Relative entropy derivation of the uncertainty principle with quantum side information

Patrick J. Coles,\textsuperscript{1} Li Yu,\textsuperscript{1} and Michael Zwolak\textsuperscript{2}

\textsuperscript{1}Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA
\textsuperscript{2}Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545

We give a simple proof of the uncertainty principle with quantum side information, as in [Berta et al. Nature Physics 6, 659 (2010)], invoking the monotonicity of the relative entropy. Our proof shows that the entropic uncertainty principle can be viewed as a data-processing inequality, a special case of the notion that information cannot increase due to evolution in time. This leads to a systematic method for finding the minimum uncertainty states of various entropic uncertainty relations; interestingly such states are intimately connected with the reversibility of time evolution.

PACS numbers: 03.67.-a, 03.67.Hk

Classical information theory, pioneered by Shannon \cite{Shannon1948}, addresses the question of how information storage, processing, and transmission tasks can be performed with macroscopic, decohered resources. The more general question of what can be done with resources that may or may not be decohered is the subject of quantum information theory. All of Shannon’s quantitative tools, such as entropy and mutual information, apply perfectly well in the quantum domain provided that one focuses on a single type of information \cite{Coles2007}, associated with a particular measurement on the quantum system of interest. What makes quantum information theory different from its classical counterpart is precisely the existence of multiple types of information or properties of a quantum system, e.g. the x and z components of an electron’s spin, and the notion that these properties can be incompatible in the sense that one cannot simultaneously know both types of information. This purely quantum idea is captured quantitatively in the uncertainty principle.

Formulations of the uncertainty principle have become progressively stronger over the years. Variances \cite{Bertlmann2006} have been replaced by entropies \cite{Coles2009} as measures of uncertainty, and recent formulations allow the observer to possess background or “side” information about the quantum observables, i.e. either classical \cite{Berta2010} or quantum \cite{Coles2009} side information. These latter formulations for two bases, and their generalization to two POVMs \cite{Berta2010, Coles2009} and to smooth entropies \cite{Coles2009} \cite{Coles2009}, represent the strongest versions of the uncertainty principle for two observables to date.

Thusfar, the uncertainty principle with quantum side information (UPQSI) in terms of Shannon entropies has only been proven as a corollary to a similar formulation in terms of smooth entropies \cite{Coles2009} \cite{Coles2009}, so there is the question as to whether the machinery of smooth entropies is necessary to understand the UPQSI. While smooth entropies have operational meanings \cite{Coles2009} and show great promise for quantum cryptography \cite{Coles2009}, one still yearns for the intuition behind the UPQSI. In this article, we derive the UPQSI using the properties of the relative entropy, which plays a central role in quantum information theory \cite{Holevo2011} and is, thus, familiar to many researchers in this field. In particular, we find that the UPQSI is connected to the fact that the relative entropy, which roughly acts like a distance between two density operators, does not increase over time, a principle called the monotonicity of the relative entropy. This approach allows us to generalize the UPQSI to a state-dependent bound, which strengthens it when the measurement(s) are complementary to one’s prior knowledge of the state.

This approach also leads us to a systematic method for answering the question: when is the uncertainty principle satisfied with equality? Such states are called minimum uncertainty states (MUS). Squeezed states of the harmonic oscillator, which have application in high-sensitivity interferometry and gravity-wave detection \cite{Bertlmann2006} \cite{Coles2009}, are well-known MUS of the variance uncertainty relation \cite{Bertlmann2006}. Very little is known about the MUS of entropic uncertainty relations, though see \cite{Coles2009} \cite{Coles2009}. Knowledge of such MUS may help in optimizing recently proposed applications of the UPQSI to entanglement witnessing and quantum cryptography \cite{Coles2009}. In this article we find necessary and sufficient conditions for a state to be a MUS, for several entropic uncertainty relations.

\textit{Conditional entropy.} The uncertainty or missing information about a POVM $P_a = \{P_{a,j}\}$ on system a is given by Shannon’s entropy of the associated probability distribution $\{p_j\}$: $H(P_a) = H(\{p_j\}) = -\sum_j p_j \log p_j$. Classical side information, e.g. given by a POVM $Q_b$ on system b, only reduces one’s uncertainty about $P_a$:

$$H(P_a|Q_b) = H(P_a) - H(P_a : Q_b) \leq H(P_a)$$  \hspace{1cm} (1)

where $H(P_a : Q_b) = H(P_a) + H(Q_b) - H(P_a, Q_b)$ is the mutual information. A quantum analog:

$$H(P_a|b) := H(P_a) - \chi(P_a, b),$$  \hspace{1cm} (2)

comes from replacing $H(P_a : Q_b)$ in (1) with the Holevo quantity $\chi(P_a, b) = S(\rho_b) - \sum_j p_j S(\rho_{b,j})$, where $\rho_b = \text{Tr}_a(\rho_{ab})$, $\rho_{b,j} = \frac{\text{Tr}_a(P_{a,j} \rho_{ab})}{p_j}$, $\rho_{ab}$ is the quantum state of $ab$, and $S(\rho) = -\text{Tr}(\rho \log \rho)$ is von Neumann’s entropy. By Holevo’s bound $H(P_a|b) \leq H(P_a|Q_b)$, and
by analogy to $H(P_a; Q_b)$ we say that $\chi(P_a, b)$ measures the “quantum side information” about $P_a$ located in system $b$. Also, $H(P_a|b) \geq 0$ and equals zero iff $b$ perfectly contains the $P_a$ information. Henceforth we drop the $a$ subscript from $P_a$. We note that another quantum analog of (11) is $S(a|b) = S(\rho_{ab}) - S(\rho_b)$, which can be negative for entangled $\rho_{ab}$.

Uncertainty relation for bases. The UPQSI strongly constrains the possible correlations in a tripartite state $\rho_{abc}$, stating that if $b$ knows something about an observable of $a$, then $c$ cannot know too much about about a complementary observable of $a$. The proof of the UPQSI is simplest for two bases $v = \{|v_j\rangle\}$ and $w = \{|w_k\rangle\}$ of $H_a$ that are mutually unbiased bases (MUBs): $|\langle v_j | w_k \rangle|^2 = 1/d$ for all $j, k$, where $d$ is the dimension of $H_a$. We wish to show that:

$$H(v|c) + H(w|b) \geq \log d,$$

(3)

noting that the proof for pure $\rho_{abc}$ immediately implies the proof for mixed $\rho_{abc}$ by the concavity of conditional entropy (p. 520 of [14]). We exploit the connection, proved in [9], between the conditional entropy and relative entropy $S(\rho||\sigma) = \text{Tr}(\rho \log \rho) - \text{Tr}(\rho \log \sigma)$: let $\rho_{abc}$ be a pure state, then

$$H(v|c) = S(\rho_{ab}||\sum_j |v_j\rangle \langle v_j|\rho_{ab}|v_j\rangle),$$

(4)

where we use the notation $|\psi\rangle := |\psi\rangle\langle\psi|$ for a rank-1 projector. It follows that

$$H(v|c) = S(\rho_{ab}||\sum_j |v_j\rangle \langle v_j|\rho_{ab}|v_j\rangle) \geq S(\sum_k |w_k\rangle \langle w_k|\rho_{ab}||) \log d + S(\sum_k |w_k\rangle \langle w_k|\rho_{ab}||) = S(\sum_k |w_k\rangle \langle w_k|\rho_{ab}||).$$

(5)

FIG. 1: The UPQSI states that the relative entropy after the $w$ measurement is never larger than that just before it.

