Quantum conditional entropy from information-theoretic principles

Sarah Brandsen,1,2, * Isabelle Jianing Geng,2 Mark M. Wilde,3 and Gilad Gour2

1Department of Physics, Duke University, Durham, NC, USA 27708
2Department of Mathematics and Statistics, Institute for Quantum Science and Technology, University of Calgary, AB, Canada T2N 1N4
3Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

We introduce an axiomatic approach for characterizing quantum conditional entropy. Our approach relies on two physically motivated axioms: monotonicity under conditional majorization and additivity. We show that these two axioms provide sufficient structure that enable us to derive several key properties applicable to all quantum conditional entropies studied in the literature. Specifically, we prove that any quantum conditional entropy must be negative on certain entangled states and must equal $-\log d$ on $d \times d$ maximally entangled states. We also prove the non-negativity of conditional entropy on separable states, and we provide a generic definition for the dual of a quantum conditional entropy. Finally, we develop an operational approach for characterizing quantum conditional entropy via games of chance, and we show that, for the classical case, this complementary approach yields the same ordering as the axiomatic approach.

I. INTRODUCTION

Conditional entropy quantifies the uncertainty of a bipartite state shared between Alice and Bob, given access to Bob’s system only. Previous work on conditional entropy for quantum bipartite states has demonstrated that, unlike classical conditional entropy, quantum conditional entropy can be negative [1] and achieves its minimum on maximally entangled states. The initial discovery that quantum conditional entropy could be negative was so surprising that various conjectures about its meaning were put forth by the authors of [1], such as its potential relation to “anti-qubits,” which are quanta of negative information. The most satisfying information theoretic interpretation of quantum conditional entropy and its negativity is given by the state merging protocol of [2], which demonstrates that conditional entropy is the rate at which maximally entangled Bell states need to be consumed in order to transfer Alice’s share of a bipartite state to Bob, who possesses the other share of the state. The importance of the existence of negative conditional entropy is also reflected in the fact that there is an information quantity, called coherent information [3], dedicated to measuring how negative the conditional entropy of a bipartite quantum state is. Multiple previous works have investigated topics related to conditional entropy [4–12] and have found additional applications of conditional entropy to quantum cryptographic protocols [13].

The central role of conditional entropy in quantum information theory motivates the need for an axiomatic approach that defines the set of all possible quantum conditional entropies. Previous literature has extensively studied axiomatic derivations for other entropies [14, 15], beginning with approaches deriving the Shannon entropy [16] and progressing to a complete axiomatic approach for classical entropies and relative entropies [17]. More recently, an axiomatic approach for entropy based on core-concavity was developed and connections were made between such core-concave functions and classical conditional entropy [18].

In this work, we start with a minimal set of assumptions with strong information-theoretic motivation and demonstrate that these simple assumptions are sufficient for recovering key properties of quantum conditional entropy, such as its negativity for maximally entangled states, non-negativity for separable bipartite states, invariance under local unitaries, and reduction to an entropy function on uncorrelated systems. At the core of our approach, we fully characterise the set of channels that are conditional entropy non-decreasing, and we require that conditional entropy be monotonic under the action of such channels. More specifically, we require that any conditional entropy non-decreasing channel maps a state of maximal entropy on Alice’s side to an output state that also has maximal entropy on Alice’s side, a condition which is satisfied by conditionally unital channels. Likewise, given that conditional entropy can be understood as measuring how much information Bob has about Alice’s system, any conditional entropy non-decreasing channel cannot “leak” information from Alice’s system to Bob’s system. Thus, conditional entropy non-decreasing channels must also satisfy the semi-causality condition outlined in [19]. In other words, we identify the set of conditional entropy non-decreasing channels with channels that are both conditionally unital and semi-causal (CUSC).

This gives rise to a natural majorization relation where $\sigma_{AB} \preceq_A \rho_{AB}$ (i.e. $\rho_{AB}$ has less uncertainty than $\sigma_{AB}$ with respect to system $A$) if $\sigma_{AB}$ can be obtained by applying a CUSC channel to $\rho_{AB}$. Any conditional entropy is thus required to be monotonic under this partial ordering which we call conditional majorization. In [11] a different definition of conditional majorization (restricted
to classical systems) was given, however, we prove that
the quantum conditional majorization defined in this pa-
er holds the exact same ordering as in [11] on (bipartite)
classical systems.

