Entropic uncertainty relations for general SIC-POVMs and MUMs

Shan Huang,† Zeng-Bing Chen,† and Shengjun Wu‡

†Department of Physics, Nanjing University, Nanjing 210093, China
‡Institute for Brain Sciences and Kuang Yaming Honors School, Nanjing University, Nanjing 210023, China

(Dated: November 3, 2020)

We construct inequalities between Rényi-α entropy and the indexes of coincidence of probability distributions, based on which we obtain improved state-dependent entropic uncertainty relations for general symmetric informationally complete positive operator-valued measures (SIC-POVM) and mutually unbiased measurements (MUM) on finite dimensional systems. We show that our uncertainty relations for general SIC-POVMs and MUMs can be tight for sufficiently mixed states, moreover, comparisons to the numerically optimal results are made via information diagrams.

PACS numbers:

I. INTRODUCTION

Incompatible observables like momentum and position along the same direction can’t be both be measured with certainty on any quantum system, and the more certain of an observable generally implies the more uncertain of observables incompatible with it, though contrary to general cognition to the physical world based on macroscopic experience, this is a fundamental element to quantum mechanics leading to inherent unpredictability or uncertainty about outcomes of incompatible measurements.

Heisenberg was the first to realize this kind of unpredictability of quantum mechanical world by stating the famous uncertainty principle [1], where he proposed the first uncertainty relation in terms of product of standard deviations of momentum and position along the same direction

\[ \Delta P \Delta Q \geq \frac{\hbar}{2}, \]  

where \( \hbar \) is the reduced Planck constant. Robertson generalized it to arbitrary two observables [2]

\[ \Delta X \Delta Y \geq \frac{1}{2} |\langle \psi | [X,Y] | \psi \rangle|, \]  

where \( \Delta X \) and \( \Delta Y \) denote the standard deviations of \( X \) and \( Y \) respectively when measured on the state \( | \psi \rangle \), and \( [X,Y] \) denotes the commutator between \( X \) and \( Y \).

Inequalities like above successfully captured some nonclassical characters of quantum measurements and have a far-reaching influence on people’s understanding of quantum mechanics. However, the standard deviation way of expressing uncertainty sometimes can be quite strange and counterintuitive [3–5], for example, as pointed out by Deutsch [3], the right hand side of Eq. (2) can be trivially zero even for incompatible observables and standard deviations are also variant under simply relabeling of measurement outcomes, which violates our intuitive requirement for uncertainty from the perspective of quantum information theory.

On the other hand, the sum of entropies of probability distributions induced by different measurements is found to be a more universal and effective measure of uncertainty [3–8]. Following Deutsch [3] and Kraus [9], Maassen and Uffink proved the famous state-independent uncertainty relation for two non-degenerate observables in terms of Shannon entropy [10]. Since then entropic uncertainty relations (EURs) for multiple mutually unbiased bases have been largely investigated [11–16], and generalizations to mutually unbiased measurements (MUM) [17] as well as symmetric informationally complete positive operator-valued measures (SIC-POVM) [18] in terms of Rényi entropy [19] are also explored [20–22]. In two recent works entropic uncertainty relations are constructed from quantum designs for the first time [23, 24].

Entropic way of measuring uncertainty have many applications in quantum information theory such as quantum randomness [25, 26], quantum cryptography [26, 27], entanglement witnessing [28, 29], etc. (See more in the review [30] and references there in).

This article is structured as follows. In Sec. II we introduce some notations that will be used throughout this article as well as necessary concept concerning entropy and generalized measurement, especially, SIC-POVM and MUM. In Sec. III we propose entropic uncertainty relation for general SIC-POVMs, based on which in Sec. IV uncertainty relations for MUMs are constructed. In Sec V, we propose a conjecture and draw a brief conclusion.

II. PRELIMINARIES

A positive-operator-valued-measure (POVM) \( \mathcal{P} \) on a \( d \)-dimensional Hilbert space \( \mathcal{H}_d \) is a set of positive semi-definite operators that sum up to identity: \( \mathcal{P} = \{P_i | P_i \geq 0, \sum_i P_i = 1_d \} \), where \( 1_d \) denotes the identity. The probability distribution of performing \( \mathcal{P} \) on a quantum state \( \rho \) is \( \mathcal{P} = \{p_i = \text{Tr}(P_i \rho) \} \) is the probability of obtaining the \( i \)-th result. The index of coincidence of \( \mathcal{P} \) is denoted by \( IC(\mathcal{P}) = \sum_i p_i^2 \), and the sum of indices of coincidence of probability distributions induced by a finite set \( \mathcal{P} \) of POVMs is denoted by \( IC(\mathcal{P}) = \sum_{m=1} IC(\mathcal{P}^m | \rho) \). Following [31], we call ranges of the map \( IC(\mathcal{P}) \)-entropy information dia-
grams.

The Shannon entropy of $\mathcal{P}$ is defined by $H(\mathcal{P}|\rho) = -\sum_{i=1}^{d} p_i \log_2 p_i$. Rényi generalized it to a family of entropies [19]

$$H_\alpha(\mathcal{P}|\rho) = \frac{1}{1-\alpha} \log_2 \left( \sum_{i=1}^{d} p_i^\alpha \right), \quad (\alpha > 0, \alpha \neq 1)$$

which returns to Shannon entropy in the limitation $\lim_{\alpha \to 1} H_\alpha(\mathcal{P}|\rho) = H(\mathcal{P}|\rho)$.

