arXiv:2208.14424v3 [quant-ph] 14 Nov 2024

Inevitability of knowing less than nothing

Gilad Gour[1], Mark M. Wilde[2], S. Brandsen[3], and Isabelle Jianing Geng[4]

[1]Faculty of Mathematics, Technion - Israel Institute of Technology, Haifa 3200003, Israel
[2]School of Electrical and Computer Engineering, Cornell University, Ithaca, New York 14850, USA
[3]Department of Physics, Duke University, Durham, North Carolina 27708, USA
[4]Department of Mathematics and Statistics, Institute for Quantum Science and Technology, University of Calgary, Alberta, Canada T2N 1N4

A colloquial interpretation of entropy is that it is the knowledge gained upon learning the outcome of a random experiment. Conditional entropy is then interpreted as the knowledge gained upon learning the outcome of one random experiment after learning the outcome of another, possibly statistically dependent, random experiment. In the classical world, entropy and conditional entropy take only non-negative values, consistent with the intuition that one has regarding the aforementioned interpretations. However, for certain entangled states, one obtains negative values when evaluating commonly accepted and information-theoretically justified formulas for the quantum conditional entropy, leading to the confounding conclusion that one can know less than nothing in the quantum world. Here, we introduce a physically motivated framework for defining quantum conditional entropy, based on two simple postulates inspired by the second law of thermodynamics (non-decrease of entropy) and extensivity of entropy, and we argue that all plausible definitions of quantum conditional entropy should respect these two postulates. We then prove that all plausible quantum conditional entropies take on negative values for certain entangled states, so that it is inevitable that one can know less than nothing in the quantum world. All of our arguments are based on constructions of physical processes that respect the first postulate, the one inspired by the second law of thermodynamics.

Contents

1 Introduction **2**
1.1 Summary of results 2
1.2 Notation 3

2 Uncertainty and entropy **3**
2.1 Uncertainty measures 3
2.2 Entropy 4

3 Conditional uncertainty and entropy **5**
3.1 Conditional uncertainty measures 5
3.2 Conditional entropy 7

4 Main result: Inevitability of negative quantum conditional entropy **7**

5 Conditional entropy from relative entropy **8**

6 Conclusion **9**

References **9**

A Proofs of properties of entropy **12**

B Choi matrix formulations **12**
B.1 Choi matrix formulation of conditional unitality 13
B.2 Choi matrix formulation of conditionally mixing channels 14

C Quantum conditional majorization reduces to classical conditional majorization **14**

D Proofs of properties of quantum conditional entropy **15**

E A subclass of conditional entropies **16**
E.1 Proof of Theorem 5 17
E.2 Proof of Theorem 1 19

F Proof of Theorem 2 **20**
F.1 Conditional min-entropy is a conditional entropy 20

G Entropy inequalities **21**

H Example of a non-negative, weakly additive monotonic function **22**

Gilad Gour: giladgour@technion.ac.il
Mark M. Wilde: wilde@cornell.edu

1 Introduction

Quantum information science is famously known for having features that are not present in classical information science. Among these, perhaps the feature that distinguishes it the most from the classical case is that our knowledge of the state of two particles can be greater than our knowledge of the individual states of the particles. Indeed, for an Einstein–Podolsky–Rosen (EPR) state $|\Phi^{(2)}\rangle := (|00\rangle+|11\rangle)/\sqrt{2}$ [1], said to be entangled, our knowledge of the joint state of the two particles is complete. However, the state of an individual particle is described by an unbiased probabilistic ensemble (i.e., $|0\rangle$ or $|1\rangle$ with equal probability), and we thus have incomplete knowledge of the state of an individual particle [2].

Given this peculiar situation, one of the earliest puzzles in quantum information science was to devise a quantum generalization of the concept of conditional entropy. In the classical case, conditional entropy is meant to capture the uncertainty about the state of one particle given access to a second particle that is potentially statistically dependent on the first [3]. If there is no statistical dependence whatsoever, then the conditional entropy reduces to the usual entropy. However, if there is dependence, then there is a reduction in the uncertainty of the first particle when given access to the second. When generalizing the definition of conditional entropy in a straightforward way to the quantum case (by means of this reduction), one finds that quantum conditional entropy is negative when evaluated for the EPR state mentioned above [4]. Since conditional entropy is never negative in classical information science, this situation represents a radical departure from the classical case and has been popularly described as "knowing less than nothing" [5].

In quantum information theory [6, 7, 8, 9, 10], the traditional approach to resolving such puzzles is to devise a physically meaningful task for which an information quantity is an optimal rate for accomplishing the task. For the case of quantum conditional entropy, this approach was successfully applied and led to the introduction of the quantum state merging protocol [11, 12]. In this protocol, the goal is for one party A to use entanglement and classical communication to transfer her share A of a joint state to another party B. If the conditional entropy $H(A|B)$ is non-negative, then the protocol consumes entanglement to accomplish the task, at a rate given by $H(A|B)$ EPR states per copy of the original state. Whereas, if the conditional entropy $H(A|B)$ is negative, then the protocol generates entanglement at a rate of $H(A|B)$ EPR states per copy of the original state, while transferring A to B. This protocol inspired further developments regarding conditional entropy in the context of the uncertainty principle [13] and thermodynamics [14], in which its negativity plays an essential role.

Although these works brought great insight to the peculiar situation mentioned above, they mainly focus on a particular definition of the conditional entropy, based on the formula from [4]. It is also well known that there are an infinite number of ways to define quantum conditional entropy, with particular examples being the conditional min-entropy [15], conditional max-entropy [16], conditional Rényi entropies [17, 18, 19], and those based more broadly on generalized divergences (see, e.g., [10]). Several of these alternate definitions have physical meanings as well [16, 20]. However, all of these definitions are based on particular mathematical formulas for quantum conditional entropy that hitherto have not been derived on the basis of simple physical principles. As such, the aforementioned contributions do not address the question of why any plausible quantum conditional entropy can be negative when defining it in the most general way possible, which is consistent with some simple physical principles.

1.1 Summary of results

In this paper, it is our aim to address this pressing foundational question, and we do so by introducing a physically motivated framework for defining quantum conditional entropy, based on two simple postulates. The two postulates are related to the second law of thermodynamics and the extensivity of entropy, widely accepted in scenarios of physical interest. This novel framework is essentially the simplest way of defining any plausible quantum conditional entropy, and one of our main conclusions is that any such quantum conditional entropy must take on a negative value when evaluated for the EPR state. Following the popular terminology introduced in [5], we can summarize this finding as the "inevitability of knowing less than nothing" in the quantum world. As such, our approach answers the aforementioned question, i.e., why the quantum conditional entropy can be negative. We additionally note here that this finding is in stark contrast to the conclusion, also presented here, that any plausible unconditional quantum entropy takes on a non-negative value for all quantum states (here quantum entropy is defined by a similar physically motivated framework related to that from [21, 22] and that generalizes the classical case in [23]).

One may wonder which of the two postulates is responsible for conditional entropy becoming negative in the quantum world. Interestingly, we identify that it is the extensivity postulate that does so. To arrive at this conclusion, we provide an example of a non-negative

measure of conditional entropy that satisfies the monotonicity postulate of conditional entropy (i.e., the one related to the second law) and exhibits weak additivity for tensor-product states. This measure, essentially a regularized version of the "measured" conditional entropy, does not exhibit full additivity under arbitrary tensor-product states. This distinction is important, as it highlights that fully quantum additivity under tensor-product states commands negativity. See Appendix H for further details.

As additional findings, we show that all plausible quantum conditional entropies cannot be smaller than the quantum conditional min-entropy, thus justifying once and for all the name of the latter quantity as the smallest plausible quantum conditional entropy. We also establish a logical equivalence between the non-negativity of all quantum conditional entropies and the well known reduction criterion [24] from entanglement theory. As a corollary, it follows that all separable states have non-negative quantum conditional entropy.

In what follows, we present details of our claims. We first generalize the classical entropy framework from [23] in order to define quantum entropy, and then we introduce our framework for defining quantum conditional entropy. After that, we indicate several properties of quantum conditional entropy that follow from the two postulates, before giving arguments for our main results. All arguments for our main results are based on constructions of physical processes that respect the first aforementioned postulate for quantum conditional entropy. The only background needed to understand the rest of our paper is basic knowledge of density operators and quantum channels (physical processes), available in various textbooks on quantum information [6, 7, 8, 9, 10].

1.2 Notation

Throughout our paper, we employ standard notation used in quantum information [6, 7, 8, 9, 10]. We use symbols like ρ, σ, ω to refer to quantum states (density operators) and $\mathfrak{D}(A)$ to denote the set of density operators acting on a Hilbert space A. We also identify a quantum system A with its corresponding Hilbert space A. Let $\mathbf{u}_A \coloneqq I_A/|A|$ denote the uniform state of a system A, where I_A is the identity operator acting on A and $|A|$ denotes the dimension of A. We use system labels to denote bipartite states as $\rho_{AB}, \sigma_{AB}, \omega_{AB} \in \mathfrak{D}(AB)$, and $\rho_A \equiv \mathrm{Tr}_B[\rho_{AB}]$ denotes the marginal state after performing a partial trace over system B. A quantum channel $\mathcal{N}_{A\to B}$ is a completely positive and trace-preserving map that takes an input system A to an output system B. If the input and output systems are the same, then we simply write $\mathcal{N}_A$. A particular kind of quantum channel is an isometric channel, typically denoted by $\mathcal{U}_{A\to B}$ or $\mathcal{V}_{A\to B}$, and is realized as $\mathcal{U}_{A\to B}(\cdot) = U_{A\to B}(\cdot)(U_{A\to B})^\dagger$, where $U_{A\to B}$ an isometry (i.e., it satisfies $(U_{A\to B})^\dagger U_{A\to B} = I_A$).

2 Uncertainty and entropy

2.1 Uncertainty measures

Let us first define what it means for a function to be an uncertainty measure, following the approach of [23] for the classical case. What we expect for any measure of uncertainty is that it does not decrease under the action of a mixing operation. Thus, we demand that an uncertainty measure is a function that is monotone with respect to a set of mixing operations, similar to how one variant of the second law of thermodynamics states that the entropy of an isolated system left to spontaneous evolution does not decrease. This approach is also consistent with that taken in quantum resource theories when defining resource monotones [25], in turn inspired by the second law of thermodynamics. See also [26, 27].

In the classical case of finite and discrete probability distributions, such a mixing operation is described by a random relabeling of values. For example, for a six-sided die described by a probability distribution, one can randomly relabel or permute the numbers on the faces of the die to realize a mixing operation, and the outcome of the die toss becomes more difficult to guess (and intuitively, more uncertain). Furthermore, the odds of winning a game of chance are lower after performing a mixing operation [28], providing a precise way of measuring uncertainty.

Under such mixing operations, the uniform distribution remains invariant. Furthermore, our intuition is that the uniform distribution should take the largest value of an uncertainty measure for all probability distributions of the same size, given that it is an unbiased distribution with equally likely outcomes and thus the unique distribution for which it is the most difficult to guess which outcome will occur (more generally, for which one is the least likely to win a game of chance [28]). Given everything stated above, one thus defines a function to be an uncertainty measure in the classical case if it is not equal to the zero function (i.e., a function that trivially takes all inputs to the value zero) and if it does not decrease under the action of a mixing operation [23].

In order to generalize this notion to the quantum case, the key previously stated property that we use to define a mixing operation is the preservation of the uniform state $\mathbf{u}$. In this way, it is guaranteed that a mixing operation does not increase the value of an uncertainty measure for the uniform state; rather, it remains

constant. Indeed, such a state is already intuitively maximally uncertain, as argued above, and so it cannot be any more uncertain, among all states of a fixed dimension. At the very least, the mixing operations we consider for defining an uncertainty measure should be quantum channels because all physical processes are described by quantum channels. The additional property of preserving the uniform state thus constrains the mixing operations to be unital channels (i.e., channels that preserve the identity operator). Formally, recall that a channel $\mathcal{N}_A$ is unital if $\mathcal{N}_A(I_A) = I_A$, or equivalently, if $\mathcal{N}_A(\mathbf{u}_A) = \mathbf{u}_A$ [6, 7, 8, 9, 10]. We thus define a function of quantum states to be an uncertainty measure if it is not equal to the zero function and if it does not decrease under the action of a unital channel $\mathcal{N}_A$, i.e.,

$$f(\mathcal{N}_A(\rho_A)) \geq f(\rho_A) \tag{1}$$

for every state ρ_A and unital channel $\mathcal{N}_A$.

A state ρ_A majorizes another state σ_A, denoted by $\rho_A \succ \sigma_A$, if the following inequalities hold for all $k \in \{1, \ldots, |A|\}$:

$$\sum_{i=1}^{k} \lambda_i^{\downarrow}(\rho_A) \geq \sum_{i=1}^{k} \lambda_i^{\downarrow}(\sigma_A), \tag{2}$$

where $\lambda_i^{\downarrow}(\omega_A)$ denotes the components of the vector of eigenvalues of a density operator ω_A sorted in decreasing order. It is well known that ρ_A majorizes σ_A if and only if there exists a unital channel $\mathcal{N}_A$ such that $\sigma_A = \mathcal{N}_A(\rho_A)$ [29, 30, 31] (see also [8, Theorem 4.32]). Thus, our definition of an uncertainty measure f implies that f reverses the majorization pre-order, i.e.,

$$\rho_A \succ \sigma_A \quad \Rightarrow \quad f(\rho_A) \leq f(\sigma_A). \tag{3}$$

It is ideal and natural to generalize the definition of an uncertainty measure just slightly in order to compare the uncertainties of states of different dimensions. Suppose that we have a state ρ_A with an uncertainty value given by $f(\rho_A)$ for some uncertainty measure f. Now consider the density matrix $\rho_{A'} \coloneqq \rho_A \oplus \mathbf{0}$, where $\mathbf{0}$ denotes the zero square matrix, so that $\rho_{A'}$ represents an embedding of ρ_A into a larger system A'. There is no reason that the uncertainty measure should change at all under this embedding, because, intuitively, the extra matrix entries corresponding to $\mathbf{0}$ neither increase nor decrease uncertainty. Indeed, when flipping a die on a table, one could add an extra degree of freedom corresponding to whether the table is upside down or right side up, and under normal circumstances, the table will never be upside down, so that the uncertainty of the die toss will be unaffected by this embedding. As such, we impose the additional constraint that the uncertainty measure f satisfies the equality $f(\rho_A) = f(\rho_{A'})$ for every state ρ_A and every embedding into a larger system A'.

In this spirit, we now generalize the standard definition of majorization for quantum states from [29, 30, 31] (see also [8]).

Definition 1 (Majorization) *Given two states ρ_A and $\sigma_{A'}$, we say that ρ_A majorizes $\sigma_{A'}$, and write*

$$\rho_A \succ \sigma_{A'} \tag{4}$$

if there exist a system A'' such that $|A''| \geq \max\{|A|, |A'|\}$, isometric channels $\mathcal{V}_{A\to A''}$ and $\mathcal{U}_{A'\to A''}$, and a unital channel $\mathcal{N}_{A''}$ such that

$$\mathcal{U}_{A'\to A''}(\sigma_{A'}) = (\mathcal{N}_{A''} \circ \mathcal{V}_{A\to A''})(\rho_A)\,. \tag{5}$$

The isometric channels introduced in the definition of majorization above enable us to compare states of different dimensions, as desired. Also, note that majorization between two probability vectors of the same dimension can be extended to vectors of different dimensions in this way (it is a standard approach to add zero components to the vector with the smaller dimension so that the resulting vectors have equal dimension). As discussed in the previous paragraph, such an extension is motivated by the fact that uncertainty measures should not change under embeddings.