An alternative, operationally motivated means for charac-
terizing uncertainty is to utilize games of chance. It
was found in [20] that games of chance provide an op-
erational interpretation to conditional majorization for
classical bipartite states. Games of chance are ideal for
studying uncertainty, as the probability of winning a
game with a given physical system depends solely on the
uncertainty of the system’s output. Such games of chance
does lead to a natural partial ordering between channels,
where state $\rho_{AB}$ is “less uncertain” than state $\sigma_{AB}$
if $\rho_{AB}$ performs at least as well as $\rho_{AB}$ for any game
of chance. In this work, we extend previous results by
providing games of chance capable of characterising un-
certainty of quantum bipartite states.

Lastly, we provide an alternative approach for charac-
terising the entropy of a quantum channel [21], which
is operationally motivated by games of chance. Our
results reflect that entanglement is a resource [22, 23],
as entanglement-preserving unitary channels strictly out-
perform entanglement-breaking channels. We addition-
ally compute the reward function for key quantum gam-
bling games with several special classes of quantum chan-
nels, such as unitary, amplitude damping, depolarizing,
measurement, and dephasing channels.

II. NOTATION

We denote quantum channels in uppercase calligraphic
letters, where $\mathcal{N} \in \text{CPTP}(AB \rightarrow AB')$ denotes a com-
pletely positive, trace preserving map (i.e. quantum
channel) that takes as an input a bipartite state of sys-
tems $A$ and $B$ and outputs a bipartite state of systems
$A'$ and $B'$. At times, we use subscripts to indicate the
input and output system of a channel such as $\mathcal{N}_{AB \rightarrow AB'}$.
The dimension of a system is denoted as $|A|$, $|B|$, e.t.c.,
We also use the tilde symbol to indicate a replica of a
system. For example, $\tilde{A}$ or $\tilde{B}$ represents copies of $A$ and
$B$, and in particular $|\tilde{A}| = |A|$ and $|\tilde{B}| = |B|$.

Likewise, we denote a classical channel by $\mathcal{T}$, whose
corresponding transition matrix is a column stochastic
matrix $T = (t_{w|z'})_{w,z'}$ such that each entry $t_{w|z'} \in [0,1]$
and $\sum_{w} t_{w|z'} = 1$ for each $z'$.

In the case in which the bipartite state has a trivial com-
ponent (e.g., $|B| = |B'| = 1$), we omit the letters that
refer to the trivial system and denote the channel
$\mathcal{N}$ as $\mathcal{N} \in \text{CPTP}(A \rightarrow A')$. Throughout this work, the
computational basis of a given system with dimension $d$
is denoted as $\{|1\rangle, \ldots , |d\rangle\}$. We denote the maximally
entangled Bell state for systems $A$ and $B$ (where $|A| =
|B|$) as

$$|\phi_{AB}^+\rangle \triangleq \frac{1}{\sqrt{|A|}} \sum_{j=1}^{|A|} |jj\rangle,$$

where $|jj\rangle \triangleq |j\rangle \otimes |j\rangle$, and denote the corresponding un-
normalized state as

$$|\Phi_{AB}\rangle \triangleq \sum_{j=1}^{|A|} |jj\rangle.$$

For a matrix $M$, the sum of its $w$ largest singular val-
ues will be denoted as the Ky-fan $w$ norm $\|M\|_w$. We
denote with $I_A$ the identity operator on system $A$, and
denote with $u_A = \frac{1}{|A|} I_A$ the maximally mixed state in
system $A$. When the systems on which the operators act
are clear from the context, we might omit the subscripts
indicating the systems.

The Choi matrix $J_{AB}^\mathcal{N}$ for a given channel $\mathcal{N}_{A \rightarrow B}$ is
defined by the action of $\mathcal{N}$ on the maximally entangled
operator as

$$J_{AB}^\mathcal{N} = \sum_{i,j=1}^{|A|} |i\rangle\langle j| \otimes \mathcal{N}(|i\rangle\langle j|).$$

III. AXIOMATIC APPROACH TO QUANTUM
CONDITIONAL ENTROPY

We begin by completely characterising the set of bi-
partite quantum channels which are conditional entropy
non-decreasing. A bipartite quantum channel $\mathcal{N}$ is con-
tional entropy non-decreasing if for any bipartite state
$\rho_{AB}$, the state $\mathcal{N}(\rho_{AB})$ has at least as much conditional
entropy.