We follow the same strategy as for MUBs, from [10]:

$$H(v|c) + H(w|b) \geq -\log r(v, w),$$

(11)

one can directly prove (11) starting with $H(v) = S(|\psi\rangle\langle\psi|) = \sum_j |\langle v_j | w_k \rangle|^2$ in [16] with $r(v, w)$ makes the overall operator larger.

While it is clear that (10) implies the well-known uncertainty relation of Maassen and Uffink [19]:

$$H(v) + H(w) \geq -\log r(v, w),$$

(12)

here we used the fact [13] that $S(\rho||\sigma) = S(\rho||\tau)$ if $\tau \geq \sigma$; i.e. replacing each $|\langle v_j | w_k \rangle|^2$ in (10) with $r(v, w)$ makes the overall operator larger.

State-dependent bound. More generally the UPQSI can be written for two POVMs $P$ and $Q$ [3, 9]:

$$H(P|b) + H(Q|c) \geq -\log r(P, Q).$$

(12)

where $r(P, Q) = \max_{j,k} \|\sqrt{P_j} \sqrt{Q_k}\|_\infty$ and $\| \cdot \|_\infty$ denotes the supremum norm (the maximum singular value). One can even formulate an UPQSI for a single POVM [9]:

$$H(P|b) \geq -\log \max_j \|P_j\|_\infty.$$  

(13)

Here we generalize (12) and (13), replacing the right-hand-sides with (possibly) state-dependent bounds [21].

**Theorem 1.** Let $P = \{P_j\}$ and $Q = \{Q_k\}$ be arbitrary POVMs, and let $\Pi$ be any projector on $a$ that projects onto a space that contains the support of $\rho_a$, then

$$H(P|b) + H(Q|c) \geq -\log r(P, Q; \Pi),$$

(14)

where $r(P, Q; \Pi) = \max_{j,k} \|\sqrt{Q_k} \Pi \sqrt{P_j}\|_\infty^2$ and each $H(\cdot|\cdot)$ term is bounded by, e.g.

$$H(P|b) \geq -\log \max_j \|P_j\|_\infty^2.$$  

(15)

\square
Let $\Pi = I_a$ reduces \((14)\) and \((15)\) to \((12)\) and \((16)\). In many cases choosing a $\Pi$ with a lower rank than $I_a$ in \((14)\) leads to a stronger bound (examples below), though this is not a general rule. On the other hand, the strongest bound in \((15)\) always results from choosing $\Pi$ to have the smallest possible rank, i.e. the projector onto the support of $\rho_a$ (see \((25)\)). We remark that all the results in \((1)\) hold if one replaces $r(P, Q)$ with $r(P, Q; \Pi)$. For example, if $P$ is any POVM on $a$ and $N$ is a rank-1 POVM on $a$, then

$$H(P|b) + H(N|b) \geq - \log r(P, N; \Pi) + S(a|b),$$

which is obtained from \((14)\) applied to pure $\rho_{abc}$ by adding $H(N|b) - H(N|c) = S(a|b)$ (see \((3)\)) to both sides.

**Example 1.** Let $\rho_a = [\psi]$ be unbiased w.r.t. both the $v$ and $w$ bases. Then $- \log r(v, w; [\psi]) = 2 \log d$, which is precisely stronger than the bound $- \log r(v, w) \leq \log d$. Likewise, \((15)\) gives $H(v|b) \geq \log d$ whereas \((13)\) gives $H(v|b) \geq 0$. A similar strengthening of the bounds occurs if $\rho_a$ is approximately unbiased w.r.t. $v$ and $w$. Thus the state-dependent bounds account for the complementarity between the state and the measurement(s) of interest.

**Example 2.** Consider a qutrit ($d = 3$) with bases $v = \{0, 1, 2\}$ and $w = \{0, 1 + 2, 1 - 2\}$. Since $r(v, w) = 1$, \((10)\) gives a trivial bound. But if $\rho_a$ lives only in the space spanned by $|1\rangle$ and $|2\rangle$, then set $\Pi = [1]+[2]$, and obtain: $H(v|b) + H(w|c) \geq \log 2$. This reveals the hidden complementarity between $v$ and $w$.

**Minimum uncertainty states.** Because the UPQSI is intimately connected to the monotonicity of the relative entropy, states that satisfy the former with equality are precisely states that satisfy the latter with equality. Petz showed \((20), (21)\) that $S(\rho||\sigma) = S(\mathcal{E}(\rho)||\mathcal{E}(\sigma))$ if and only if there exists a quantum channel $\mathcal{E}$ that undoes the action of $\mathcal{E}$ on $\rho$ and $\sigma$:

$$\mathcal{E}\rho = \rho, \quad \mathcal{E}\sigma = \sigma.$$

The construction given for this is \((21)\):

$$\mathcal{E}(\rho) = \sqrt{\sigma} \mathcal{E}^\dagger(\mathcal{E}(\sigma)^{-1/2}\rho \mathcal{E}(\sigma)^{-1/2})\sqrt{\sigma},$$

which automatically satisfies $\mathcal{E}\rho = \rho$, so one just needs to solve $\mathcal{E}\sigma = \sigma$. We take this approach to finding the MUS for particular uncertainty relations.

In what follows we consider a special pair of MUBs, the $x$ and $z$ bases, which are related by the Fourier transform:

$$|z_k\rangle = \sum_j \frac{\omega^{jk}}{\sqrt{d}} |x_j\rangle, \quad |x_j\rangle = \sum_k \frac{\omega^{-jk}}{\sqrt{d}} |z_k\rangle,$$

where $\omega = e^{2\pi i/d}$. Consider the uncertainty relations $\mathcal{E} = \mathcal{E}^\dagger$:

$$H(x) + H(z) \geq \log d, \quad H(x) + H(z) \geq \log d + S(\rho_a),$$

which are shown in order of increasing generality: \((21)\) becomes \((20)\) for unipartite pure states, \((22)\) becomes \((21)\) for bipartite product states $\rho_ab = \rho_a \otimes \rho_b$, and \((23)\) becomes \((22)\) for states with $\chi(x, b) = 0$. We have found all MUS associated with \((20), (21),\) and \((22)\), and we discuss the MUS for \((23)\).