A. Mutually unbiased measurements

We say two orthonormal bases $\{|b_i^1\rangle\}$ and $\{|b_j^2\rangle\}$ ($1 \leq i, j \leq d$) in $\mathcal{H}_d$ are mutually unbiased bases (MUBs) [32-35] if the inner product between their basis vectors satisfy $\langle b_i^1 | b_j^2 \rangle = \frac{1}{d} \quad (\forall i, j \leq d)$. MUBs can be equivalently understood under POVM notation, let $B_{i}^{\rho}$ denote the rank-1 projector $|b_i^\rho\rangle\langle b_i^\rho|$, then the Hilbert-Schmidt product between elements in $B^1 = \{ b_i^1 \}$ and that in $B^2 = \{ b_j^2 \}$ is $\text{Tr}(B_i^1 B_j^2) = \frac{1}{d}$.

In $\mathcal{H}_d$, one can find at least 3 MUBs for any $d \geq 2$ and at most $d + 1$ MUBs, which is called a complete set of MUBs. A complete set of MUBs exists if $d$ is power of a prime number, while it is still an open set of MUBs. A complete set of MUBs exists if $d$ and at most $d + 1$ MUBs, which is called a complete set of MUBs. A complete set of MUBs exists if $d$ is power of a prime number.

According to [10], for the set of MUBs in $\mathcal{H}_d$

$$IC(B|\rho) \leq C(B|\rho) = \text{Tr}(\rho^2) + \frac{|B| - 1}{d}$$

Mutually unbiased measurements (MUM) [17], introduced as generalization of MUBs, are POVMs that each containing $d$ elements $P^m = \{ P_i \}$ ($1 \leq i \leq d$) and satisfy

$$\text{Tr}(P_i^m) = 1$$
$$\text{Tr}(P_i^m P_j^{m'}) = \delta_{ij} \delta_{mm'}$$

and

$$+ (1 - \delta_{ij}) \delta_{mm'} \frac{1 - \kappa}{d - 1} + (1 - \delta_{mm'}) \frac{1}{d}$$

where $\kappa (\frac{1}{2} < \kappa \leq 1)$ is called the efficiency parameter.

MUBs can be viewed as MUMs with $\kappa = 1$. A complete set of $d + 1$ MUMs exists for any $d$ with proper efficiency parameter [17].

For any set of MUMs on $\mathcal{H}_d$ there is [17, 22]

$$IC(P|\rho) \leq C(P|\rho)$$

$$= \frac{\|P\|^2}{d} + \frac{\kappa d - 1}{d(d-1)} \left[ d \text{Tr}(\rho^2) - 1 \right],$$

and if $P$ is complete

$$IC(P|\rho) = \frac{d + 1}{d} + \frac{\kappa d - 1}{d(d-1)} \left[ d \text{Tr}(\rho^2) - 1 \right].$$

B. Symmetric informationally complete POVM

A POVM on $\mathcal{H}_d$ is said to be symmetric informationally complete (SIC-POVM) [18] if it consists of

d rank-1 operators $S = \{ S_i \}$ such that $\text{Tr}(S_i S_j) = \frac{\delta_{ij} + 1}{d(d+1)}$. From the geometric point of view, with $S_i = \frac{1}{d} |\phi_i\rangle\langle \phi_i|$, SIC-POVM is comprised of $d^2$ subnormalized equiangular vectors $\{ \frac{1}{\sqrt{d}} |\phi_i\rangle \}$ in $\mathbb{C}^d$ as $|\langle \phi_i | \phi_j \rangle|^2 = \frac{\delta_{ij} + 1}{d}$ and $\sum_{i=1}^{d^2} |\langle \phi_i | \phi_j \rangle|^2 = 1.$

Although it is still an ongoing research to prove or disprove the existence of SIC-POVM for general $d$, analytic and numerical results confirmed its existence for dimensions up to 67 [27].

By generalizing the method proposed in [16], Rastegin obtained [20]

$$IC(S|\rho) = \sum_{i=1}^{d^2} p_i^2 = \frac{1 + \text{Tr}(\rho^2)}{d(d + 1)},$$

where $p_i = \text{Tr}(\rho S_i)$.

Generalizations of SIC-POVM to that with elements of any rank are explored in Refs. [38, 39]. In Ref. [40] the authors proved the existence of general SIC-POVMs in all dimensions by giving the explicit construction. The general SIC-POVM $S_g = \{ S_i \}$ ($i = 1, 2, \ldots, d^2$) is defined by

$$\text{Tr}(S_i S_j) = a_i, (\forall i, 1/d^3 \leq a_i \leq 1/d^2)$$
$$\text{Tr}(S_i S_j) = \frac{1 - ad}{d(d^2 - 1)} \quad (\forall i \neq j).$$

It is shown in Ref. [21] that

$$IC(S_g|\rho) = \frac{(ad^3 - 1)\text{Tr}(\rho^2) + d(1 - ad)}{d(d^2 - 1)}$$

III. UNCERTAINTY RELATION FOR GENERAL SIC-POVMS

Let $S^d_c$ denote the set of $d$-dimensional probability distributions with the same index of coincidence:

$$S^d_c = \{ \mathcal{P} | \text{size}(\mathcal{P}) = d, \ IC(\mathcal{P}) = c \} .$$

We show Lemma 1 and Theorem 1 in Appendix A.