Since conditional entropy measures the uncertainty of
system $A$ when one is given access to system $B$, evidently
allowing information to leak from system $A$ to system $B$
could decrease conditional entropy. This gives rise to the
semi-causality requirement, which is defined as follows

**Definition 1.** A bipartite channel $\mathcal{N}_{AB \rightarrow AB'}$ is $A \leftrightarrow B'$

semi-causal if any channel $\mathcal{M}_{A \rightarrow A}$ that Alice applies to
her system cannot be detected by Bob. Formally, this
requirement can be stated as follows:

$$\mathcal{N}_{AB \rightarrow AB'} \circ \mathcal{M}_{A \rightarrow A} = \mathcal{N}_{AB \rightarrow B'},$$

for all $\mathcal{M} \in \text{CPTP}(A \rightarrow A)$ and where $\mathcal{N}_{AB \rightarrow B'}$ indi-
cates that the channel output in system $A$ is traced out.

See Figure 1 for a depiction of the semi-causality re-
requirement.
satisfies \( u \) in which maps any input to the maximally mixed state.

Let \( A \) and \( N \)

\[ F \in \mathbb{R}^{B} \]

can only flow from only on \( A \) terms of Choi matrices.

Proof. Refer to [24] for a proof.

\[ \rho \text{ for every state} \]

Finally, we reframe the definition of semi-causality in terms of Choi matrices.

Lemma 2. Let \( N \in \text{CPTP}(AB \to AB') \). If \( N_{AB\to\tilde{A}B'} \rightarrow \tilde{A}B' \) is semicausal, then its Choi matrix \( J_{AB\to\tilde{A}B'}^{N} \) satisfies

\[ J_{AB\to\tilde{A}B'}^{N} = u_{A} \otimes J_{BB'}^{N} \]  

Proof. Let \( R_{A'} \rightarrow A \) be the completely randomising channel which maps any input to the maximally mixed state \( u_{A} \).

Since \( N \) is semi-causal, then

\[ \text{Tr}_{\tilde{A}}[N_{AB\to\tilde{A}B'} \circ R_{A'} \rightarrow A] = \text{Tr}_{\tilde{A}}[N_{AB\to\tilde{A}B}] \]

and \( J_{AB\to\tilde{A}B'}^{N} = J_{AB\to\tilde{A}B'}^{N} \). We note that the Choi matrix for \( N_{AB\to\tilde{A}B'} \circ R \) is then equivalent to the Choi matrix for \( N_{AB\to\tilde{A}B'} \) and find

\[ J_{AB\to\tilde{A}B'}^{N} = \sum_{i,j=1}^{[A]} \sum_{k,l=1}^{[B]} |i\rangle \langle j|_{AB} \otimes |k\rangle \langle l|_{AB'} (R_{i,j} \otimes |k\rangle \langle l|) \]

\[ = u_{A} \otimes \sum_{k,l=1}^{[B]} |k\rangle \langle l|_{B} \otimes N_{AB\to\tilde{A}B'} (I_{A} \otimes |k\rangle \langle l|) \]

\[ = u_{A} \otimes M \]  

where \( M = \sum_{k,l=1}^{[B]} |k\rangle \langle l|_{B} \otimes N_{AB\to\tilde{A}B'} (I_{A} \otimes |k\rangle \langle l|) \). Upon taking the partial trace of \( A \) over both sides, then \( M = J_{BB'}^{N} \) and the theorem follows. \( \square \)

In addition to semi-causality, any entropy non-decreasing channel must map states of maximal conditional entropy (i.e. states of the form \( u_{A} \otimes \rho_{B} \)) to other states of maximal conditional entropy. We define channels satisfying this condition to be conditionally unital.

Definition 2. A bipartite channel \( N_{AB\to\tilde{A}B'} \rightarrow \tilde{A}B' \) is a conditionally unital channel if for all \( \rho \in \mathcal{D}(B) \), there exists a state \( \sigma_{B'} \in \mathcal{D}(B') \) such that

\[ \langle N(u_{A} \otimes \rho_{B}) \rangle = u_{A} \otimes \sigma_{B'} \]

We can also reframe the definition of a conditional unital channel in terms of Choi matrices.