In conclusion, our final definition of an uncertainty measure is a function f that is not equal to the zero function and reverses the majorization pre-order from (4); i.e., for all states ρ_A and $\sigma_{A'}$,

$$\rho_A \succ \sigma_{A'} \quad \Rightarrow \quad f(\rho_A) \leq f(\sigma_{A'}). \tag{6}$$

2.2 Entropy

Following the approach of [23] for the classical case, an uncertainty measure is not necessarily an entropy. To be an entropy, we also require that the uncertainty measure possess an extensivity or additivity property. That is, f is an entropy if it is an uncertainty measure and it is additive on tensor-product states; i.e., for all states ρ_A and $\sigma_{A'}$,

$$f(\rho_A \otimes \sigma_{A'}) = f(\rho_A) + f(\sigma_{A'}). \tag{7}$$

The additivity postulate of entropy is motivated by the basic expectation that the disorder of isolated, non-interacting systems should be the sum of the individual disorders. Indeed, if two physical systems are separate from each other, not having interacted in any way whatsoever previously, we expect that the disorder characterizing the overall system should simply be the sum of the individual disorders (i.e., there would not be a

meaningful alternative way of quantifying the overall disorder in this scenario).

The motivation for additivity of entropy is also related to the second law of thermodynamics, in particular, the Clausius and Kelvin–Planck statements. In those statements, cyclic processes are considered, in which a system of interest undergoes a thermodynamic transition, while all other systems, including the environment, heat baths, etc., all return to their original state. In recent developments, using a quantum information-theoretic approach to small-scale thermodynamics [32], these cyclic processes were identified as catalytic processes. That is, consider a thermodynamical evolution of a physical system A, in a state ρ_A, to a physical system A' in a state $\sigma_{A'}$. The thermal machine that corresponds to this transition includes all the other systems (e.g., heat baths, environment, etc.) that can be represented with an additional system M in a state τ_M. Therefore, for cyclic processes, the thermodynamical transition takes the form

$$\rho_A \otimes \tau_M \to \sigma_{A'} \otimes \tau_M \, . \tag{8}$$

In this context, the second law not only states that the entropy of system A is no greater than that of system A', but also that this property holds if and only if the entropy of the joint state $\rho_A \otimes \tau_M$ increases (or remains unchanged) in such a thermodynamic cyclic process that returns τ_M intact. Clearly, if the entropy is measured with an additive function (under tensor products) then the entropy of ρ_A is never greater than that of $\sigma_{A'}$ if and only if the same relation holds between $\rho_A \otimes \tau_M$ and $\sigma_{A'} \otimes \tau_M$. Thus, the second law of thermodynamics also motivates the additivity postulate for entropy.

With all of this reasoning in place, we now provide a formal definition of entropy:

Definition 2 (Entropy) *A function*

$$\mathbf{H} : \bigcup_A \mathfrak{D}(A) \to \mathbb{R} \tag{9}$$

is a quantum entropy if it is not equal to the zero function and satisfies the following two postulates for all systems A and A' and states ρ_A and $\sigma_{A'}$:

1. *Monotonicity:*

$$\rho_A \succ \sigma_{A'} \quad \Rightarrow \quad \mathbf{H}(\rho_A) \leq \mathbf{H}(\sigma_{A'}). \tag{10}$$

2. *Additivity:*

$$\mathbf{H}(\rho_A \otimes \sigma_{A'}) = \mathbf{H}(\rho_A) + \mathbf{H}(\sigma_{A'}). \tag{11}$$

As stated previously, the requirement that an entropy not be equal to the zero function is to avoid a trivial situation in which the entropy of every state takes on the value zero.

By employing the two basic postulates of entropy and the same reasoning given in [23, Lemma 3], some important properties follow (see Appendix A for proofs). The entropy of every state ρ is non-negative, i.e., $\mathbf{H}(\rho) \geq 0$, and it is equal to zero for every pure state $|\psi\rangle\langle\psi|$, i.e., $\mathbf{H}(|\psi\rangle\langle\psi|) = 0$. For every state σ of a fixed dimension d, the entropy does not exceed the entropy of the uniform state of dimension d, i.e., $\mathbf{H}(\mathbf{u}^{(d)}) \geq \mathbf{H}(\sigma)$. Finally, the entropy of the uniform state of dimension two is strictly positive, i.e., $\mathbf{H}(\mathbf{u}^{(2)}) > 0$. This allows for setting the normalization of entropy to be in units of bits, such that $\mathbf{H}(\mathbf{u}^{(2)}) = 1$, and henceforth, we thus set the normalization of entropy in this way. It then follows that $\mathbf{H}(\mathbf{u}^{(d)}) = \log_2 d$.

3 Conditional uncertainty and entropy

3.1 Conditional uncertainty measures

Our goal now is to define what it means for a function of a bipartite state ρ_{AB} to be a measure of conditional uncertainty. Inspired by the criterion of monotonicity under mixing operations in the previous case of unconditional uncertainty, we are interested in identifying a criterion for a bipartite channel $\mathcal{N}_{AB\to A'B'}$ to be a conditionally mixing operation. For the unconditional case, we ultimately stated that a quantum channel is a mixing operation if it preserves the uniform state, and this choice led to identifying unital channels as the mixing operations that define an uncertainty measure. The reason for this, as previously stated, is that the uniform state is the least informative state, it thus gives the least likelihood of winning a game of chance, and should thus be a state of maximal uncertainty (for which a mixing operation cannot increase uncertainty).

For the setting of conditional uncertainty, it remains the case that the uniform state $\mathbf{u}_{AB} = \mathbf{u}_A \otimes \mathbf{u}_B$ is the least informative state, for reasons similar to those given before. However, for conditional uncertainty, we are interested in the uncertainty of system A in the presence of system B. That is, the system B can be used as an aid or side information to help win a game of chance performed on the system A. With this in mind, a state of the form $\mathbf{u}_A \otimes \rho_B$ is still least informative, with ρ_B an arbitrary state, because the system B is independent of system A and thus cannot help at all in winning a game of chance performed on A. At the same time, the state of system A is still uniform and thus the least informative. As such, we identify the set of states having the form $\mathbf{u}_A \otimes \rho_B$, where ρ_B is an arbitrary state of system B, as being the set of states with maximum conditional uncertainty.

Thus, we have argued that one requirement for a function to be a measure of conditional uncertainty is that it should not decrease under the action of a bipartite channel that preserves the set of states with maximum conditional uncertainty. We call such channels conditionally unital channels because they generalize unital channels. Formally, a bipartite channel $\mathcal{N}_{AB\to AB'}$ is a conditionally unital channel if for every $\rho_B \in \mathfrak{D}(B)$, there exists a state $\sigma_{B'} \in \mathfrak{D}(B')$ such that

$$\mathcal{N}_{AB\to AB'}(\mathbf{u}_A \otimes \rho_B) = \mathbf{u}_A \otimes \sigma_{B'}. \tag{12}$$

See Appendix B.1 for an equivalent formulation of this condition in terms of the Choi matrix of $\mathcal{N}_{AB\to AB'}$. We note here that conditionally unital channels were independently defined in our companion paper [33] and in [34]. See also [15, Lemma 3.1.12] for an implicit definition of sub-conditionally-unital channels.

In the setting of conditional uncertainty, it is important to consider a further constraint on the class of transformations allowed. Intuitively, conditional uncertainty measures the uncertainty of system A when one is given access to system B. Evidently, by allowing transformations in which information can leak from system A to system B, Bob could gain information that could decrease the uncertainty associated with system A. Thus, it is natural to assume that channels that do not decrease conditional uncertainty do not allow for communication or signaling from Alice to Bob. This requirement is also motivated from a cryptographic perspective, in which conditional uncertainty measures the amount of uncertainty that an eavesdropper has about another system that is intended to be secure [35] (with respect to our system labeling, the eavesdropper would have system B and the secure system would be A). With this motivation in mind, let us recall the definition of $A \not\to B$ semi-causal channels [36]. A bipartite channel $\mathcal{N}_{AB\to AB'}$ is $A \not\to B'$ semi-causal if for every channel $\mathcal{M}_A$

$$\mathcal{N}_{AB\to B'} \circ \mathcal{M}_A = \mathcal{N}_{AB\to B'}, \tag{13}$$

where

$$\mathcal{N}_{AB\to B'} \coloneqq \mathrm{Tr}_A \circ \mathcal{N}_{AB\to AB'}. \tag{14}$$

See Figure 1 for a depiction of the semi-causality requirement.

The authors of [37] proved that every semi-causal channel can be written as a local channel on B that feeds an output (corresponding to a quantum system R) into a channel acting only on A and R. See also [38] and Figure 2 for a visual depiction. As such, for these channels, information can only flow from B to A, and it never leaks from A to B. Formally, if $\mathcal{N}_{AB\to AB'}$ is an $A \not\to B'$ semi-causal channel, then there exists a

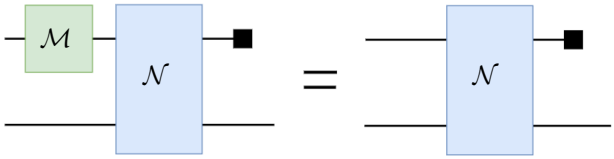

Figure 1: Depiction of the definition of a semi-causal channel $\mathcal{N}$, where the black square represents the discarding channel.

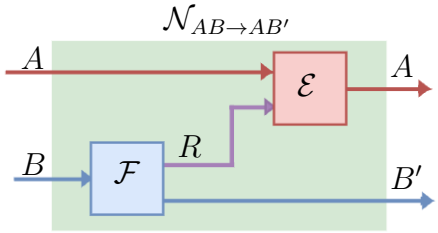

Figure 2: Eq. (15) asserts that a semi-causal bipartite channel $\mathcal{N}_{AB\to AB'}$ can be implemented in the depicted fashion.

quantum system R, a quantum channel $\mathcal{E}_{AR\to A}$, and an isometric channel $\mathcal{F}_{B\to RB'}$ such that

$$\mathcal{N}_{AB\to AB'} = \mathcal{E}_{RA\to A} \circ \mathcal{F}_{B\to RB'}\,. \tag{15}$$

We use the term conditionally mixing to describe a channel $\mathcal{N}_{AB\to AB'}$ that is both conditionally unital and $A \not\to B'$ semi-causal. Note that the properties of such a channel are invariant under the exchange of the A input and A output systems. This is most readily observed using the formalism of Choi matrices and we discuss it further in Appendix B.2. Importantly, we use conditionally mixing channels to define measures of conditional uncertainty.

Generalizing how unital channels define the majorization pre-order on quantum states, we now define a conditional majorization pre-order on bipartite quantum states, based on conditionally mixing channels.

Definition 3 (Conditional majorization) *Given two bipartite states ρ_{AB} and $\sigma_{A'B'}$, we say that ρ_{AB} conditionally majorizes $\sigma_{A'B'}$ with respect to system A, and write*

$$\rho_{AB} \succ_A \sigma_{A'B'} \tag{16}$$

if there exist a system A'' such that $|A''| \geq \max\{|A|, |A'|\}$, isometric channels $\mathcal{V}_{A\to A''}$ and $\mathcal{U}_{A'\to A''}$, and a conditionally mixing channel $\mathcal{N}_{A''B\to A''B'}$ such that

$$\mathcal{U}_{A'\to A''}(\sigma_{A'B'}) = (\mathcal{N}_{A''B\to A''B'} \circ \mathcal{V}_{A\to A''})(\rho_{AB})\,. \tag{17}$$

The isometric channels introduced in the definition above enable us to compare bipartite states of different dimensions, as was the case before with the generalized definition of majorization in (4)–(5). We note

here that conditional majorization, as defined here, reduces to classical conditional majorization, as defined in [39], when the the states involved are classical (see Appendix C for a proof). Specifically, we show that in the classical domain, conditionally mixing channels coincide with the conditionally doubly stochastic (CDS) channels that were defined in [39].

We define a function f to be a conditional uncertainty measure if it is not equal to the zero function and if it reverses the conditional majorization pre-order, i.e.,

$$\rho_{AB} \succ_A \sigma_{A'B'} \quad \Rightarrow \quad f(\rho_{AB}) \leq f(\sigma_{A'B'}). \tag{18}$$

3.2 Conditional entropy

Having defined measures of conditional uncertainty, we now follow the approach from before and define a function to be a conditional entropy if it is not only a conditional uncertainty measure but also additive on tensor-product states. Our motivation for the additivity postulate is the same as before.

Definition 4 (Conditional entropy) *A function*

$$\mathbf{H} : \bigcup_{A,B} \mathfrak{D}(AB) \to \mathbb{R} \tag{19}$$

is a quantum conditional entropy if it is not equal to the zero function when $|B| = 1$ and satisfies the following postulates for all systems A, B, A', B' and states ρ_{AB} and $\sigma_{A'B'}$:

1. *Monotonicity:*

$$\rho_{AB} \succ_A \sigma_{A'B'} \quad \Rightarrow \quad \mathbf{H}(A|B)_\rho \leq \mathbf{H}(A'|B')_\sigma. \tag{20}$$

2. *Additivity:*

$$\mathbf{H}(AA'|BB')_{\rho\otimes\sigma} = \mathbf{H}(A|B)_\rho + \mathbf{H}(A'|B')_\sigma. \tag{21}$$

As indicated in the definition, a conditional entropy $\mathbf{H}$ is defined for all density matrices in all finite dimensions. As a very special case, if $\rho \in \mathfrak{D}(AB)$ and B is a trivial system for which $|B| = 1$, then we write

$$\mathbf{H}(A|B)_\rho = \mathbf{H}(A)_\rho \,, \tag{22}$$

where $\mathbf{H}(A)_\rho$ is an entropy, as defined before.

Here we state some basic properties of quantum conditional entropy before presenting our main result in the next section (see Appendix D for proofs). First, conditional entropy reduces to entropy when evaluated on a tensor-product state. That is, for every product state $\omega_A \otimes \tau_B$, the following equality holds:

$$\mathbf{H}(A|B)_{\omega\otimes\tau} = \mathbf{H}(A)_\omega. \tag{23}$$

Next, conditional entropy is invariant under the action of local isometric channels. That is, for a bipartite state ρ_{AB} and isometric channels $\mathcal{U}_{A\to A'}$ and $\mathcal{V}_{B\to B'}$,

$$\mathbf{H}(A|B)_\rho = \mathbf{H}(A'|B')_\omega, \tag{24}$$

where $\omega_{AB'} \coloneqq (\mathcal{U}_{A\to A'} \otimes \mathcal{V}_{B\to B'})(\rho_{AB})$. Finally, consider the product state $\mathbf{u}_A^{(2)} \otimes \rho_B$. Similar to what we found previously for unconditional entropy, the following inequality holds

$$\mathbf{H}(A|B)_{\mathbf{u}^{(2)}\otimes\rho} = \mathbf{H}(A)_{\mathbf{u}^{(2)}} > 0. \tag{25}$$

This strict inequality allows us to set a normalization factor for conditional entropy. To be consistent with the normalization convention for unconditional entropy, we set $\mathbf{H}(A|B)_{\mathbf{u}^{(2)}\otimes\rho} = 1$, which in turn implies that

$$\mathbf{H}(A|B)_{\mathbf{u}^{(d)}\otimes\rho} = \log_2 d, \tag{26}$$

where the state being evaluated is $\mathbf{u}_A^{(d)} \otimes \rho_B$.

