusion relation (12) to (13). Indeed, in order to obtain a state-independent bound for uncertainty relations, we must consider the subspace with the least complementarity. On the other hand, in information exclusion relations it is the overall complementarity, i.e. with respect to the various subspaces that compose the space, that matters. The reason our approach is better suited to capture information complementarity is that r measures the overall complementarity, averaged over the whole space, of X and Z. Notice that to obtain our improved *state-independent* information exclusion relation of Corollary 3 we had to tap into the strength of our *state-dependent* uncertainty relation of Theorem 2. Finally, to better appreciate the difference between the complementarity of uncertainty and the complementarity of information, it is instructive to consider the conditions under which our state-independent bounds become trivial, i.e., $q = 0$ and $r = 2\log d$. Let U be the unitary relating X and Z; we have $q = 0$ iff at least one entry of U has magnitude 1. In contrast, $r = 2\log d$ iff U is of the form $U = \sum_j e^{i\phi_j}|P(j)\rangle\langle j|$ for some permutation function P and phase factors $e^{i\phi_j}$. These are vastly different conditions, with the latter one implying that U must be trivial *over the entire space*, whereas the former condition says that only one row or column of U need be trivial.

V. APPLICATIONS

The relevance of (3) for witnessing of entanglement (WoE) and security analysis for quantum key distribution was discussed in [7] and implemented experimentally for WoE in [8, 9]. Since our bound Eq. (8) is an improvement over (3), it enables a tighter analysis. To use our bound $q(\rho_A)$ the only information about ρ_A needed is the probability distributions $\{p_j^x\}$ and $\{p_k^z\}$. In the case of WoE using the uncertainty relation with quantum memory as in [8, 9], Alice already determines these probability distributions experimentally, so no extra effort is needed to use our bound.

Likewise, our Eq. (13) is relevant to *witnessing of good quantum channels* [15, 24]. Consider a channel $\mathcal{E}$ from Alice to Bob. To show that $\mathcal{E}$ is good, Alice can send the X basis states with equal probability through $\mathcal{E}$, and Bob measures the output in basis X_B. Alice does the same for Z and Bob measures Z_B. They compare their results over a classical communication line and estimate $I(X:X_B)$ and $I(Z:Z_B)$. Then they can lower bound the quantum capacity of $\mathcal{E}$, denoted $Q(\mathcal{E})$, using

$$Q(\mathcal{E}) \geqslant I(X:X_B) + I(Z:Z_B) - r, \tag{15}$$

which follows from applying (13) to $\rho_{AB} = (\mathcal{I}\otimes\mathcal{E})(|\phi\rangle\langle\phi|)$ where $|\phi\rangle$ is maximally entangled, and using $Q(\mathcal{E}) \geqslant -H(A|B)$ [25]. Thus, showing that $\mathcal{E}$ has a positive quantum capacity amounts to showing that the r.h.s. of (15) is positive.

Closely related to quantum cryptography are ideas of monogamy or decoupling, whereby strong quantum correlations between A and B guarantee weak correlations between A and any third system C. Equation (3) has been used [15, 26] to give sufficient conditions for which C is decoupled from A, in terms of Bob's uncertainty about X and Z. Our results, (8) and (13), allow these quantitative statements of monogamy to be tightened.

VI. GENERALISATION TO POVMS

A. Results in tripartite form

Our previous results can be rewritten in a form that considers a tripartite state on ABC rather than a bipartite state on AB. The tripartite formulation is equivalent to the bipartite one, i.e., one formulation implies the other [7, 15]. In what follows, we state the tripartite formulation of our results since this form allows us to generalise our results to POVMs in a straightforward way.

Our first main result was Eq. (8). This says that, for any tripartite state ρ_{ABC} and any orthonormal bases

$X = \{|x_j\rangle\}$ and $Z = \{|z_k\rangle\}$ on $\mathcal{H}_A$,

$$H(X|B) + H(Z|C) \geqslant q(\rho_A) \tag{16}$$

where $q(\rho_A)$ was defined in (9). Notice that the term $H(A|B)$ that appeared in (8) has now disappeared since we have changed $H(Z|B)$ to $H(Z|C)$.

Our second main result was Eq. (13). It says that, for any tripartite state ρ_{ABC} and any orthonormal bases $X = \{|x_j\rangle\}$ and $Z = \{|z_k\rangle\}$ on $\mathcal{H}_A$,

$$I(X:B) + I(Z:C) \leqslant r, \tag{17}$$

where r was defined in (14).

We used our second result to prove a conjecture by Grudka et al. [16], which strengthened Hall's information exclusion principle [12]. Hall's scenario considered the case where Y is a classical register and we want to bound the sum $I(X:Y) + I(Z:Y)$. Our second result implied the following bound on this sum:

$$I(X:Y) + I(Z:Y) \leqslant r, \tag{18}$$

which in turn implied Grudka et al.'s conjecture.

In what follows, we will generalise all of these results, Eqs. (16), (17), and (18), to the case where X and Z are arbitrary POVMs (assuming they contain a finite number of POVM elements) on system A.