Lemma 3. \( N \in \text{CPTP}(AB \to \tilde{A}B') \) is conditionally unital if and only if its Choi matrix \( J_{AB\to\tilde{A}B'}^{N} \) satisfies

\[ J_{AB\to\tilde{A}B'}^{N} = J_{BB'}^{N} \otimes u_{A} \]

Proof. Recall that the definition of a conditional unital channel can be written as

\[ \langle N(u_{A} \otimes \rho_{B}) \rangle = \text{Tr}_{\tilde{A}}[J_{AB\to\tilde{A}B'}^{N} (u_{A} \otimes (\rho_{B})_{T} \otimes I_{\tilde{A}B'})] \]

\[ = \frac{1}{|A|} \text{Tr}_{\tilde{A}}[J_{BB'}^{N} ((\rho_{B})_{T} \otimes I_{\tilde{A}B'})] \]

\[ \triangleq u_{A} \otimes \sigma_{B'} \]

We now multiply both sides by \( |A| \eta_{\tilde{A}} \otimes \omega_{B'} \) where \( \eta_{\tilde{A}} \) is traceless and Hermitian and \( \omega_{B'} \) is a Hermitian matrix. It follows that

\[ 0 = \text{Tr}[(I_{A} \otimes \sigma_{B'}) (\eta_{\tilde{A}} \otimes \omega_{B'})] \]

\[ = \text{Tr}_{\tilde{A}B'}[J_{BB'}^{N} (\rho_{B} \otimes I_{\tilde{A}B'}) (\eta_{\tilde{A}} \otimes \omega_{B'})] \]

\[ = \text{Tr}[J_{BB'}^{N} (\rho_{B} \otimes \eta_{\tilde{A}} \otimes \omega_{B'})] \]

From this, it follows that for any orthonormal basis \( \{ \rho_{B}^{(i)} \}_{i} \) of Hermitian operators, any orthonormal basis \( \{ \eta_{\tilde{A}}^{(j)} \}_{j} \) of traceless Hermitian operators, and any orthonormal basis \( \{ \omega_{B'}^{(k)} \}_{k} \) of Hermitian operators, then

\[ S = \{ \rho_{B}^{(i)} \otimes \eta_{\tilde{A}}^{(j)} \otimes \omega_{B'}^{(k)} \}_{i,j,k} \]

spans the orthogonal complement to \( J_{BB'}^{N} \).
We note that \( u_\tilde{A} \) is the only matrix (up to a scalar) which is orthogonal to all traceless Hermitian matrices \( \{|\eta^{(j)}_A\} \). Thus, \( J^N_{AB'} \) is spanned by \( \{ \rho_B \otimes u_\tilde{A} \otimes \omega_{B'} \} \), so
\[
J^N_{AB'} = \sum_{i,j,k} c_{i,j,k} \rho_B^{(i)} \otimes \omega_{B'}^{(k)} \otimes u_\tilde{A} \quad (16)
\]
\[
= J^N_{BB'} \otimes u_\tilde{A} \quad (17)
\]
This concludes the proof. □

Having motivated the need for a conditional entropy non-decreasing channel to be semi-causal and conditional uncertain than state \( \sigma \) (Conditional Majorization) which is orthogonal to all traceless Hermitian matrices or \( \rho \) by conditional majorization. Further discussion of it in [17]. We are now ready to define entropy functions don’t change under such embedding (see discussion). Such extension is motivated by the fact that entropy functions don’t change under tensor products and monotonic under mixing (i.e. majorization).

If a given function \( H \) satisfies the above conditions, then \( \alpha H \) is also a conditional entropy for any \( \alpha > 0 \). To eliminate this extra degree of freedom due to scaling, we consider only normalised conditional entropy functions such that when \( |A| = 2 \) and \( |B| = 1 \) (i.e. the bipartite state is a single qubit), then
\[
H(A|B)_\frac{1}{2} = H\left( \frac{1}{2} \right) = 1. \quad (22)
\]
where \( H \) is the Shannon entropy. The above three axioms (i.e. including the normalization condition) are sufficient to guarantee several key properties of conditional entropy. We start by showing that conditional entropies reduce to entropies on uncorrelated states.