4 Main result: Inevitability of negative quantum conditional entropy

We now have almost everything set up to state our main result. Before doing so, let us define the conditional min-entropy [15] of a bipartite state ρ_{AB} as

$$H_{\min}(A|B)_\rho \coloneqq -\inf_{\lambda\geq 0} \log_2\{\lambda : \rho_{AB} \leq \lambda I_A \otimes \rho_B\}, \tag{27}$$

The conditional min-entropy is indeed a conditional entropy according to (20)–(21) (see Appendix F.1 for a proof). This quantity was previously given the name conditional min-entropy because it is known to be the least among all Rényi conditional entropies [15, 17, 18]. As part of our main result, we strengthen this observation by proving that all plausible quantum conditional entropies are not smaller than the conditional min-entropy; that is, the conditional min-entropy is the smallest plausible conditional entropy.

The following theorem (proved in Appendix E.2) states that every plausible quantum conditional entropy is bounded from below by the conditional min-entropy and also that this lower bound is saturated for the maximally entangled state. As a consequence, every plausible conditional entropy takes on a negative value for the maximally entangled state.

Theorem 1 *Let $\mathbf{H}$ be a quantum conditional entropy, as defined in (19)–(21). Then, for every state ρ_{AB},*

$$\mathbf{H}(A|B)_\rho \geq H_{\min}(A|B)_\rho. \tag{28}$$

Equality is attained in (28) if ρ_{AB} is equal to a maximally entangled state $\Phi^{(k)}_{A'B'}$ by the action of local isometric channels. Thus,

$$\mathbf{H}(A|B)_{\Phi^{(k)}} = -\log_2 k \tag{29}$$

for every conditional entropy $\mathbf{H}$.

The following statement is a direct consequence of Theorem 1, indicating that every plausible conditional entropy being non-negative is equivalent to the reduction criterion [24], well known in entanglement theory. Thus, Corollary 1 provides a direct link between every conditional entropy (including the conditional min-entropy) and the reduction criterion.

Corollary 1 *Let ρ_{AB} be a bipartite state. Then the following are equivalent:*

1. *For every conditional entropy* $\mathbf{H}$,

$$\mathbf{H}(A|B)_\rho \geq 0. \tag{30}$$

2. *The reduction criterion holds: $I_A \otimes \rho_B \geq \rho_{AB}$.*

Proof. The proof of Corollary 1 is immediate from Theorem 1. Indeed, if the first condition holds, then the inequality $H_{\min}(A|B)_\rho \geq 0$ holds in particular (since $H_{\min}(A|B)_\rho$ is a conditional entropy), and the second condition is then a consequence of the definition in (27). Now suppose that the second condition holds. This then means that there exists $\lambda \leq 1$ such that $\rho_{AB} \leq \lambda I_A \otimes \rho_B$. According to the definition in (27), it follows that $H_{\min}(A|B)_\rho \geq 0$ and then Theorem 1 implies the first condition. ■

Suppose that a state ρ_{AB} is separable [40], meaning that it can be written as a probabilistic mixture of product states (i.e., $\rho_{AB} = \sum_x p(x)\sigma^x_A \otimes \tau^x_B$ for some probability distribution $\{p(x)\}_x$ and sets $\{\sigma^x_A\}_x$ and $\{\tau^x_B\}_x$ of states). Then it follows that the second condition in Corollary 1 holds [24], so that every conditional entropy is non-negative for all separable states.

Although all plausible conditional entropies are non-negative on all separable states, this does not mean that the conditional entropies are only non-negative on separable states. In fact, some entangled states have non-negative conditional entropy on *all* choices of conditional entropy functions. In light of this, Corollary 1 is helpful because it provides a simple sufficient criterion for a bipartite state to have non-negative quantum conditional entropy.

In Appendix E, we discuss a subclass of conditional entropies and report results that are similar to those given in Theorem 1 and Corollary 1. The difference is that the lower bound in (28) is expressed in terms of a different definition of conditional min-entropy and the equivalent condition in Corollary 1 becomes a variation of the reduction criterion, rather than the reduction criterion itself.

In addition to the inequality (28), one can obtain other entropy inequalities that are based on the formalism developed here. For example, in Appendix G we show that for every bipartite state ρ_{AB},

$$\mathbf{H}(AB)_\rho \geq \mathbf{H}(A|B)_\rho - \log_2 |B| \,. \tag{31}$$

This inequality follows from the fact that the teleportation protocol of communicating a quantum system from Bob to Alice is a conditionally mixing channel (see Appendix G for more details and other entropy inequalities).

5 Conditional entropy from relative entropy

In this section, we prove that one can construct a quantum conditional entropy from a quantum relative entropy, based on two simple postulates that define a quantum relative entropy [23, 41]. Before giving this construction, let us first recall the definition of a quantum relative entropy [23, 41].

Definition 5 (Relative entropy) *A function*

$$\mathbf{D} : \bigcup_A \{\mathfrak{D}(A) \times \mathfrak{D}(A)\} \to \mathbb{R}_+ \cup \{\infty\} \tag{32}$$

is a quantum relative entropy if it is not equal to the zero function and satisfies the following postulates for all systems A and B, all states ρ_A, σ_A, ω_B, τ_B, and every channel $\mathcal{N}_{A\to B}$:

1. *Monotonicity:*

$$\mathbf{D}(\rho_A\|\sigma_A) \geq \mathbf{D}(\mathcal{N}_{A\to B}(\rho_A)\|\mathcal{N}_{A\to B}(\sigma_A)). \tag{33}$$

2. *Additivity:*

$$\mathbf{D}(\rho_A \otimes \omega_B\|\sigma_A \otimes \tau_B) = \mathbf{D}(\rho_A\|\sigma_A) + \mathbf{D}(\omega_B\|\tau_B). \tag{34}$$

Now we state our final result (see Appendix F for a proof):

Theorem 2 *Let* $\mathbf{D}$ *be a quantum relative entropy. Then the function*

$$\log_2|A| - \mathbf{D}(\rho_{AB}\|\mathbf{u}_A \otimes \rho_B) \tag{35}$$

is a quantum conditional entropy.

Suitable relative entropies include the Umegaki (standard) [42], Belavkin–Staszewski [43], Petz–Rényi [44, 45], sandwiched Rényi [18, 46], α-z [47], and geometric Rényi [48, 49] relative entropies, for which we point to [10] for a summary. The construction in (35), for generating a conditional entropy from a relative entropy, has been known for some time, but the difference with our result here is that we prove, starting from the simple postulates for relative entropy in (32)–(34), that the function in (35) satisfies the two simple postulates for conditional entropy in (19)–(21).

6 Conclusion

In summary, we have established a physically motivated, minimal framework for defining conditional entropies, based on two simple postulates (non-decrease under conditionally mixing channels and additivity on tensor-product states). We then explored consequences of these postulates, which include reduction to entropy for tensor-product states, invariance under local isometric channels, a lower bound in terms of the conditional min-entropy, and taking on a negative value for the maximally entangled state. We consider the latter property to be the most profound result of this work, and all of our arguments rely on constructing physical processes that respect the first postulate of non-decrease under conditionally mixing channels.

Returning to the point raised in the second paragraph of Section 1.1, one may wonder if it might be possible to replace the additivity postulate of conditional entropy with a weaker criterion. For example, suppose we replace the additivity postulate (21) with a weaker postulate of the form

$$\mathbf{H}(A^n|B^n)_{\rho^{\otimes n}} = n\mathbf{H}(A|B)_\rho \quad \forall n \in \mathbb{N}. \tag{36}$$

In Appendix H we show that there exist examples of functions that satisfy both the monotonicity postulate and the above weak additivity, but take a non-negative value on *all* states. Therefore, the negativity of conditional-entropies (as defined in this paper) on the maximally entangled state is closely related to the (strong) additivity postulate (21).

Every relative entropy $\mathbf{D}$ can be used as in (35) to define a conditional entropy. One can also attempt to define another type of conditional entropy of the form

$$\log_2 |A| - \min_{\sigma \in \mathfrak{D}(B)} \mathbf{D}(\rho_{AB} \| \mathbf{u}_A \otimes \sigma_B) \,. \tag{37}$$

While the function above behaves monotonically under conditionally unital channels, and therefore satisfies the monotonicity postulate (20) it is not clear if this function is additive under tensor products. It is known to be additive if $\mathbf{D}$ is taken to be the Petz Rényi relative entropy [50, Lemma 3] or the sandwiched Rényi relative entropy [51], but it is not known if it is additive in general.

From its definition, the set of all conditional entropies is convex. Hence, not every conditional entropy has to take the form of a difference of logarithmic dimension and divergence. For example, if $\mathbf{D}$ is one of the quantum Rényi relative entropies, then every convex combination of the functions (35) and (37) is also a conditional entropy. It is an interesting open problem if there exists conditional entropies that cannot be expressed in terms of divergences. Another interesting open problem is whether or not all quantum conditional entropies are upper bounded by the conditional max-entropy. We leave these open problems to future investigations.

Acknowledgments. The authors thank Nilanjana Datta, Kaiyuan Ji, and Henry Pfister for helpful discussions and Jonathan Sorce for feedback on our manuscript. GG and IG acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC). SB acknowledges support from the National Science Foundation (NSF) under Grant Nos. 1908730 and 1910571. MMW acknowledges support from the NSF under Grant No. 1907615. Any opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of these sponsors.

Author Contributions. The following describes the different contributions of all authors of this work, using roles defined by the CRediT (Contributor Roles Taxonomy) project [52]:

GG: Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing - original draft, Writing - review & editing.

MMW: Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing.

SB: Conceptualization, Visualization

IJG: Conceptualization

References

[1] Albert Einstein, Boris Podolsky, and Nathan Rosen. "Can quantum-mechanical description of physical reality be considered complete?". Physical Review **47**, 777–780 (1935).

[2] Erwin Schrödinger. "Discussion of probability relations between separated systems". Proceedings of the Cambridge Philosophical Society **31**, 555–563 (1935).

[3] Claude E. Shannon. "A mathematical theory of communication". The Bell System Technical Journal **27**, 379–423 (1948).

[4] Nicolas J. Cerf and Christoph Adami. "Negative entropy and information in quantum mechanics". Physical Review Letters **79**, 5194–5197 (1997).

[5] Ian Turner. "Scientist knows less than nothing" (2005). Screenshot available at https://twitter.com/markwilde/status/1558125303569846278.

[6] Masahito Hayashi. "Quantum information theory: Mathematical foundation". Springer. (2017). Second edition.

[7] Mark M. Wilde. "Quantum information theory". Cambridge University Press. (2017). Second edition.

[8] John Watrous. "The theory of quantum information". Cambridge University Press. (2018).

[9] Alexander S. Holevo. "Quantum systems, channels, information: A mathematical introduction". Walter de Gruyter. (2019). Second edition.

[10] Sumeet Khatri and Mark M. Wilde. "Principles of quantum communication theory: A modern approach" (2024). arXiv:2011.04672v2.

[11] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. "Partial quantum information". Nature **436**, 673–676 (2005).

[12] Michal Horodecki, Jonathan Oppenheim, and Andreas Winter. "Quantum state merging and negative information". Communications in Mathematical Physics **269**, 107–136 (2007).

[13] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. "The uncertainty principle in the presence of quantum memory". Nature Physics **6**, 659–662 (2010).

[14] Lidia del Rio, Johan Aberg, Renato Renner, Oscar Dahlsten, and Vlatko Vedral. "The thermodynamic meaning of negative entropy". Nature **474**, 61–63 (2011).

[15] Renato Renner. "Security of quantum key distribution". PhD thesis. ETH Zurich. (2005). arXiv:quant-ph/0512258.

[16] Robert Koenig, Renato Renner, and Christian Schaffner. "The operational meaning of min- and max-entropy". IEEE Transactions on Information Theory **55**, 4337–4347 (2009).

[17] Marco Tomamichel, Roger Colbeck, and Renato Renner. "A fully quantum asymptotic equipartition property". IEEE Transactions on Information Theory **55**, 5840–5847 (2009).

[18] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. "On quantum Rényi entropies: a new generalization and some properties". Journal of Mathematical Physics **54**, 122203 (2013). arXiv:1306.3142.

[19] Marco Tomamichel, Mario Berta, and Masahito Hayashi. "Relating different quantum generalizations of the conditional Rényi entropy". Journal of Mathematical Physics **55**, 082206 (2014). arXiv:1311.3887.

[20] Masahito Hayashi and Marco Tomamichel. "Correlation detection and an operational interpretation of the Rényi mutual information". Journal of Mathematical Physics **57**, 102201 (2016).

[21] Elliott H. Lieb and Jakob Yngvason. "The physics and mathematics of the second law of thermodynamics". Physics Reports **310**, 1–96 (1999).

[22] Elliott H. Lieb and Jakob Yngvason. "A guide to entropy and the second law of thermodynamics". Pages 353–363. Springer Berlin Heidelberg. Berlin, Heidelberg (2004). arXiv:math-ph/9805005.

[23] Gilad Gour and Marco Tomamichel. "Entropy and relative entropy from information-theoretic principles". IEEE Transactions on Information Theory **67**, 6313–6327 (2021).

[24] Michał Horodecki and Paweł Horodecki. "Reduction criterion of separability and limits for a class of distillation protocols". Physical Review A **59**, 4206–4216 (1999).

[25] Eric Chitambar and Gilad Gour. "Quantum resource theories". Reviews of Modern Physics **91**, 025001 (2019).

[26] Michał Horodecki, Karol Horodecki, Paweł Horodecki, Ryszard Horodecki, Jonathan Oppenheim, Aditi Sen(De), and Ujjwal Sen. "Local information as a resource in distributed quantum systems". Physical Review Letters **90**, 100402 (2003).

[27] Michał Horodecki, Paweł Horodecki, and Jonathan Oppenheim. "Reversible transformations from pure to mixed states and the unique measure of information". Physical Review A **67**, 062104 (2003).

[28] Sarah Brandsen, Isabelle Jianing Geng, and Gilad Gour. "What is entropy? A perspective from games of chance". Physical Review E **105**, 024117 (2022).

[29] Armin Uhlmann. "Sätze über dichtematrizen". Wissenschaftliche Zeitschrift der Karl-Marx-Universitat Leipzig. Mathematisch-naturwissenschaftliche Reihe **20**, 633–653 (1971).

[30] Armin Uhlmann. "Endlich-dimensionale dichtematrizen I". Wissenschaftliche Zeitschrift der Karl-Marx-Universitat Leipzig. Mathematisch-naturwissenschaft-liche Reihe **21**, 421–452 (1972).

[31] Armin Uhlmann. "Endlich-dimensionale dichtematrizen II". Wissenschaftliche Zeitschrift der Karl-Marx-Universitat Leipzig Mathematisch-naturwissenschaft-liche Reihe **22**, 139–177 (1973).

[32] Fernando G. S. L. Brandão, Michal Horodecki, Nelly Ng, Jonathan Oppenheim, and Stephanie Wehner. "The second laws of quantum thermodynamics". Proceedings of the National Academy of Sciences **112**, 3275–3279 (2015).

[33] Sarah Brandsen, Isabelle J. Geng, Mark M. Wilde, and Gilad Gour. "Quantum conditional entropy from information-theoretic principles" (2021). arXiv:2110.15330.

[34] Mahathi Vempati, Saumya Shah, Nirman Ganguly, and Indranil Chakrabarty. "A-unital Operations and Quantum Conditional Entropy". Quantum **6**, 641 (2022).

[35] Christopher Portmann and Renato Renner. "Security in quantum cryptography". Reviews of Modern Physics **94**, 025008 (2022).

[36] David Beckman, Daniel Gottesman, Michael A. Nielsen, and John Preskill. "Causal and localizable quantum operations". Physical Review A **64**, 052309 (2001).

[37] T. Eggeling, D. Schlingemann, and Reinhard F. Werner. "Semicausal operations are semilocalizable". Europhysics Letters **57**, 782–788 (2002).