Lemma 1. For $H_\alpha(\mathcal{P})$ to attain local extreme value in $S^d_c$ with any $c \in [\frac{1}{2}, 1]$, the nonzero probabilities of $\mathcal{P}$ arranged in descending order must be in the form $(p_0, \ldots, p_{a}, p_{b}, \ldots, p_a)$, here $p_0 \geq p_a$.

We can parameterize this form of distribution with three parameters: $N$, the number of nonzero probabilities; $N_a$, the number of probabilities equal to $p_a$; $c$, the index of coincidence. And we represent it formally as

$$\mathcal{P}[c, N, N_a] = \left\{ N_a \otimes p_a, (N - N_a) \otimes p_b \right\},$$

where $N_a \otimes p_a$ means there are $N_a$ probabilities being $p_a$. Combined with the condition that $N_a p_a + (N - N_a) p_b = 1$ and $IC(\mathcal{P}[c, N, N_a]) = c$, we have $p_a = \frac{1 + \sqrt{(Nc - 1)(N - N_a)c}}{N - N_a}$ and $p_b = \frac{1 - \sqrt{(Nc - 1)(N - N_a)c}}{N - N_a}$.

Note that for any probability distribution in the form $\mathcal{P}[c, N, N - 1]$ there is $N = \lceil \frac{1}{c} \rceil$. The following notation will be used throughout the rest of this article

$$\mathcal{P}_2[c] = \mathcal{P}[c, d, 1]; \quad \mathcal{P}_g[c] = \mathcal{P}[c, N, N - 1].$$
Theorem 1. \((2 - \alpha)H_\alpha(\mathcal{P})\) attains minimum value and maximum value in \(S_d^c\) respectively at \(\mathcal{P} = \mathcal{P}[c, N, N-1]\) and \(\mathcal{P} = \mathcal{P}[c, d, 1]\).

We stress that Theorem 1 is a generalization of the Shannon entropic bounds obtained earlier in Refs. [12 31] to Rényi entropy. Without loss of generality, information diagrams of Shannon entropy (\(\alpha = 1\)) and Rényi-5 entropy for \(d = 5\) are shown in Fig. 1 as examples.

The function \(H(\mathcal{P}[c])\) is differentiable with respect to \(c\) except when \(c \in \{\frac{1}{d}, \frac{1}{d-1}, \ldots, 1\}\), thus the points \((\frac{1}{d}, \log_2 k)\) \((2 \leq k \leq d - 1)\) divide the graph of \(H(\mathcal{P}[c])\) into \(d - 1\) sections and the number of nonzero probabilities of \(\mathcal{P}[c]\) is \(k\) for \(c \in \{\frac{1}{d}, \frac{1}{d-1}\}\). More detailed properties of \(H(\mathcal{P}[c])\) and \(H(\mathcal{P}[x][c])\) are discussed in Ref. [31].

Apply Theorem 1 to entropy for general SIC-POVM performed on \(d\)-dimensional systems, immediately

\[(2 - \alpha)H_\alpha(S_d[\rho]) \leq (2 - \alpha)H_\alpha(\mathcal{P}[x][IC(S_d[\rho])]),\]  
\[(2 - \alpha)H_\alpha(S_d[\rho]) \geq (2 - \alpha)H_\alpha(\mathcal{P}[x][IC(S_d[\rho])]).\]  

Where \(IC(S_d[\rho])\) is given by [7]. Now we show [11] and [12] is tight for \(tr(\rho^2) \in [\frac{1}{d}, \frac{d - 2 + nd^2}{(d-1)^2}]\) and \(tr(\rho^2) \in [\frac{1}{d}, d^2a]\) respectively. We only need to show the probability distributions \(\mathcal{P}[x][IC(S_d[\rho])]\) can be achieved by some matrix in the form \(\rho = \sum_i x_i S_i\), where \(\{x_i\}\) are real parameters determined by \(tr(\rho) = 1\) and \(tr(\rho S_i) = p_i \in \mathcal{P}[x][IC(S_d[\rho])]\).

For [11], we have \(x_1 = \cdots = x_{d-1} \geq x_d\), as \(\sum_i S_i = 1\), then \(\forall \phi \in H_d, \langle \phi | \rho | \phi \rangle = \langle \phi | \sum_i x_i S_i | \phi \rangle = x_1 + (x_d - x_1)tr(S_2 | \phi | \phi \rangle) \geq x_1 + (x_d - x_1)/d \geq 0\), thus \(\rho\) is a density matrix. As for [12], when \(t \leq d^2 a\) we have \(x_1 \geq x_2 = \cdots = x_d \geq 0\), obviously \(\rho \geq 0\).

It’s not a surprise to see that our entropic lower bound for SIC-POVM, as shown in Fig. 2, is not tight when \(tr(\rho^2) \geq \frac{1}{d}\) for \(d = 3\) as shown in [11] and [12] are based on Eq. [7] only, but interestingly, the tight bound agrees with \(H(\mathcal{P}[c, 8, 6])\).