B. Notation for POVMs

In the general case where $X = \{X_j\}$ and $Z = \{Z_k\}$ are POVMs on A, we consider the isometries $V_X : \mathcal{H}_A \to \mathcal{H}_{XX'A}$ and $V_Z : \mathcal{H}_A \to \mathcal{H}_{ZZ'A}$ defined by [10]

$$V_X = \sum_j |j\rangle_X \otimes |j\rangle_{X'} \otimes \sqrt{X_j}, \tag{19a}$$

$$V_Z = \sum_k |k\rangle_Z \otimes |k\rangle_{Z'} \otimes \sqrt{Z_k}, \tag{19b}$$

where $|j\rangle$ and $|k\rangle$ are elements of the standard (orthonormal) basis on the appropriate spaces. For some initial tripartite state ρ_{ABC} we denote the alternative post-measurement states as:

$$\widehat{\rho}_{XX'ABC} = V_X \rho_{ABC} V_X^\dagger, \tag{20a}$$

$$\overline{\rho}_{ZZ'ABC} = V_Z \rho_{ABC} V_Z^\dagger. \tag{20b}$$

Then we define

$$H(X|B) = H(\widehat{\rho}_{XB}) - H(\rho_B), \tag{21a}$$

$$H(Z|C) = H(\overline{\rho}_{ZC}) - H(\rho_C), \tag{21b}$$

which are the conditional entropies of the classical quantum states $\widehat{\rho}_{XB} = \mathrm{Tr}_{X'AC}(\widehat{\rho}_{XX'ABC})$ and $\overline{\rho}_{ZC} = \mathrm{Tr}_{Z'AB}(\overline{\rho}_{ZZ'ABC})$, respectively. For example, notice that we can write

$$\begin{aligned}\widehat{\rho}_{XB} &= \sum_j |j\rangle\langle j|_X \otimes \mathrm{Tr}_A(X_j \rho_{AB}) \\ &= (\mathcal{X} \otimes \mathcal{I})(\rho_{AB})\end{aligned} \tag{22}$$

for the quantum channel $\mathcal{X} : \rho_A \mapsto \sum_j |j\rangle\langle j|_X \mathrm{Tr}(X_j \rho_A)$. Also, we denote the probabilities associated with these two POVMs as $p_j^x = \mathrm{Tr}(X_j \rho_A)$ and $p_k^z = \mathrm{Tr}(Z_k \rho_A)$.

C. Uncertainty relation for POVMs

Generalising the results to POVMs essentially amounts to finding an appropriate generalisation of the complementarity factor that appears in our bounds, such as $q(\rho_A)$ and r. In what follows, we will use the factors:

$$h_j(X, Z) = \Big\| \sum_k Z_k X_j Z_k \Big\|_\infty, \tag{23a}$$

$$h_k(Z, X) = \Big\| \sum_j X_j Z_k X_j \Big\|_\infty, \tag{23b}$$

where the infinity norm (or operator norm) $\|M\|_\infty$ is the largest singular value of M, or in the case of (23) it is the largest eigenvalue since the arguments are positive semi-definite matrices. We discuss in the next subsection why we chose this complementarity factor - the reason being that it gives a stronger bound than an alternative, as discussed below.

Now we generalise (16) to the case of arbitrary POVMs with the following result, proved in App. A 1.

Theorem 5. *Let $X = \{X_j\}$ and $Z = \{Z_k\}$ be arbitrary POVMs on A. Then for any tripartite state ρ_{ABC},*

$$H(X|B) + H(Z|C) \geqslant q(\rho_A) \tag{24}$$

where we define

$$q(\rho_A) = \max\{q(\rho_A, X, Z), q(\rho_A, Z, X)\}, \tag{25a}$$

$$q(\rho_A, X, Z) = -\sum_j p_j^x \log h_j(X, Z), \tag{25b}$$

$$q(\rho_A, Z, X) = -\sum_k p_k^z \log h_k(Z, X). \tag{25c}$$

Notice that our definition of $q(\rho_A)$ reduces to that given in (9) when we specialise to the case of orthonormal bases (in other words, rank-one projective POVMs). This is because, when Z is projective, then $h_j(X, Z) = \max_k \|Z_k X_j Z_k\|_\infty$ and further specialising to X and Z being composed of rank-one projectors reduces the formula to $h_j(X, Z) = \max_k c_{jk}$, which is the formula appearing in (9).

While we have taken the tripartite view to give a simple statement of our results for POVMs, it is possible rewrite (24) in a bipartite form, using an approach similar to that in [18]. We obtain:

$$H(X|B) + H(Z|B) \geqslant q(\rho_A) + H(A|B) - f \tag{26}$$

where $f := \min\{H(A|BX)_{\widehat{\rho}}, H(A|BZ)_{\overline{\rho}}\}$, and where $H(A|BX)_{\widehat{\rho}}$ and $H(A|BZ)_{\overline{\rho}}$ denote the conditional entropies of $\widehat{\rho}_{XAB}$ and $\overline{\rho}_{ZAB}$, respectively. For the case of orthonormal bases considered earlier, $f = 0$.

D. Choice of complementarity factor

The following technical lemma, proved in App. A 2, is relevant to our choice of complementarity factor for POVMs [42].