[38] Marco Piani, Michal Horodecki, Pawel Horodecki, and Ryszard Horodecki. "Properties of quantum nonsignaling boxes". Physical Review A **74**, 012305 (2006).

[39] Gilad Gour, A. Grudka, M. Horodecki, W. Klobus, J. Lodyga, and V. Narasimhachar. "The conditional uncertainty principle". Physical Review A **97**, 042130 (2018).

[40] Reinhard F. Werner. "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model". Physical Review A **40**, 4277–4281 (1989).

[41] Gilad Gour and Marco Tomamichel. "Optimal extensions of resource measures and their applications". Physical Review A **102**, 062401 (2020).

[42] Hisaharu Umegaki. "Conditional expectations in an operator algebra IV (entropy and information)". Kodai Mathematical Seminar Reports **14**, 59–85 (1962).

[43] V. P. Belavkin and P. Staszewski. "C*-algebraic generalization of relative entropy and entropy". Annales de l'I.H.P. Physique théorique **37**, 51–58 (1982). url: http://www.numdam.org/item/AIHPA_1982__37_1_51_0/.

[44] Dénes Petz. "Quasi-entropies for States of a von Neumann Algebra". Publications of the Research Institute for Mathematical Sciences **21**, 787–800 (1985).

[45] Dénes Petz. "Quasi-entropies for finite quantum systems". Reports in Mathematical Physics **23**, 57–65 (1986).

[46] Mark M. Wilde, Andreas Winter, and Dong Yang. "Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy". Communications in Mathematical Physics **331**, 593–622 (2014).

[47] Koenraad M. R. Audenaert and Nilanjana Datta. "α-z-Rényi relative entropies". Journal of Mathematical Physics **56**, 022202 (2015). arXiv:1310.7178.

[48] Keiji Matsumoto. "A new quantum version of f-divergence" (2013). arXiv:1311.4722.

[49] Keiji Matsumoto. "A new quantum version of f-divergence". In Masanao Ozawa, Jeremy Butterfield, Hans Halvorson, Miklós Rédei, Yuichiro Kitajima, and Francesco Buscemi, editors, Reality and Measurement in Algebraic Quantum Theory. Volume 261, pages 229–273. Singapore (2018). Springer Singapore.

[50] Naresh Sharma and Naqueeb Ahmad Warsi. "Fundamental bound on the reliability of quantum information transmission". Physical Review Letters **110**, 080501 (2013).

[51] Salman Beigi. "Sandwiched Rényi divergence satisfies data processing inequality". Journal of Mathematical Physics **54**, 122202 (2013). arXiv:1306.5920.

[52] NISO. "Credit – contributor roles taxonomy". https://credit.niso.org/, Accessed 2024-09-04.

[53] Xin Wang. "Semidefinite optimization for quantum information". PhD thesis. University of Technology Sydney. Centre for Quantum Software and

Information, Faculty of Engineering and Information Technology (2018). url: http://hdl.handle.net/10453/127996.

[54] Gilad Gour. "Comparison of quantum channels by superchannels". IEEE Transactions on Information Theory **65**, 5880–5904 (2019).

[55] Robert Salzmann, Nilanjana Datta, Gilad Gour, Xin Wang, and Mark M. Wilde. "Symmetric distinguishability as a quantum resource". New Journal of Physics **23**, 083016 (2021).

[56] Robert König and Renato Renner. "Sampling of min-entropy relative to quantum knowledge". IEEE Transactions on Information Theory **57**, 4760–4787 (2011).

[57] Gilad Gour. "Resources of the quantum world (Volume 1)" (2024). arXiv:2402.05474v1.

[58] Nilanjana Datta. "Min- and max-relative entropies and a new entanglement monotone". IEEE Transactions on Information Theory **55**, 2816–2826 (2009).

A Proofs of properties of entropy

Let us employ the postulates in (9)–(11) to derive some basic properties of every entropy. The proofs given here are similar to those given for the classical case in [23, Lemma 3]. First, let us recall an often overlooked fact, that the number 1 is a trivial quantum state of dimension one. Indeed, the underlying Hilbert space is of dimension one, so that the only possible density "operator" for this space is the number 1. We now prove that the entropy of this state is equal to zero, as a direct consequence of the additivity postulate. Consider that, for a state ρ,

$$\mathbf{H}(\rho) = \mathbf{H}(\rho \otimes 1) = \mathbf{H}(\rho) + \mathbf{H}(1) \;\Rightarrow\; \mathbf{H}(1) = 0. \quad (38)$$

The first equality follows as a trivial assertion, because $\rho = \rho \otimes 1$. The next equality is a consequence of the additivity postulate. The conclusion that $\mathbf{H}(1) = 0$ then follows by canceling $\mathbf{H}(\rho)$ on both sides of the equality.

It follows that the entropy of every pure state $|\psi\rangle\langle\psi|$ is equal to zero, as a consequence of (38). Indeed, the state vector $|\psi\rangle$ is an isometry from the trivial Hilbert space of dimension one to the Hilbert space that contains $|\psi\rangle$. As such, it follows that $1 \succ |\psi\rangle\langle\psi|$ and $1 \prec |\psi\rangle\langle\psi|$ according to the definitions in (4)–(5) and (9)–(11). So we conclude that

$$\mathbf{H}(|\psi\rangle\langle\psi|) = 0 \quad (39)$$

for every pure state $|\psi\rangle\langle\psi|$.

Next we conclude that the entropy of every state is non-negative. This follows because there exists a unital channel taking a pure state $|\psi\rangle\langle\psi|$ to an arbitrary state ρ, so that $|\psi\rangle\langle\psi| \succ \rho$. Indeed, let $\rho = \sum_x p_x |\psi_x\rangle\langle\psi_x|$ be a pure state decomposition of ρ. Let U_x be a unitary that transforms $|\psi\rangle$ to $|\psi_x\rangle$. Thus, $(\cdot) \to \sum_x p_x U_x (\cdot) U_x^\dagger$ is a unital channel that transforms $|\psi\rangle\langle\psi|$ to ρ. By applying (10) and (39), we conclude that

$$\mathbf{H}(\rho) \geq 0 \quad (40)$$

for every state ρ.

For a fixed dimension d, the uniform state $\mathbf{u}^{(d)} = I/d$ is the state of maximal entropy, which follows from (10) and because every d-dimensional state ρ can be taken to $\mathbf{u}^{(d)}$ by means of the unital channel $(\cdot) \to \mathrm{Tr}[(\cdot)]\mathbf{u}^{(d)}$. Thus,

$$\mathbf{H}(\mathbf{u}^{(d)}) \geq \mathbf{H}(\rho) \quad (41)$$

for every state ρ of dimension d.

Let us prove that $\mathbf{H}(\mathbf{u}^{(2)}) > 0$. By the inequality in (40) and the assumption that entropy is not equal to the zero function, there exists a state ρ of some dimension d such that $\mathbf{H}(\rho) > 0$. By applying (41), we conclude that $\mathbf{H}(\mathbf{u}^{(d)}) \geq \mathbf{H}(\rho)$. By taking n sufficiently large (i.e., such that $2^n > d$), we conclude that $\mathbf{u}^{(d)} \succ (\mathbf{u}^{(2)})^{\otimes n}$, which implies that $\mathbf{H}((\mathbf{u}^{(2)})^{\otimes n}) > \mathbf{H}(\mathbf{u}^{(d)})$. Finally, by applying the additivity postulate, we conclude that $n\mathbf{H}(\mathbf{u}^{(2)}) = \mathbf{H}((\mathbf{u}^{(2)})^{\otimes n})$. Combining all inequalities, we conclude that $\mathbf{H}(\mathbf{u}^{(2)}) > 0$.

After establishing the normalization convention that $\mathbf{H}(\mathbf{u}^{(2)}) = 1$, the proof that $\mathbf{H}(\mathbf{u}^{(d)}) = \log_2 d$ is the same as that given in the proof of [23, Lemma 3].

B Choi matrix formulations

In this appendix, we provide Choi matrix formulations of conditionally unital channels and conditionally mixing channels. These formulations are useful conceptually and also for numerical optimization if one wishes to use semi-definite programming to optimize over these sets of channels. See [8, 53, 10] for examples and discussions of semi-definite programming in quantum information.

Before giving these formulations, let us first recall the definition of the Choi matrix. The Choi matrix $J_{CD}^{\mathcal{M}}$ of a linear map $\mathcal{M}_{C\to D}$ is defined by its action on the maximally entangled operator, as follows:

$$J_{CD}^{\mathcal{M}} \coloneqq \sum_{i,j=1}^{|C|} |i\rangle\langle j| \otimes \mathcal{M}(|i\rangle\langle j|). \quad (42)$$

In general, the matrix $J_{CD}^{\mathcal{M}}$ is the Choi matrix of a trace-preserving map if and only if

$$\mathrm{Tr}_D[J_{CD}^{\mathcal{M}}] = I_C, \quad (43)$$

and it is the Choi matrix of a completely positive map if and only if

$$J^{\mathcal{M}}_{CD} \geq 0. \tag{44}$$

Thus, $J^{\mathcal{M}}_{CD}$ is the Choi matrix of a quantum channel if and only if (43) and (44) both hold.

B.1 Choi matrix formulation of conditional unitality

Here we prove that a conditionally unital channel has an equivalent formulation in terms of its Choi matrix. We then see that a channel being conditionally unital represents a semi-definite constraint on its Choi matrix, which is convenient for optimizing over such channels using semi-definite programming.

Lemma 1 *A channel $\mathcal{N}_{AB\to\tilde{A}B'}$ is conditionally unital if and only if its Choi matrix $J^{\mathcal{N}}_{AB\tilde{A}B'}$ satisfies*

$$J^{\mathcal{N}}_{B\tilde{A}B'} = J^{\mathcal{N}}_{BB'} \otimes \mathbf{u}_{\tilde{A}}, \tag{45}$$

where

$$J^{\mathcal{N}}_{B\tilde{A}B'} \coloneqq \mathrm{Tr}_A[J^{\mathcal{N}}_{AB\tilde{A}B'}], \tag{46}$$
$$J^{\mathcal{N}}_{BB'} \coloneqq \mathrm{Tr}_{A\tilde{A}}[J^{\mathcal{N}}_{AB\tilde{A}B'}], \tag{47}$$

are marginals of $J^{\mathcal{N}}_{AB\tilde{A}B'}$ and we employ this kind of shorthand in what follows.

Proof. We begin by proving that the channel $\mathcal{N}_{AB\to\tilde{A}B'}$ is conditionally unital if its Choi matrix has the form in (45). To this end, let $\rho_B \in \mathfrak{D}(B)$ and consider that

$$\begin{aligned}
&\mathcal{N}\left(\mathbf{u}_A \otimes \rho_B\right) \\
&= \mathrm{Tr}_{AB}\left[J^{\mathcal{N}}_{AB\tilde{A}B'}\left(\mathbf{u}_A \otimes (\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] && (48)\\
&= \frac{1}{|A|}\,\mathrm{Tr}_B\left[J^{\mathcal{N}}_{B\tilde{A}B'}\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] && (49)\\
&= \frac{1}{|A|}\,\mathrm{Tr}_B\left[\mathbf{u}_{\tilde{A}} \otimes J^{\mathcal{N}}_{BB'}\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] && (50)\\
&= \mathbf{u}_{\tilde{A}} \otimes \frac{1}{|A|}\,\mathrm{Tr}_B\left[J^{\mathcal{N}}_{BB'}\left((\rho_B)^T \otimes I_{B'}\right)\right]. && (51)
\end{aligned}$$

The first equality is a well known identity in quantum information, indicating how the action of a channel on an input state can be understood in terms of a propagation formula using its Choi matrix (see Eq. (2.66) of [8]). The third equality above follows from applying (45), which holds by assumption. It then follows that $\frac{1}{|A|}\,\mathrm{Tr}_B\left[J^{\mathcal{N}}_{BB'}\left((\rho_B)^T \otimes I_{B'}\right)\right]$ is a state on the output system B' (i.e., the choice for the state $\sigma_{B'}$ in (12)) because $\frac{1}{|A|}J^{\mathcal{N}}_{BB'}$ is the Choi operator of the channel

$$\tau_B \to (\mathrm{Tr}_{\tilde{A}} \circ \mathcal{N}_{AB\to\tilde{A}B'})(\mathbf{u}_A \otimes \tau_B). \tag{52}$$

We now prove that the Choi matrix of $\mathcal{N}_{AB\to\tilde{A}B'}$ has the form in (45) if $\mathcal{N}_{AB\to\tilde{A}B'}$ is conditionally unital. Recall from (12) that the defining property of a conditionally unital channel is that for every state ρ_B, there exists a state $\sigma_{B'}$ such that

$$\mathcal{N}_{AB\to\tilde{A}B'}(\mathbf{u}_A \otimes \rho_B) = \mathbf{u}_{\tilde{A}} \otimes \sigma_{B'}. \tag{53}$$

By taking the trace over $\tilde{A}$ on both sides of the equation above, we find that

$$\begin{aligned}
\sigma_{B'} &= \mathcal{N}_{AB\to B'}(\mathbf{u}_A \otimes \rho_B) && (54)\\
&= \mathrm{Tr}_{AB}\left[J^{\mathcal{N}}_{ABB'}\left(\mathbf{u}_A \otimes (\rho_B)^T \otimes I_{B'}\right)\right] && (55)\\
&= \frac{1}{|A|}\mathrm{Tr}_B\left[J^{\mathcal{N}}_{BB'}\left((\rho_B)^T \otimes I_{B'}\right)\right]. && (56)
\end{aligned}$$

On the other hand, observe that

$$\begin{aligned}
&\mathcal{N}_{AB\to\tilde{A}B'}(\mathbf{u}_A \otimes \rho_B) \\
&= \mathrm{Tr}_{AB}\left[J^{\mathcal{N}}_{AB\tilde{A}B'}\left(\mathbf{u}_A \otimes (\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] && (57)\\
&= \frac{1}{|A|}\,\mathrm{Tr}_B\left[J^{\mathcal{N}}_{B\tilde{A}B'}\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right]. && (58)
\end{aligned}$$

Therefore, from the two expressions above for σ_B and $\mathcal{N}(\mathbf{u}_A \otimes \rho_B)$, we conclude that (53) can be expressed as

$$\begin{aligned}
&\mathrm{Tr}_B\left[J^{\mathcal{N}}_{B\tilde{A}B'}\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] \\
&\quad= \mathrm{Tr}_B\left[(\mathbf{u}_{\tilde{A}} \otimes J^{\mathcal{N}}_{BB'})\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right], && (59)
\end{aligned}$$

holding for all $\rho_B \in \mathfrak{D}(B)$. Let

$$\eta_{B\tilde{A}B'} \coloneqq J^{\mathcal{N}}_{B\tilde{A}B'} - \mathbf{u}_{\tilde{A}} \otimes J^{\mathcal{N}}_{BB'}, \tag{60}$$

and observe that (59) can be written as

$$\mathrm{Tr}_B\left[\eta_{B\tilde{A}B'}\left((\rho_B)^T \otimes I_{\tilde{A}B'}\right)\right] = 0 \qquad \forall\, \rho_B \in \mathfrak{D}(B). \tag{61}$$

Due to the existence of bases of density operators that span the space of linear operators acting on B (see Example 2.7 of [8]), we conclude from (61) that for every Hermitian operator $X_B \in \mathfrak{L}(B)$, we have

$$\mathrm{Tr}_B[\eta_{B\tilde{A}B'}\,(X_B \otimes I_{\tilde{A}B'})] = 0\,. \tag{62}$$

Note that by multiplying both sides of the equation above by an arbitrary element $Y_{\tilde{A}B'} \in \mathfrak{L}(\tilde{A}B')$ and taking the trace, we get that

$$\mathrm{Tr}\left[\eta_{B\tilde{A}B'}\,(X_B \otimes Y_{\tilde{A}B'})\,\right] = 0\,. \tag{63}$$

Since the equation above holds for all $X_B \in \mathfrak{L}(B)$ and all $Y_{\tilde{A}B'} \in \mathfrak{L}(\tilde{A}B')$ it also holds for all linear combinations of matrices of the form $X_B \otimes Y_{\tilde{A}B'}$. Since matrices of the form $X_B \otimes Y_{\tilde{A}B'}$ span the whole space $\mathfrak{L}(B\tilde{A}B')$ we conclude that

$$\mathrm{Tr}\left[\eta_{B\tilde{A}B'}W_{B\tilde{A}B'}\right] = 0 \quad \forall\, W_{B\tilde{A}B'} \in \mathfrak{L}(B\tilde{A}B')\,. \tag{64}$$

Therefore, we conclude that $\eta_{B\tilde{A}B'} = 0$, which is equivalent to (45). ∎

B.2 Choi matrix formulation of conditionally mixing channels

Here we provide a Choi matrix formulation for conditionally mixing channels. Before doing so, let us first recall [38, Eq. (40)], which indicates how semi-causality can be expressed in terms of the Choi matrix.

