

Witnessing Negative Conditional Entropy

by

Mahathi Vempati, Nirman Ganguly, Indranil Chakrabarty, Arun K Pati

in

PHYSICAL REVIEW A

Report No: IIIT/TR/2021/-1



Centre for Security, Theory and Algorithms
International Institute of Information Technology
Hyderabad - 500 032, INDIA
July 2021

Witnessing Negative Conditional Entropy

Mahathi Vempati,^{1,*} Nirman Ganguly,^{2,†} Indranil Chakrabarty,^{3,4,‡} and Arun K Pati^{4,§}

¹*Center for Computational Natural Sciences and Bioinformatics,*

International Institute of Information Technology-Hyderabad, Gachibowli, Telangana-500032, India.

²*Department of Mathematics, Birla Institute of Technology and Science Pilani, Hyderabad Campus, Telangana-500078, India.*

³*Center for Security, Theory and Algorithmic Research,*

International Institute of Information Technology-Hyderabad, Gachibowli, Telangana-500032, India.

⁴*Quantum Information and Computation Group,*

Harish-Chandra Research Institute, HBNI, Allahabad 211019, India.

(Dated: July 1, 2021)

Quantum states that possess negative conditional von Neumann entropy provide quantum advantage in several information-theoretic protocols including superdense coding, state merging, distributed private randomness distillation and one-way entanglement distillation. While entanglement is an important resource, only a subset of entangled states have negative conditional von Neumann entropy. In this work, we characterize the class of density matrices having non-negative conditional von Neumann entropy as convex and compact. This allows us to prove the existence of a Hermitian operator (a witness) for the detection of states having negative conditional entropy for bipartite systems in arbitrary dimensions. We show two constructions of such witnesses. For one of the constructions, the expectation value of the witness in a state is an upper bound to the conditional entropy of the state. We pose the problem of obtaining a tight upper bound to the set of conditional entropies of states in which an operator gives the same expectation value. We solve this convex optimization problem numerically for a two qubit case and find that this enhances the usefulness of our witnesses. We also find that for a particular witness, the estimated tight upper bound matches the value of conditional entropy for Werner states. We explicate the utility of our work in the detection of useful states in several protocols.

I. INTRODUCTION

Entanglement is an indispensable resource for several information processing tasks [1]. However, not all entangled states qualify to be a resource in some information processing tasks. For example, under the standard teleportation scheme, only entangled states whose fully entangled fraction is above a threshold value are considered to be useful [2]. Therefore, even within the set of entangled states one needs to identify distinctive features from an operational perspective.

One such distinctive feature is given by conditional von Neumann entropy, also known as the quantum conditional entropy [3]. The quantum conditional entropy (henceforth referred to as conditional entropy when there is no ambiguity) for a quantum state ρ_{AB} is defined as

$$S_{A|B}(\rho_{AB}) = S(\rho_{AB}) - S(\rho_B), \quad (1)$$

where $S(\rho) = -\text{Tr}(\rho \log \rho)$ is the von Neumann entropy of the state. Unlike its classical counterpart, conditional entropy in the quantum realm can be negative [3] and thus can be exploited for quantum information tasks [3, 4]. Its operational interpretation is given in the context of quantum state merging [5, 6], where negative conditional entropy is an indication of resources for future

communication [5, 6]. Negative conditional entropy provides quantum advantage in superdense coding [7–9] and is a characteristic of states for which one-way entanglement distillation is possible [10]. It also helps in the reduction in the uncertainty in predicting the outcomes of two incompatible measurements [11] and in the maximization of the rates of distributed private randomness distillation [12].

The significance of negative conditional entropy has a wide-ranging impact beyond quantum information theory. In [13, 14], a modification of the Bekenstein-Hawking area law for a Schwarzschild black hole that resolves previously present paradoxes is shown, where simple entropy in the original law is replaced with the negative of conditional entropy of the purifying system of the black hole with respect to the positive-energy particles in the black hole. According to the modified law, when the change in the above-mentioned conditional entropy is negative, there is an increase in the surface area of the black hole. In the context of thermodynamics, it is shown that the work cost of erasing a system is proportional to the conditional entropy of the state of the system with respect to that of the observer [15]. In particular, when this conditional entropy is negative, the observer gains work when erasing the system, thereby cooling the environment rather than heating it up.

While all states possessing negative conditional entropy are entangled, the converse is not true. This fact, along with the vast range of applications like those mentioned above warrants that quantum states are characterized with respect to their conditional entropy. To this end, in [16], we have characterized states whose condi-

* mahathi.vempati@research.iiit.ac.in

† nirmanganguly@hyderabad.bits-pilani.ac.in

‡ indranil.chakrabarty@iiit.ac.in

§ akpati@hri.res.in

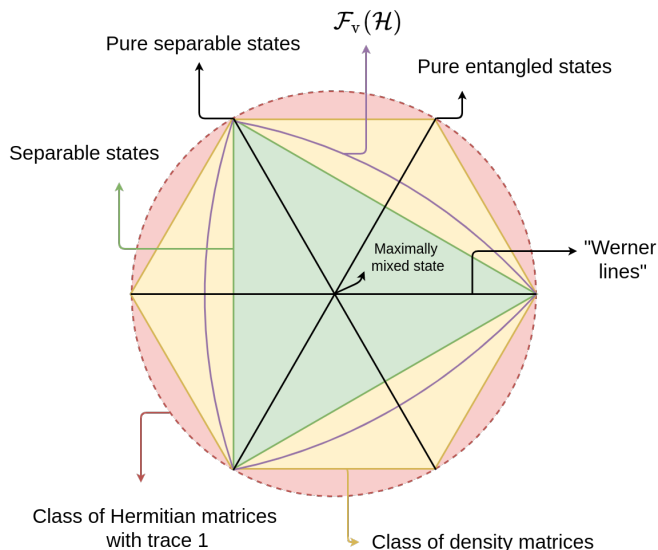


FIG. 1. Schematic diagram of $\mathcal{F}_v(\mathcal{H})$: The outer class is the set of all Hermitian matrices with Trace 1. It extends beyond the figure as indicated by the dotted lines. Within that is the class of density matrices. The pure states (entangled or separable) lie farthest from the maximally mixed state. The density matrix set is the set of all convex combinations of pure state projectors – but since there are infinite such projectors, it does not form a polytope, unlike what it may seem like from the figure. Within the density matrix space is the convex class of separable states. $\mathcal{F}_v(\mathcal{H})$ is also a convex class and it contains the class of separable states.

tional entropy remains non-negative even after the application of global unitary operations. Conditional entropy is also not a direct observable, and this calls for practical methods to detect states that possess this resource.

This paper is organized as follows: In Section II, we characterize the class of states with non-negative conditional entropy as convex and compact. This allows for the existence of witnesses for all states outside this class, and we construct such a witness. We also discuss examples of this witness for Werner and Isotropic states. In Section III, we prove that the expectation value of any of the witnesses in a state is an upper bound to the state's conditional entropy. We then pose the problem of obtaining a tight upper bound to the set of conditional entropies of states in which an operator gives the same expectation value. We solve this convex optimization problem numerically, and discuss several interesting results. In Section IV, we provide an alternative witness formulation exploiting geometric methods. In Section V, we discuss some applications of our work. In Section VI, we conclude.