Lemma 6. *Let σ be an arbitrary operator—that is, an arbitrary square matrix, although we will be interested mostly in the case in which σ is positive semidefinite, hence the choice of notation—and let $Z = \{Z_k\}$ be any POVM. Then*

$$\Big\| \sum_k Z_k \sigma Z_k \Big\|_\infty \leqslant \max_k \Big\| \sqrt{Z_k}\sigma\sqrt{Z_k} \Big\|_\infty. \tag{27}$$

Our choice of complementarity factor was inspired by Refs. [18, 19]. In particular, in Chapter 7 of [18] Tomamichel conjectures that, for any two POVMs X and Z,

$$\max_j \Big\| \sum_k Z_k X_j Z_k \Big\|_\infty \leqslant \max_{j,k} c_{jk}, \tag{28}$$

where

$$c_{jk} = \Big\| \sqrt{Z_k} X_j \sqrt{Z_k} \Big\|_\infty = \Big\| \sqrt{X_j}\sqrt{Z_k} \Big\|_\infty^2. \tag{29}$$

Clearly our Lemma 6 implies Eq. (28) and hence resolves an outstanding conjecture. The reason this conjecture was interesting was because the factors on the left- and right-hand-sides of (28) were alternative complementarity factors that could potentially be used as bounds in the uncertainty relation. Indeed the r.h.s. of (28) was used in several uncertainty relations [10, 15, 28], so proving that the l.h.s. of (28) is smaller, as we have done here, shows that the l.h.s. provides a better bound for POVM uncertainty relations. (This issue is only of concern for general POVMs, since the two factors in (28) are equal when X and Z are orthonormal bases.)

This discussion has relevance to the present article since our derived bound in Theorem 5 involves quantities $q(\rho_A, X, Z)$ and $q(\rho_A, Z, X)$ defined in terms of $h_j(X, Z)$ and $h_k(Z, X)$ given in (23). But from Lemma 6, these quantities are bounded by

$$h_j(X,Z) \leqslant \max_k \big\| \sqrt{Z_k} X_j \sqrt{Z_k} \big\|_\infty = \max_k c_{jk}, \tag{30a}$$

$$h_k(Z,X) \leqslant \max_j \big\| \sqrt{X_j} Z_k \sqrt{X_j} \big\|_\infty = \max_j c_{jk}. \tag{30b}$$

Hence our bound involving $h_j(X, Z)$ and $h_k(Z, X)$ is stronger than the one obtained from replacing them with the quantities on the right-hand-sides of (30). This provides justification for our choice of complementarity factor.

E. Information exclusion relation for POVMs

Here we use Theorem 5 to derive an information exclusion relation that is generalised to the POVM case. Again, we note that the following definition of r reduces to that in (14) when X and Z are specialised to be orthonormal bases.

Corollary 7. *Let $X = \{X_j\}$ and $Z = \{Z_k\}$ be arbitrary POVMs on A. Then for any tripartite state ρ_{ABC},*

$$I(X:B) + I(Z:C) \leqslant r, \tag{31}$$

where we define

$$r = \min\{r(X,Z), r(Z,X)\}, \tag{32a}$$

$$r(X,Z) = \log[|Z| \sum_j h_j(X,Z)], \tag{32b}$$

$$r(Z,X) = \log[|X| \sum_k h_k(Z,X)]. \tag{32c}$$

where $|Z|$ and $|X|$ denote the number of POVM elements.

Proof. Write $H(X|B) = H(X) - I(X:B)$ and $H(Z|C) = H(Z) - I(Z:C)$, then rearrange (24) and use $H(Z) \leqslant \log|Z|$ to get

$$I(X:B) + I(Z:C) \leqslant \log|Z| + H(X) - q(\rho_A, X, Z).$$

Now write

$$\begin{aligned} H(X) - q(\rho_A, X, Z) &= \sum_j p_j^x \log[h_j(X,Z)/p_j^x] \\ &\leqslant \log[\sum_j h_j(X,Z)], \end{aligned} \tag{33}$$

where we used the concavity of the log. Bringing $|Z|$ inside the log completes the proof, and by symmetry the same bound holds where one interchanges X and Z. □

Finally, we generalise (18) to the POVM case. The following result is applicable to the same scenario that Hall considered in his information exclusion principle, except we have generalised it to the case where X and Z are POVMs.

Corollary 8. *Let $X = \{X_j\}$ and $Z = \{Z_k\}$ be arbitrary POVMs on A. Let Y be a classical register that may be correlated to A, i.e., ρ_{AY} is an arbitrary quantum-classical state. Then,*

$$I(X:Y) + I(Z:Y) \leqslant r \tag{34}$$

where r is defined by Eq. (32).

Proof. Apply (31) to the tripartite state $\rho_{AYY'}$ where system Y' is an exact copy of system Y, such that $\rho_{AY'} = \mathrm{Tr}_Y(\rho_{AYY'})$ is of the same form as $\rho_{AY} = \mathrm{Tr}_{Y'}(\rho_{AYY'})$. (Note: the fact that Y is classical allows us to copy its correlations with A.) In this case we have $I(X:Y') = I(X:Y)$, hence proving (34). □

VII. STATE-INDEPENDENT BOUND FOR UNCERTAINTY RELATION

A. Computable expression

Now let us consider the state-independent version of our bound, defined by

$$q = \min_{\rho_A} q(\rho_A).$$

In Sec. II we noted that this bound can be rewritten in an alternative form that may be easier to calculate. Here we derive this alternative form.

Let us first rewrite $q(\rho_A)$ as follows:

$$\begin{aligned} q(\rho_A) &= \max\{q(\rho_A, X, Z), q(\rho_A, Z, X)\} \\ &= \max_{0\leqslant p\leqslant 1} [p\, q(\rho_A, X, Z) + (1-p) q(\rho_A, Z, X)] \\ &= \max_{0\leqslant p\leqslant 1} [p \mathrm{Tr}[\rho_A \cdot \sum_j X_j \log(1/h_j(X,Z))] \\ &\quad + (1-p)\mathrm{Tr}[\rho_A \cdot \sum_k Z_k \log(1/h_k(Z,X))]] \\ &= \max_{0\leqslant p\leqslant 1} \mathrm{Tr}[\rho_A \Delta(p)], \end{aligned} \tag{35}$$

where we define

$$\begin{aligned} \Delta(p) &= p\Delta_{XZ} + (1-p)\Delta_{ZX}, \\ \Delta_{XZ} &= \sum_j \log(1/h_j(X,Z)) \cdot X_j, \\ \Delta_{ZX} &= \sum_k \log(1/h_k(Z,X)) \cdot Z_k. \end{aligned} \tag{36}$$

From $h_j(X,Z) \leqslant 1$ and $h_k(Z,X) \leqslant 1$, it follows that $\Delta_{XZ} \geqslant 0$ and $\Delta_{ZX} \geqslant 0$, and hence $\Delta(p) \geqslant 0$.