Lemma 2 ([38, 54]) *The channel $\mathcal{N}_{AB\to AB'}$ is $A \not\to B'$ semi-causal if and only if the marginals of its Choi matrix $J^{\mathcal{N}}_{AB\tilde{A}B'}$ satisfy*

$$J^{\mathcal{N}}_{ABB'} = \mathbf{u}_A \otimes J^{\mathcal{N}}_{BB'}. \tag{65}$$

As a direct corollary of (43)–(44) and Lemmas 1 and 2, we have the following simple semi-definite characterization of a conditionally mixing channel in terms of its Choi matrix.

Corollary 2 *A matrix $J^{\mathcal{N}}_{AB\tilde{A}B'}$ is the Choi matrix of a conditionally mixing channel $\mathcal{N}_{AB\to\tilde{A}B'}$ if and only if*

$$J^{\mathcal{N}}_{AB\tilde{A}B'} \geq 0, \tag{66}$$

$$\mathrm{Tr}_{\tilde{A}B'}[J^{\mathcal{N}}_{AB\tilde{A}B'}] = I_{AB}, \tag{67}$$

$$J^{\mathcal{N}}_{ABB'} = \mathbf{u}_A \otimes J^{\mathcal{N}}_{BB'}, \tag{68}$$

$$J^{\mathcal{N}}_{B\tilde{A}B'} = J^{\mathcal{N}}_{BB'} \otimes \mathbf{u}_{\tilde{A}}. \tag{69}$$

Corollary 2 simplifies the task of optimizing an objective function over the set of conditionally mixing channels by making use of semi-definite programming. Additionally, by inspecting (68)–(69), we see that the properties of these channels are invariant under the exchange of the input system A and the output system $\tilde{A}$.

C Quantum conditional majorization reduces to classical conditional majorization

In [39], a pre-order $\succ_C$ between two joint probability distributions (i.e., classical states) is called *classical conditional majorization*. In this appendix, we prove that the approach to conditional majorization defined in Section 3.1 is a generalization of classical conditional majorization. To this end, we demonstrate that quantum conditional majorization is equivalent to the classical conditional majorization when we restrict the underlying bipartite quantum states to classical states.

Suppose the system $A \coloneqq X$ is classical. We call a channel $\mathcal{M}_{XB\to XB'}$ *conditionally doubly stochastic* (CDS for short) if it has the following form:

$$\mathcal{M}_{XB\to XB'} = \sum_j \mathcal{D}^{(j)}_{X\to X} \otimes \mathcal{F}^{(j)}_{B\to B'}, \tag{70}$$

where each $\mathcal{D}^{(j)}$ is a classical doubly stochastic channel and each $\mathcal{F}^{(j)}_{B\to B'}$ is a completely positive (CP) map such that $\sum_j \mathcal{F}^{(j)}_{B\to B'}$ is a quantum channel. CDS channels with classical B and B' systems were introduced and studied in [39], and CDS channels with quantum B and B' systems were introduced and studied in [55].

Definition 6 ([39]) *Let ρ_{XY} and $\sigma_{XY'}$ be classical bipartite states such that $A = X$, $B = Y$, and $B' = Y'$ are all classical systems. We say that ρ_{XY} conditionally majorizes $\sigma_{XY'}$ and write $\rho_{XY} \succ_C \sigma_{XY'}$ if there exists a CDS channel $\mathcal{N}_{XY\to XY'}$ such that $\sigma_{XY'} = \mathcal{N}_{XY\to XY'}(\rho_{XY})$.*

Theorem 3 *Suppose all the systems involved are classical with $A \coloneqq X$, $B \coloneqq Y$, and $B' \coloneqq Y'$. Then a quantum channel $\mathcal{N}_{XY\to XY'}$ is a CDS channel if and only if it is conditionally mixing.*

Proof. Suppose first that $\mathcal{N}_{XY\to XY'}$ is a CDS channel, having the form of (70). Then for every input of the form $\mathbf{u}_A \otimes \sigma_B$, it follows that

$$\begin{aligned}
&\mathcal{N}_{AB\to AB'}(\mathbf{u}_A \otimes \sigma_B) \\
&= \sum_j \mathcal{D}^{(j)}_A(\mathbf{u}_A) \otimes \mathcal{F}^{(j)}_{B\to B'}(\sigma_B) &(71)\\
&= \mathbf{u}_A \otimes \sum_j \mathcal{F}^{(j)}_{B\to B'}(\sigma_B) &(72)
\end{aligned}$$

and therefore $\mathcal{N}$ is conditionally unital. (Note that the third line of the above follows from the fact that $\mathcal{D}^{(j)}$ is doubly stochastic, so that $\mathcal{D}^{(j)}(\mathbf{u}_A) = \mathbf{u}_A$ for all j.) Next, we show that $\mathcal{N}$ is semi-causal. Let $\mathcal{T}_A$ be an arbitrary classical channel. Then consider that

$$\begin{aligned}
&\mathrm{Tr}_A \circ \mathcal{N} \circ \mathcal{T}_A \\
&= \mathrm{Tr}_A \circ \left[\sum_j (\mathcal{D}^{(j)}_A \circ \mathcal{T}_A) \otimes \mathcal{F}^{(j)}_{B\to B'}\right] &(73)\\
&= \sum_j (\mathrm{Tr}_A \circ \mathcal{D}^{(j)}_A \circ \mathcal{T}_A) \otimes \mathcal{F}^{(j)}_{B\to B'} &(74)\\
&= \sum_j (\mathrm{Tr}_A \circ \mathcal{D}^{(j)}_A) \otimes \mathcal{F}^{(j)}_{B\to B'} &(75)\\
&= \mathrm{Tr}_A \circ \mathcal{N}_{AB\to AB'}. &(76)
\end{aligned}$$

Given that $\mathcal{N}$ is conditionally unital and semi-causal, it is by definition conditionally mixing.

Conversely, suppose $\mathcal{N}_{XY\to XY'}$ is conditionally mixing. Since $\mathcal{N}$ is semi-causal, it has the form given in (15) (see also Figure 2). Moreover, since systems A, B, and B' are all classical, we can assume without loss of generality that system R in (15) is also classical. Hence, $\mathcal{N}$ has the form

$$\mathcal{N}_{XY\to XY'} = \sum_{j=1}^{k} \mathcal{E}^{(j)}_{X\to X} \otimes \mathcal{R}^{(j)}_{Y\to Y'}\,, \tag{77}$$

where $k \in \mathbb{N}$, and for each $j \in [k]$, $\mathcal{E}^{(j)} \in \mathrm{CPTP}(X \to X)$ and $\mathcal{R}^{(j)} \in \mathrm{CP}(Y \to Y')$ with $\mathcal{R} \coloneqq \sum_j \mathcal{R}^{(j)} \in \mathrm{CPTP}(Y \to Y')$. The key idea of the proof is to decompose each $\mathcal{R}^{(j)}$ into "extreme" maps. We do that in the following way. We first denote by $\{r_{jy'|y}\}$ the components of the transition matrix of $\mathcal{R}^{(j)}$. That is, for all $j \in [k]$ and $y \in [n]$

$$\mathcal{R}^{(j)}_{Y\to Y'}\left(|y\rangle\langle y|_Y\right) = \sum_{y'=1}^{n'} r_{jy'|y}|y'\rangle\langle y'|_{Y'} \ . \tag{78}$$

We also denote by $r_{y'|y} \coloneqq \sum_{j\in[k]} r_{jy'|y}$ the components of the transition matrix of $\mathcal{R}$, and by $q_{j|yy'} \coloneqq r_{jy'|y}/r_{y'|y}$. Finally, for any $y \in [n]$ and $y' \in [n']$ we define the CP map $\mathcal{F}^{(yy')}$ as

$$\mathcal{F}^{(yy')}_{Y\to Y'}\left(|w\rangle\langle w|_Y\right) = \delta_{yw} r_{y'|y}|y'\rangle\langle y'|_{Y'} \ . \tag{79}$$

By definition, $\sum_{y\in[n]}\sum_{y'\in[n']} \mathcal{F}^{(yy')}$ is a (CPTP) classical channel, and

$$\mathcal{R}^{(j)}_{Y\to Y'} = \sum_{y\in[n]}\sum_{y'\in[n']} q_{j|yy'}\mathcal{F}^{(yy')}_{Y\to Y'} \ . \tag{80}$$

With these definitions we get that

$$\begin{aligned}\mathcal{N}_{XY\to XY'} &= \sum_{y\in[n]}\sum_{y'\in[n']}\sum_{j\in[k]} q_{j|yy'}\mathcal{E}^{(j)}_{X\to X}\otimes\mathcal{F}^{(yy')}_{Y\to Y'} \\ &= \sum_{y\in[n]}\sum_{y'\in[n']} \mathcal{D}^{(yy')}_{X\to X}\otimes\mathcal{F}^{(yy')}_{Y\to Y'} \ , \end{aligned} \tag{81}$$

where

$$\mathcal{D}^{(yy')}_{X\to X'} \coloneqq \sum_{j\in[k]} q_{j|yy'}\mathcal{E}^{(j)}_{X\to X} \ . \tag{82}$$

Observe that each $\mathcal{D}^{(yy')}_{X\to X'}$ is a classical channel since a convex combination of classical channels is a classical channel. It is therefore left to show that each $\mathcal{D}^{(yy')}_{X\to X'}$ is unital. To show that, we use the fact that $\mathcal{N}$ is a conditionally unital channel.

Since $\mathcal{N}$ is a conditionally unital channel, for all $w \in [n]$ we have that

$$\mathbf{u}_X \otimes \sigma^{(w)}_{Y'} = \mathcal{N}_{XY\to XY'}\left(\mathbf{u}_X \otimes |w\rangle\langle w|_Y\right), \tag{83}$$

where $\sigma^{(w)} \in \mathfrak{D}(Y')$ is some state (in fact, $\sigma^{(w)}_{Y'} = \mathcal{N}_{XY\to Y'}\left(\mathbf{u}_X \otimes |w\rangle\langle w|_Y\right)$). Therefore, using the form of $\mathcal{N}$ in (81) we get

$$\begin{aligned}&\mathbf{u}_X \otimes \sigma^{(w)}_{Y'} \\ &= \sum_{y\in[n]}\sum_{y'\in[n']} \mathcal{D}^{(yy')}_{X\to X'}\left(\mathbf{u}_X\right)\otimes\mathcal{F}^{(yy')}_{Y\to Y'}\left(|w\rangle\langle w|_Y\right) \qquad (84)\\ &= \sum_{y'\in[n']} r_{y'|w}\mathcal{D}^{(wy')}_{X\to X}\left(\mathbf{u}_X\right)\otimes|y'\rangle\langle y'|_{Y'} \ , \qquad (85)\end{aligned}$$

where the last equality follows from (79). By taking the expectation value $\langle w'|(\cdot)|w'\rangle$ of system Y' on both sides of the equation above, we get that for all $w' \in [n']$

$$\mathbf{u}^X\langle w'|\sigma^{(w)}_{Y'}|w'\rangle = r_{w|w'}\mathcal{D}^{(ww')}_{X\to X}\left(\mathbf{u}_X\right) \ . \tag{86}$$

By taking the trace on both sides of the equation above we get that $\langle w'|\sigma^{(w)}_{Y'}|w'\rangle = r_{w|w'}$ (recall that $\mathcal{D}^{(ww')}_{X\to X}$ is a channel and in particular trace preserving). Therefore, the equation above implies that

$$\mathbf{u}_X = \mathcal{D}^{(ww')}_{X\to X}\left(\mathbf{u}_X\right) \ ; \tag{87}$$

that is, each $\mathcal{D}^{(ww')}_{X\to X}$ is a doubly stochastic matrix. ■

Remark 4 *Observe that the theorem above implies in particular that $\sigma_{XY'} \succ_X \rho_{XY}$ if and only if $\sigma_{XY'} \succ_C \rho_{XY}$, where $\succ_C$ is the classical conditional majorization defined in [39]. In other words, the pre-orders $\succ_X$ and $\succ_C$ coincide in the classical domain.*

D Proofs of properties of quantum conditional entropy

Suppose that $\rho_{AB} = \omega_A \otimes \tau_B$ is a product state. Let $\mathcal{E}_B$ be the completely randomizing channel acting on system B, which traces out system B and replaces with the uniform state $\mathbf{u}_B$. Then it follows that $\mathrm{id}_A \otimes \mathcal{E}_B$ is a conditionally unital channel and that

$$\left(\mathrm{id}_A \otimes \mathcal{E}_B\right)\left(\omega_A \otimes \tau_B\right) = \omega_A \otimes \mathbf{u}_B. \tag{88}$$

Moreover, let $\mathcal{R}_B$ be the replacement channel on system B, which takes every quantum state to a fixed state τ_B. Then, $\mathrm{id}_A \otimes \mathcal{R}_B$ is conditionally unital and

$$\left(\mathrm{id}_A \otimes \mathcal{R}_B\right)\left(\omega_A \otimes \mathbf{u}_B\right) = \omega_A \otimes \tau_B \ . \tag{89}$$

Thus,

$$\omega_A \otimes \tau_B \succ_A \omega_A \otimes \mathbf{u}_B \text{ and } \omega_A \otimes \mathbf{u}_B \succ_A \omega_A \otimes \tau_B, \tag{90}$$

which implies, by the monotonicity property of the conditional entropy $\mathbf{H}$, that

$$\mathbf{H}(A|B)_{\omega_A\otimes\tau_B} = \mathbf{H}(A|B)_{\omega_A\otimes\mathbf{u}_B}. \tag{91}$$

Thus, $\mathbf{H}(A|B)_{\omega_A\otimes\tau_B}$ depends on ω_A only. Moreover, the function $\omega_A \mapsto \mathbf{H}(A|B)_{\omega_A\otimes\tau_B}$ satisfies the two postulates of entropy in (10)–(11) and therefore is an entropy of ω_A. Note that the same statement also holds when system $|B|$ is trivial, i.e., when $|B| = 1$.