FIG. 1. Information diagrams of Shannon entropy (See also in Refs. [12 31]) and Rényi-5 entropy as well as the corresponding upper bound (UB) and lower bound (LB) on entropy.

The Shannon entropy case of Lemma 2 has already been proved in Ref. [31] via a different method. Similar to [12], from Lemma 2 we have for the set \(\mathcal{P}\) of MUMs with efficient parameter \(\kappa\) and when \(tr(\rho^2) \in [\frac{1}{d}, \frac{d - 2 + nd^2}{(d-1)^2}]\), the tight lower bound on Rényi-\(\alpha\) \((0 < \alpha \leq 1)\) entropy is

\[(|\mathcal{P}|-1) \log_2 d + H_\alpha(\mathcal{P}[\rho](C(\rho)|\rho) - (|\mathcal{P}|-1)/d).\]  

For Shannon entropy, [13] can be generalized further.

Theorem 2. The sum of Shannon entropies of \(M\) \(d\)-dimensional probability distributions under the restriction \(\sum_{m=1}^{M} IC(\mathcal{P}^m) = c (\frac{M}{d} \leq c \leq M)\) is minimum when the probability distributions are \(\{\mathcal{P}_m[\rho]\}_{\mathcal{P}}\) where 

\[n = \lfloor \frac{M}{c} \rfloor, k = \lfloor \frac{M - d - 1}{n+1} \rfloor, M - k - 1 \otimes \mathcal{P}_m[\rho], m \in [\{\mathcal{P}_m[\rho]\}, X \otimes \mathcal{P}_m[\rho]]\].

Here \(n = \lfloor \frac{M}{c} \rfloor\), \(k = \lfloor \frac{M - d - 1}{n+1} \rfloor\) and \(X \otimes \mathcal{P}_m[\rho]\) is shorthand for \(X\) copies of \(\mathcal{P}_m[\rho]\).

Despite the complex expression, this theorem can be understood in a simple way as is discussed in Appendix C, and we have for the set \(\mathcal{P}\) of MUMs

\[\sum_{m=1}^{M} H(\mathcal{P}_m[\rho]) \geq H(\mathcal{P}[c_0]) + k \log_2 n + (|\mathcal{P}|-k-1) \log_2(n+1),\]

where \(c_0 = c - \frac{k}{n}\) and \(c = C(\mathcal{P}[\rho])\) (defined by Eq. [1]).
We can linearize the first term on the right hand side of Eq. (14) based on its convexity with respect to \( \alpha \) as follows: 
\[
H(\mathcal{P}_x[c]) \geq H(\mathcal{P}_x[1]) + \left[H(\mathcal{P}_x[1]) - H(\mathcal{P}_x[1/2])\right]
\]
where
\[
H(\mathcal{P}_x[1]) = -\sum_{i=1}^{d} p_i \log_2 p_i
\]
and
\[
H(\mathcal{P}_x[1/2]) = -\sum_{i=1}^{d} \frac{1}{2} \log_2 \frac{1}{2}
\]
with
\[
p_i = \frac{1}{d}
\]
and
\[
p_{i/2} = \frac{1}{2d}
\]
which would then reduce to the result of Wu et al. [16] for the set \( \mathcal{B} \) of MUBs
\[
\sum_{m=1}^{\lceil \frac{\log N}{d} \rceil} H(\mathcal{B}^m|\rho) \geq \left| \mathcal{B} \right| - nC(\mathcal{B}|\rho) \left( n + 1 \right) \log_2(n + 1) - \left| \mathcal{B} \right| - (n + 1)C(\mathcal{B}|\rho)n \log_2 n,
\]
where \( n = \left\lfloor \frac{\left| \mathcal{B} \right|}{C(\mathcal{B}|\rho)} \right\rfloor \). For MUBs, Eq. (14) is equivalent to Eq. (15) when \( \alpha = \frac{1}{n} \) or \( \frac{1}{n+1} \), while more generally it is improved than Eq. (15).

The following are good approximations of tight upper bound on Shannon entropy for complete MUBs, at least in the first several dimensions, when \( \alpha(\rho^2) \approx 1 \) and \( \alpha(\rho^3) \approx \frac{1}{3} \) respectively
\[
(d + 1)H(\mathcal{P}_x[1]C(\mathcal{B}|\rho)/(d + 1)) = \frac{d \log_2 d + H(\mathcal{P}_x[1]C(\mathcal{B}|\rho) - 1)}{d}.
\]
Eq. (14) is not tight when \( Tr(\rho^2) > \frac{1}{2} \) for \( d = 3 \), and the tight bound for complete MUBs, according to Fig. 3 agrees with
\[
H(\mathcal{B}|\rho) \geq 1 + 3H(\mathcal{P}_y[1 + \frac{16}{3} Tr(\rho^2)])
\]
where \( c = IC(\mathcal{P}) \), \( p_a = 1 + \sqrt{(d-1)(d-1)} \) and \( p_b = 1 - \sqrt{(d-1)(d-1)} \).

Theorem 3. For a finite set of mutually unbiased measurements \( \mathcal{P} \) on \( \mathcal{H}_d \), we have
\[
\frac{1}{|\mathcal{P}|} \sum_{m=1}^{|\mathcal{P}|} H_\alpha(\mathcal{P}^m|\rho) \geq E_\alpha(\mathcal{P}_x[c]), \quad (\alpha \geq 2)
\]
where \( \bar{c} = \frac{1}{|\mathcal{P}|} IC(\mathcal{P}|\rho) \).