II. EXISTENCE AND CONSTRUCTION OF A WITNESS FOR NEGATIVE CONDITIONAL ENTROPY STATES

A. Characterization of non-negative conditional entropy states

In this work, the set of all quantum states (density matrices) over a Hilbert space \mathcal{H} is denoted by $\mathcal{S}(\mathcal{H})$, and \mathcal{H} has the dimensionality $d \otimes d$. We refer to the class of states possessing non-negative conditional entropy as $\mathcal{F}_v(\mathcal{H})$. A schematic representation of $\mathcal{F}_v(\mathcal{H})$ is shown in FIG. 1. We now prove that $\mathcal{F}_v(\mathcal{H})$ is convex and compact.

Theorem 1. $\mathcal{F}_v(\mathcal{H})$ is convex and compact.

Proof. $\mathcal{F}_v(\mathcal{H})$ is convex— Let $\sigma_1, \sigma_2 \in \mathcal{F}_v(\mathcal{H})$, i.e., $S_{A|B}(\sigma_1), S_{A|B}(\sigma_2) \geq 0$. Let $\sigma = \lambda\sigma_1 + (1-\lambda)\sigma_2$, where $\lambda \in [0, 1]$. From the concavity of $S_{A|B}$ [17], we have $S_{A|B}(\sigma) \geq \lambda S_{A|B}(\sigma_1) + (1-\lambda) S_{A|B}(\sigma_2)$. Therefore, $S_{A|B}(\sigma) \geq 0$ and $\sigma \in \mathcal{F}_v(\mathcal{H})$, implying $\mathcal{F}_v(\mathcal{H})$ is convex.

$\mathcal{F}_v(\mathcal{H})$ is compact— For $\sigma_{AB} \in \mathcal{F}_v(\mathcal{H})$ we have $S_{A|B}(\sigma_{AB}) \geq 0$ by definition. The value 0 is attained for several states in $\mathcal{F}_v(\mathcal{H})$, for example, all pure separable states. From the subadditivity of $S_{A|B}$ [17], we have $S(\sigma_{AB}) \leq S(\sigma_A) + S(\sigma_B)$, implying $S_{A|B}(\sigma_{AB}) \leq S(\sigma_A) \leq \log_2 d$. This value is attained at I/d^2 . Thus, the image set of $\mathcal{F}_v(\mathcal{H})$ under $S_{A|B}$ is a closed set, i.e. $S_{A|B}(\mathcal{F}_v(\mathcal{H})) = [0, \log_2 d]$. As $S_{A|B}$ is a continuous function under trace norm [18, 19], we conclude that $\mathcal{F}_v(\mathcal{H})$ is also closed [20]. $\mathcal{F}_v(\mathcal{H})$ is bounded under trace norm as every density matrix has a bounded spectrum. For our finite dimensional Hilbert space, this implies that $\mathcal{F}_v(\mathcal{H})$ is compact. ■

By the Hahn-Banach theorem [20], this convexity and compactness of $\mathcal{F}_v(\mathcal{H})$ implies that for every state with negative conditional entropy, there exists a Hermitian operator that separates it from $\mathcal{F}_v(\mathcal{H})$ (known as a witness). We provide an explicit expression for a family of such witnesses below.

B. Witness to detect negative conditional entropy

To prove that a Hermitian operator W is a witness for states possessing negative conditional entropy, it is sufficient to prove (i) $\exists \rho \notin \mathcal{F}_v(\mathcal{H}) \mid \text{Tr}(W\rho) < 0$, and (ii) $\forall \sigma \in \mathcal{F}_v(\mathcal{H}), \text{Tr}(W\sigma) \geq 0$.

Theorem 2. A witness for any state $\rho_{AB} \notin \mathcal{F}_v(\mathcal{H})$ is $W_\rho = -\log(\rho_{AB}) + I \otimes \log(\rho_B)$.

Proof. W_ρ is Hermitian due to the Hermiticity of $\log(\rho_{AB})$ and $\log(\rho_B)$. To prove (i), we have $\text{Tr}(W_\rho \rho_{AB}) = S_{A|B}(\rho_{AB}) < 0$. To prove (ii), using the monotonicity of the relative entropy [17, 21] we have $\forall \sigma_{AB} \in \mathcal{F}_v(\mathcal{H}), S(\sigma_B || \rho_B) \leq$

$S(\sigma_{AB}||\rho_{AB})$. On substituting the expression for relative entropy, we have $\text{Tr}(\sigma_B \log \sigma_B) - \text{Tr}(\sigma_B \log \rho_B) \leq \text{Tr}(\sigma_{AB} \log \sigma_{AB}) - \text{Tr}(\sigma_{AB} \log \rho_{AB})$. On rearranging, we have $-\text{Tr}(\sigma_{AB} \log \rho_{AB}) + \text{Tr}(\sigma_B \log \rho_B) \geq S_{A|B}(\sigma_{AB}) \geq 0$ which implies that $\text{Tr}(W_\rho \sigma_{AB}) \geq 0$. Hence, W_ρ is a witness operator. ■

The witness given in Theorem 2 is closely related to the conditional amplitude operator [3, 22, 23]. Note that this witness can only be constructed for full-rank density matrices [24]. In Sec IV, we have also given an alternative prescription of a witness exploiting geometrical considerations that can be constructed for any state. We also note that it is possible to construct witnesses employing the uncertainty principle in the presence of quantum memory [11]. We now provide examples of witness construction using Theorem 2.

C. Examples

1. Werner states

Let us consider the Werner state [25] $\gamma = 0.99|\phi^+\rangle\langle\phi^+| + 0.01\frac{I}{4}$. According to the prescription in Theorem 2 we construct a Hermitian operator $W_\gamma = -\log(\gamma_{AB}) + I \otimes \log(\gamma_B)$ which is given by

$$W_\gamma = \begin{bmatrix} a & 0 & 0 & c \\ 0 & b & 0 & 0 \\ 0 & 0 & b & 0 \\ c & 0 & 0 & a \end{bmatrix}, \quad (2)$$

where $a \approx 3.3274$, $b \approx 7.6439$ and $c \approx -4.3165$. We thus obtain

$$\text{Tr}(W_\gamma \gamma) \approx -0.9244. \quad (3)$$

We observe that the witness gives us a negative value, indicating that γ possesses negative conditional entropy. Consider another state that W_γ witnesses: $\delta = 0.9|\phi^+\rangle\langle\phi^+| + 0.1\frac{I}{4}$. We have

$$\text{Tr}(W_\gamma \delta) \approx -0.3417. \quad (4)$$

Once again, this certifies the negative conditional entropy. For implementation in a laboratory, it is important that the witness is decomposed in terms of local observables [26]. If we consider the witness (2), then we obtain the decomposition in terms of Pauli matrices given by

$$W_\gamma = \frac{a+b}{2}I \otimes I + \frac{a-b}{2}Z \otimes Z + \frac{c}{2}X \otimes X - \frac{c}{2}Y \otimes Y. \quad (5)$$