**Lemma 4.** Conditional entropy reduces to an entropy function on uncorrelated states. That is, for any product state \( \omega_A \otimes \tau_B \),
\[
H(A|B)_{\omega_A \otimes \tau_B} = H(A)_{\omega}. \quad (23)
\]
Proof. Suppose that \( \rho_{AB} = w_A \otimes \tau_B \) is a product state. Let \( \mathcal{E}_B \in \text{CPTP} (B \rightarrow B) \) be the completely randomizing channel on system \( B \). Then it follows that \( \mathcal{I}_{A \rightarrow A} \otimes \mathcal{E}_{B \rightarrow B} \in \text{CUSC} (AB \rightarrow AB) \) and that
\[
\mathcal{I}_{A \rightarrow A} \otimes \mathcal{E}_{B \rightarrow B} (w_A \otimes \tau_B) = w_A \otimes u_B. \quad (24)
\]
Moreover, let \( \mathcal{R}_B \) be the replacement channel on system \( B \), which takes any quantum state to \( \tau_B \). Then
\[
\mathcal{I}_{A \rightarrow A} \otimes \mathcal{R}_{B \rightarrow B} (w_A \otimes u_B) = w_A \otimes \tau_B. \quad (25)
\]
Thus,
\[
w_A \otimes \tau_B \gtrsim_A w_A \otimes u_B \quad \text{and} \quad w_A \otimes u_B \gtrsim_A w_A \otimes \tau_B \quad (26)
\]
which implies, by the monotonicity property of the conditional entropy, that
\[
H(A|B)_{w_A \otimes \tau_B} = H(A|B)_{w_A \otimes u_B}. \quad (27)
\]
Thus, \( H(A|B)_{w_A \otimes \tau_B} \) depends on \( w_A \) only. Moreover, the function \( w_A \mapsto H(A|B)_{w_A \otimes \tau_B} \) satisfies the three axioms of entropy [7] and therefore can be considered as an entropy of \( w_A \).

The same statement also holds when system \( |B| \) is trivial, i.e., when \( |B| = 1 \). Unlike entropy, quantum conditional entropy can be negative. In the following theorem, we find a lower bound for the conditional entropy and show that only entangled states can have negative conditional entropy.
Theorem 1. Let $H$ be any conditional entropy, and let $A$ and $B$ be two Hilbert spaces. Then for any $\rho \in \mathcal{D}(AB)$ the conditional entropy is lower bounded as,

$$\tag{28} - \log \min (|A|, |B|) \leq H(A|B)_{\rho},$$

The lower bound is achieved when $\rho_{AB} = \phi_{AB}^+$ is the maximally entangled state. Additionally, if $\rho_{AB}$ is a separable (non-entangled) state, then $H(A|B)_{\rho} \geq 0$.

Proof. Given that $\phi_{AB}^+$ can be used to teleport an arbitrary quantum state, it follows that there exists some channel $\mathcal{N} \in \text{CUSC}(AB \rightarrow AB)$ such that $\rho_{AB} = \mathcal{N}(\phi_{AB}^+)$. It immediately follows that for any $\rho \in \mathcal{D}(AB)$, then

$$H(A|B)_{\phi_{AB}^+} \leq H(A|B)_{\rho_{AB}} \tag{29}$$

Thus, it is sufficient to demonstrate that $H(A|B)_{\phi_{AB}^+}$ is lower bounded by $-\log (\min (|A|, |B|))$. To this aim, we define $d \triangleq |A| = |B|$ and consider the state $\phi_{AB}^+ \otimes u_{A_2}$ where $|A_2| = d$. Then

$$H(\mathcal{A}_2|B)_{\phi_{AB}^+ \otimes u_{A_2}} = H(\mathcal{A}_2|B)_{\phi_{AB}^+ \otimes u_{A_2} \otimes u_{B_2}} \tag{30}$$

$$= H(A|B)_{\phi_{AB}^+} + H(A_2|B_2)_{u_{A_2} \otimes u_{B_2}} \tag{31}$$

$$= H(A|B)_{\phi_{AB}^+} + H(u_{A_2}) \tag{32}$$

$$= H(A|B)_{\phi_{AB}^+} + \log (d), \quad \tag{33}$$

where the first line follows by introducing a trivial system $B_2$ such that $|B_2| = 1$ and the second follows from conditional entropy.