**Theorem 2.** Let $d$ be prime, then

(i) A state $\rho_a$ is a MUS of \((20)\) if and only if it is (pure and) a basis state from either the $z$ or $x$ basis.

(ii) A state $\rho_a$ is a MUS of \((21)\) if and only if it is diagonal in either the $x$ or $z$ basis.

(iii) A state $\rho_{ab}$ is a MUS of \((22)\) if and only if $\rho_{ab} = \sum_k [z_k]_\rho [z_k]_\sigma$ or $\rho_{ab} = \sum_j [x_j]_\rho [x_j]_\sigma \otimes \rho_b$. \(\square\)

As a corollary to Theorem 2 we have found the MUSs of the uncertainty relation \((1)\) for a qubit $(d = 2)$:

$$H(x) + H(y) + H(z) \geq 2 \log 2 + S(\rho_a),$$

where $x, y,$ and $z$ are any complete set of three MUBs of the qubit.

**Corollary 3.** A state $\rho_a$ is a MUS of \((24)\) if and only if it is diagonal in either the $x$, $y$, or $z$ basis. \(\square\)

We now generalize Theorem 2 to arbitrary $d$, letting $\{s_a\}_{a=1}^d$ be the set of all factors of $d$, e.g. $\{1, 2, 4\}$ for $d = 4$. It is helpful to introduce the states:

$$|w^n_{\beta,\gamma}\rangle = \sum_{\alpha=0}^{s_a-1} \frac{\omega^{-\alpha d/s_a}}{\sqrt{s_a}} |z_{\beta+n d/s_a}\rangle,$$

where $\alpha = 1, ..., \eta; \beta = 0, ..., (d/s_a) - 1; \gamma = 0, ..., s_a - 1$. For a fixed $\alpha$, the set of $|w^n_{\beta,\gamma}\rangle$ with different $\beta, \gamma$ form an orthonormal basis, denoted the $w^\alpha$ basis. It is sometimes helpful to think of $w^\alpha$ as a tensor product of the $z$ and $x$ bases respectively on subsystems $a_1$ and $a_2$ of dimension $d/s_a$ and $s_a$, i.e. $|w^n_{\beta,\gamma}\rangle = |z_{\beta}\rangle_{a_1} |x_{\gamma}\rangle_{a_2}$. It will also be useful to introduce $p_j = \text{Tr}(|x_j\rangle \langle x_j| \rho_a)$, $q_k = \text{Tr}(|z_k\rangle \langle z_k| \rho_a)$, $\sigma_{\alpha,\beta} = \text{Tr}_a(|x_j\rangle \langle x_j| \rho_{ab})$, and $\sigma'_{\alpha,\beta} = \text{Tr}_a(|z_k\rangle \langle z_k| \rho_{ab})$.

**Theorem 4.** Let $d$ be arbitrary (with $\eta$ factors), then

(i) A state $\rho_a$ is a MUS of \((20)\) if and only if it is one of the pure states $|w^n_{\beta,\gamma}\rangle$ given in \((25)\), i.e. a basis state from one of the $w^\alpha$ bases.

(ii) The MUS of \((21)\) are diagonal in one of the $w^\alpha$ bases, with further constraints on the diagonal elements as follows. Let $\rho^n_{ab}$ denote the general solution that is diagonal in the $w^\alpha$ basis, then $\rho^n_{ab} = d \sum_{\beta,\gamma} p_{\beta} |w^n_{\beta,\gamma}\rangle \langle w^n_{\beta,\gamma}|$.

(iii) The MUS of \((22)\) are $\rho^n_{ab} = d \sum_{\beta,\gamma} p_{\beta} |w^n_{\beta,\gamma}\rangle \langle w^n_{\beta,\gamma}| \otimes \sigma_{\alpha,\beta}$

Our approach should work for other MUBs as well. For example, the following result for tensor products of $x$ and of $z$ implies Theorem 1 by setting all but one $d_c$ to 1.
Theorem 5. Let $a$ consist of $\lambda$ subsystems with $d_1, \ldots, d_\nu, \ldots, d_\nu$ such that all $d_\nu$ are pairwise coprime, with $\{s_{\alpha,\nu}^{(\nu)}\}_{\alpha=1}^{\nu}$ the set of factors of $d_\nu$. Then the MUS of

$$H(\bigotimes_{\nu=1}^{\lambda} x_\nu) + H(\bigotimes_{\nu=1}^{\lambda} z_\nu|b) \geq \log d + S(a|b)$$  \hspace{1cm} (26)$$

have the form $\rho_{ab}^{\alpha} = d_{\nu} (\bigotimes_{\nu=1}^{\lambda} |w_{\nu}^{\alpha}\rangle_{\nu}^{\nu}) \otimes \sigma_{b,\alpha}^z$, where $\alpha = (a_1, \ldots, a_\lambda)$ and likewise for $\beta$ and $\gamma$, with $\beta_\nu = 0, \ldots, d_\nu/s_{\alpha,\nu} - 1$ and $\gamma_\nu = 0, \ldots, s_{\alpha,\nu} - 1$. \hfill \Box

MUS of (23). The MUS of (24) are tripartite pure states $\rho_{abc}$ that satisfy

$$H(x|b) + H(z|c) = H(x|c) + H(z|b) = \log d. \hspace{1cm} (27)$$

Let us denote with $\Xi$ the set of all states for which at least one of the four $H(\cdot|\cdot)$ terms in (27) is zero. Renes and Boileau \cite{renes2009} noted that all states in $\Xi$ satisfy (27) and remarked that it is an open question as to whether $\Xi$ are the only states that satisfy (27). Theorem 4 shows that there are other solutions in non-prime d, e.g. the states in (25) satisfy (27) with $H(z|b) = H(z|c) = H(z) = \log s_a$ and $H(x|b) = H(x|c) = H(x) = \log (d/s_a)$. Generally, instead of just four solutions (as in $\Xi$), one should consider $2\eta$ solutions that, for some $\alpha$, have either $H(w^{\alpha}|c) = 0$ or $H(w^{\alpha}|b) = 0$, with further constraints given in [25]: denote this set of MUS as $\Upsilon$, so $\Upsilon \supseteq \Xi$.