For example, if we have polarized photons, one may take $|H\rangle = |0\rangle$, $|V\rangle = |1\rangle$, $|D\rangle = \frac{|H\rangle+|V\rangle}{\sqrt{2}}$, $|F\rangle = \frac{|H\rangle-|V\rangle}{\sqrt{2}}$, $|L\rangle = \frac{|H\rangle+i|V\rangle}{\sqrt{2}}$, $|R\rangle = \frac{|H\rangle-i|V\rangle}{\sqrt{2}}$. Using this basis, the

witness can be decomposed as follows, where the values of a , b and c are defined in Equation 2:

$$W_\gamma = a (|HH\rangle\langle HH| + |VV\rangle\langle VV|) + b (|VH\rangle\langle VH| + |HV\rangle\langle HV|) + c (|DD\rangle\langle DD| + |FF\rangle\langle FF| - |RR\rangle\langle RR| - |LL\rangle\langle LL|). \quad (6)$$

2. Isotropic states

Isotropic states of dimensionality d are states of the form:

$$\rho_{Iso} = \alpha|\Phi^+\rangle\langle\Phi^+| + (1-\alpha)\frac{I}{d^2}, \quad (7)$$

where $|\Phi^+\rangle$ is a maximally entangled state of two qudits:

$$|\Phi^+\rangle = \frac{1}{\sqrt{d}} \sum_{i=1}^d |i\rangle \otimes |i\rangle, \quad (8)$$

$$\alpha \in \left[-\frac{1}{d^2-1}, 1 \right].$$

The witness for such states is given by:

$$W_{\rho_{Iso}} = -\log(\rho_{Iso}) + I \otimes \log(\rho_{Iso_B}) = -\log(\rho_{Iso}) + I \otimes \log\left(\frac{I}{d}\right) = -\log(\rho_{Iso}) - \log(d).I. \quad (9)$$

Here, we take an example of a 3×3 system. We construct the analytical witness using the following state $\rho_i = 0.8|\phi^+\rangle\langle\phi^+| + 0.2\frac{I}{9}$. The witness is given by

$$W_{\rho_i} = a(|00\rangle\langle 00| + |11\rangle\langle 11| + |22\rangle\langle 22|) + b(|01\rangle\langle 01| + |02\rangle\langle 02| + |10\rangle\langle 10| + |12\rangle\langle 12| + |20\rangle\langle 20| + |21\rangle\langle 21|) + c(|00\rangle\langle 11| + |00\rangle\langle 22| + |11\rangle\langle 00| + |11\rangle\langle 22| + |22\rangle\langle 00| + |22\rangle\langle 11|), \quad (10)$$

where $a \approx 2.1704$, $b \approx 3.9069$ and $c \approx -1.7365$.

We test the witness against ρ_i and notice that it is negative as expected:

$$\text{Tr}(W_{\rho_i} \rho_i) \approx -0.3764. \quad (11)$$

The witness does not give a negative value for all matrices with negative conditional entropy. Consider the state $\rho_t = 0.715|\phi^+\rangle\langle\phi^+| + 0.285\frac{I}{9}$ which has negative conditional entropy. However, the witness above (formed from the state ρ_i) gives a positive value:

$$\text{Tr}(W_{\rho_i} \rho_t) \approx 0.0172. \quad (12)$$

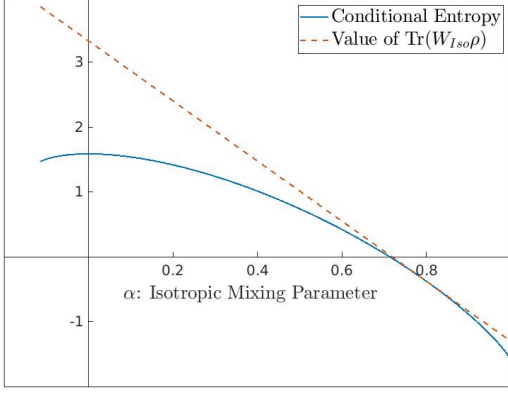


FIG. 2. The x -axis is α , the isotropic mixing parameter, and refers to three dimensional states of the form $\rho = \alpha|\phi^+\rangle\langle\phi^+| + (1 - \alpha)\frac{I}{d^2}$. The y -axis represents different quantities for the red (dashed) line, and the blue (solid) line, placed together for comparison. For the red (dashed) line, it represents the value of $\text{Tr}(W_{\rho_i}\rho)$ and for the blue (solid) line, it represents the conditional von Neumann entropy in bits. Notice that the witness gives a positive value whenever the Conditional von Neumann Entropy is positive.

The behaviour of the above witness is shown in Fig. 2. This witness can be used to detect negative conditional entropy in a lab through the decomposition in terms of Gell Mann matrices, which are:

$$\begin{aligned}
 I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_2 = \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
 \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \\
 \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
 \end{aligned} \tag{13}$$

The above witness W_{ρ_i} can easily be decomposed into the Gell-Mann basis, and is infact diagonal in the basis:

$$\begin{aligned}
 W_{\rho_i} &= -a\lambda_1 \otimes \lambda_1 + a\lambda_2 \otimes \lambda_2 - a\lambda_3 \otimes \lambda_3 \\
 &\quad - a\lambda_4 \otimes \lambda_4 + a\lambda_5 \otimes \lambda_5 - a\lambda_6 \otimes \lambda_6 \\
 &\quad + a\lambda_7 \otimes \lambda_7 - a\lambda_8 \otimes \lambda_8 + bI_3 \otimes I_3,
 \end{aligned} \tag{14}$$

where $a \approx 0.86825$ and $b \approx 3.3281$. This facilitates the identification of negativity of conditional entropy in the lab.

III. UPPER BOUND ON CONDITIONAL ENTROPY

We now use the expectation value of the witness to quantify the amount of conditional entropy in a state. To do so, we define $T_M^\chi(\mathcal{H}) := \{S_{A|B}(\sigma) \mid \sigma \in \mathcal{S}(\mathcal{H}) \wedge \text{Tr}(M\sigma) = \chi\}$ where M is a Hermitian operator and $\chi \in \mathbb{R}$. The \mathcal{H} is henceforth omitted as there is no ambiguity. In words, T_M^χ is the set of conditional entropies of all states in which the expectation value of M is χ . Obtaining a tight upper bound for T_M^χ is equivalent to asking the question: what is the maximum value of conditional entropy that an unknown state can possess, given that the expectation value of M in the state is χ ? For several applications in quantum information, as described at the end of this paper, a state is useful only if its conditional entropy is less than a certain value. Thus, an upper bound of T_M^χ could aid in the decision of whether to use a state for a given application or to discard it. In the following theorem, we prove that for a witness W_ρ constructed using Theorem 2, an upper bound of $T_{W_\rho}^\chi$ is χ itself. However, this upper bound need not be tight. We then pose the problem of obtaining a tight upper bound as a convex optimization problem, solve it numerically and discuss some interesting results.

Theorem 3. χ is an upper bound of $T_{W_\rho}^\chi$.