We now prove that $\mathbf{H}(A|B)$ is invariant under local isometric channels. It follows immediately from the definition of quantum conditional entropy (specifically, (16)–(17) and (20)) that it is invariant under local

isometric channels acting on system A. So we prove invariance under isometric channels acting on system B. The inequality

$$\mathbf{H}(A|B)_\rho \leq \mathbf{H}(A|B')_\omega \tag{92}$$

follows because $\mathcal{V}_{B\to B'}$ is a conditionally unital channel. Let τ_B be a state. The reverse inequality follows because

$$\mathcal{R}^{\mathcal{V}}_{B'\to B}(\cdot) \coloneqq \mathcal{V}^\dagger(\cdot) + \mathrm{Tr}[(\mathrm{id}_{B'} - \mathcal{V}^\dagger)(\cdot)]\tau_B \tag{93}$$

is a channel that is a left inverse of $\mathcal{V}_{B\to B'}$ [7, Section 4.6.3]. Thus, $\mathcal{R}^{\mathcal{V}}_{B'\to B}$ is a conditionally unital channel, from which we conclude that

$$\mathbf{H}(A|B')_\omega \leq \mathbf{H}(A|B)_{\mathcal{R}^{\mathcal{V}}(\omega)} = \mathbf{H}(A|B)_\rho. \tag{94}$$

This concludes the proof.

E A subclass of conditional entropies

Here we consider a subclass of conditional entropies in which the postulate in (20) is replaced by monotonicity under conditionally unital channels. Since the set of conditionally unital channels contains the set of conditionally mixing channels, every function that is a conditional entropy according to this alternative definition is also a conditional entropy $\mathbf{H}$, as defined in (19)–(21). We will later see that this approach is consistent and sufficient for deriving various properties of conditional entropy.

Let us begin by defining another version of a conditional majorization pre-order on bipartite quantum states, based on conditionally unital channels exclusively. We call this approach relaxed conditional majorization.

Definition 7 (Relaxed conditional majorization) *Given two bipartite states ρ_{AB} and $\sigma_{A'B'}$, we say that ρ_{AB} conditionally majorizes $\sigma_{A'B'}$ in the relaxed sense and with respect to system A, and write*

$$\rho_{AB} \succ^A \sigma_{A'B'} \tag{95}$$

if there exist a system A'' such that $|A''| \geq \max\{|A|, |A'|\}$, isometric channels $\mathcal{V}_{A\to A''}$ and $\mathcal{U}_{A'\to A''}$, and a conditionally unital channel $\mathcal{N}_{A''B\to A''B'}$ such that

$$\mathcal{U}_{A'\to A''}(\sigma_{A'B'}) = (\mathcal{N}_{A''B\to A''B'} \circ \mathcal{V}_{A\to A''})(\rho_{AB})\ . \tag{96}$$

We define a function f to be a conditional uncertainty measure in the relaxed sense if it is not equal to the zero function and if it reverses the relaxed conditional majorization pre-order, i.e.,

$$\rho_{AB} \succ^A \sigma_{A'B'} \quad \Rightarrow \quad f(\rho_{AB}) \leq f(\sigma_{A'B'}). \tag{97}$$

We now define this other notion of conditional entropy. Formally, a function

$$\mathbf{H}^{\blacktriangle} : \bigcup_{A,B} \mathfrak{D}(AB) \to \mathbb{R} \tag{98}$$

is a quantum conditional entropy if it is not equal to the zero function and satisfies the following postulates for all systems A, B, A', B' and states ρ_{AB} and $\sigma_{A'B'}$:

1. Monotonicity:

$$\rho_{AB} \succ^A \sigma_{A'B'} \ \Rightarrow\ \mathbf{H}^{\blacktriangle}(A|B)_\rho \leq \mathbf{H}^{\blacktriangle}(A'|B')_\sigma. \tag{99}$$

2. Additivity:

$$\mathbf{H}^{\blacktriangle}(AA'|BB')_{\rho\otimes\sigma} = \mathbf{H}^{\blacktriangle}(A|B)_\rho + \mathbf{H}^{\blacktriangle}(A'|B')_\sigma. \tag{100}$$

Let us now recall a different definition of conditional min-entropy from [15]

$$H^{\uparrow}_{\min}(A|B)_\rho \coloneqq \sup_{\sigma_B\in\mathfrak{D}(B)} H_{\min}(\rho_{AB}|\sigma_B), \tag{101}$$

where

$$H_{\min}(\rho_{AB}|\sigma_B) \coloneqq -\inf_{\lambda\geq 0} \log_2\{\lambda : \rho_{AB} \leq \lambda I_A \otimes \sigma_B\}. \tag{102}$$

Note that $H^{\uparrow}_{\min}(A|B)_\rho$ is a conditional entropy $\mathbf{H}^{\blacktriangle}$ because it is additive [16, page 4342], and [15, Lemma 3.1.12] implies that $H^{\uparrow}_{\min}(A|B)_\rho$ is non-decreasing under conditionally unital channels.

We then have the following findings, analogous to those from Theorem 1 and Corollary 1, respectively.

Theorem 5 *Let $\mathbf{H}^{\blacktriangle}$ be a quantum conditional entropy, as defined in (98)–(100). Then, for every state ρ_{AB},*

$$\mathbf{H}^{\blacktriangle}(A|B)_\rho \geq H^{\uparrow}_{\min}(A|B)_\rho. \tag{103}$$

Equality is attained in (103) if ρ_{AB} is equal to a maximally entangled state $\Phi^{(k)}_{A'B'}$ by the action of local isometric channels. Thus,

$$\mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} = -\log_2 k \tag{104}$$

for every conditional entropy $\mathbf{H}^{\blacktriangle}$.

Corollary 3 *Let ρ_{AB} be a bipartite state. Then the following are equivalent:*

1. *For every conditional entropy $\mathbf{H}^{\blacktriangle}$,*

$$\mathbf{H}^{\blacktriangle}(A|B)_\rho \geq 0. \tag{105}$$

2. *There exists a state σ_B such that $I_A \otimes \sigma_B \geq \rho_{AB}$.*

Proof. The proof of Corollary 3 is immediate from Theorem 5. Indeed, if the first condition holds, then the inequality $H_{\min}^{\uparrow}(A|B)_\rho \geq 0$ holds in particular (since $H_{\min}^{\uparrow}(A|B)_\rho$ is a conditional entropy), and the second condition is then a consequence of the definition in (101). Now suppose that the second condition holds. This then means that there exists $\lambda \leq 1$ and a state σ_B such that $\rho_{AB} \leq \lambda I_A \otimes \sigma_B$. According to the definition in (101), it follows that $H_{\min}^{\uparrow}(A|B)_\rho \geq 0$, and then Theorem 5 implies the first condition. ∎

We have opted to emphasize the conditional entropy $\mathbf{H}$ over $\mathbf{H}^{\blacktriangle}$ in the main text because

1. conditionally mixing channels have a strong motivation in terms of no information leakage from Alice to Bob, as would be expected for a measure of conditional uncertainty;

2. conditional majorization defined in this way reduces to classical conditional majorization, as defined in [39], for the classical case;

3. the conditional min-entropy in (27) is in fact the smallest conditional entropy among all conditional entropies consistent with both sets of postulates;

4. every conditional entropy defined in this other way is in fact a conditional entropy $\mathbf{H}$, because every function that is monotone under all conditionally unital channels is also monotone under the subset of all conditionally mixing channels.

E.1 Proof of Theorem 5

Let σ_B be a state that achieves the optimal value of $H_{\min}^{\uparrow}(A|B)_\rho$, in the definition in (101).

Let us first consider the case when $H_{\min}^{\uparrow}(A|B)_\rho \geq 0$. Set

$$k \coloneqq |A|\,, \tag{106}$$

$$m \coloneqq \left\lfloor 2^{H_{\min}^{\uparrow}(A|B)_\rho} \right\rfloor, \tag{107}$$

and let X be a classical system with dimension k. By the assumption that $H_{\min}^{\uparrow}(A|B)_\rho \geq 0$, and given the dimension bound $H_{\min}^{\uparrow}(A|B)_\rho \leq \log_2 k$ [56, Lemma 5.2], it follows that $m \in [k]$.

Without loss of generality, we can suppose that $k > m$; if it is not the case, then we can embed the system A of ρ_{AB} in a larger Hilbert space, so that $H_{\min}^{\uparrow}(A|B)_\rho$ stays constant but k increases. Set Π_X to be a projection onto an m-dimensional subspace of X (i.e., set $\Pi_X \coloneqq \sum_{x=1}^{m} |x\rangle\langle x|_X$, where $\{|x\rangle\}_x$ is an orthonormal basis specifying the classical system X), and define $\tau_{AB}^{(m)}$ as follows:

$$\tau_{AB}^{(m)} \coloneqq \frac{I_A \otimes \sigma_B - m\rho_{AB}}{k-m}. \tag{108}$$

The operator $\tau_{AB}^{(m)}$ is positive semi-definite because $k - m > 0$ and

$$\begin{aligned}
&\frac{1}{m} I_A \otimes \sigma_B - \rho_{AB} \\
&\geq 2^{-H_{\min}^{\uparrow}(A|B)_\rho} I_A \otimes \sigma_B - \rho_{AB} && (109)\\
&= 2^{D_{\max}(\rho_{AB} \| I_A \otimes \sigma_B)} I_A \otimes \sigma_B - \rho_{AB} && (110)\\
&\geq 0. && (111)
\end{aligned}$$

Also, $\tau_{AB}^{(m)}$ has trace equal to one and is thus a state. Let us now define the following measure-and-prepare channel:

$$\begin{aligned}
&\mathcal{N}_{X\to AB}(\omega_X) \coloneqq \\
&\qquad \operatorname{Tr}[\Pi_X \omega_X]\rho_{AB} + \operatorname{Tr}[(I_X - \Pi_X)\,\omega_X]\tau_{AB}^{(m)}. \quad (112)
\end{aligned}$$

Indeed, the action of this channel is to perform a measurement according to the POVM $\{\Pi_X, I_X - \Pi_X\}$ and prepare the state ρ_{AB} if the first outcome is obtained and the state $\tau_{AB}^{(m)}$ if the second outcome is obtained. It should be understood that Alice has access to the input system X.

This channel is conditionally unital because

$$\begin{aligned}
&\mathcal{N}_{X\to AB}(I_X) \\
&= \operatorname{Tr}[\Pi_X I_X]\rho_{AB} + \operatorname{Tr}[(I_X - \Pi_X)\, I_X]\tau_{AB}^{(m)} && (113)\\
&= m\rho_{AB} + (k-m)\,\tau_{AB}^{(m)} && (114)\\
&= I_A \otimes \sigma_B. && (115)
\end{aligned}$$

Also, if we input the uniform state $\mathbf{u}_X^{\Pi} \coloneqq \Pi_X/m$ on the subspace onto which Π_X projects, the output is given by

$$\mathcal{N}_{X\to AB}(\mathbf{u}_X^{\Pi}) = \rho_{AB}. \tag{116}$$

Putting these observations together, we conclude that

$$\begin{aligned}
\log_2 \left\lfloor 2^{H_{\min}^{\uparrow}(A|B)_\rho} \right\rfloor &= \mathbf{H}^{\blacktriangle}(X)_{\mathbf{u}^{\Pi}} && (117)\\
&\leq \mathbf{H}^{\blacktriangle}(A|B)_{\mathcal{N}(\mathbf{u}^{\Pi})} && (118)\\
&= \mathbf{H}^{\blacktriangle}(A|B)_\rho. && (119)
\end{aligned}$$

The first equality follows because the entropy of a uniform state is equal to the logarithm of its rank and from the definition of m in (107). The inequality follows from the fact that $\mathcal{N}_{X\to AB}$ is conditionally unital (see (113)–(115)) and $\mathbf{H}^{\blacktriangle}$ is a quantum conditional entropy, thus obeying the first postulate in (99). The final equality follows from (116).

Therefore, since all conditional entropies are additive for tensor-product states, we conclude that

$$\begin{aligned}
\mathbf{H}^{\blacktriangle}(A|B)_\rho &= \lim_{n\to\infty} \frac{1}{n}\mathbf{H}^{\blacktriangle}(A^n|B^n)_{\rho^{\otimes n}} && (120)\\
&\geq \lim_{n\to\infty} \frac{1}{n}\log_2\left\lfloor 2^{H^{\uparrow}_{\min}(A^n|B^n)_{\rho^{\otimes n}}}\right\rfloor && (121)\\
&= \lim_{n\to\infty} \frac{1}{n}\log_2\left\lfloor 2^{nH^{\uparrow}_{\min}(A|B)_{\rho}}\right\rfloor && (122)\\
&= H^{\uparrow}_{\min}(A|B)_\rho\,. && (123)
\end{aligned}$$

The inequality follows from (117)–(119). The second equality follows from additivity, and the final equality from the assumption that $H^{\uparrow}_{\min}(A|B)_\rho \geq 0$.

Now let us consider the case when $H^{\uparrow}_{\min}(A|B) < 0$. Let

$$\begin{aligned}
t &:= 2^{-H^{\uparrow}_{\min}(A|B)}, && (124)\\
k' &:= \lceil t \rceil\,, && (125)\\
k &:= |A|\,. && (126)
\end{aligned}$$

Observe that $t > 1$ by assumption. Also, observe that t is the smallest real satisfying $\rho_{AB} \leq tI_A \otimes \sigma_B$. Let A' be a system of dimension k', and let X be a classical system of dimension $|X| := kk'$. Now define

$$\tau^{(k,k')}_{AB} := \frac{k'I_A \otimes \sigma_B - \rho_{AB}}{kk' - 1}, \tag{127}$$

and observe that $\tau^{(k,k')}_{AB}$ is positive semi-definite because $kk' > 1$ and

$$\begin{aligned}
&k'I_A \otimes \sigma_B - \rho_{AB}\\
&\geq 2^{-H^{\uparrow}_{\min}(A|B)} I_A \otimes \sigma_B - \rho_{AB} && (128)\\
&= 2^{D_{\max}(\rho_{AB}\|I_A\otimes\sigma_B)} I_A \otimes \sigma_B - \rho_{AB} && (129)\\
&\geq 0. && (130)
\end{aligned}$$

Since $\tau^{(k,k')}_{AB}$ has trace equal to one, it follows that it is a state. We then construct the following channel:

$$\begin{aligned}
&\mathcal{N}_{X\to AA'B}(\omega_X) :=\\
&\left(\mathrm{Tr}[|1\rangle\langle 1|_X \omega_X]\rho_{AB} + \mathrm{Tr}[\Pi_X \omega_X]\tau^{(k,k')}_{AB}\right) \otimes \mathbf{u}_{A'}, \quad (131)
\end{aligned}$$

where

$$\Pi_X := I_X - |1\rangle\langle 1|_X, \tag{132}$$

so that $\mathrm{Tr}[\Pi_X] = kk' - 1$. Also, it should be understood that Alice has access to the X input system. Consider that

$$\mathcal{N}_{X\to AA'B}(|1\rangle\langle 1|_X) = \rho_{AB} \otimes \mathbf{u}_{A'}. \tag{133}$$

Furthermore, the channel is conditionally unital because

$$\begin{aligned}
&\mathcal{N}_{X\to AA'B}(I_X)\\
&= \left(\mathrm{Tr}[|1\rangle\langle 1|_X I_X]\rho_{AB} + \mathrm{Tr}[\Pi_X I_X]\tau^{(k,k')}_{AB}\right) \otimes \mathbf{u}_{A'}\\
&= \left(\rho_{AB} + (kk'-1)\,\tau^{(k,k')}_{AB}\right) \otimes \mathbf{u}_{A'} && (134)\\
&= k' I_A \otimes \sigma_B \otimes \mathbf{u}_{A'} && (135)\\
&= I_A \otimes \sigma_B \otimes I_{A'}. && (136)
\end{aligned}$$

Then it follows that

$$\begin{aligned}
0 &= \mathbf{H}^{\blacktriangle}(X)_{|1\rangle\langle 1|} && (137)\\
&\leq \mathbf{H}^{\blacktriangle}(AA'|B)_{\mathcal{N}(|1\rangle\langle 1|)} && (138)\\
&= \mathbf{H}^{\blacktriangle}(AA'|B)_{\rho\otimes\mathbf{u}} && (139)\\
&= \mathbf{H}^{\blacktriangle}(A|B)_\rho + \mathbf{H}(A')_{\mathbf{u}} && (140)\\
&= \mathbf{H}^{\blacktriangle}(A|B)_\rho + \log_2\left\lceil 2^{-H^{\uparrow}_{\min}(A|B)}\right\rceil. && (141)
\end{aligned}$$

The first equality follows because the entropy of a pure state is equal to zero. The inequality follows because $\mathcal{N}_{X\to AA'B}$ is a conditionally unital channel and $\mathbf{H}^{\blacktriangle}$ is a quantum conditional entropy, thus obeying the first postulate in (99). The second equality follows from (133). The penultimate equality follows from the additivity postulate of quantum conditional entropy. The final equality follows because the entropy of the uniform state $\mathbf{u}_{A'}$ is equal to the logarithm of its rank, which is in turn equal to $\left\lceil 2^{-H^{\uparrow}_{\min}(A|B)}\right\rceil$.