When \( \alpha > 2 \), Eq. (20) is improved than Rastegin’s lower bounds \( L_{Ras1} \) [20] and \( L_{Ras2} \) [24], and when \( \alpha = 2 \) they are all equivalent to \( H_2(\mathcal{P}_x[c]) \).

C. Entropy region

The entropies of performing a finite ordered set of generalized measurements \( \mathcal{P} = \{\mathcal{P}^m\} \) on a \( d \)-dimensional system described by \( \rho \) form an vector, the \( m \)-th element of which is \( H(\mathcal{P}^m|\rho) \). The region of all possible entropic vectors induced by \( \mathcal{P} \) is called the entropy region of \( \mathcal{P} \). The entropy region of a given measurement set contains much more information besides the entropic lower bound, and we expect it to be as meaningful in classical information theory as in the quantum counterpart.

We make a comparison here between the Shannon entropy region for 3 MUBs in \( \mathcal{H}_d \) and that of 3 \( d \)-dimensional probability distributions satisfying
\[
\frac{3}{d} \leq \sum_{m=1}^3 IC(\mathcal{P}^m) \leq \max_{\rho} \{IC(\mathcal{B}|\rho)\} = 1 + \frac{2}{d}
\]
As can be seen in Fig. 5, the entropy region of probability distributions satisfying Eq. (21) is the same to that for 3 MUBs when \( d = 2 \), while in higher dimensions clearer distinctions show up at places where the sum of entropies is relatively small, which is in accordance with the information diagrams.
V. CONJECTURE AND CONCLUSION

As a set of \( d + 1 \) MUMs are informationally complete, we can always find a set of real parameters \( \{ x^m \} \) such that an equality of our uncertainty relation for MUMs is attained at the probability distributions \( \{ P^m \} \) generated from \( \text{tr}(\rho' P^m_i) = p^m_i \in P^m \), where \( \rho' = \sum_{i,m} x^m_i P^m_i \) and \( P^m_i \) is the \( i \)-th element of the \( m \)-th MUM, and similarly for SIC-POVMs. However, \( \rho' \) may not be a density matrix as a density matrix is necessarily positive semi-definite. When our uncertainty relations for MUMs or SIC-POVMs are not tight, the tight bound can be attained only by those \( \rho' \) the minimum eigenvalue of which is 0.
On the other hand, as is shown in Figs. 3, 6, the tight lower Rényi entropic ($\alpha < 2$) bound curves for both complete MUBs and SIC-POVMs are non-differentiable at $\text{tr}(\rho^2) = \frac{1}{d}$ ($vk = 2, \cdots, d-1$), which divide the curves into $d-1$ sections. This implies different sections of the lower bound curve corresponds with density matrices with different numbers of zero-valued eigenvalues.

**Conjecture.** The tight lower bound on Shannon entropy for complete MUBs or SIC-POVMs can only be achieved by density matrices satisfying $(\lambda_1, \lambda_2, \cdots) = \mathcal{P}[\text{tr}(\rho^2)]$, where $(\lambda_i)$ are nonzero eigenvalues of $\rho$ and arranged in descending order.

Based on this conjecture, we have an equivalent form of Eq. (14) for MUBs when $\text{tr}(\rho^2) \leq \frac{1}{d-1}$

$$\sum_{m=1}^{[B]} H(P^m|\rho) \geq (|B| - 1) \log_2 d - \text{tr}[\rho \log_2 \rho],$$

which coincides with the uncertainty relation for two observables proposed by Berta et al. [41].

To conclude, in this work we firstly derive state-dependent bounds on Rényi entropy for a single measurement (Theorem 1), and further we propose lower bound on Shannon entropy for multiple generalized measurements (Theorem 2). We focus only on applications of Theorems 1 and 2 to general SIC-POVMs and MUMs respectively. We show our Rényi-$\alpha$ entropic bounds for general SIC-POVMs as well as lower bound on Shannon entropy for MUMs are tight for sufficiently mixed systems in all dimensions, especially, the lower bound for SIC-POVMs is always tight when $\alpha \geq 2$. Improved lower bound on Rényi-$\alpha$ entropy with $\alpha \geq 2$ for MUMs is also obtained, while it is almost never tight. At last, we propose a conjecture concerning the eigenvalues of density matrices when the tight state-dependent entropic bounds for MUBs and SIC-POVMs is attained.

**Acknowledgments**

This work is supported by the National Key R&D Program of China (Grant Nos. 2017YFA0303703 and 2016YFA0301801) and the National Natural Science Foundation of China (Grant No. 11475084).

---

**Appendix A: Proof of Theorem 1**

We show Lemma 1 first, Lagrangian multiplier method can be employed to find the necessary form of probability distributions at which Rényi-$\alpha$ entropy attains extreme values in $S^d_e$.

$$L = \sum_{i=1}^{d} p_i^\alpha + \lambda (\sum_{1 \leq i \leq d} p_i^2 - c) + \lambda_0 (\sum_{i=1}^{d} p_i - 1)$$

subject to $0 \leq p_i \leq p_b$, $N_a \geq N \geq 1$, $N_a N_a = 1$ and $N_a N_a = 1$.