Proof. Let $x \in T_{W_\rho}^\chi$. Therefore, $\exists \sigma_{AB} \in \mathcal{S}(\mathcal{H})$ such that $x = S_{A|B}(\sigma_{AB})$ and $\text{Tr}(W_\rho \sigma_{AB}) = \chi$. Once again, using the monotonicity of relative entropy [17, 21], we have $S(\sigma_B \parallel \rho_B) \leq S(\sigma_{AB} \parallel \rho_{AB})$. On substituting for relative entropy, we have $\text{Tr}(\sigma_B \log \sigma_B) - \text{Tr}(\sigma_B \log \rho_B) \leq \text{Tr}(\sigma_{AB} \log \sigma_{AB}) - \text{Tr}(\sigma_{AB} \log \rho_{AB})$. On rearranging, we have $S_{A|B}(\sigma_{AB}) \leq \text{Tr}(W_\rho \sigma_{AB})$ which implies that $x \leq \chi$. ■

As an example, we observe from Equation 4 that $\text{Tr}(W_\gamma \delta) \approx -0.3417$. According to Theorem 3, this value should serve as an upperbound to $S_{A|B}(\delta)$. Indeed, we compute that $S_{A|B}(\delta)$ is approximately -0.4968 .

In general, χ is not a tight upper bound for $T_{W_\rho}^\chi$. However when $\chi = \text{Tr}(W_\rho \rho) = S_{A|B}(\rho)$, We have $\chi \in T_{W_\rho}^\chi$ while also being an upper bound to $T_{W_\rho}^\chi$, and is therefore a tight upper bound. From Equation 3 we observe that -0.9244 is the tight upperbound to $T_{W_\gamma}^{-0.9244}$.

A tight upper bound for T_M^χ can be obtained in general by solving the following convex optimization problem for all $\sigma \in \mathcal{S}(\mathcal{H})$:

$$\begin{aligned}
 &\text{maximize}_{\sigma} \quad S_{A|B}(\sigma) \\
 &\text{subject to} \quad \text{Tr}(M\sigma) = \chi.
 \end{aligned} \tag{15}$$

For a small number of qubits, this is easy to solve numerically [27–30]. We simultaneously describe the procedure as well as results of a numerical experiment to solve this optimization problem for $M = W_\gamma$ defined in Equation 2. For our numerical experiment, we denote the 2-qubit

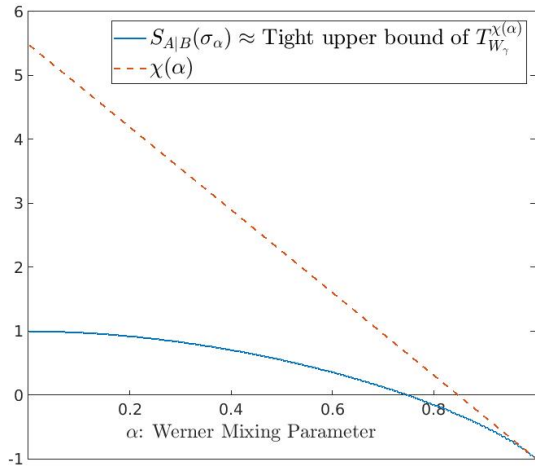


FIG. 3. Conditional entropy of Werner states and the obtained tight upper bound: The x -axis refers to the parameter α for states of the form $\sigma_\alpha = \alpha|\phi^+\rangle\langle\phi^+| + (1-\alpha)I/4$. The y -axis represents different quantities for the red (dashed) line and the blue (solid) line, placed together for comparison. For the red (dashed) line, it represents the value of $\chi(\alpha) = \text{Tr}(W_\gamma\sigma_\alpha)$. For the blue (solid) line, it represents $S_{A|B}(\sigma_\alpha)$ in bits as well the tight upper bound of $T_{W_\gamma}^{\chi(\alpha)}$, which happen to coincide.

Werner states with $\sigma_\alpha \equiv \alpha|\phi^+\rangle\langle\phi^+| + (1-\alpha)I/4$ where $\alpha \in [0, 1]$. The x -axis in Figure 3 is the value of α for the Werner states. First, for each α , we compute the value of $\chi(\alpha) = \text{Tr}(W_\gamma\sigma_\alpha)$, which is plotted with a red dashed line. For each $\chi(\alpha)$, we then solve the optimization problem (15) to obtain the tight upper bound of $T_{W_\gamma}^{\chi(\alpha)}$, which is plotted with a blue solid line. We then compute the value of $S_{A|B}(\sigma_\alpha)$, which is also the blue solid line. We observe two interesting results: Firstly, the value of the tight upper bound coincides with the value of conditional entropy for the Werner states. That is, for this particular witness W_γ , among all 2-qubit states in which the expectation value of W_γ is $\chi(\alpha)$, the corresponding Werner state $\sigma_\alpha = \alpha|\phi^+\rangle\langle\phi^+| + (1-\alpha)I/4$ possesses the maximum conditional entropy. The second interesting aspect is that the convex optimization enhances the usefulness of the witness. We observe from the graph in Figure 3 that there exist states where a positive $\chi(\alpha)$ is obtained, but the tight upper bound of $T_{W_\gamma}^{\chi(\alpha)}$ is negative, certifying that the state indeed possesses negative conditional entropy.

This optimization procedure is not limited to witnesses, but rather, any measurement. It can be used to identify the presence of negative conditional entropy, and therefore also entanglement. It is known that $(|\phi^+\rangle\langle\phi^+|)^{T_A}$ where $(\cdot)^{T_A}$ is the partial transpose with respect to the first subsystem, is an entanglement witness [31]. Consider the operator $E =$

$(0.3|\phi^+\rangle\langle\phi^+| + 0.7I/4)^{T_A}$. It is a Hermitian operator, but not a witness as it is positive definite. We find that the tight upper bound of $T_E^{0.05}$ is approximately -0.4541 , thus immediately certifying negative conditional entropy (and therefore also entanglement) for any state in which the expectation value of E is 0.05. We now give an alternate construction of a witness for conditional entropy.

IV. GEOMETRIC WITNESS FOR NEGATIVE CONDITIONAL ENTROPY

A. Construction of the witness

The witness constructed in Theorem 2 can only be constructed for states corresponding to full-rank matrices. In this section, we discuss a witness construction that is valid for any state. We show that an operator constructed in the same fashion as one that witnesses entangled states from the class of separable states ([32], [33], [34]) in fact acts as a witness for a state outside any convex and compact set, and thus also for $\mathcal{F}_v(\mathcal{H})$. In the proof that follows, $\|\cdot\|$ refers to the Frobenius norm and $\langle A, B \rangle = \text{Tr}(A^\dagger B)$. Also, arguments to $\langle A, B \rangle$ are Hermitian, hence $\langle A, B \rangle$ is always real. We have $\|A\| = \sqrt{\langle A, A \rangle}$.

Theorem 4. *A witness operator W^g separating state ρ_s from any convex and compact set $S \subseteq \mathcal{S}(\mathcal{H})$ is given by,*

$$W^g = \frac{\text{Tr}(\sigma_c \rho_s - \sigma_c^2)I + \sigma_c - \rho_s}{\sqrt{\text{Tr}(\sigma_c - \rho_s)^2}},$$

where

$$\sigma_c = \arg \min_{\sigma} \|\rho_s - \sigma\| \quad \forall \sigma \in S.$$

Since $\mathcal{F}_v(\mathcal{H})$ is convex and compact, it follows that W^g acts as a witness for a state outside $\mathcal{F}_v(\mathcal{H})$.