We now demonstrate that $H(\mathcal{A}_2|B)_{\phi_{AB}^+ \otimes u_{A_2}} \leq 0$ by demonstrating the existence of a channel $\mathcal{N} \in \text{CUSC}(\mathcal{A}_2B \rightarrow \mathcal{A}_2B)$ such that $\mathcal{N}(\phi_{AB}^+ \otimes u) = |\psi\rangle\langle\psi|_{\mathcal{A}_2}$ for some pure state $|\psi\rangle \in H(\mathcal{A}_2)$. Consider the Choi matrix

$$J_{\mathcal{A}_2B \rightarrow \mathcal{A}_2B} = \sum_{j=1}^{d^2} \phi_{AB}^{(j)} \otimes I_{A_2} \otimes \psi_{B_2}^{(j)} \tag{34}$$

where $\{|\phi_{AB}^{(j)}\rangle\}_{j=1}^{d^2}$ is an orthonormal basis of $AB$ consisting of maximally entangled states and where we set $\phi_1 = \phi_{AB}^+$. Likewise, $\{|\psi_{B_2}^{(j)}\rangle\}_{j=1}^{d^2}$ is an orthonormal basis of $A_2$. Evidently, $\mathcal{N}$ is trace-preserving and conditional unital (as $J_{\mathcal{A}_2B \rightarrow \mathcal{A}_2B} = I_{\mathcal{A}_2B} \otimes u_{\mathcal{A}_2B_2}$). Additionally, $\mathcal{N}$ is trivially semi-causal as subsystem $B$ is traced out. From this, it follows that $\mathcal{N} \in \text{CUSC}(\mathcal{A}_2B \rightarrow \mathcal{A}_2B)$ and therefore

$$H(\mathcal{A}_2|B)_{\phi_{AB}^+ \otimes u_{A_2}} \leq H(\mathcal{A}_2B) \mathcal{N}(\phi_{AB}^+ \otimes u_{A_2}) \tag{35}$$

$$= H(\mathcal{N}(\phi_{AB}^+ \otimes u_{A_2})) \tag{36}$$

where $\mathcal{B}$ is a trivial subsystem. Note that

$$\mathcal{N}(\phi_{AB}^+ \otimes u_{A_2}) \tag{37}$$

$$= \text{Tr}_{\mathcal{A}_2B} \left[ J_{\mathcal{A}_2B \rightarrow \mathcal{A}_2B}^N (\phi_{AB}^+ \otimes u_{A_2} \otimes I_{\mathcal{A}_2B_2}) \right] \tag{38}$$

$$= \text{Tr}_{\mathcal{A}_2B} \left[ \left( \sum_{j=1}^{d^2} \phi_{AB}^{(j)} \otimes I_{A_2} \otimes \psi_{A_2B_2}^{(j)} \right) \phi_{AB}^+ \right. \tag{39}$$

$$= \text{Tr}_{\mathcal{A}_2B} \left[ \phi_{AB}^+ \phi_{AB}^+ \right. \tag{40}$$

$$= \phi_{A_2B_2}^{(1)} \tag{41}$$

which is evidently a pure state. Hence,

$$H(\mathcal{A}_2B)_{\phi_{AB}^+ \otimes u_{A_2}} \leq H(|\psi\rangle\langle\psi|_{\mathcal{A}_2}) \tag{42}$$

$$= 0. \tag{43}$$

Upon combining everything, we see that

$$0 \leq H(\mathcal{A}_2B)_{\phi_{AB}^+ \otimes u_{A_2}} \tag{44}$$

$$\leq H(\mathcal{A}_2B)_{\phi_{AB}^+} + \log |A| \tag{45}$$

$$\leq 0, \tag{46}$$

from which it follows that $H(\mathcal{A}_2B)_{\phi_{AB}^+} = -\log |A|$ and therefore $-\log (\min(|A|, |B|))$ lower bounds any density matrix $\rho_{AB}$. Finally, we demonstrate that if $\rho \in \mathcal{D}(AB)$ is a separable state, then $H(\mathcal{A}_2B)_{\rho} \geq 0$. This follows from noting that any separable state $\rho_{AB}$ may be written as

$$\rho_{AB} = \sum_{j=1}^{d^2} p_j \psi_A^{(j)} \otimes \phi_B^{(j)} \tag{47}$$

where $\{\psi_A^{(j)}\}_j$ and $\{\phi_B^{(j)}\}_j$ are sets of pure states. It follows that the CUSC channel

$$N_{AB \rightarrow AB} \triangleq \sum_{j=1}^{d^2} p_j U_A^{(j)} \otimes V_B^{(j)} \tag{48}$$

satisfies

$$H(0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B) = \rho_{AB} \tag{49}$$

for a choice of unitaries $\{U_A^{(j)}\}$ such that $U_A^{(j)} |0\rangle = |\phi_A^{(j)}\rangle$ and a choice of unitaries $\{V_B^{(j)}\}$ such that $V_B^{(j)} |0\rangle = |\psi_B^{(j)}\rangle$. Then

$$0 = H(\mathcal{A}_2B)_{|0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B} \tag{50}$$

$$\leq H(\mathcal{A}_2B)_{|0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B} \tag{51}$$

$$\leq H(\mathcal{A}_2B)_{\rho_{AB}}. \tag{52}$$

This concludes the proof. \qed
Although conditional entropies are positive on all separable states, this does not mean that the conditional entropies are only positive on separable states. In fact, some entangled states have positive conditional entropy on all choice of conditional entropy functions. The following lemma provides a simple criterion for finding states with positive conditional entropy.