Next, thanks to the linearity in the arguments, we can use the minimax theorem to interchange the min and max in q as follows:

$$\begin{aligned} q &= \min_{\rho_A} \max_{0\leqslant p\leqslant 1} \mathrm{Tr}[\rho_A \Delta(p)] \\ &= \max_{0\leqslant p\leqslant 1} \min_{\rho_A} \mathrm{Tr}[\rho_A \Delta(p)] \\ &= \max_{0\leqslant p\leqslant 1} \lambda_{\min}[\Delta(p)]. \end{aligned} \tag{37}$$

The formula in (37) makes it possible to numerically calculate q. Given the POVM elements of X and Z, it is straightforward to numerically diagonalise $\Delta(p)$ for a fixed p; then the maximisation over p can be plotted graphically. For example, Fig. 1 shows this plot for Example 1 given in the main text, yielding a value of $q \approx 0.64$.

It is also worth noticing that, since $\lambda_{\min}[\Delta_{XZ}] = \lambda_{\min}[\Delta_{ZX}] = q_{\mathrm{MU}}$, Eq. (37) is another way of seeing that $q \geqslant q_{\mathrm{MU}}$. Also, since the smallest eigenvalue satisfies $\lambda_{\min}[A+B] \geqslant \lambda_{\min}[A] + \lambda_{\min}[B]$ for any two Hermitian matrices [30], we have that $q = q_{\mathrm{MU}}$ iff the function $\lambda_{\min}[\Delta(p)]$ is independent of p and hence is equal to q_{MU} for all p.

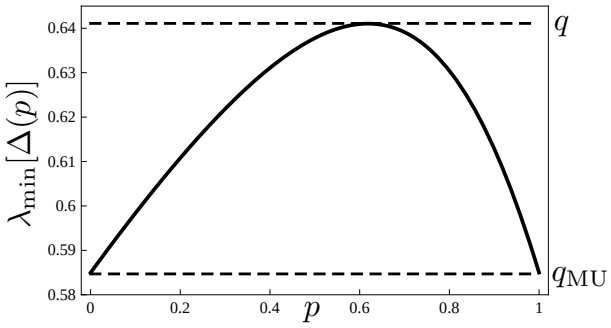

FIG. 1: Plot of the minimum eigenvalue of $\Delta(p)$ as a function of p, for Example 1 given in the main text. The maximum in the plot corresponds to $q \approx 0.64$, which is an improvement over the old bound $q_{\mathrm{MU}} \approx 0.58$ corresponding to the value at $p = 0$ and $p = 1$.

B. Analytical bound

While q is our strongest state-independent bound, we can find a slightly weaker state-independent bound q' that is given by a simple, analytical expression and is still an improvement over q_{MU}. In Cor. 1, we gave the form of q' in terms of the largest and second-largest entries of the matrix c_{jk}. To state this result for general POVMs, we define c_{jk} according to (29), which reduces to the expression in (2) in the case of orthonormal bases.

Now we generalise Cor. 1 to POVMs as follows, with the proof in App. A 3.

Theorem 9. *Let $(\hat{j}, \hat{k})$ be a pair of indices such that $c_{\hat{j}\hat{k}} = \max_{jk} c_{jk} = c_{\max}$, where c_{jk} is defined in (29), so that $c_{\hat{j}\hat{k}} = \|\sqrt{X_{\hat{j}}}\sqrt{Z_{\hat{k}}}\|_\infty^2$. Let c_2 be the second-largest entry of the matrix c_{jk} (possibly equal to $c_{\max}$). It holds that $q(\rho_A) \geqslant q'$ where q' is a state-independent parameter given by*

$$q' = q_{\mathrm{MU}} + \frac{1}{2}(1 - \sqrt{c_{\max}}) \log\left(\frac{c_{\max}}{c_2}\right). \tag{38}$$

Eq. (38) allows us to argue that, if $d \geqslant 3$, our bound q' (and hence also q) is an improvement over the standard bound q_{MU} for almost all pairs of orthonormal bases (X, Z). To argue this, we will need the following lemma, proved in App. A 4, kindly provided by N. Johnston [33].

Lemma 10. *For any dimension $d \geqslant 3$, the entries U_{ij} of a generic d-dimensional unitary U satisfy*

$$|U_{ij}| \neq |U_{kl}|, \quad \forall (i,j) \neq (k,l). \tag{39}$$

That is, for any dimension $d \geqslant 3$ the set of unitaries that violate (39) has vanishing measure with respect to the Haar measure.

Combining Theorem 9 with Lemma 10 immediately leads to the following conclusion.