However, there is an entirely different class, $\Omega$, of states that satisfy (27). Consider the tripartite state with $0 < g < 1$: $|\psi\rangle_{abc} = \sqrt{g}|x_0\rangle|0\rangle|0\rangle + \sqrt{1-g}|z_k\rangle|1\rangle|1\rangle$, for which $\rho_{ab} = \rho_{ac} = g|x_0\rangle\langle x_0| + (1-g)|z_k\rangle\langle z_k| \otimes (1|1\rangle\langle 1|)$. Since $H(x|b) = H(x|c) = (1-g)\log d$ and $H(z|b) = H(z|c) = g\log d$, this is a solution to (27) that is not in $\Upsilon$. (Note that this sort of MUS works for arbitrary MUBs, not just $x$ and $z$.) More generally, $\Omega$ contains:

$$\rho_{ab} = \sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma}|w_{\beta,\gamma}^{\alpha}\rangle\langle w_{\beta,\gamma}^{\alpha}| \otimes \rho_{\alpha,\beta,\gamma}, \hspace{1cm} (28)$$

where the different $\rho_{\alpha,\beta,\gamma}$ are all orthogonal and $0 \leq g_{\alpha,\beta,\gamma} \leq 1$.

Finally, we believe there is a third class of MUS, $\Lambda$, that is neither in $\Upsilon$ nor $\Omega$. For example in $d=2$, any state of the form $|\psi\rangle_{abc} = (|0\rangle|\phi_b\rangle|\phi_c\rangle + |1\rangle|\phi_b\rangle|\phi_c\rangle)/\sqrt{2}$, where $|\phi_b\rangle, |\phi_c\rangle$ are arbitrary kets, has $H(z|b) = \log 2 - S(\rho_b)$ and $H(x|c) = S(\rho_b)$. The three classes are seen as distinct as follows: in $\Upsilon$, either $\rho_{ab}$ or $\rho_{ac}$ has zero discord; in $\Omega$, $\rho_{ab}$ and $\rho_{ac}$ are separable with non-zero discord; in $\Lambda$, $\rho_{ab}$ and $\rho_{ac}$ are entangled.

Berta et al. \cite{berta2009} outlined methods for using the UPQSI \cite{uhlmann2002} for witnessing entanglement and for quantum cryptography. For both applications, one essentially lowers the entanglement of $\rho_{ab}$ with, e.g. $-S(\rho_b|b) \geq \log d - H(x|b) - H(z|b)$, where Alice and Bob find upper-bounds: $H(x|b) \leq H(x|x)$ and $H(z|b) \leq H(z|z)$ by comparing their measurement results in the $x$ and $z$ bases on an unknown state $\rho_{ab}$. The MUS are precisely the states for which this method should work best (otherwise the bound on the entanglement would be loose); providing motivation for further studying MUS.

In summary, the entropic uncertainty principle can be viewed as a data-processing inequality, expressing the notion that information cannot increase in the process shown in Fig. 1. Finding minimum uncertainty states then maps onto the question of whether this process is reversible, or whether information is irreversibly lost.

We thank Robert Griffiths for helpful conversations. This research is supported by the Office of Naval Research and the U.S. Department of Energy through the LANL/LDRD Program.

\begin{thebibliography}{99}

[1] C. E. Shannon, Bell system technical journal 27 (1948).
[2] R. B. Griffiths, Phys. Rev. A 76, 062320 (2007).
[3] H. F. Robertson, Phys. Rev. 34, 163 (1929).
[4] S. Wehner and A. Winter, New Journal of Physics 12, 025009 (2010), 0907.3704.
[5] M. J. W. Hall, Phys. Rev. Lett. 74, 3307 (1995).
[6] J. M. Renes and J.-C. Boileau, Phys. Rev. Lett. 103, 020402 (2009).
[7] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nature Physics 6, 659 (2010).
[8] M. Tomamichel and R. Renner, Phys. Rev. Lett. 106, 110506 (2011).
[9] P. J. Coles, L. Yu, V. Gheorghiu, and R. B. Griffiths, Information theoretic treatment of tripartite systems and quantum channels, eprint arXiv:1006.4859v4 [quant-ph].
[10] R. Konig, R. Renner, and C. Schaffner, IEEE Trans. Inf. Theory 55, 4337 (2009).
[11] V. Vedral, Rev. Mod. Phys. 74, 197 (2002).
[12] R. S. Bondurant and J. H. Shapiro, Phys. Rev. D 30, 2548 (1984).
[13] D. F. Walls, Nature 306, 141 (1983).
[14] J. M. Renes, Proc. R. Soc. A (2010).
[15] I. Devetak and A. Winter, Phys. Rev. A 68, 042301 (2003).
[16] M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information (Cambridge University Press, Cambridge, 2000), 5th ed.
[17] R. B. Griffiths, Consistent Quantum Theory (Cambridge University Press, Cambridge, 2002).
[18] M. Oya and D. Petz, Quantum Entropy and Its Use (Springer, 1993), 1st ed.
[19] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. 60, 1103 (1988).
[20] D. Petz, Rev. Math. Phys. 15, 79 (2003).
[21] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Commun. Math. Phys. 246, 359 (2004), URL http://dx.doi.org/10.1007/s00220-004-1049-z.
[22] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. 88, 017901 (2001).
[23] R. A. Horn and C. R. Johnson, Matrix Analysis (Cambridge University Press, 1985).
[24] We remark that (4) is connected to the quantum discord (23) since $S(\rho_{ab}|b) = \sum \{\nu_j|\nu_j\rangle\langle \nu_j|\} \rho_{ab} |\nu_j\rangle\langle \nu_j|$ measures the “distance”
\end{thebibliography}
SUPPLEMENTAL MATERIAL

Here we give the proofs of Theorems 1 through 5 of the main manuscript, and also elaborate on the MUS of (23). (We preserve the numbering of the equations and theorems in the main manuscript, and add a prefix “S” to such objects appearing in this supplemental material.) Let us first state the following useful result, proved in [2], that relates the conditional entropy to the relative entropy. This will allow us to rewrite the UPQSI in terms of relative entropy.

**Lemma S1.** Let \( \Pi = \{\Pi_j\} \) be a projective decomposition of \( I_a \) and let \( P = \{P_j\} \) be a POVM on \( a \).