Proof. To find a witness operator W^g separating a state ρ_s from S , given that σ_c is the closest state in S to ρ_s , we proceed as follows: We consider a function $f(\chi) = \langle \chi - \sigma_c, \sigma_c - \rho_s \rangle$ defined for all density matrices χ . We show that (i) $\forall \sigma \in S$, $f(\sigma) \geq 0$ and (ii) $\exists \rho \notin S$ such that $f(\rho) < 0$. We then find the W^g such that $\text{Tr}(W^g \chi) = f(\chi) \quad \forall \chi$, and thus find the witness W^g .

(i) To prove that the $f(\sigma)$ is non-negative for all $\sigma \in S$, we assume otherwise:

$$\exists \sigma' \in S \mid \langle \sigma' - \sigma_c, \sigma_c - \rho_s \rangle < 0. \quad (16)$$

Within this assumption, we consider two cases. Intuitively, these two cases correspond to having an (i) obtuse or right angle between $\sigma_c - \sigma'$ and $\rho_s - \sigma'$ and having an (ii) acute angle between them as shown in Fig 4 and Fig 5.

Case 1: $\langle \rho_s - \sigma', \sigma_c - \sigma' \rangle \leq 0$. Here, we show that σ' is closer to ρ_s than σ_c , hence leading to a contradiction. Consider the expression,

$$\langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle - \langle \rho_s - \sigma', \rho_s - \sigma' \rangle, \quad (17)$$

which can be rewritten as follows:

$$\begin{aligned} & \langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle - \langle \rho_s - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle - \langle \rho_s - \sigma' + \sigma_c - \sigma_c, \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle - \langle \rho_s - \sigma_c, \rho_s - \sigma' \rangle - \langle \sigma_c - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma_c, \rho_s - \sigma_c - \rho_s + \sigma' \rangle - \langle \sigma_c - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma_c, \sigma' - \sigma_c \rangle - \langle \sigma_c - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma_c + \sigma' - \sigma', \sigma' - \sigma_c \rangle - \langle \sigma_c - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \rho_s - \sigma', \sigma' - \sigma_c \rangle + \langle \sigma' - \sigma_c, \sigma' - \sigma_c \rangle - \langle \sigma_c - \sigma', \rho_s - \sigma' \rangle \\ &= \langle \sigma' - \sigma_c, \sigma' - \sigma_c \rangle - 2\langle \rho_s - \sigma', \sigma_c - \sigma' \rangle. \end{aligned} \quad (18)$$

We know that $\sigma' \neq \sigma_c$, as that would make the initial assumption 16 untrue. Hence, the first term is always positive by definition of inner product and the second term is non-positive by the assumption in Case 1.

Hence, we have $\langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle - \langle \rho_s - \sigma', \rho_s - \sigma' \rangle > 0$ or $\sqrt{\langle \rho_s - \sigma_c, \rho_s - \sigma_c \rangle} > \sqrt{\langle \rho_s - \sigma', \rho_s - \sigma' \rangle}$. Therefore, $\|\rho_s - \sigma_c\| > \|\rho_s - \sigma'\|$, which is a contradiction.

Case 2: $\langle \rho_s - \sigma', \sigma_c - \sigma' \rangle > 0$. Here, we show that there will exist a point $\sigma'' \in S$ such that $\langle \rho_s - \sigma'', \sigma_c - \sigma'' \rangle = 0$. Thus, Case 1 is satisfied and following the argument present there, σ'' will be closer to ρ_s than σ_c , and a contradiction is attained.

Since, $\sigma' \in S$ and $\sigma_c \in S$, by the convexity of the set S , we have:

$$\forall \lambda \in [0, 1], \lambda \sigma_c + (1 - \lambda) \sigma' \in S. \quad (19)$$

Consider $\lambda' = \frac{\langle \sigma' - \rho_s, \sigma' - \sigma_c \rangle}{\langle \sigma' - \sigma_c, \sigma' - \sigma_c \rangle}$. By the assumption in Case 2, the numerator is positive, and by the definition of inner product, the denominator is positive. Therefore $\lambda' > 0$. We also have $\lambda' = \frac{\langle \sigma' - \sigma_c + \sigma_c - \rho_s, \sigma' - \sigma_c \rangle}{\langle \sigma' - \sigma_c, \sigma' - \sigma_c \rangle} = 1 - \frac{\langle \sigma' - \sigma_c, \rho_s - \sigma_c \rangle}{\langle \sigma' - \sigma_c, \sigma' - \sigma_c \rangle}$, and by the initial contradiction assumption 16, the second term is positive, hence $\lambda' < 1$.

Therefore, $\lambda' \in (0, 1)$. Consider a state $\sigma'' = \lambda' \sigma_c + (1 - \lambda') \sigma'$. Notice that this point belongs to the set S and $\langle \rho_s - \sigma'', \sigma_c - \sigma'' \rangle = 0$, therefore satisfying Case 1. Thus both cases lead to a contradiction rendering the initial assumption 16 untrue. Thus, all points σ in S satisfy $\langle \sigma - \sigma_c, \sigma_c - \rho_s \rangle \geq 0$.

(ii) To show that $\exists \rho \notin S \mid f(\rho) < 0$, we consider the point ρ_s itself, which we wish to separate. It is easy to see that $f(\rho_s)$ gives a negative value:

$$\langle \rho_s - \sigma_c, \sigma_c - \rho_s \rangle = -\|\rho_s - \sigma_c\|^2 < 0. \quad (20)$$

We find the witness operator W^g such that $\text{Tr}(W^g \chi) = \langle \chi - \sigma_c, \sigma_c - \rho_s \rangle = \text{Tr}((\chi - \sigma_c)(\sigma_c - \rho_s))$. On solving the above, $W^g = \text{Tr}(\sigma_c \rho_s - \sigma_c^2) I + \sigma_c - \rho_s$. On normalising, we get W^g as mentioned in Theorem 4. ■

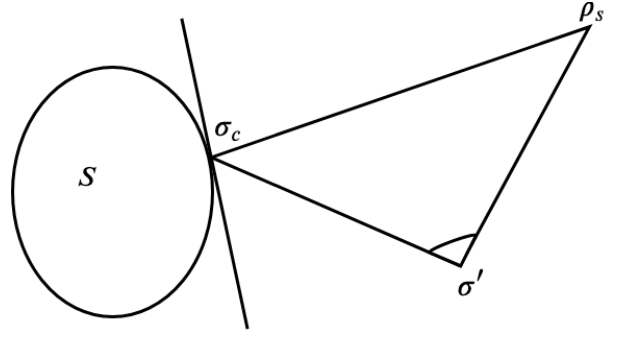


FIG. 4. Case 1: Corresponds to angle at σ' being obtuse.