**Lemma 5.** Let $\mathbf{H}$ be a conditional entropy, and let $\rho \in \mathcal{D}(A | B)$ with $d := |A| = |B|$ be a density matrix whose eigenvalues are less than or equal to $\frac{1}{d}$, and whose marginal state on system $B$ is a maximally mixed state, i.e., $\rho_B = u_B$. Then, the conditional entropy on $\rho$ is non-negative, i.e.,

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

**Proof.** Suppose $N_{\tilde{A} \rightarrow AB}$, is a $\tilde{A} \rightarrow B$ semi-causal channel as shown in Fig. 2 and suppose $|B| = 1$. For a semi-causal channel $N_{\tilde{A} \rightarrow AB'}$ such that

$$N_{\tilde{A} \rightarrow AB'} = \mathcal{E}_{\tilde{R} \rightarrow \tilde{A}} \circ \mathcal{F}_{B \rightarrow RB'}, (55)$$

$\mathcal{F}$ is required to be an isometry. When $|B| = 1$, $\mathcal{F}$ is always an isometry. In this case, the channel $\mathcal{F}$ could be viewed as a bipartite state on system $RB'$. Let this state be the maximally entangled state $\phi_{RB'}$, and let $|R| = |B'|$. Then the resulting channel $N_{\tilde{A} \rightarrow AB'}$ takes the form

$$N_{\tilde{A} \rightarrow AB'}(\omega) = \mathcal{E}_{\tilde{A} \rightarrow AB'}(\omega \otimes \phi_{RB'}), \quad \forall \omega \in \mathcal{D}(\tilde{A}).$$

(56)

Note that the channel $N_{\tilde{A} \rightarrow AB'}$ is not necessarily conditional unitil. It can be shown that channel $N_{\tilde{A} \rightarrow AB'}$ is conditional unitil if the channel $\mathcal{E}_{\tilde{A} \rightarrow AB}$ satisfies

$$\mathcal{E}_{\tilde{A} \rightarrow AB'}(u_A \otimes \phi_{RB'}) = u_{AB'}.$$  

(57)

Let $\mathcal{E}_{\tilde{A} \rightarrow AB'}$ be a quantum channel that satisfies (57). Let $\sigma_{AB'}$ be the quantum state such that

$$\sigma_{AB'} = N_{\tilde{A} \rightarrow AB'}(|1\rangle|1\rangle_A),$$

(58)

where $|1\rangle$ is a pure state in $\tilde{A}$. We then have

$$H(A|B')_\sigma = H(A|B')_{N(|1\rangle|1\rangle)}$$

$$\geq H(|1\rangle|1\rangle_A)$$

(60)

$$= 0.$$  

(61)

To simplify the notation, we rename system $B'$ as $B$. Since $|R| = |B| = |\tilde{B}|$, we rename $R$ as $B$. Then, the channel $\mathcal{E}_{\tilde{R} \rightarrow \tilde{A}}$ becomes $\mathcal{E}_{\tilde{A} \rightarrow A}$, and its Choi matrix is denoted by $J^{\mathcal{E}_{\tilde{A} \rightarrow A}}_{\tilde{A} \rightarrow A}$. By (57), we have

$$J^{\mathcal{E}_{\tilde{A} \rightarrow A}}_{\tilde{A} \rightarrow A} = I_{AB}$$

and

$$J^{\mathcal{E}_{\tilde{A} \rightarrow A}}_{ BA} = I_{BA}.$$  

(62)

Now, let $\rho_{AB}$ be as in the lemma, and define the channel $\mathcal{E}_{\tilde{A} \rightarrow A}$ via its Choi matrix

$$J^{\mathcal{E}_{\tilde{A} \rightarrow A}}_{AB} := |1\rangle|1\rangle_A \otimes (d \rho_{AB}) + \sum_{x=2}^d |x\rangle \langle x|_A \otimes \frac{I_{AB} - d \rho_{AB}}{d-1}.$$  