Similar method can be utilized to prove that (A1) is also valid for Shannon entropy. Note here $\mathcal{P}[c_2, N, N_a]$ majorizes $\mathcal{P}[c_2, N, N_a]$ if $c_1 \geq c_2$, thus $H_a(\mathcal{P}[c, N, N_a])$ is a monotonic decreasing function of $c$.

To show Theorem 1, note that with $u, c, \alpha > 0$ and $N (N_a)$ being real constants, the solution to (A2) (namely, the value of $p_a, p_b$ and $N_a (N)$), if there exists, is **unique**.

$$\begin{cases} 0 \leq p_b \leq p_a, N \geq N_a \geq 1, \\ N_a p_a + (N - N_a) p_b = 1 \\ N_a p_a^2 + (N - N_a) p_b^2 = c \\ N_a p_a^3 + (N - N_a) p_b^3 = u. \end{cases} (A2)$$

As $H_a(\mathcal{P}[\frac{1}{N}, N, N_a]) = \log_2 N$ is independent of $N_a$ and lim$_{c \to 1} \frac{1}{N_a} H_a(\mathcal{P}[c, N, N_a]) = \log_2 N_a$ is independent of $N$, $H_a(\mathcal{P}[c, N, N_a])$ is monotonic of both $N$ and $N_a$.

\begin{equation} H_a(\mathcal{P}[\frac{1}{N} + s, N, N_a]) = \log_2 N - \frac{\alpha N s}{2\ln 2} + \frac{\alpha(\alpha - 2)N^\frac{3}{2} N_a s^2}{4\ln 2(N - N_a)} + o(s^2), \quad (0 < s \ll \frac{1}{N}) \end{equation} (A3)

more over, take (A3) into consideration we have

\begin{equation} (2 - \alpha)H_a(\mathcal{P}[c, N, 1]) \geq (2 - \alpha)H_a(\mathcal{P}[c, N, 2]) \geq \cdots \geq (2 - \alpha)H_a(\mathcal{P}[c, N, N - 1]) \end{equation} (A4)

\begin{equation} (2 - \alpha)H_a(\mathcal{P}[c, N_1, N_a]) \leq (2 - \alpha)H_a(\mathcal{P}[c, N_2, N_a]), \quad (N_a < N_1 \leq N_2) \end{equation} (A5)

$$\begin{align} \int \mathcal{P}[c, N, N_a] = \frac{N a_{\alpha}}{d(N - 1)!} a_{\alpha}^{2(N - 1)} + \frac{1}{d(N - 1)!} N a^{2(N - 1)} a_{\alpha} \end{align}$$
We can conclude from (A4) and (A5) that for any \( P \in S_d^n \left( \frac{1}{N} \leq c < \frac{1}{N-1} \right) \)

\[
(2 - \alpha)H_\alpha(P[c, d, 1]) \geq (2 - \alpha)H_\alpha(P) \geq (2 - \alpha)H_\alpha(P[c, N, N - 1])
\]

(A6)

This completes the proof of Theorem 1.

**Appendix B: Proof of Lemma 2**

Let’s reparameterize \( P[c, N_i, \theta] \) as \( P[c, \theta] \), where \( \theta = 2 \arccos \sqrt{N_c/N} \) and \( \theta \in [0, \pi) \). We have

\[
H_\alpha(P[c, N, \theta]) = \frac{1}{1 - \alpha} \log_2 \left[ N \cos^2 \left( \frac{\theta}{2} \right) \left( \frac{1 + \sqrt{N_c - 1} \tan \frac{\theta}{2}}{N} \right)^\alpha + N \sin^2 \left( \frac{\theta}{2} \right) \left( \frac{1 - \sqrt{N_c - 1} \cot \frac{\theta}{2}}{N} \right)^\alpha \right] \\
= \frac{1}{1 - \alpha} \log_2 M_\alpha(P[c, N, \theta])
\]

\[
(\alpha - 1) \frac{\partial^2}{\partial c^2} H_\alpha(P[c, \theta]) = f(\alpha, z, \theta) \frac{\alpha N^{\alpha+1} \sin \frac{\theta}{2} \cos^3 \frac{\theta}{2}}{4 \ln 2 (N_c - 1)^{3/2} M^2_\alpha(P[c, N, \theta])} \left( 1 + \sqrt{N_c - 1} \tan \frac{\theta}{2} \right)^{2\alpha - 2}
\]

here, with \( z = \frac{1 - \sqrt{N_c - 1} \cot \frac{\theta}{2}}{1 + \sqrt{N_c - 1} \tan \frac{\theta}{2}} \) (0 < \( z \leq \tan \frac{\theta}{2} \)), \( f(\alpha, z, \theta) \)

\[
= \frac{2 \tan^2 \frac{\theta}{2}}{1 + z \tan^2 \frac{\theta}{2}} (z^{\alpha-1} - 1)^2 + z^{\alpha-1} \left[ -\tan^2 \frac{\theta}{2} z^{\alpha-1} + z^{1-\alpha} + (\alpha - 1)(z \tan^2 \frac{\theta}{2} - 1) + (2 - \alpha)(\tan^2 \frac{\theta}{2} - 1) \right]
\]

when \( 0 < \alpha < 1 \) or \( \alpha \geq 2 \) and \( 0 < \tan \frac{\theta}{2} \leq 1 \)

\[
f(\alpha, z, \theta) \geq z^{\alpha-1} \left[ -z^{\alpha-1} + z^{1-\alpha} + (\alpha - 1)(z - 1) \right] \geq 0 \implies (\alpha - 1) \frac{\partial^2}{\partial c^2} H_\alpha(P[c, N, \theta]) \geq 0 \quad (B1)
\]