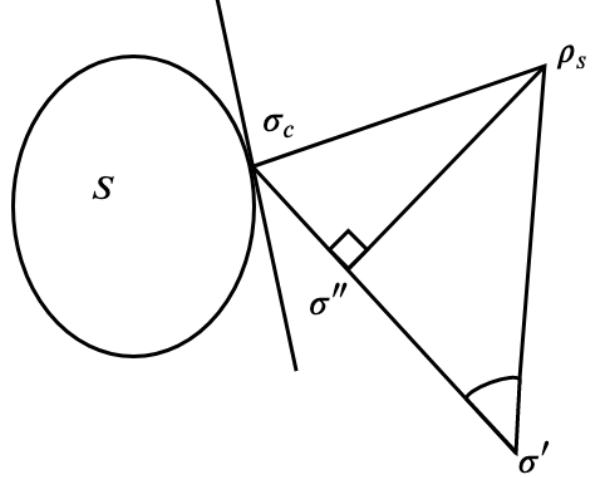


FIG. 5. Case 2: Corresponds to the angle at σ' being acute.

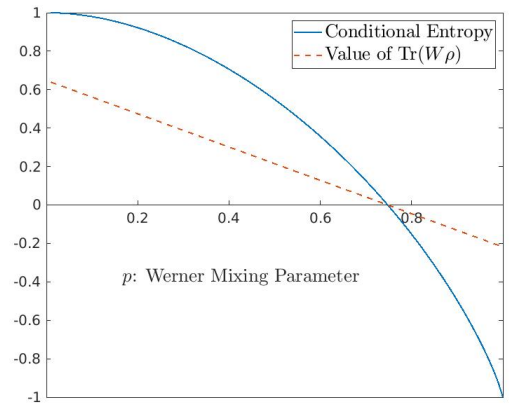


FIG. 6. Variation of conditional entropy and the value of $\text{Tr}(W\rho)$ with the mixing parameter p . The x -axis is p , the Werner mixing parameter, and refers to two dimensional states of the form $\rho = p|\phi^+\rangle\langle\phi^+| + (1-p)\frac{I}{2^2}$. The y -axis represents different quantities for the red (dashed) line and the blue (solid) line, placed together for comparison. For the red (dashed) line, it represents the value of $\text{Tr}(W\rho)$ where W is the geometric witness calculated in Eq.21 and for the blue (solid) line, it represents the conditional entropy in bits.

Now, it remains that we find σ_c , the closest state in $\mathcal{F}_v(\mathcal{H})$ to ρ_s . Note that trace distance $\|\rho_s - \sigma\|$ is a convex function in σ for a fixed ρ_s , and it needs to be optimised over all $\sigma \in \mathcal{F}_v(\mathcal{H})$, a convex set. Also, unlike the entanglement problem, the conditional entropy expression can be evaluated for any given dimension at any state, and membership in $\mathcal{F}_v(\mathcal{H})$ can be tested. Thus, this is a convex optimization problem and there exist several solvers to tackle this. We use MATLAB's CVX solver [27, 28], a package for specifying and solving convex programs and a few quantum function libraries [29, 30] to find the closest state in $\mathcal{F}_v(\mathcal{H})$ for several 2-qubit and 3-qubit states.

B. Example: Werner class of states

In the 2×2 dimensional space, consider the Bell state $|\phi^+\rangle\langle\phi^+|$. We find that the closest state in $\mathcal{F}_v(\mathcal{H})$ to $|\phi^+\rangle\langle\phi^+|$ is the following:

$$\begin{bmatrix} a & 0 & 0 & c \\ 0 & b & 0 & 0 \\ 0 & 0 & b & 0 \\ c & 0 & 0 & a \end{bmatrix}$$

where $a \approx 0.4369$, $b \approx 0.0631$, $c \approx 0.3738$. We note that this is the state where the Werner line $p|\phi^+\rangle\langle\phi^+| + (1-p)\frac{I}{4}$ intersects $\mathcal{F}_v(\mathcal{H})$. Therefore, this state at $p \approx 0.7476$ is the closest state in $\mathcal{F}_v(\mathcal{H})$ for all states with negative conditional entropy that lie on the the given Werner line. Hence the witness created using this state witnesses all negative conditional entropy states of the form $p|\phi^+\rangle\langle\phi^+| + (1-p)\frac{I}{4}$ where $0 \leq p \leq 1$.

Applying Theorem 4, we find the witness to this state is:

$$W = \begin{bmatrix} a & 0 & 0 & c \\ 0 & b & 0 & 0 \\ 0 & 0 & b & 0 \\ c & 0 & 0 & a \end{bmatrix} \quad (21)$$

where $a \approx 0.3588$, $b \approx 0.9361$ and $c \approx -0.5774$. To check that this operator indeed acts as a witness, consider the state $\rho_w = 0.75|\phi^+\rangle\langle\phi^+| + 0.25\frac{I}{4}$ which lies on the Werner line just outside $\mathcal{F}_v(\mathcal{H})$:

$$\text{Tr}(W\rho_w) \approx -0.0021 \quad (22)$$

which is a negative value as it is outside the class, as expected. Fig 6 shows the variation of $\text{Tr}(W\rho)$ with mixing parameter p , as well as the corresponding conditional entropy.

The witness operator W can readily be decomposed in the form of Pauli matrices as follows and be implemented in the laboratory for measurement:

$$W = \frac{a+b}{2}I \otimes I + \frac{a-b}{2}Z \otimes Z + \frac{c}{2}X \otimes X - \frac{c}{2}Y \otimes Y$$

C. Proof of weak optimality

An entanglement witness A is ‘‘weakly optimal’’ if there exists a separable state σ_s such that $\langle\sigma_s, A\rangle = 0$ [34, 35]. Similarly, we define weak optimality for witnesses of negative conditional entropy: A witness of negative conditional entropy W is ‘‘weakly optimal’’ if there exists a state σ_o such that (i) $S_{A|B}(\sigma_o) \geq 0$ and (ii) $\langle\sigma_o, W\rangle = 0$. To show that W^g is weakly optimal, we use the state σ_c . Using the definitions in the proof of Theorem 4, we have (i) $S_{A|B}(\sigma_c) \geq 0$ and (ii):

$$\langle\sigma_c, W^g\rangle = \text{Tr}(\sigma_c W^g) = f(\sigma_c) = \langle\sigma_c - \sigma_c, \sigma_c - \rho_s\rangle = 0.$$

Hence, W^g is a weakly optimal witness for negative conditional entropy. Such a witness is also known as a tangent functional [33, 34]. We now discuss some applications of witnessing negative conditional entropy.

V. APPLICATIONS

A. Quantum state merging

When a quantum state ρ_{AB} is shared between A and B , we ask: How much quantum communication is needed from A to transfer the full information of the state to B ? This process of merging the missing information from one system to another is called state merging [5, 6]. The quantification of this process is given by the conditional entropy $S_{A|B}(\rho_{AB})$ (if it is from A to B) of the system. If $S_{A|B}(\rho_{AB})$ is positive, it means that the sender needs to communicate this number of quantum bits to the receiver and if it is 0, there is no need for such communication. However, when it is negative, the sender and receiver gain corresponding potential for future quantum communication between them even after the merging process. The witness provided by us detects such states, and is therefore useful for the several applications of the quantum state merging primitive detailed in [6].