Now, since $\mathbf{H}^{\blacktriangle}(A|B)_\rho$ is additive for tensor-product states, we find that

$$\begin{aligned}
&\mathbf{H}^{\blacktriangle}(A|B)_\rho\\
&= \lim_{n\to\infty} \frac{1}{n}\mathbf{H}^{\blacktriangle}(A^n|B^n)_{\rho^{\otimes n}} && (142)\\
&\geq \lim_{n\to\infty} -\frac{1}{n}\log_2\left\lceil 2^{-H^{\uparrow}_{\min}(A^n|B^n)_{\rho^{\otimes n}}}\right\rceil && (143)\\
&= \lim_{n\to\infty} -\frac{1}{n}\log_2\left\lceil 2^{-nH^{\uparrow}_{\min}(A|B)_{\rho}}\right\rceil && (144)\\
&= \lim_{n\to\infty} -\frac{1}{n}\log_2\lceil t^n\rceil && (145)\\
&\geq \lim_{n\to\infty} -\frac{1}{n}\log_2(t^n+1) && (146)\\
&= -\log_2 t && (147)\\
&= H^{\uparrow}_{\min}(A|B)_\rho\,. && (148)
\end{aligned}$$

The first inequality follows from (137)–(141). The second equality follows from additivity, and the penultimate equality from the assumption that $t > 1$. This completes the proof of the lower bound.

It is left to prove the equality on maximally entangled states, i.e., (104). Since conditional entropy is invariant under local isometries, we can assume without loss of generality that $k := |A| = |B|$. Now, consider the state $\Phi^{(k)}_{AB} \otimes \mathbf{u}_{\tilde{A}}$ where $|\tilde{A}| = k$, and observe that by introducing a trivial system B_2 such that $|B_2| = 1$ we get

$$\mathbf{H}^{\blacktriangle}(A\tilde{A}|B)_{\Phi^{(k)}_{AB}\otimes\mathbf{u}_{\tilde{A}}}$$

$$= \mathbf{H}^{\blacktriangle}(A\tilde{A}|BB_2)_{\Phi_{AB}^{(k)}\otimes \mathbf{u}_{\tilde{A}B_2}} \tag{149}$$

$$= \mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} + \mathbf{H}^{\blacktriangle}(\tilde{A}|B_2)_{\mathbf{u}_{\tilde{A}B_2}} \tag{150}$$

$$= \mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} + \mathbf{H}^{\blacktriangle}(\mathbf{u}_{\tilde{A}}) \tag{151}$$

$$= \mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} + \log_2 k, \tag{152}$$

where we used the additivity property of conditional entropy and the fact that the entropy of the uniform state $\mathbf{u}_{\tilde{A}}$ is equal to $\log_2 k$. It is therefore sufficient to show that

$$\mathbf{H}^{\blacktriangle}(A\tilde{A}|B)_{\Phi_{AB}^{(k)}\otimes \mathbf{u}_{\tilde{A}}} \leq 0, \tag{153}$$

because the above equation implies that

$$\mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} \leq -\log_2 k = H_{\min}^{\uparrow}(A|B)_{\Phi^{(k)}}\ . \tag{154}$$

For this purpose, let X be a classical system of dimension $|X| = k^2$ and let $\mathcal{N}_{A\tilde{A}B\to X}$ be a quantum channel defined by

$$\mathcal{N}_{A\tilde{A}B\to X} \coloneqq \mathcal{E}_{AB\to X} \circ \operatorname{Tr}_{\tilde{A}}, \tag{155}$$

where

$$\mathcal{E}_{AB\to X}(\omega_{AB}) \coloneqq \operatorname{Tr}[\Phi_{AB}^{k}\omega_{AB}]|1\rangle\langle 1|_X + \operatorname{Tr}[(I_{AB} - \Phi_{AB}^{(k)})\omega_{AB}]\frac{I_X - |1\rangle\langle 1|_X}{k^2-1}. \tag{156}$$

This channel is conditionally unital because

$$\mathcal{N}_{A\tilde{A}B\to X}(I_A \otimes I_{\tilde{A}} \otimes \sigma_B)$$

$$= (\mathcal{E}_{AB\to X} \circ \operatorname{Tr}_{\tilde{A}})(I_A \otimes I_{\tilde{A}} \otimes \sigma_B) \tag{157}$$

$$= k\mathcal{E}_{AB\to X}(I_A \otimes \sigma_B) \tag{158}$$

$$= k\begin{pmatrix} \operatorname{Tr}[\Phi_{AB}^{(k)}(I_A \otimes \sigma_B)]|1\rangle\langle 1|_X \\ + \operatorname{Tr}[(I_{AB} - \Phi_{AB}^{(k)})(I_A \otimes \sigma_B)]\frac{I_X - |1\rangle\langle 1|_X}{k^2-1} \end{pmatrix} \tag{159}$$

$$= k\left(\frac{1}{k}|1\rangle\langle 1|_X + \left(k - \frac{1}{k}\right)\frac{I_X - |1\rangle\langle 1|_X}{k^2-1}\right) \tag{160}$$

$$= |1\rangle\langle 1|_X + \left(k^2-1\right)\frac{I_X - |1\rangle\langle 1|_X}{k^2-1} \tag{161}$$

$$= I_X. \tag{162}$$

Also, consider that

$$\mathcal{N}_{A\tilde{A}B\to X}(\Phi_{AB}^{(k)} \otimes \mathbf{u}_{\tilde{A}}) = |1\rangle\langle 1|_X. \tag{163}$$

Then

$$\mathbf{H}^{\blacktriangle}(A|B)_{\Phi^{(k)}} + \log_2 k = \mathbf{H}^{\blacktriangle}(A\tilde{A}|B)_{\Phi^{(k)}\otimes \mathbf{u}} \tag{164}$$

$$\leq \mathbf{H}^{\blacktriangle}(X)_{\mathcal{N}(\Phi^{(k)}\otimes \mathbf{u})} \tag{165}$$

$$= \mathbf{H}^{\blacktriangle}(X)_{|1\rangle\langle 1|} \tag{166}$$

$$= 0. \tag{167}$$

The first equality follows from (149)–(152). The inequality follows because the channel $\mathcal{N}_{A\tilde{A}B\to X}$ is conditionally unital. The second equality follows from (163) and the last because the entropy of every pure state is equal to zero. This concludes the proof.

E.2 Proof of Theorem 1

Here the idea is essentially the same as that given in Appendix E.1, except that we need to construct channels that are conditionally mixing (i.e., both conditionally unital and $A \not\to B$ semi-causal).

First, throughout, we make the following replacement in the proof:

$$H_{\min}^{\uparrow}(A|B)_\rho \quad \to \quad H_{\min}(A|B)_\rho. \tag{168}$$

To handle the case when $H_{\min}(A|B)_\rho \geq 0$, we employ the same channel in (112), with the exception that we set

$$\tau_{AB}^{(m)} \coloneqq \frac{I_A \otimes \rho_B - m\rho_{AB}}{k-m}. \tag{169}$$

All of the steps in (106)–(123) go through with the replacements in (168) and (169). We now verify that the $A \not\to B$ semi-causal constraint holds for this channel. Consider, for the modified channel, that

$$\operatorname{Tr}_A[\mathcal{N}_{X\to AB}(\omega_X)]$$

$$= \operatorname{Tr}[\Pi_X\omega_X]\operatorname{Tr}_A[\rho_{AB}] + \operatorname{Tr}[(I_X - \Pi_X)\,\omega_X]\operatorname{Tr}_A[\tau_{AB}^{(m)}] \tag{170}$$

$$= \operatorname{Tr}[\Pi_X\omega_X]\rho_B + \operatorname{Tr}[(I_X - \Pi_X)\,\omega_X]\rho_B \tag{171}$$

$$= \rho_B. \tag{172}$$

Thus, the output state of B has no dependence on the input state ω_X when system A is traced out, so that the semi-causal constraint holds.

Now let us consider the case when $H_{\min}(A|B)_\rho < 0$. We employ the same channel in (131), with the exception that we set

$$\tau_{AB}^{(k,k')} \coloneqq \frac{k' I_A \otimes \rho_B - \rho_{AB}}{kk' - 1}. \tag{173}$$

Then all of the steps in (124)–(148) go through, with the replacements in (168) and (173). We now verify that the $A \not\to B$ semi-causal constraint holds for this channel. Consider, for the modified channel, that

$$\operatorname{Tr}_{AA'}[\mathcal{N}_{X\to AA'B}(\omega_X)]$$

$$= \Big(\operatorname{Tr}[|1\rangle\langle 1|_X\omega_X]\operatorname{Tr}_A[\rho_{AB}] + \operatorname{Tr}[\Pi_X\omega_X]\operatorname{Tr}_A[\tau_{AB}^{(k,k')}]\Big) \otimes \operatorname{Tr}_{A'}[\mathbf{u}_{A'}] \tag{174}$$

$$= \operatorname{Tr}[|1\rangle\langle 1|_X\omega_X]\rho_B + \operatorname{Tr}[\Pi_X\omega_X]\rho_B \tag{175}$$

$$= \rho_B. \tag{176}$$

Thus, the output state of B has no dependence on the input state ω_X when system A is traced out, so that the semi-causal constraint holds. This establishes the inequality in (103).

Finally, we consider the equality in (104). The argument here is actually exactly the same as in (149)–(167),

except using the replacement in (168). We just need to verify that the channel in (155) is $A \not\to B$ semi-causal. However, the only output is system X, which belongs to Alice. Thus, after tracing it out, there is no remaining system, and the channel is then semi-causal.

F Proof of Theorem 2

First, we demonstrate that $\mathbf{H}$ satisfies the monotonicity property of conditional entropy from Definition 3. Let $\mathcal{N}_{AB\to A'B'}$ be a conditionally mixing channel, and consider the bipartite density matrix ρ_{AB}. We begin by considering the case where $A = A'$ and

$$\mathbf{H}(A|B)_{\mathcal{N}(\rho)} = \log_2 |A| - \mathbf{D}(\mathcal{N}(\rho_{AB})\|\mathbf{u}_A \otimes \mathrm{Tr}_A[\mathcal{N}(\rho_{AB})]) \quad (177)$$

and from using the fact that $\mathcal{N}$ is semi-causal and conditionally unital, we find that

$$\begin{aligned} &\mathbf{u}_A \otimes \mathrm{Tr}_A[\mathcal{N}(\rho_{AB})] \\ &= \mathbf{u}_A \otimes \mathrm{Tr}_A[\mathcal{N}(\mathbf{u}_A \otimes \rho_B)] && (178)\\ &= \mathcal{N}(\mathbf{u}_A \otimes \rho_B), && (179)\end{aligned}$$

from which we have

$$\begin{aligned} &\mathbf{H}(A|B)_{\mathcal{N}(\rho)} \\ &= \log_2 |A| - \mathbf{D}(\mathcal{N}(\rho_{AB})\|\mathcal{N}(\mathbf{u}_A \otimes \rho_B)) && (180)\\ &\geq \log_2 |A| - \mathbf{D}(\rho_{AB}\|\mathbf{u}_A \otimes \rho_B), && (181)\end{aligned}$$

where the last line follows from the data-processing inequality.

We also need to prove that it is invariant under the action of a local isometric channel acting on system A. Let $\mathcal{V}_{A\to A'}$ be an arbitrary isometric channel, and let C represent a Hilbert space such that $|C| = |A'| - |A|$. Then

$$\mathbf{H}(A|B)_{\mathcal{V}(\rho)} = \log_2 |A'| - \mathbf{D}(\mathcal{V}_{A\to A'}(\rho_{AB})\|\mathbf{u}_{A'} \otimes \rho_B). \quad (182)$$

Clearly, if $\mathcal{V}$ is a unitary channel such that $A \simeq A'$, then $\mathbf{H}(A'|B)_{\mathcal{V}(\rho)} = \mathbf{H}(A|B)_\rho$. Then without loss of generality, we assume that

$$\begin{aligned} \mathcal{V}_{A\to A'}(\rho_{AB}) &= \rho_{AB} \oplus \mathbf{0}_{CB} && (183)\\ &= \begin{pmatrix} \rho_{AB} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{CB} \end{pmatrix}, && (184)\end{aligned}$$

because the conditional entropy is invariant under local unitaries. We can additionally rewrite

$$\mathbf{u}_{A'} = t\mathbf{u}_A \oplus (1-t)\mathbf{u}_C, \quad (185)$$

where $t \coloneqq \frac{|A|}{|A'|}$. From this, it follows that

$$\begin{aligned} &\mathbf{D}(\mathcal{V}_{A\to A'}(\rho_{AB})\|\mathbf{u}_{A'} \otimes \rho_B) \\ &= \mathbf{D}\big(\rho_{AB} \oplus \mathbf{0}_{CB}\big\|\, (t\mathbf{u}_A \otimes \rho_B) \oplus ((1-t)\,\mathbf{u}_C \otimes \rho_B)\big) \\ &= \mathbf{D}\big(\rho_{AB}\big\|t\mathbf{u}_A \otimes \rho_B\big) \\ &= \mathbf{D}\big(\rho_{AB}\big\|\mathbf{u}_A \otimes \rho_B\big) - \log_2 t, && (186)\end{aligned}$$

where the second equality follows from a property of quantum relative entropy [57, Theorem 6.3.1]. Thus, upon substitution, we have that

$$\begin{aligned} &\mathbf{H}(A'|B)_{\mathcal{V}(\rho)} \\ &= \log_2 |A'| - \mathbf{D}\big(\rho_{AB}\big\|\mathbf{u}_A \otimes \rho_B\big) + \log_2 t && (187)\\ &= \mathbf{H}(A|B)_\rho. && (188)\end{aligned}$$

Putting the above two observations together, we now apply them to the definition given in Definition 3. That is, suppose there exists a system A'' such that $|A''| \geq \max\{|A|, |A'|\}$, isometric channels $\mathcal{V}_{A\to A''}$ and $\mathcal{U}_{A'\to A''}$, and a conditionally mixing channel $\mathcal{N}_{A''B\to A''B'}$ such that

$$\mathcal{U}_{A'\to A''}(\sigma_{A'B'}) = (\mathcal{N}_{A''B\to A''B'} \circ \mathcal{V}_{A\to A''})(\rho_{AB}) \quad (189)$$

Then

$$\begin{aligned} \mathbf{H}(A'|B')_\sigma &= \mathbf{H}(A''|B')_{\mathcal{U}(\sigma)} && (190)\\ &= \mathbf{H}(A''|B')_{(\mathcal{N}\circ\mathcal{V})(\rho)} && (191)\\ &\leq \mathbf{H}(A''|B)_{\mathcal{V}(\rho)} && (192)\\ &= \mathbf{H}(A|B)_\rho. && (193)\end{aligned}$$

In the above steps, we employed the invariance of $\mathbf{H}$ under local isometric channels acting on the systems labeled by A, as well as the monotonicity under conditionally mixing channels.