(i) Let \( \rho_{abc} \) be a pure state, then

\[
H(P|b) = S(\rho_{ac}|| \sum_j \Pi_j \rho_{ac} \Pi_j). \tag{S1}
\]

(ii) Let \( \rho_{abc} \) be any state, then

\[
H(P|b) \geq S(\rho_{ac}|| \sum_j P_j \rho_{ac} P_j). \tag{S2}
\]

\( \square \)

**Proof of Theorem 1**

First, consider the single-POVM UPQSI in (15). We remarked in the main manuscript that the strongest bound in (15) results from choosing \( \Pi \) to have the smallest possible rank, i.e. the projector onto the support of \( \rho_a \). One can see this by considering two projectors \( \Pi \) and \( \Pi' \) where the latter has a higher rank than the former and \( \Pi' = \Pi + \Phi \) where \( \Phi \) is also a projector, and note that \( G_j' = \sqrt{\Pi'} G_j \sqrt{\Pi} = \sqrt{\Pi} G_j \sqrt{\Pi} \). It follows [23] that the spectrum of \( G_j' \) weakly majorizes that of \( G_j \) and thus \( \|\Pi' \sqrt{\Pi_j}\|^2 \geq \|\Pi \sqrt{\Pi_j}\|^2 \). Now let us prove (15).

**Proof**. The important properties [11, 18] of \( S(||\cdot||) \) we use are:

\[
S(\rho||\sigma) \geq S(\mathcal{E}(\rho)||\mathcal{E}(\sigma)) \tag{S3}
\]

for any quantum channel \( \mathcal{E} \); and for positive operators \( \rho, \sigma, \tau \), if \( \tau \geq \sigma \), then

\[
S(\rho||\sigma) \geq S(\rho||\tau). \tag{S4}
\]

Let \( \lambda_{\text{max}}(A) \) denote the maximum eigenvalue of \( A \), let \( G_j = \sqrt{\Pi_j} G_j \sqrt{\Pi} \), note \( \lambda_{\text{max}}(G_j) = \| \Pi \sqrt{\Pi_j} \|^2 \), then from (S2):

\[
H(P|b) \geq S(\rho_{ac}|| \sum_j P_j \rho_{ac} P_j) \geq S(\rho_{ac}|| \sum_j \Pi P_j \rho_{ac} P_j \Pi) \geq S(\rho_{c}|| \sum_j \sum_j \lambda_{\text{max}}(G_j) \Pi a \{P_j \rho_{ac}\}) \geq S(\rho_{c}|| \max_j \lambda_{\text{max}}(G_j) \rho_{ac}) = - \log \lambda_{\text{max}}(G_j). \tag{S9}
\]

We invoked (S3) for (S9) with the channel \( \rho \rightarrow \Pi \rho \Pi + (I - \Pi)\rho (I - \Pi) \), and for (S6) with the channel \( \rho \rightarrow Tr_a \rho \).

Now we prove (14).

**Proof**. Let \( e \) be an auxiliary system that acts as a register for the \( Q \) measurement. Consider the quantum channel \( \mathcal{E}_Q : ab \rightarrow eb \) defined by \( \mathcal{E}_Q(\rho_{ab}) = \sum_k |e_k\rangle \otimes Tr_a(Q_k \rho_{ab}) \), where \( \{ |e_k\rangle \} \) is an orthonormal basis of \( e \). Also, define \( G_{jk} = \sqrt{P_j} \Pi Q_k \Pi \sqrt{P_j} \), and note \( G_{jk} \leq \lambda_{\text{max}}(G_{jk}) I_a \), and \( r(P, Q; \Pi) = \max_{j,k} \lambda_{\text{max}}(G_{jk}) \). Then, starting from (S6) (swapping labels \( b \) and \( c \)),

\[
H(P|c) \geq S(\rho_{ab}|| \sum_j \Pi P_j \rho_{ab} P_j \Pi) \geq S(\mathcal{E}_Q(\rho_{ab})|| \sum_j \mathcal{E}_Q(\Pi P_j \rho_{ab} P_j \Pi)) \tag{S10}
\]

\[
= S(\mathcal{E}_Q(\rho_{ab})|| \sum_j \mathcal{E}_Q(\Pi P_j \rho_{ab} P_j \Pi)) \geq S(\sum_k |e_k\rangle \otimes Tr_a(Q_k \rho_{ab})|| \sum_j |e_k\rangle \otimes Tr_a(G_{jk} \sqrt{P_j} \rho_{ab} \sqrt{P_j}) \geq S(\sum_k |e_k\rangle \otimes Tr_a(Q_k \rho_{ab})|| \sum_j \lambda_{\text{max}}(G_{jk} |e_k\rangle \otimes Tr_a(P_j \rho_{ab})) \geq S(\sum_k |e_k\rangle \otimes Tr_a(Q_k \rho_{ab})|| r(P, Q; \Pi) I_e \otimes \rho_b \) \tag{S13}
\]

\[
= - \log r(P, Q; \Pi) - H(Q|b). \tag{S15}
\]

We invoked (S3) for step (S11), (S4) for steps (S13) and (S14), and Eq. (11.58) of [16] for step (S15). \( \square \)
Proof of Theorem 2

Proof. This theorem can be viewed as a corollary to Theorem 5. Set $d$ to be prime, so that $\eta = 2$ and $\{s_\alpha\} = \{1, d\}$. For $s_\alpha = 1$, $w^\alpha$ is the $z$-basis, and for $s_\alpha = d$, $w^\alpha$ is the $x$-basis. Thus, part (i) of Theorem 4 clearly reduces to part (i) of Theorem 2. Part (ii) of Theorem 4 reduces to part (ii) of Theorem 2 since there are no constraints on the diagonal elements of $\rho_\alpha^\alpha$ for $s_\alpha$ equal to 1 or $d$. Likewise, setting $s_\alpha$ equal to 1 or $d$ in $\rho_\alpha^\alpha = d \sum_{\beta, \gamma} p_{\gamma} [w^\beta_{\beta, \gamma}] \otimes \sigma^\beta_{\beta, \gamma}$ gives the two solutions in part (iii) of Theorem 2.

Proof of Corollary 3

Proof. Define $\zeta := H(x) + H(y) + H(z) - 2 \log 2 - S(\rho_a)$. First, consider (possibly mixed) states $\rho_a$ in the $xy$ plane of the Bloch sphere; such states have $H(z) = \log 2$. For these states, $\zeta = 0$ if and only if $H(x) + H(y) = \log 2 + S(\rho_a)$. But from Theorem 2, this is true if and only if either $x$ or $y$ is the eigenbasis of $\rho_a$, i.e. the state lies on either the $x$ or $y$ axis of the Bloch sphere. Any other state in the $xy$ plane will strictly have $H(x) + H(y) > \log 2 + S(\rho_a)$. Now consider taking a vertical path in the Bloch sphere up from some point in the $xy$ plane. Such a path will never decrease the value of $\zeta$ (See Appendix F of [3]). Thus, the only states that could possibly satisfy $\zeta = 0$ are those in the $xz$ plane and the $yz$ plane. But we already know that the territory between the $x$ and $y$ axes in the $xy$ plane cannot have $\zeta = 0$, so by symmetry, the territory between the $x$ and $z$ axes in the $xz$ plane cannot have $\zeta = 0$, and likewise for the $yz$ plane. So the only states that satisfy $\zeta = 0$ are those along the $x$, $y$, and $z$ axes.