B. Superdense coding

If two parties, Alice and Bob share an entangled state, then by sending a qudit to Bob, Alice may communicate more information than possible with a classical dit, which is known as superdense coding. The superdense coding capacity for a mixed state ρ_{AB} in is given by [8]

$$C_{AB} = \max\{\log_2 d, \log_2 d + S(\rho_B) - S(\rho_{AB})\}. \quad (23)$$

Here, C_{AB} is the amount of classical information that can be sent from A to B using ρ_{AB} as a quantum resource. Note that when $S_{A|B}(\rho_{AB})$ is negative we can use the shared state to transfer classical communication greater than the classical limit of $\log_2 d$ bits. Therefore, states

with negative $S_{A|B}(\sigma_{AB})$ give quantum advantage while performing superdense coding and our result is useful in detecting such states. In addition, dense coding capacity also obeys the exclusion principle in the case of multipartite systems in which negative conditional entropy plays an important role [9].

C. Distributed private randomness distillation

The inherent probabilistic nature of quantum measurement outcomes provides a natural arena for generation of randomness [36]. The generated randomness is considered private if it is completely unpredictable in advance [12]. For example, measurement of the $|+\rangle$ state in the computational basis is private as a pure state is not correlated with any other state, including one with a possible Eavesdropper. However, this may not always be the case. In [36], the problem of extracting randomness from a state where part of the state is in the possession of the Eavesdropper is considered. In [12], a distributed setting of the above problem is considered: if two parties, Alice and Bob trust each other and share a state ρ_{AB} and Eve possesses the purification of the state, then what is the maximum rate of randomness, that Alice and Bob (R_A and R_B respectively) can extract private from Eve?

In the case that Alice and Bob cannot communicate with each other and can only perform local unitary operations, while each having access to free noise (maximally mixed states), the bound turns out to be: $R_A \leq \log |A| - S_{A|B}(\rho_{AB})$, $R_B \leq \log |B| - S_{B|A}(\rho_{AB})$ and $R_A + R_B \leq R_G$ where $R_G = \log |AB| - S(\rho_{AB})$ and $|X|$ is the dimensionality of the quantum system X .

Alice and Bob can use the witness provided by us to obtain an upper bound on $S_{A|B}(\rho_{AB})$ and $S_{B|A}(\rho_{AB})$. For a two qubit system, we have provided a local decomposition of the witness to accomplish this (5). Note that classical communication is necessary for Alice and Bob to estimate these upper bounds.

D. One-way entanglement distillation

Entanglement distillation is the process of transformation of N copies of an arbitrary entangled state ρ_{AB} into some number of approximately pure EPR pairs, using only local operations and classical communication (LOCC) at optimal rates [37]. In one-way entanglement distillation, the goal is also to transform a given shared state into maximally entangled states, albeit in an asymmetric setting as described in [10]. Here, the one-way (or forward) entanglement capacity of ρ_{AB} , which is the maximal achievable rate of entanglement distillation, is given by the Hashing inequality [10]: $D_{\rightarrow}(\rho) \geq -S_{A|B}(\rho)$. The witness provided by us detects states with a positive one-way entanglement capacity, and the subsequent optimization procedure can be utilized to provide a tight lower bound on the same.

E. Coherent information

Consider a quantum system A and a reference system P . Initially, the quantum system A and the reference system P are in a pure entangled state, denoted by $|\psi\rangle_{PA}$. The density matrix ρ_A of the initial system A is given by $\rho_A = \text{Tr}_P(|\psi\rangle_{PA}\langle\psi|)$. Consider an unknown quantum channel N through which the state ρ_A of the system A is transmitted while the system P remains isolated from the environment and does not get affected by the channel. The combined output state ρ_{PB} is given by $\rho_{PB} = (I_P \otimes N_A)(|\psi\rangle_{PA}\langle\psi|)$ as shown in Fig 7.

As a result of this transmission, both the system A and the entanglement between the system A and P get affected. The quantum capacity $Q(N)$ of the channel is associated with a correlation measure known as *quantum coherent information* [38]. The entropy exchange between the initial system PA and the output system PB is a measure of the extent to which the initially pure state $|\psi\rangle_{PA}$ becomes mixed as a result of the channel, and is defined as $S_E(\rho_A; N(\rho_A)) = S(\rho_{PB})$. A lower bound for the quantum capacity $Q(N)$ is given by the quantum coherent information which is expressed as $S(N(\rho_A)) - S_E(\rho_A; N(\rho_A)) = S(\rho_B) - S(\rho_{PB})$ [39]. The expression $S(\rho_B) - S(\rho_{PB})$ is the negative of the conditional entropy $S_{P|B}(\rho_{PB})$. Thus, if we have a negative value of $S_{P|B}(\rho_{PB})$, then we have a positive value of quantum coherent information, and therefore of $Q(N)$. Our witness operator can detect whether the state ρ_{PB} has negative conditional entropy. This implies that we can detect whether the unknown quantum channel N is useful in transmitting quantum information.

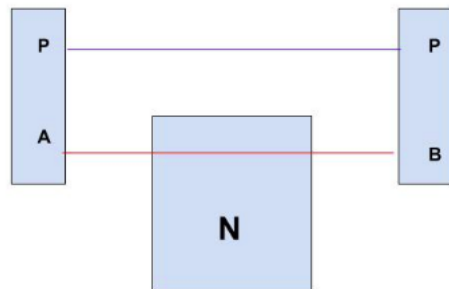


FIG. 7. The dynamics of the system A when sent through the quantum channel N with P as the reference system. B is the output of the system A after the channel transmission.

F. Uncertainty relations

The entropic uncertainty relation [40] for two incompatible measurements X and Y performed on a quantum

state is given by [41, 42]:

$$H(X) + H(Y) \geq \log_2 \frac{1}{c}. \quad (24)$$

Here, $H(X)$ and $H(Y)$ denote the Shannon entropy of the outcomes of measurement X and Y , respectively when performed on any quantum state, and $c = \max_{i,j} |\langle x_i | y_j \rangle|^2$ where $|x_i\rangle, |y_j\rangle$ are eigenvectors of X and Y , respectively.

The operational interpretation of Eq. 24 is as follows: Alice and Bob agree on two measurements X and Y . Bob prepares a state and sends it to Alice who measures, according to her choice, X or Y and reveals her measurement choice (X or Y). Bob then guesses her outcome. Bob's aim is to prepare the initial state such that he reduces the uncertainty about the outcome revealed by Alice. By Eq. 24, whatever state Bob prepares, $H(X) + H(Y)$ will be at least $\log_2 \frac{1}{c}$.

However, note that if Bob has access to quantum memory, he can prepare a state that is maximally entangled with the memory and send it to Alice, then apply the same measurement that Alice picks on his own memory, and can guess Alice's outcome with full certainty. Hence, the above inequality does not hold in the presence of quantum memory. A new uncertainty relation taking into consideration quantum memory [11] is the following:

$$H_A(X|B) + H_A(Y|B) \geq \log_2 \frac{1}{c} + S_{A|B}(\rho_{AB}), \quad (25)$$

where ρ_{AB} is the combined state of Bob's quantum memory and the subsystem he sends to Alice. $H_A(X|B)$ (or $H_A(Y|B)$) quantifies the uncertainty about the outcome of measurement X (or Y), given information stored in a quantum memory B [11].