Corollary 11. *In any dimension $d \geqslant 3$, for almost all choices of two orthonormal bases one has $c_2 < c_{\max} < 1$, hence $q' > q_{\mathrm{MU}}$.*

Proof. The bases are related by a unitary transformation U, represented in the first basis by entries U_{ij}. The parameter $c_{\max}$ corresponds to the modulus square of the largest entry. Because of Lemma 10, we have $c_{\max} < 1$ generically. Indeed, if $c_{\max} = 1$, all the remaining entries in the same row or column must vanish, violating (39). Finally, also the condition $c_2 = c_{\max}$ corresponds to a violation of (39). The use of (38) completes the claim. □

C. Arbitrarily large gap between q_{MU} and our bound

We show here that the gap between q_{MU} and our state-independent bound can grow unboundedly, and more precisely logarithmically in the dimension of the system involved. In the case of q, this can be seen by observing that it is additive on tensor copies, i.e., $q(\{X^{\otimes n}, Z^{\otimes n}\}) = nq(\{X, Z\})$ where X and Z are arbitrary POVMs. Since q_{MU} is also additive on tensor copies, any gap between q_{MU} and q for a single copy of X and Z will get multiplied by n.

Our simple analytical bound q' is not additive on tensor copies. Nonetheless, we construct the following example for which $\delta := q' - q_{\mathrm{MU}}$ grows as $\log d$. Consider a Hilbert space $\mathcal{H}_A = \mathcal{H}_{A_1} \oplus \mathcal{H}_{A_2}$ with $\dim(\mathcal{H}_A) = d$, $\dim(\mathcal{H}_{A_1}) = 1$, $\dim(\mathcal{H}_{A_2}) = d - 1$. Let

$$U_0 = \mathbb{1}_1 \oplus F_{d-1} \tag{40}$$

be a unitary acting on $\mathcal{H}_A$ where $\mathbb{1}_1$ is the 1×1 identity matrix (acting on $\mathcal{H}_{A_1}$). Also, $F_{d-1} = \sum_{j,k} \frac{\omega^{jk}}{\sqrt{d-1}} |j\rangle\langle k| = \sum_j |t_j\rangle\langle j|$, with $\omega = e^{2\pi i/(d-1)}$, is the Fourier matrix of dimension $d - 1$, which acts on $\mathcal{H}_{A_2}$ by mapping the standard basis $S = \{|j\rangle\}$ to the basis $T = \{|t_j\rangle\}$. We suppose that the orthonormal bases on $\mathcal{H}_A$ of interest (X and Z) for the uncertainty relation are related by a unitary U that is the product of U_0 and a slight rotation U_r, i.e.,

$$U = U_r U_0. \tag{41}$$

Now let $|y_0\rangle \in \mathcal{H}_{A_2}$ be a state that is unbiased with respect to both the S basis and the T basis on $\mathcal{H}_{A_2}$. (It is always possible to find such a state regardless of the Hilbert space dimension.) We define U_r by

$$U_r = e^{-iH_r\theta}, \quad H_r = |y_0'\rangle\langle 0| + |0\rangle\langle y_0'|, \tag{42}$$

where $|y_0'\rangle = V|y_0\rangle$ and $V : \mathcal{H}_{A_2} \to \mathcal{H}_A$ is an isometry that embeds the $d-1$ dimensional space $\mathcal{H}_{A_2}$ into the d dimensional space $\mathcal{H}_A$ defined by $V = \sum_{j=0}^{d-2} |j+1\rangle\langle j|$. We choose the rotation angle $0 < \theta < \pi/2$ to be a constant, i.e., independent of d. Note that

$$H_r^2 = |0\rangle\langle 0| + |y_0'\rangle\langle y_0'|, \quad H_r^3 = H_r, \quad H_r^4 = H_r^2,$$

which implies that

$$\sin(H_r\theta) = H_r \cdot \sin\theta,$$
$$\cos(H_r\theta) = (\mathbb{1} - H_r^2) + H_r^2 \cdot \cos\theta,$$

and

$$\begin{aligned} U_r &= \cos(H_r\theta) - i\sin(H_r\theta) \\ &= (\mathbb{1} - H_r^2) + H_r^2 \cdot \cos\theta - iH_r \cdot \sin\theta. \end{aligned} \tag{43}$$

For $j \neq 0$ and $k \neq 0$, we have

$$\begin{aligned} &\langle 0|U|0\rangle = \cos(\theta), \\ &\langle 0|U|j\rangle = -i\sin(\theta)\langle y_0'|j\rangle, \\ &\langle j|U|0\rangle = -i\sin(\theta)\langle j|y_0'\rangle, \\ &\langle j|U|k\rangle = \langle j|F_{d-1}'|k\rangle + (\cos(\theta) - 1)\langle j|y_0'\rangle\langle y_0'|F_{d-1}'|k\rangle, \end{aligned}$$

where we write $F_{d-1}' = VF_{d-1}V^\dagger$ for clarity. In the limit of large d, this gives

$$\begin{aligned} c_{00} &= |\langle 0|U|0\rangle|^2 = \cos^2\theta, \\ c_{0j} &= |\langle 0|U|j\rangle|^2 \approx (1/d)\sin^2\theta, \\ c_{j0} &= |\langle j|U|0\rangle|^2 \approx (1/d)\sin^2\theta, \\ c_{jk} &= |\langle j|U|k\rangle|^2 \approx 1/d. \end{aligned} \tag{44}$$

Thus, in this limit, we have

$$c_{\max} = \cos^2\theta, \quad c_2 \approx 1/d. \tag{45}$$

So for large d the gap is given by

$$\begin{aligned} \delta &:= q' - q_{\mathrm{MU}} \\ &= \frac{1}{2}(1 - \sqrt{c_{\max}}) \log \frac{c_{\max}}{c_2} \\ &\approx \frac{1}{2}(1 - \cos\theta) \log(d\cos^2\theta). \end{aligned} \tag{46}$$

So δ grows with $\log d$ in this example.