Proof of Theorem 4

Proof. Even though this is a corollary of Theorem 5, it is instructive to see the direct proof as it is simpler than that of Theorem 5. We discuss below that parts (i) and (ii) follow from part (iii).

(i) Clearly from (21) the only states that can satisfy (20) with equality are pure states $S(\rho_a) = 0$. Thus, the MUS of (20) are a subset of the MUS of (21), precisely the subset with $S(\rho_a) = 0$. Assuming part (ii) of this theorem is true, then the only states that can be MUS of (21) are diagonal in a $w^\alpha$ basis, and thus the only states that can be MUS of (20) are (pure) basis vectors from a $w^\alpha$ basis, and indeed it is easily verified that all such basis vectors are MUS of (20).

(ii) Likewise part (ii) follows from part (iii) of this theorem. The MUS of (21) are a subset of the MUS of (22), precisely the subset with $\rho_{ab} = \rho_a \otimes \rho_b$. Imposing this condition on $\rho_{ab}^\alpha = d \sum_{\beta, \gamma} p_{\gamma} [w^\beta_{\beta, \gamma}] \otimes \sigma^\beta_{\beta, \gamma}$ and tracing over $b$ gives $\rho_{\alpha}^\alpha = d \sum_{\beta, \gamma} p_{\gamma} [w^\beta_{\beta, \gamma}] \otimes \sigma^\beta_{\beta, \alpha}$. (It turns out we did not need to impose the condition $\rho_{ab} = \rho_a \otimes \rho_b$ since all MUS of (22) have a $\rho_{\alpha}^\alpha$ of this form.)

(iii) It remains only to prove part (iii). Using (17) and (18) with $\rho = \rho_{ab}$, $\sigma = \sum_j [x_j] \rho_{ab} [x_j]$, $\mathcal{E}(\cdot) = \sum_k [z_k] \mathcal{E}^1(\cdot)$, gives:

$$\rho_{ab} = \sum_{j,j', k} \omega^{(j-j')k} [x_j] [x_{j'}] \otimes \sqrt{\sigma_{b,j}^\alpha \sigma_{b,k}^\beta} \sigma_{b,k+\mu}^\beta \sigma_{b,j}^\alpha \sigma_{b,j}^\alpha.$$  

Now specializing to $\chi(x, b) = 0$, meaning $\sigma_{b,j}^\alpha = p_j \rho_b$ for each $j$, becomes:

$$\rho_{ab} = \sum_{j, j'} \sqrt{p_j p_{j'}} [x_j] [x_{j'}] \otimes \mathcal{T}_{a} (Z^{j-j'} \rho_{ab}),$$

where $Z = \sum_k \omega^k [z_k]$. Computing $\mathcal{T}_{a} (Z^\mu \rho_{ab})$ from (16) for $\mu = 1, \ldots, d - 1$, one arrives at a system of equations (one for each $\mu$):

$$f_{\mu}(z) g_{\mu}(x) = 0,$$

where

$$f_{\mu}(z) := \mathcal{T}_{a} (Z^\mu \rho_{ab}) = \sum_k \omega^k \sigma_{b,k}^\alpha,$$

$$g_{\mu}(x) := 1 - \sum_j \sqrt{p_j p_{j+\mu}}.$$  

One can show that $g_{\mu}(x) = 0$ if and only if $p_{j} = p_{j+\mu}$ for all $j, m \in \mathbb{Z}_d$, as follows. Using the method of Lagrange multipliers, the Lagrangian is $L = 1 - \sum_j \sqrt{p_j p_{j+\mu}} + \lambda (1 - \sum_j p_j)$. Taking $\partial L / \partial p_{j+\mu} = 0$ gives $-2 \lambda \sqrt{p_j p_{j+\mu}} = \lambda (1 - \sum_j p_j)$, and summing this over all $k$ shows that $\lambda = -1$. Thus rearranging: $\sqrt{p_k - \mu p_{k-\mu}} = \sqrt{p_{k+\mu} - \mu p_{k+\mu}}$, which must also equal $\sqrt{p_{k+2\mu} - \mu p_{k+2\mu}}$, etc. Since each stepwise difference is the same and doing $d$ steps brings us back to the same point ($p_{k} = p_{k+m\mu}$), it must be that $p_k = p_{k+\mu}$ for all $m = 0, \ldots, d - 1$.

Now note that $g_{\mu}(x)$ is 0 implies that $g_{m\mu}(x) = 0$. This fact implies that there are $\eta$ and only $\eta$ different ways to set some $g_{\mu}(x)$ terms to zero, each way corresponding to setting $g_{\mu}(x) = g_{m\mu}(x) = 0$, thus $p_j = p_{j+m\mu}$ for $m = 0, \ldots, d - 1$, and $g_{\mu}(x) \neq 0$ for $m \neq m\mu$. Thus we have $\eta$ solutions where the $\alpha$-th solution, denoted $\rho_{\alpha}^\alpha$, has the properties that $p_j = p_{j+m\alpha}$, and $\sigma_{b,k}^\alpha = \sigma_{b,k+m\alpha}^\alpha$. Now we rewrite the $\rho_{ab}$ in (17), letting $j = \alpha + m\alpha$, $j' = \alpha' + m'\alpha$, $\beta = \beta + nd\alpha$, with...
0 \leq \gamma, \gamma', n \leq s\alpha - 1 and 0 \leq \beta, m, m' \leq d/s\alpha - 1, giving:
\[
\rho_{ab} = \sum_{\gamma, \gamma', m, m', \beta, n} \omega((\beta + nd/s\alpha)(\gamma - \gamma' + ms\alpha - m's\alpha), \chi) \gamma, \gamma', m, m', \beta, n
\]
\[
\sqrt{P_{\gamma + ms\alpha} P_{\gamma' + m's\alpha}} |x_{\gamma + ms\alpha}(x_{\gamma' + m's\alpha})| \otimes \sigma_{b, \beta + nd/s\alpha}^\gamma.
\] (S20)

So for the \( \alpha \)-th solution this reduces to:
\[
\rho_{ab}^\alpha = \sum_{\gamma, \gamma', m, m', \beta, n} \omega((\beta + nd/s\alpha)(\gamma - \gamma' + ms\alpha - m's\alpha), \chi) \gamma, \gamma', m, m', \beta, n
\]
\[
\sqrt{P_{\gamma + ms\alpha} P_{\gamma' + m's\alpha}} |x_{\gamma + ms\alpha}(x_{\gamma' + m's\alpha})| \otimes \sigma_{b, \beta}^\gamma.
\] (S21)