As for Shannon entropy, when \( 0 < z < \tan \frac{\theta}{2} \leq 1 \)

\[
\frac{\partial^2}{\partial c^2} H(P[c, \theta]) = \log_2(1 + z \tan \frac{\theta}{2}) - \log_2(1 - z \cot \frac{\theta}{2}) - \frac{1}{1 - z \cot \frac{\theta}{2}} + \frac{1}{1 + z \tan \frac{\theta}{2}} \leq 0.
\]

(B2)

Lemma 2 is equivalent to (B1) combined with (B2).

**Appendix C: Proof of Theorem 2**

Let \( g = \{ P^g \} \) denote the probability distributions at which \( \sum_{m=1}^M H(P^m) \) is minimum under the restriction

\[
\forall 1 \leq m \leq M, \dim(P^m) = d; \sum_{m=1}^M P^m = c (c \text{ is a constant, } c \in \left[ \frac{M}{d}, M \right]) \quad (C1)
\]

For simplicity, we use \( P^g[c] \) and \( P^g[c] \) instead of \( P[c, N, N - 1] \) and \( P[c, d, 1] \) respectively, then according to (A6) and (B2)

**Property 1.** \( P^g \) must be in the form \( P^g = P^g[c^g] \) for any \( g \).

**Property 2.** \( H(P^g[c^g]) \) is convex with respect to \( c^g \), thus at most one element in \( g \), \( P^k \) say, is not uniform in its nonzero part.

Moreover, as is required by (C3), with \( N_g \) denoting the number of nonzero probabilities of \( P^g \), \( g \) must satisfy

\[
\begin{align*}
1. \ max_{g \neq g'} \{ |N_g - N_{g'}| \} & \leq 1 \\
2. \ if \ N_k = \min_g \{ N_g \}, \ then \ \forall g, N_g - N_k = 0
\end{align*}
\]

(C2)

\[
\begin{align*}
1. \ H(P^g[1/n]) + H(P^g[1/m + s]) & > H(P^g[1/m]) + H(P^g[1/n + s]), \ 0 \leq s \leq 1/n/(n - 1) \\
2. \ H(P^g[1/n - s]) + H(P^g[1/m]) & > H(P^g[1/m - s]) + H(P^g[1/n]), \ 0 \leq s \leq 1/n/(n + 1)
\end{align*}
\]

(C3)
Here $1 \leq m < n$ ($n, m \in N^+$), and note that

$$IC(\mathcal{P}_y[1/n]) + IC(\mathcal{P}_y[1/m + s]) = IC(\mathcal{P}_y[1/n + s]) + IC(\mathcal{P}_y[1/m]).$$

$$IC(\mathcal{P}_y[1/n - s]) + IC(\mathcal{P}_y[1/m]) = IC(\mathcal{P}_y[1/m - s]) + IC(\mathcal{P}_y[1/n]).$$

To show the first inequality of (C3), we show $\log_2 N - H(\mathcal{P}[1/N + s, N, 0])$ is an increasing function of $N$

$$\log_2 N - H(\mathcal{P}[1/N + s, N, 0]) = \cos^2 \frac{\theta}{2} \left(1 + \sqrt{Ns} \tan \frac{\theta}{2}\right) \log_2 \left[ \cos^2 \frac{\theta}{2} \left(1 + \sqrt{Ns} \tan \frac{\theta}{2}\right) \right]$$

$$\sin^2 \frac{\theta}{2} \left(1 - \sqrt{Ns} \cot \frac{\theta}{2}\right) \log_2 \left[ \sin^2 \frac{\theta}{2} \left(1 - \sqrt{Ns} \cot \frac{\theta}{2}\right) \right] = h(s, N, \theta)$$

let $\theta_s(N) = 2 \arctan \frac{1}{\sqrt{N-1}}$, we have $h(s, N, \theta_s(N))$ is an increasing function of $N$ for

$$\frac{\partial}{\partial s} h(s, N, \theta_s(N)) > 0, \quad \frac{\partial}{\partial \theta} h(s, N, \theta) < 0, \quad \frac{\partial \theta_s(N)}{\partial N} < 0 \quad \frac{\partial \theta_s(N)}{\partial N} \theta = \theta_s(N) \geq 0$$