VIII. CONCLUSIONS

We gave two main results: we strengthened the bound in the uncertainty principle with quantum memory, and we formulated an information exclusion relation (a bound on complementary mutual information terms) that also allows for quantum memory. The latter is a major improvement over previously known information exclusion relations, with a much stronger bound that even provides qualitatively new insight into the complementarity of information and how it differs from that of uncertainty. Our results have applications in, e.g., quantum cryptography, entanglement verification and quantum communication. It would be interesting to see if our results extend to smooth entropies or smooth mutual informations that are relevant to non-asymptotic information theory [18, 27].

IX. ACKNOWLEDGMENTS

We thank Marco Tomamichel, Koenraad Audenaert and Maris Ozols for helpful discussions. We thank Nathaniel Johnston for providing the proof of Lemma 10. PJC 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). MP acknowledges support from NSERC, CIFAR, DARPA, and Ontario Centres of Excellence.

Appendix A: Proofs of Technical Results

1. Proof of Thm. 5

Let us first state the following lemma that is used in proving the uncertainty relation. The lemma was given in [15], but we reproduce its proof here for completeness.

Lemma 12. *[15] Let $Z = \{Z_k\}$ be any POVM on system A, then for any tripartite state ρ_{ABC},*

$$H(Z|C) \geqslant D(\rho_{AB}||\sum_k Z_k \rho_{AB} Z_k). \tag{A1}$$

Proof. Consider the state $\overline{\rho}_{ZZ'ABC}$ defined in (20). Applying strong subadditivity to this state gives $H(Z|C) + H(Z|Z'AB) \geqslant 0$. Now note that conditional entropy can be rewritten in terms of relative entropy with the formula $-H(A|B)_\sigma = D(\sigma_{AB}||\mathbb{1} \otimes \sigma_B)$. So we have:

$$\begin{aligned} H(Z|C) &\geqslant -H(Z|Z'AB) && \text{(A2)}\\ &= D(\overline{\rho}_{ZZ'AB}||\mathbb{1} \otimes \overline{\rho}_{Z'AB}) && \text{(A3)}\\ &\geqslant D(\overline{\rho}_{ZZ'AB}||V_Z V_Z^\dagger(\mathbb{1} \otimes \overline{\rho}_{Z'AB})V_Z V_Z^\dagger) && \text{(A4)}\\ &= D(\rho_{AB}||V_Z^\dagger(\mathbb{1} \otimes \overline{\rho}_{Z'AB})V_Z) && \text{(A5)}\\ &= D(\rho_{AB}||\sum_k Z_k \rho_{AB} Z_k). && \text{(A6)} \end{aligned}$$

The third line used the property $D(\rho||\sigma) \geqslant D(\rho||\Pi_\rho \sigma \Pi_\rho)$ where Π_ρ is a projector onto a space that includes the support of ρ; in this case we chose $\Pi_\rho = V_Z V_Z^\dagger$. The fourth line used the invariance of relative entropy under isometries. It is straightforward to verify the fifth line using $\overline{\rho}_{Z'AB} = \sum_k |k\rangle\langle k| \otimes \sqrt{Z_k}\rho_{AB}\sqrt{Z_k}$. □

Now we prove Thm. 5, which implies Thm. 2 as a special case.

Proof. Starting from Lemma 12 we invoke the data-processing inequality for the quantum channel $\mathcal{X}$ in (22), as follows

$$\begin{aligned} H(Z|C) &\geqslant D(\rho_{AB}||\sum_k Z_k \rho_{AB} Z_k)\\ &\geqslant D(\widehat{\rho}_{XB}||\sum_{j,k} |j\rangle\langle j| \otimes \mathrm{Tr}_A(Z_k X_j Z_k \rho_{AB})) && \text{(A7)}\\ &\geqslant D(\widehat{\rho}_{XB}||\sum_j h_j(X,Z)|j\rangle\langle j| \otimes \rho_B) && \text{(A8)}\\ &= -H(\widehat{\rho}_{XB}) - \mathrm{Tr}_{XB}[\widehat{\rho}_{XB} \log \sum_j h_j(X,Z)|j\rangle\langle j| \otimes \rho_B] && \text{(A9)}\\ &= -H(X|B) - \mathrm{Tr}_X[\rho_X \log \sum_j h_j(X,Z)|j\rangle\langle j|] && \text{(A10)}\\ &= -H(X|B) + q(\rho_A, X, Z) && \text{(A11)} \end{aligned}$$

where the fifth line used the additivity of the log for tensor products. The third line invoked the property $D(S||T) \geqslant D(S||T')$ if $T' \geqslant T$, where we note that

$$\begin{aligned} &\sum_{j,k} |j\rangle\langle j| \otimes \mathrm{Tr}_A(Z_k X_j Z_k \rho_{AB})\\ &\quad = \sum_j |j\rangle\langle j| \otimes \mathrm{Tr}_A[(\sum_k Z_k X_j Z_k)\rho_{AB}]\\ &\quad \leqslant \sum_j h_j(X,Z)|j\rangle\langle j| \otimes \rho_B \end{aligned} \tag{A12}$$

since $\sum_k Z_k X_j Z_k \leqslant \|\sum_k Z_k X_j Z_k\|_\infty \mathbb{1}$.