The sum over \( n \) gives a \( \delta_{\gamma, \gamma'} \) and we arrive at:
\[
\rho_{ab}^\alpha = \sum_{\gamma, m, m', \beta} \omega^{(ms\alpha - m's\alpha)}(x_{\gamma + ms\alpha}, x_{\gamma + m's\alpha}) \otimes \sigma_{b, \beta}^\gamma.
\] (S22)

Using \( \sqrt{d} |w_{\beta, \gamma}^{|s\alpha|} = \sqrt{\omega^{(s\alpha)} s\alpha = \sum_m \omega^{(ms\alpha)} |x_{\gamma + ms\alpha}| , \) we arrive at \( \rho_{ab}^\alpha = d \sum_{\beta, \gamma} p_{\gamma} |w_{\beta, \gamma}^{|s\alpha|} \otimes \sigma_{b, \beta}^\gamma. \)

\[\square\]

**Proof of Theorem 5**

**Proof.** This proof mirrors that of Theorem 4, except now we use a vector notation for all quantities, e.g. \( j = (j_1, ..., j_\alpha) \) and \( \bar{\mu} = (\mu_1, ..., \mu_\alpha) \), where each component refers to a particular subsystem. From (17) and (18), the MUS of (26) are:
\[
\rho_{ab} = \sum_j \sqrt{p_j} |j\rangle \langle j\rangle |x_{\alpha}(x_{\beta})| \otimes \text{Tr}_a \{ \sum_{l=1} Z_{\nu}^{ab} \rho_{ab} \},
\] (S23)

where \( Z_{\nu} = \sum_{k_\nu} \omega^{k_\nu} |z_{k_\nu}\rangle \) and \( \omega_{\nu} = e^{2\pi i / d\nu} \). Now let \( \mu_\nu = 0, ..., d\nu - 1 \), compute \( \text{Tr}_a \{ \sum_{l=1} Z_{\nu}^{ab} \rho_{ab} \} \) and using \( \text{Tr}_a \{ Z_{\nu}^{ab} |x_{\alpha}(x_{\beta})| \} = \delta_{j_\alpha, j_\beta + \mu_\nu} \) arrive at a system of equations:
\[
f_{\bar{\mu}}(z) g_{\bar{\mu}}(x) = 0
\] (S24)

where
\[
f_{\bar{\mu}}(z) := \text{Tr}_a \{ \sum_{l=1} Z_{\nu}^{ab} \} = \sum_{l=1} \omega^{l \nu} \sigma_{b, \bar{\mu} + l}^z
\]
\[
g_{\bar{\mu}}(x) := 1 - \sum_j \sqrt{p_j} |\bar{\mu} + j\rangle.
\] (S25)

Consider the following rules. Rule (1): \( g_{\bar{\mu}}(x) = 0 \) if and only if \( p_j = z_{\bar{\mu}, j} \) for all \( j \). This implies the following rules. Rule (2): If \( g_{\bar{\mu}}(x) = 0 \) then \( g_{m\bar{\mu}}(x) = 0 \) for all \( m = 0, ..., d\nu - 1 \), where \( m\bar{\mu} = \bar{\mu} + \mu + ... \) (m times).

Rule (3): If the \( d\nu \) are pairwise coprime and if \( g_{\bar{\mu}}(x) = 0 \) then \( g_{m\bar{\mu}}(x) = 0 \) for all \( m = (m_1, ..., m_\lambda) \), where \( m_\nu = 0, ..., d\nu - 1 \) and \( m\bar{\mu} = (m_1 \mu_1, ..., m_\lambda \mu_\lambda) \).

Rule (1) follows by the method of Lagrange multipliers, as in the proof of Theorem 4. Rule (2) follows from Rule (1) in a straightforward way. Rule (3) follows from Rule (2) by the Chinese Remainder Theorem, which implies that the ring \( Z_{d\nu} \) is isomorphic to the ring \( Z_{d_1} \times ... \times Z_{d_\lambda} \). The bijection relating \( m \in Z_{d\nu} \) to \( \bar{m} \in Z_{d_1} \times ... \times Z_{d_\lambda} \) is defined through \( m_\nu = (m \mod d_\nu) \). By this bijection and the ring isomorphism, \( (g_{m\bar{\mu}}(x); m = 0, ..., d\nu - 1) = (g_{\bar{\mu}}(x); m_\nu = 0, ..., d\nu - 1) \), and so Rule (3) follows.

From the above rules and letting \( \{ s_{\nu}^{(\alpha)} \}_{\alpha=1}^{7} \) be the set of factors of \( d\nu \), there are only \( \prod_{\nu=1}^{\lambda} r\nu \) different ways to set some of the \( g_{\bar{\mu}}(x) \) terms to zero, one for each \( \alpha \). The way corresponding to a particular \( \alpha \) involves setting \( g_{\bar{\mu}}(x) = g_{m\bar{\mu}}(x) = 0, \forall \bar{m} \) where \( \bar{s}_\alpha = (s_\alpha^{(1)}, ..., s_\alpha^{(\lambda)}) \), and \( g_{\bar{\mu}}(x) \neq 0 \) for \( \bar{\mu} \neq \bar{m} s_\alpha \). Of course, to solve \( (S24) \) we must set \( p_{\bar{\mu}}(z) = 0 \) for \( \bar{\mu} \neq \bar{m} s_\alpha \). From the latter condition, it follows that \( \sigma_{b, \bar{k}} = \sigma_{b, \bar{k} + \bar{m} s_\alpha} \), \forall \bar{m} \), and from Rule (1), we have \( p_{\bar{j}} = \sum_{j} p_{\bar{j} + \bar{m} s_\alpha} \). Plugging these two conditions into \( (S23) \), make the variable changes (like in the proof of Theorem 4)
\[
\bar{j} = \bar{j} + \bar{m} s_\alpha, \bar{j}' = \bar{j}' + \bar{m} s_\alpha, \text{ and } \bar{k} = \bar{k} + \bar{m} s_\alpha, \text{ then sum over } \bar{m} \text{ to get } \delta_{\bar{j}, \bar{j}'} \text{ and change the } z_{\nu} \text{ bases to the } w_{\alpha \nu} \text{ bases to arrive at } \rho_{ab}^\alpha = d \sum_{\beta, \gamma} p_{\gamma} (\sum_{\nu=1}^{\lambda} |w_{\beta, \gamma}^{(\alpha \nu)} \rangle \langle w_{\beta, \gamma}^{(\alpha \nu)}| \otimes \sigma_{b, \beta}^\gamma.
\]

\[\square\]

**MUS of (23)**

Here we discuss different classes of MUS of (23). We remind that reader that discord is a measure of the non-classicality of bipartite correlations. All of our discussion will refer to the one-way discord, as originally defined in [22], that is asymmetric under interchanging the two systems; in particular, the discord that uses projectors on system \( a \).