The second inequality of (C3) can be proved similarly. (C2) combined with Properties 1,2 is enough to determine $g$ (Theorem 2). It turns out that $g$ can be also obtained with the method of locally steepest descent. Consider $c = \frac{M}{d}$ (this is when probability distributions are all uniform) at the beginning and then let $c$ increase, according to Properties 1,2 and (A3) obviously the steepest descent of Shannon entropy is given by

$$\begin{align*}
\{(M - 1) \otimes \mathcal{P}_y[1/d], \mathcal{P}_y[c - \frac{M - 1}{d}]\}, & \quad \frac{M}{d} \leq c \leq \frac{M - 1}{d} + \frac{1}{d - 1} \\
\{(M - 2) \otimes \mathcal{P}_y[1/d], \mathcal{P}_y[c - \frac{M - 2}{d} - \frac{1}{d - 1}]\}, & \quad \frac{M - 1}{d} + \frac{1}{d - 1} \leq c \leq \frac{M - 2}{d} + \frac{2}{d - 1} \\
\ldots \ldots & \\
\{(M - 1) \otimes \mathcal{P}_y[1/d], \mathcal{P}_y[c - \frac{M - 1}{d} - 1]\}, & \quad \frac{M}{d} \leq c \leq \frac{M - 1}{d} + \frac{1}{d - 2} \\
\ldots \ldots &
\end{align*}$$

[1] W. Heisenberg, Z. Phys. 43, 172 (1927).
[2] H. P. Robertson, Phys. Rev. 34, 163 (1929).
[3] D. Deutsch, Phys. Rev. Lett. 50, 631 (1983).
[4] L. Dammeier, R. Schwonnek, and R. F. Werner, New J. Phys. 17, 093046 (2015).
[5] I. Bialynicki-Birula, and L. Rudnicki, 2011, in Statistical Complexity, edited by K. Sen (Springer Netherlands, Dordrecht), pp. 1-34.
[6] P. J. Coles, R. Colbeck, L. Yu, and M. Zwolak, Phys. Rev. Lett. 108, 210405 (2012).
[7] S. Friedland, V. Gheorghiu, and G. Gour, Phys. Rev. Lett. 111, 230401 (2013).
[8] J. B. M. Uffink and J. Hilgevoord, Found. Phys. 15, 925 (1985).
[9] K. Kraus, Phys. Rev. D 35 3070 (1987).
[10] H. Maassen and J. B. M. Uffink, Phys. Rev. Lett. 60, 1103 (1988).
[11] I. D. Ivanovic, J. Phys. A 25, L363 (1992).
[12] J. Sánchez, Phys. Lett. A 173, 233 (1993).
[13] J. Sánchez-Ruiz, Phys. Lett. A 201, 125 (1995).
[14] M. A. Ballester and S. Wehner, Phys. Rev. A 75, 022319 (2007).
[15] S. Wehner and A. Winter, New J. Phys. 12, 025009 (2010).
[16] S. Wu, S. Yu, and K. Molmer, Phys. Rev. A 79, 022104 (2009).
[17] A. Kalev and G. Gour, New J. Phys. 16, 053038 (2014).
[18] J. M. Renes, R. Blume-Kohout, A. J. Scott, and C. M. Caves, J. Math. Phys. 45 2171 (2004).
[19] A. Rényi, 1961, in Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1 (University of California Press, Berkeley, CA), pp. 547-561.
[20] A. E. Rastegin, Eur. Phys. J. D 67, 269 (2013).
[21] A. E. Rastegin, Phys. Scr. 89, 085101 (2014).
[22] B. Chen, and S. Fei, Quantum Inf. Process. 14, 2227-2238 (2015).
[23] A. Ketterer and O. Gühne, Phys. Rev. Research 2, 023130 (2020).
[24] A. E. Rastegin, J. Phys. A: Math. Theor., Vol. 53, 405301 (2020).
[25] G. Vallone, D. G. Marangon, M. Tomasin, and P. Villaro, Phys. Rev. A 90, 052327 (2014).
[26] R. König, S. Wehner, and J. Wullschleger, IEEE Trans. Inf. Theory 58, 1962 (2012).
[27] F. Dupuis, O. Fawzi, and S. Wehner, IEEE Trans. Inf. Theory 61, 1093 (2015).
[28] V. Giovannetti, Phys. Rev. A 70, 012102 (2004).
[29] O. Gühne, and M. Lewenstein, Phys. Rev. A 70, 022316 (2004).
[30] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, Rev. Mod. Phys. 89, 015002 (2017).
[31] P. Harremoës and F. Topsøe, IEEE Trans. Inf. Theory
[32] I. D. Ivanovic, J. Phys. A 14, 3241 (1981).
[33] W. K. Wootters and B. D. Fields, Ann. Phys. (N.Y.) 191, 363 (1989).
[34] A. Klappenecker and M. Rötteler, Finite Fields and Applications (Springer, Berlin-Heidelberg, 2004), pp. 137-144.
[35] A. O. Pittenger and M. H. Rubin, Linear Algebra. Appl. 390, 255 (2004).
[36] I. Bengtsson et al., J. Math. Phys. 48, 052106 (2007).
[37] A. J. Scott, M. Grassl, J. Math. Phys. 51, 042203 (2010).
[38] D. M. Appleby, Opt. Spectrosc. 103, 416-428 (2007).
[39] A. Kalev, J. Phys. A: Math. Theor. 47, 265301 (2014).
[40] G. Gour and A. Kalev, J. Phys. A: Math. Theor. 47, 335302 (2014).
[41] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, Nat. Phys. 6, 659 (2010).
[42] J. Sánchez-Ruiz, J. Phys. A 27, L843 (1994).
[43] P. Wocjan and T. Beth, Quantum Inf. Comput. 5, 93 (2005).