Finally, by symmetry, one can interchange X and Z in the bound and hence use $q(\rho_A)$. □

2. Proof of Lem. 6

Proof. First notice that

$$\max_k \left\| \sqrt{Z_k}\sigma\sqrt{Z_k} \right\|_\infty = \left\|\rho\right\|_\infty,$$

where

$$\rho := \sum_k |k\rangle\langle k| \otimes \sqrt{Z_k}\sigma\sqrt{Z_k}$$

and $\{|k\rangle\}$ is the standard basis on an auxiliary space. Now consider the isometry $V = \sum_k |k\rangle \otimes \sqrt{Z_k}$ and notice that

$$\sum_k Z_k \sigma Z_k = V^\dagger \rho V.$$

So we wish to show that

$$\left\|\rho\right\|_\infty \geqslant \left\|V^\dagger \rho V\right\|_\infty.$$

Consider the projector $\Pi = VV^\dagger$ and the channel $\mathcal{E}(\cdot) = \Pi(\cdot)\Pi + (\mathbb{1} - \Pi)(\cdot)(\mathbb{1} - \Pi)$ that pinches with respect to this projector. It is a standard result in matrix analysis

that the infinity norm never increases upon pinching the argument [30]. So we have

$$\begin{aligned}\|\rho\|_\infty &\geqslant \|\mathcal{E}(\rho)\|_\infty \\ &= \max\{\|\Pi\rho\Pi\|_\infty, \|(\mathbb{1}-\Pi)\rho(\mathbb{1}-\Pi)\|_\infty\} \\ &\geqslant \|\Pi\rho\Pi\|_\infty = \|V^\dagger\rho V\|_\infty, \end{aligned} \tag{A13}$$

where the last equality uses the invariance of the norm under isometries. □

3. Proof of Thm. 9

We first note the following useful lemma, shown, e.g., in Refs. [31, 32].

Lemma 13. *For any positive semi-definite operators $S \geqslant 0$ and $T \geqslant 0$, we have*

$$\|S+T\|_\infty \leqslant \max\{\|S\|_\infty, \|T\|_\infty\} + \|\sqrt{S}\sqrt{T}\|_\infty. \tag{A14}$$

Now we prove Thm. 9, which implies Cor. 1 as a special case.

Proof. From the definition (25) of $q(\rho_A)$, and using (30), we have

$$\begin{aligned} q(\rho_A) &\geqslant \max\{-\sum_j p_j^x \log(\max_k c_{jk}), -\sum_k p_k^z \log(\max_j c_{jk})\} \\ &\geqslant \frac{1}{2}[-\sum_j p_j^x \log(\max_k c_{jk}) - \sum_k p_k^z \log(\max_j c_{jk})] \\ &\geqslant \frac{1}{2}[-p_{\hat{j}}^x \log c_{\max} - (1-p_{\hat{j}}^x)\log c_2 \\ &\quad - p_{\hat{k}}^z \log c_{\max} - (1-p_{\hat{k}}^z)\log c_2] \\ &= q_{\mathrm{MU}} + \frac{1}{2}\log\left(\frac{c_{\max}}{c_2}\right)[2-(p_{\hat{j}}^x + p_{\hat{k}}^z)]. \end{aligned} \tag{A15}$$

Since $\log(c_{\max}/c_2) \geqslant 0$ by assumption, to bound $q = \min_{\rho_A} q(\rho_A)$ we need to evaluate

$$\begin{aligned} \max_{\rho_A}(p_{\hat{j}}^x + p_{\hat{k}}^z) &= \max_{\rho_A} \mathrm{Tr}[\rho_A(X_{\hat{j}} + Z_{\hat{k}})] \\ &= \|X_{\hat{j}} + Z_{\hat{k}}\|_\infty \\ &\leqslant \max\{\|X_{\hat{j}}\|_\infty, \|Z_{\hat{k}}\|_\infty\} + \|\sqrt{X_{\hat{j}}}\sqrt{Z_{\hat{k}}}\|_\infty \\ &\leqslant 1 + \sqrt{c_{\max}}, \end{aligned} \tag{A16}$$

where in the first inequality we have used (A14) from Lemma 13, and in the second inequality the fact that $X_{\hat{j}}$ and $Z_{\hat{k}}$, being POVM elements, both have operator norm less than unity. Plugging this into (A15) proves (38). □

4. Proof of Lem. 10

Proof. The proof relies on concepts of algebraic geometry [34]. The set of unitaries in dimension d has real dimension d^2, that is, one has to specify d^2 real parameters to specify a unitary U. On the other hand, unitaries can be seen as forming a real algebraic variety $\mathcal{U}$ in $\mathbb{R}^{2d^2}$. Indeed, let the real numbers x_{kl} and y_{kl} be the real and imaginary components of the matrix entry U_{kl}, i.e., $U_{kl} = x_{kl} + iy_{kl}$. Then the condition $U^\dagger U = \mathbb{1}$ corresponds to a system of quadratic equations in the x_{kl}'s and y_{kl}'s. Since the unitaries form a connected group, the algebraic variety $\mathcal{U}$ is irreducible [34]. In particular, if $\mathcal{Z}$ is another algebraic variety, either $\mathcal{U} \cap \mathcal{Z}$ is equal to $\mathcal{U}$ (if $\mathcal{U} \subseteq \mathcal{Z}$) or $\mathcal{U} \cap \mathcal{Z}$ has real dimension strictly smaller than d^2. For any choice of two ordered pairs (i,j) and (k,l), consider the algebraic variety $\mathcal{Z}_{(i,j)(k,l)} \subseteq \mathbb{R}^{2d^2}$ defined by $(x_{ij}^2 + y_{ij}^2) - (x_{kl}^2 + y_{kl}^2) = 0$. It is easy to check that for $d \geqslant 3$, $\mathcal{U} \not\subseteq \mathcal{Z}_{(i,j)(k,l)}$ for every choice of $(i,j) \neq (k,l)$. This is because, when $d \geqslant 3$, for any $(i,j) \neq (k,l)$ it is possible to find a unitary that does not belong to $\mathcal{Z}_{(i,j)(k,l)}$. Notice that, on the other hand, $\mathcal{U} = \mathcal{Z}_{(1,1)(2,2)} = \mathcal{Z}_{(1,2)(2,1)}$ for $d = 2$. Thus, for $d \geqslant 3$ and $(i,j) \neq (k,l)$, $\mathcal{U} \cap \mathcal{Z}_{(i,j)(k,l)}$ has real dimension strictly less than d^2, hence vanishing Haar measure. Given that there is a finite number of sets $Z_{(i,j)(k,l)}$, it also holds that the union of all $\mathcal{U} \cap \mathcal{Z}_{(i,j)(k,l)}$'s has real dimension strictly less than d^2 and vanishing Haar measure. The claim follows. □