Finally, additivity of $\mathbf{H}$ follows immediately from additivity of $\mathbf{D}$.

F.1 Conditional min-entropy is a conditional entropy

Here we justify that the conditional min-entropy is indeed a conditional entropy according to the definition in (19)–(21). Indeed, we use the definition of conditional min-entropy in (27) to rewrite it in terms of the max-relative entropy $D_{\max}(\rho\|\sigma)$ as

$$H_{\min}(A|B)_\rho = \log_2 |A| - D_{\max}(\rho_{AB}\|\mathbf{u}_A \otimes \rho_B), \quad (194)$$

where the max-relative entropy of states ρ and σ is defined as [58]:

$$D_{\max}(\rho\|\sigma) \coloneqq \log_2 \inf_{\lambda\geq 0}\{\rho \leq \lambda\sigma\}. \quad (195)$$

Then we apply Theorem 2.

G Entropy inequalities

Entropy inequalities play a major role in information theory. For example, the strong subadditivity of the von Neumann entropy is a deep result in quantum Shannon theory that can be expressed as an inequality involving the von Neumann conditional entropy:

$$H(A|B)_\rho \geq H(A|BC)_\rho \quad \forall \rho \in \mathfrak{D}(ABC)\,, \tag{196}$$

where $H(A|B)_\rho \coloneqq H(AB)_\rho - H(B)_\rho$ and $H(A)_\rho \coloneqq -\mathrm{Tr}[\rho_A \log_2 \rho_A]$ is the von Neumann entropy. The above inequality holds for all conditional entropies: namely, for every conditional entropy $\mathbf{H}$, we have

$$\mathbf{H}(A|B)_\rho \geq \mathbf{H}(A|BC)_\rho \quad \forall \rho \in \mathfrak{D}(ABC)\,. \tag{197}$$

The above inequality is a simple consequence of the fact that tracing the subsystem C is a conditionally mixing channel and therefore

$$\rho_{ABC} \succ_A \rho_{AB}\,. \tag{198}$$

It is therefore natural to ask if there are other entropy inequalities that can be applied to all entropies. In this section we present a few such inequalities.

We show here that the teleportation protocol of communicating a quantum system from Bob to Alice is a conditionally mixing channel. This property implies in particular that (see arguments in Proposition 1 below)

$$\rho_{AB} \otimes \Phi_{\tilde{A}\tilde{B}} \succ_{A\tilde{A}} \rho_{A\tilde{A}}\,, \tag{199}$$

where $\Phi_{\tilde{A}\tilde{B}}$ is a maximally entangled state, $\rho_{A\tilde{A}}$ is the same state as ρ_{AB} with system B relabeled as $\tilde{A}$, and we assume for simplicity that all systems have the same dimension. The above relation immediately implies that every conditional entropy $\mathbf{H}$ satisfies the following:

$$\begin{aligned}
\mathbf{H}(AB)_\rho &= \mathbf{H}(A\tilde{A})_\rho && (200)\\
&\geq \mathbf{H}(A\tilde{A}|B\tilde{B})_{\rho\otimes\Phi} && (201)\\
&= \mathbf{H}(A|B)_\rho + \mathbf{H}(\tilde{A}|\tilde{B})_\Phi\,, && (202)
\end{aligned}$$

where the equality follows from additivity. Now, from Theorem 1 we get that $\mathbf{H}(\tilde{A}|\tilde{B})_\Phi = -\log_2|B|$, so that we arrive at the following entropy inequality:

$$\mathbf{H}(AB)_\rho \geq \mathbf{H}(A|B)_\rho - \log_2|B|\,. \tag{203}$$

Proposition 1 *Let ρ_{AB} be a bipartite state. Then*

$$\mathbf{H}(A|B)_\rho \leq \mathbf{H}(AB)_\rho + \log_2|B|\,, \tag{204}$$

$$\mathbf{H}^{\blacktriangle}(A|B)_\rho \leq \mathbf{H}(AB)_\rho + \log_2|B|\,. \tag{205}$$

Proof. Let $\Phi_{A'B'}$ be a maximally entangled state of Schmidt rank $|B|$, with $|A'| = |B'| = |B|$. We perform teleportation to transmit system A to A'. In more detail, the teleportation channel we consider is

$$\mathcal{N}_{A'BB'\to A'}(\omega_{A'BB'}) \coloneqq \sum_x \mathcal{U}^x_{A'}\left(\langle\Phi^x|_{BB'}\omega_{A'BB'}|\Phi^x\rangle_{BB'}\right), \tag{206}$$

where $\{|\Phi^x\rangle_{BB'}\}_x$ is the standard Bell basis and $\{\mathcal{U}^x_{A'}\}_x$ is the set of unitary correction channels in teleportation. The channel $\mathcal{N}_{A'BB'\to A'}$ is conditionally unital because

$$\begin{aligned}
&\mathcal{N}_{A'BB'\to A'}(I_{A'}\otimes\sigma_{BB'})\\
&= \sum_x \mathcal{U}^x_{A'}\left(\langle\Phi^x|_{BB'}\left(I_{A'}\otimes\sigma_{BB'}\right)|\Phi^x\rangle_{BB'}\right) && (207)\\
&= \sum_x \mathcal{U}^x_{A'}(I_{A'})\otimes\langle\Phi^x|_{BB'}\left(\sigma_{BB'}\right)|\Phi^x\rangle_{BB'} && (208)\\
&= \sum_x I_{A'}\otimes\langle\Phi^x|_{BB'}\left(\sigma_{BB'}\right)|\Phi^x\rangle_{BB'} && (209)\\
&= I_{A'}\otimes\sum_x\langle\Phi^x|_{BB'}\left(\sigma_{BB'}\right)|\Phi^x\rangle_{BB'} && (210)\\
&= I_{A'}. && (211)
\end{aligned}$$

The channel $\mathcal{N}_{A'BB'\to A'}$ is also semi-causal because it is realized as a one-way LOCC channel from BB' to A'. Thus, $\mathcal{N}_{A'BB'\to A'}$ is conditionally mixing. Then consider that

$$\begin{aligned}
&\mathbf{H}(A|B)_\rho - \log_2|B|\\
&= \mathbf{H}(A|B)_\rho + \mathbf{H}(A'|B')_\Phi && (212)\\
&= \mathbf{H}(AA'|BB') && (213)\\
&\leq \mathbf{H}(AA')_{\mathcal{N}(\rho\otimes\Phi)} && (214)\\
&= \mathbf{H}(AA')_\rho && (215)\\
&= \mathbf{H}(AB)_\rho. && (216)
\end{aligned}$$

The first equality follows from Theorem 1. The second equality follows from the additivity postulate. The inequality follows because $\mathcal{N}_{A'BB'\to A'}$ is conditionally mixing, as argued above. The final equality is just a relabeling of systems.

The inequality in (205) follows similarly, using the fact that $\mathcal{N}_{A'BB'\to A'}$ is conditionally unital. ■

Proposition 2 *Let ρ_{AB} be a bipartite state. Then*

$$\mathbf{H}^{\blacktriangle}(A|B)_\rho \geq \mathbf{H}(AB)_\rho - \log_2|B|\,. \tag{217}$$

Proof. Consider the initial state

$$\rho_{AA'}\otimes|0\rangle\langle 0|_{B'}, \tag{218}$$

where $|A'| = |B'| = |B|$ and we have relabeled B of ρ_{AB} as A' (physically, the system B begins on Alice's side, and so we label it as A'). Then perform the swap and prepare channel from [34, Definition 8], which was shown there to be a conditionally unital channel. That is, this channel is realized as $\mathcal{D}_{A'}\circ\mathcal{F}_{A'B'}$, where $\mathcal{F}_{A'B'}$ is the unitary swap channel and $\mathcal{D}_{A'}$ is the completely

depolarizing channel (traces out its input and replaces with the uniform state). This is conditionally unital because

$$\begin{aligned}(\mathcal{D}_{A'} \circ \mathcal{F}_{A'B'})(I_{A'} \otimes \sigma_{B'}) &= \mathcal{D}_{A'}(\sigma_{A'} \otimes I_{B'}) && (219)\\ &= \mathcal{D}_{A'}(\sigma_{A'}) \otimes I_{B'} && (220)\\ &= \mathbf{u}_{A'} \otimes I_{B'} && (221)\\ &= I_{A'} \otimes \mathbf{u}_{B'}. && (222)\end{aligned}$$

However, this channel is not semi-causal because it allows for communication from A' to B'. That is,

$$\begin{aligned}&(\mathcal{D}_{A'} \circ \mathcal{F}_{A'B'})(\omega_{A'} \otimes \sigma_{B'})\\ &= \mathcal{D}_{A'}(\sigma_{A'} \otimes \omega_{B'}) && (223)\\ &= \mathbf{u}_{A'} \otimes \omega_{B'}. && (224)\end{aligned}$$

Now consider that

$$\begin{aligned}&(\mathcal{D}_{A'} \circ \mathcal{F}_{A'B'})(\rho_{AA'} \otimes |0\rangle\langle 0|_{B'})\\ &= \mathcal{D}_{A'}(\rho_{AB'} \otimes |0\rangle\langle 0|_{A'}) && (225)\\ &= \rho_{AB'} \otimes \mathcal{D}_{A'}(|0\rangle\langle 0|_{A'}) && (226)\\ &= \rho_{AB'} \otimes \mathbf{u}_{A'}. && (227)\end{aligned}$$

Then

$$\begin{aligned}&\mathbf{H}(AB)_\rho\\ &= \mathbf{H}(AA')_\rho && (228)\\ &= \mathbf{H}^{\blacktriangle}(AA'|B')_{\rho_{AA'} \otimes |0\rangle\langle 0|_{B'}} && (229)\\ &\leq \mathbf{H}^{\blacktriangle}(AA'|B')_{(\mathcal{D}_{A'} \circ \mathcal{F}_{A'B'})(\rho_{AA'} \otimes |0\rangle\langle 0|_{B'})} && (230)\\ &= \mathbf{H}^{\blacktriangle}(AA'|B')_{\rho_{AB'} \otimes \mathbf{u}_{A'}} && (231)\\ &= \mathbf{H}^{\blacktriangle}(A|B')_{\rho_{AB'}} + \mathbf{H}(A')_{\mathbf{u}_{A'}} && (232)\\ &= \mathbf{H}^{\blacktriangle}(A|B)_{\rho_{AB}} + \log_2 |B|. && (233)\end{aligned}$$

The first equality follows from relabeling systems. The second equality follows because conditional entropy is invariant under tensoring with another state for the conditioning system, in turn being a special kind of conditionally unital channel. The inequality follows because the swap and prepare channel is conditionally unital, as argued above. The penultimate equality follows from additivity, and the final equality because the entropy of the uniform state of dimension $|B|$ is equal to $\log_2 |B|$. ■

Proposition 3 *Let ρ_{XA} be a classical–quantum state of the following form:*

$$\rho_{XA} = \sum_x p(x)|x\rangle\langle x|_X \otimes \rho_A^x. \quad (234)$$

Then

$$\mathbf{H}(X)_\rho \leq \mathbf{H}(AX)_\rho. \quad (235)$$

Proof. For every state ρ_A^x, there is a unital channel $\mathcal{N}_A^x$ that prepares it from the pure state $|0\rangle\langle 0|_A$. Then we take the overall channel to be one that measures the classical system X in the standard basis $\{|x\rangle\}_x$ and conditioned on the outcome, performs the channel $\mathcal{N}_A^x$ on the system A. Thus, the overall channel is realized as

$$\mathcal{N}_{XA}(\omega_{XA}) = \sum_x |x\rangle\langle x|_X \otimes \mathcal{N}_A^x(\langle x|_X \omega_{XA} |x\rangle_X). \quad (236)$$

This channel is unital because

$$\begin{aligned}&\mathcal{N}_{XA}(I_{XA})\\ &= \mathcal{N}_{XA}(I_X \otimes I_A) && (237)\\ &= \sum_x |x\rangle\langle x|_X \otimes \mathcal{N}_A^x(\langle x|_X (I_X \otimes I_A) |x\rangle_X) && (238)\\ &= \sum_x |x\rangle\langle x|_X \otimes \mathcal{N}_A^x(I_A) && (239)\\ &= \sum_x |x\rangle\langle x|_X \otimes I_A && (240)\\ &= I_X \otimes I_A. && (241)\end{aligned}$$

Furthermore,

$$\mathcal{N}_{XA}(\rho_X \otimes |0\rangle\langle 0|_A) = \rho_{XA}. \quad (242)$$

This implies that

$$\begin{aligned}\mathbf{H}(X)_{\rho_X} &= \mathbf{H}(XA)_{\rho_X \otimes |0\rangle\langle 0|_A} && (243)\\ &\leq \mathbf{H}(XA)_{\mathcal{N}_{XA}(\rho_X \otimes |0\rangle\langle 0|_A)} && (244)\\ &= \mathbf{H}(XA)_{\rho_{XA}}. && (245)\end{aligned}$$

This concludes the proof. ■

H Example of a non-negative, weakly additive monotonic function

In this appendix, we prove that there exists a non-zero function that satisfies the following three properties:

1. Monotonicity under conditional majorization; i.e., satisfying (20).
2. Weak additivity of the form (36).
3. Non-negative.

The third property above demonstrates that the negativity of conditional entropy is closely related to the full additivity postulate in (21). The construction of the example is based on a technique developed in [41] for the extension of resource monotones from one domain to another.

Let $\mathbb{H}$ be a classical conditional entropy. In [41] it was shown that every such function has two optimal extensions to the quantum domain. Here we are interested in the following extension. For every $\rho \in \mathfrak{D}(AB)$, we define

$$f(A|B)_\rho \coloneqq \inf_{\rho_{AB} \succ_A \sigma_{XY}} \mathbb{H}(X|Y)_\sigma \ , \tag{246}$$

where the infimum above is over all classical systems X and Y and every classical state σ_{XY} with the property that $\rho_{AB} \succ_A \sigma_{XY}$. The function f has the following key properties [41]:

1. Reduction: For every classical state $\rho \in \mathfrak{D}(XY)$

$$f(X|Y)_\rho = \mathbb{H}(X|Y)_\rho \ . \tag{247}$$

2. Monotonicity: For all $\rho \in \mathfrak{D}(AB)$ and $\omega \in \mathfrak{D}(A'B')$ such that $\rho_{AB} \succ_A \omega_{A'B'}$,

$$f(A'|B')_\omega \geq f(A|B)_\rho \ . \tag{248}$$

3. Subadditivity: The function f is subadditive under tensor-product states.

4. Positivity: The function f is non-negative.

The last property follows from the definition in (246). Since f is subadditive, the limit

$$g(A|B)_\rho \coloneqq \lim_{n\to\infty} \frac{1}{n} f(A^n|B^n)_{\rho^{\otimes n}} \tag{249}$$

exists. Therefore, the function g is weakly additive. Moreover, the function g satisfies the first two properties above, namely, reduction to $\mathbb{H}$ on classical states and behaves monotonically under conditional majorization. Observe that the reduction of g to $\mathbb{H}$ on classical states implies that it is not equal to the zero function since $\mathbb{H}$ is not the zero function. Since g is non-negative it cannot be a conditional entropy. This means that g cannot be fully additive, as this is the only property that does not follow directly from its definition.