Generally, any bipartite state can be classified as either zero-discord (ZD), separable with non-zero discord (SNZD), or entangled (E) [22]. We shall classify MUS of (23) by classifying the reduced density operators \( \rho_{ab} \) and \( \rho_{ac} \) of the tripartite pure state \( \rho_{abc} \) into one of these three categories, i.e. by giving an ordered pair of form \( \rho_{ab} \) (ZD category, \rho_{ac} \) (SNZD category), for example (ZD, E) means \( \rho_{ab} \) is ZD and \( \rho_{ac} \) is E. Naively this would give \( 3 \times 3 = 9 \) possible ordered pairs, but if \( \rho_{ab} \) is ZD then \( \rho_{ac} \) cannot be SNZD, and vice-versa. (The proof for this is as follows: If \( \rho_{ab} \) is ZD, then there exists a basis \( w \) for which \( H(w|c) = 0 \). In turn, if \( H(w|b) = 0 \) then \( \rho_{ac} \) is ZD, otherwise if \( H(w|b) > 0 \) then \( H(w|c) - H(w|b) = S(a|c) < 0 \) implying that \( \rho_{ac} \) is E. So the only possibilities are for \( \rho_{ac} \) to be ZD or E, it cannot be SNZD.) So there are only seven possible ordered pairs, and all seven are physically possible.

Below we find three classes of MUS of (23): one class denoted \( \Lambda \) for which both \( \rho_{ab} \) and \( \rho_{ac} \) are E, so (E,E); one
These represent the $2\eta$ states for which either $H$ is the number of factors of $d$, e.g. $\eta = 3$ for $d = 4$ described in the main manuscript that compose the set $\Upsilon$. Setting $s_\alpha = 1$ or $s_\alpha = d$ in $\text{(S27)}$ and $\text{(S28)}$ shows that $\Upsilon$ contains all states for which either $H(z|c)$, $H(x|b)$, $H(z|c)$, $H(x|b)$ equals zero, and so $\Upsilon$ contains the set $\Xi$ defined in the main manuscript.

Let us consider a second class $\Omega$ of MUS of the form:

$$\rho_{ab} = \sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma} [w_{\beta,\gamma}^a \otimes \sigma_{\alpha,\beta,\gamma}] ,$$

where the different $\rho_{ab}$ are all orthogonal to each other and $0 \leq g_{\alpha,\beta,\gamma} \leq 1$. For these states $S(a|b) = 0$, $H(z|b) = H(z|c) = \sum g_{\alpha,\beta,\gamma} H(z|w_{\beta,\gamma}^a) = \sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma} \log s_{\alpha}$, and $H(x|b) = H(x|c) = \sum g_{\alpha,\beta,\gamma} H(x|w_{\beta,\gamma}^a) = \sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma} \log (d/s_{\alpha})$. So they satisfy $\text{(27)}$ since $\sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma} = 1$. Also, one can show (with a Schmidt decomposition across the $ab/c$ cut) that if $\rho_{ab}$ is given by $\text{(S29)}$, then $\rho_{ac}$ has the same form:

$$\rho_{ac} = \sum_{\alpha,\beta,\gamma} g_{\alpha,\beta,\gamma} [w_{\beta,\gamma}^a \otimes \sigma_{\alpha,\beta,\gamma}],$$

where the different $\sigma_{\alpha,\beta,\gamma}$ are all orthogonal to each other. Thus, both $\rho_{ab}$ and $\rho_{ac}$ are separable, and as long as more than one $w^a$ basis appears in the sums in $\text{(S29)}$ and $\text{(S30)}$, then they both have non-zero discord.

Finally, the main manuscript gives an example for $d = 2$ of MUS that are neither in $\Upsilon$ nor in $\Omega$. The tripartite state:

$$|\psi\rangle_{abc} = (|0\rangle|\phi_b\rangle|\phi_c\rangle + |1\rangle|\varphi_b\rangle|\varphi_c\rangle)/\sqrt{2}$$

where $|\phi_b\rangle$, $|\phi_c\rangle$, $|\varphi_b\rangle$, $|\varphi_c\rangle$ are arbitrary kets with $\langle\phi_b|\varphi_b\rangle$, $\langle\phi_c|\varphi_c\rangle \in \mathbb{R}$, satisfies $\text{(27)}$ with $H(z|b) = \log 2 - S(\rho_b)$, $H(z|c) = \log 2 - S(\rho_c)$, $H(x|b) = S(\rho_b)$, and $H(x|c) = S(\rho_c)$. Likewise, replacing the $z$ states $\{|0\rangle, |1\rangle\}$ in $\text{(S31)}$ with the $x$ states $\{|+, |-\rangle\}$, the tripartite state:

$$|\psi\rangle_{abc} = (|+\rangle|\phi_b\rangle|\phi_c\rangle + |-\rangle|\varphi_b\rangle|\varphi_c\rangle)/\sqrt{2}$$

satisfies $\text{(27)}$ with $H(x|b) = \log 2 - S(\rho_b)$, $H(x|c) = \log 2 - S(\rho_c)$, and $H(x|c) = S(\rho_b)$. Except for the extreme cases where $S(\rho_b) = S(\rho_c) = 0$ or $\log 2$, the states described by $\text{(S31)}$ and $\text{(S32)}$ are clearly not in $\Upsilon$, and the fact that they are not in $\Upsilon$ follows from $S(b|a) = -S(b|c) < 0$ and $S(c|a) = -S(c|b) < 0$, implying that both $\rho_{ab}$ and $\rho_{ac}$ are entangled, in contrast to the separable states in $\Omega$. There is reason to believe that there are MUS for $d > 2$ of a similar nature to the qubit examples given here (with both $\rho_{ab}$ and $\rho_{ac}$ entangled), as we have found such MUS for $d = 3$. For example:

$$|\psi\rangle_{abc} = (|z_0\rangle|0\rangle|0\rangle + |z_1\rangle|+\rangle|+) + |z_2\rangle|y\rangle|y\rangle)/\sqrt{3},$$

where $b$ and $c$ are qubits and $|y\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$, has $H(z|b) = H(z|c) = \log 3 - S(\rho_b)$ and $H(x|b) = H(x|c) = S(\rho_b)$.