[1] W. Heisenberg, Zeitschrift für Physik **43**, 172 (1927).
[2] E. Kennard, Z. Phys **44**, 326 (1927).
[3] H. P. Robertson, Phys. Rev. **34**, 163 (1929).
[4] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. **60**, 1103 (1988).
[5] S. Wiesner, SIGACT News **15**, 78 (1983).
[6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[7] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nature Physics **6**, 659 (2010).
[8] R. Prevedel, D. R. Hamel, R. Colbeck, K. Fisher, and K. J. Resch, Nature Physics **7**, 757 (2011).
[9] C.-F. Li, J.-S. Xu, X.-Y. Xu, K. Li, and G.-C. Guo, Nature Physics **7**, 752 (2011).

[10] M. Tomamichel and R. Renner, Phys. Rev. Lett. **106**, 110506 (2011).

[11] M. Tomamichel, C. C. W. Lim, N. Gisin, and R. Renner, Nature Communications **3**, 634 (2012).

[12] M. J. W. Hall, Phys. Rev. Lett. **74**, 3307 (1995).

[13] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley, New York, 2005), 2nd ed.

[14] M. J. W. Hall, Phys. Rev. A **55**, 100 (1997).

[15] P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths, Phys. Rev. A **83**, 062338 (2011).

[16] A. Grudka, M. Horodecki, P. Horodecki, R. Horodecki, W. Kłobus, and Ł. Pankowski, Phys. Rev. A **88**, 032106 (2013), 1210.8317.

[17] A. K. Pati, M. M. Wilde, A. R. U. Devi, A. K. Rajagopal, and Sudha, Phys. Rev. A **86**, 042105 (2012).

[18] M. Tomamichel, Ph.D. thesis, ETH Zürich (2012), URL `http://arxiv.org/abs/1203.2142`.

[19] M. Tomamichel and E. Hänggi, Journal of Physics A: Mathematical and Theoretical **46**, 055301 (2013).

[20] J. I. de Vicente and J. Sánchez-Ruiz, Phys. Rev. A **77**, 042110 (2008).

[21] S. Friedland, V. Gheorghiu, and G. Gour, ArXiv e-prints (2013), 1304.6351.

[22] Z. Puchała, Ł. Rudnicki, and K. Życzkowski, Journal of Physics A: Mathematical and Theoretical **46**, 272002 (2013).

[23] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, Rev. Mod. Phys. **84**, 1655 (2012).

[24] M. Christandl and A. Winter, IEEE Trans. Inf. Theory **51**, 3159 (2005).

[25] S. Lloyd, Phys. Rev. A **55**, 1613 (1997).

[26] J. M. Renes and J.-C. Boileau, Phys. Rev. Lett. **103**, 020402 (2009).

[27] R. Renner, Ph.D. thesis, ETH Zürich (2005), URL `http://arxiv.org/abs/quant-ph/0512258`.

[28] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, Phys. Rev. Lett. **108**, 210405 (2012).

[29] P. J. Coles, Phys. Rev. A **85**, 042103 (2012).

[30] R. Bhatia, *Matrix analysis*, vol. 169 (Springer, 1997).

[31] M. Tomamichel, S. Fehr, J. Kaniewski, and S. Wehner, ArXiv e-prints (2012), 1210.4359.

[32] C. Schaffner, Ph.D. thesis, University of Aarhus (2007), URL `http://arxiv.org/abs/0709.0289`.

[33] N. Johnston, private communication.

[34] J. E. Humphreys and J. E. Humphreys, *Linear algebraic groups*, vol. 430 (Springer New York, 1975).

[35] S. Wehner and A. Winter, New J. Phys. **12**, 025009 (2010).

[36] D. Deutsch, Physical Review Letters **50**, 631 (1983).

[37] I. Bialynicki-Birula and L. Rudnicki, *Entropic uncertainty relations in quantum physics*, e-print arXiv:1001.4668 [quant-ph].

[38] See [35] for a historical review, and [36, 37] for reasons why standard deviation is an inadequate uncertainty measure.

[39] See Cor. 1 for a precise definition of $H(X|B)$.

[40] Except in applications involving transmission over quantum channels [15, 24], where information is a more a natural quantity than uncertainty.

[41] This is the largest possible dependence on d, since $H(X|B) + H(Z|B) \leqslant 2 \log d$.

[42] Lemma 6 was proved in a collaborative discussion with M. Tomamichel, and approval to publish it in this paper was granted by M. Tomamichel.