Here we notice that whenever $S_{A|B}(\rho_{AB})$ is negative, by Eq. 25, Bob's prepared state beats the original uncertainty bound of Eq. 24. In other words, by our characterization of $\mathcal{F}_v(\mathcal{H})$, we make it possible to experimentally witness states that have the ability to utilize the available quantum memory in overcoming the original uncertainty bound. Among these states, are the states with $S_{A|B}(\rho_{AB}) \leq \log_2 c$, for which the uncertainty bound becomes trivial: $H_A(X|B) + H_A(Y|B) \geq 0$. Fig. 8 depicts this classification of states for a given measurement set $\{X, Y\}$.

VI. CONCLUSION

In the present work, we prove that the states having non-negative conditional entropy form a convex and compact set. This allows for the construction of witnesses to detect states with negative conditional entropy. We provide two constructions of such witnesses and for one of them, we prove that the expectation value of the witness in a state is an upper bound to the conditional entropy of the state. We pose the problem of obtaining a tight

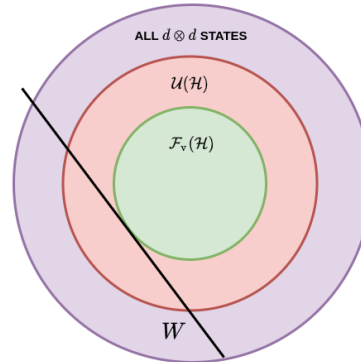


FIG. 8. The inner set represents $\mathcal{F}_v(\mathcal{H})$ — the set of states that give no advantage with quantum memory over the original scenario. Beyond it, labelled $\mathcal{U}(\mathcal{H})$, are the set of states that do give an advantage, but still have a degree of uncertainty. And beyond those (outermost) are the states with $S_{A|B}(\rho_{AB}) \leq \log_2 c$, for which the uncertainty bound is trivial.

upper bound to the set of conditional entropies of states in which an operator gives the same expectation value as a convex optimization problem, and discuss several interesting results. We believe that further insight on conditional entropy as a resource distinct from entanglement can be gained by studying it in a resource-theoretic framework. In such a study, the set of states with non-negative conditional entropy characterized in this work would be the set of free states.

ACKNOWLEDGMENTS

We acknowledge Prof. John Watrous, Dr. Luca Innocenti and Mr. Siddharth Bhat for having useful discussions at various stages of completion of this manuscript. We also thank the anonymous referees for their valuable suggestions, which were immensely useful in improving our work. N.G. would like to acknowledge support from the Research Initiation Grant of BITS-Pilani, Hyderabad vide letter no. BITS/GAU/RIG/2019/H0680 dated 22nd April, 2019.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).

[2] N. Ganguly, S. Adhikari, A. S. Majumdar, and J. Chatterjee, *Phys. Rev. Lett.* **107**, 270501 (2011).

- [3] N. J. Cerf and C. Adami, *Phys. Rev. Lett.* **79**, 5194 (1997).
- [4] S. Abe and A. Rajagopal, *Physica A: Statistical Mechanics and its Applications* **289**, 157 (2001).
- [5] M. Horodecki, J. Oppenheim, and A. Winter, *Nature* **436**, 673 (2005).
- [6] M. Horodecki, J. Oppenheim, and A. Winter, *Communications in Mathematical Physics* **269**, 107 (2007).
- [7] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
- [8] D. Bruß, G. M. D’Ariano, M. Lewenstein, C. Macchiavello, A. Sen(De), and U. Sen, *Phys. Rev. Lett.* **93**, 210501 (2004).
- [9] R. Prabhu, A. K. Pati, A. Sen(De), and U. Sen, *Phys. Rev. A* **87**, 052319 (2013).
- [10] I. Devetak and A. Winter, *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **461**, 207 (2005).
- [11] M. Berta, M. Christandl, R. Colbeck, J. M. Renes, and R. Renner, *Nature Physics* **6**, 659 (2010).
- [12] D. Yang, K. Horodecki, and A. Winter, *Phys. Rev. Lett.* **123**, 170501 (2019).
- [13] K. Azuma and S. Subramanian, arXiv preprint arXiv:1807.06753 (2018).
- [14] K. Azuma and G. Kato, arXiv preprint arXiv:2001.02897 (2020).
- [15] L. del Rio, J. Åberg, R. Renner, O. Dahlsten, and V. Vedral, *Nature* **474**, 61 (2011).
- [16] S. Patro, I. Chakrabarty, and N. Ganguly, *Phys. Rev. A* **96**, 062102 (2017).
- [17] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, USA, 2011).
- [18] R. Alicki and M. Fannes, *Journal of Physics A: Mathematical and General* **37**, L55 (2004).
- [19] A. Winter, *Communications in Mathematical Physics* **347**, 291 (2016).
- [20] R. B. Holmes, *Geometric Functional Analysis and its Applications* (Springer New York, 1975).
- [21] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
- [22] N. J. Cerf and C. Adami, *Phys. Rev. A* **60**, 893 (1999).
- [23] N. Friis, S. Bulusu, and R. A. Bertlmann, *Journal of Physics A: Mathematical and Theoretical* **50**, 125301 (2017).
- [24] This is because $\log 0$ is undefined.
- [25] R. F. Werner, *Physical Review A* **40**, 4277 (1989).
- [26] O. Gühne and G. Tóth, *Physics Reports* **474**, 1 (2009).
- [27] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx> (2014).
- [28] M. Grant and S. Boyd, in *Recent Advances in Learning and Control*, Lecture Notes in Control and Information Sciences, edited by V. Blondel, S. Boyd, and H. Kimura (Springer-Verlag Limited, 2008) pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.
- [29] T. Cubitt, “Quantinf package, version 0.5.1,” <https://www.dr-qubit.org/matlab.html> (2013).
- [30] H. Fawzi, J. Saunderson, and P. A. Parrilo, Foundations of Computational Mathematics (2018), package cvxquad at <https://github.com/hfawzi/cvxquad>.
- [31] M. M. Wolf, *Partial transposition in quantum information theory* (2003).
- [32] R. A. Bertlmann and P. Krammer, *Annals of Physics* **324**, 1388 (2009).
- [33] R. A. Bertlmann, H. Narnhofer, and W. Thirring, *Phys. Rev. A* **66**, 032319 (2002).
- [34] R. A. Bertlmann, K. Durstberger, B. C. Hiesmayr, and P. Krammer, *Phys. Rev. A* **72**, 052331 (2005).
- [35] B.-H. Wang, H.-R. Xu, S. Campbell, and S. Severini, *Quantum Info. Comput.* **15**, 1109–1121 (2015).
- [36] M. Berta, O. Fawzi, and S. Wehner, in *Lecture Notes in Computer Science* (Springer Berlin Heidelberg, 2012) pp. 776–793.
- [37] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
- [38] L. Gyongyosi, S. Imre, and H. V. Nguyen, *IEEE Communications Surveys Tutorials* **20**, 1149 (2018).
- [39] G. Smith, arXiv e-prints, arXiv:1007.2855 (2010), arXiv:1007.2855 [cs.IT].
- [40] H. P. Robertson, *Physical Review* **34**, 163 (1929).
- [41] K. Kraus, *Phys. Rev. D* **35**, 3070 (1987).
- [42] H. Maassen and J. B. M. Uffink, *Phys. Rev. Lett.* **60**, 1103 (1988).