(63)

Note that the matrix above is positive semidefinite since $I_{AB} \geq d \rho_{AB}$ and it satisfies the two conditions in (62) since $\rho_B = u_B$. With this choice of $\mathcal{E}$ we get

$$\sigma_{AB} = \mathcal{E}_{\tilde{A} \rightarrow A}(|1\rangle|1\rangle_A \otimes \phi_{RB'})$$

$$= Tr_{\tilde{A}B} \left( (J^{\mathcal{E}_{\tilde{A} \rightarrow A}}_{AB} \otimes I_B) \left( |1\rangle \langle 1|_A \otimes (\phi_{BB})^{T_B} \otimes I_A \right) \right)$$

(65)

$$= Tr_{\tilde{B}} \left( (\rho_{AB} \otimes I_B) (\phi_{BB})^{T_B} \otimes I_A \right)$$

(66)

$$= \rho_{AB}.$$  

(67)

where $T_B$ is the partial transposition map on system $\tilde{B}$. This completes the proof that the conditional entropy of $\sigma_{AB}$ is non-negative. 

**IV. EXAMPLES OF CONDITIONAL ENTROPIES AND DUALITY RELATIONS**

We now demonstrate that any quantum relative entropy can be used to generate a corresponding quantum conditional entropy.

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

$$\mathbf{H}^+(A|B)_{\rho} = \log |A| - D(\rho_{AB})\|u_A \otimes \rho_B)$$

is a quantum conditional entropy.

**Proof.** First, we demonstrate that $\mathbf{H}$ satisfies the monotonicity property of conditional entropy. Let $\mathcal{N} \in \text{CUSC}(AB \rightarrow A'B')$ and consider the bipartite density matrix $\rho_{AB}$. We begin by considering the case where $A = A'$ and

$$\mathbf{H}^+(A|B)_{\mathcal{N}(\rho)} = \log |A| - D(\mathcal{N}(\rho_{AB}))\|u_A \otimes Tr_A[\mathcal{N}(\rho_{AB})])$$

(68)

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

$$u_A \otimes Tr_A[\mathcal{N}(\rho_{AB})] = u_A \otimes Tr_A[\mathcal{N}(u_A \otimes \rho_B)]$$

(69)

$$= \mathcal{N}(u_A \otimes \rho_B)$$

(70)

from which we have

$$\mathbf{H}^+(A|B)_{\mathcal{N}(\rho)} = \log |A| - D(\mathcal{N}(\rho_{AB})\|\mathcal{N}(u_A \otimes \rho_B))$$

(71)

$$\geq \log |A| - D(\rho_{AB})\|u_A \otimes \rho_B)$$

(72)

where the last line follows from the data processing inequality.

We now prove that any entropy $\mathbf{H}$ satisfying the form given in the theorem statement is invariant under isometries. Consider any isometry channel $\mathcal{V} \in \text{CPTP}(A \rightarrow$
A') and let $C$ represent a Hilbert space where $|C| = |A'| - |A|$. Then
\[
H^I (A| B)_{\mathcal{V}(\rho)} = \log |A'| - D(\mathcal{V}_{A \rightarrow A'}(\rho_{AB})\|u_A \otimes \rho_B) \tag{73}
\]
Clearly, if $\mathcal{V}$ is unitary such that $A \cong A'$, then $H(A'|B)_{\mathcal{V}(\rho)} = H(A|B)_{\rho}$. Then w.l.o.g. we assume that
\[
\mathcal{V}_{A \rightarrow A'}(\rho_{AB}) = \rho_{AB} \oplus 0_{CB} \tag{74}
\]
\[= \begin{pmatrix} \rho_{AB} & 0 \\ 0 & 0_{CB} \end{pmatrix} \tag{75}
\]
since the conditional entropy is invariant under unitaries. We can additionally rewrite $u_{A'} = tu_A \oplus (1-t)u_C$ where $t \triangleq \frac{|A|}{|A'|}$. From this, it follows that
\[
D(\mathcal{V}_{A \rightarrow A'}(\rho_{AB})\|u_{A'} \otimes \rho_B) = D\left(\rho_{AB}\|u_A \otimes \rho_B\right) + ((1-t)u_C \otimes \rho_B) \tag{76}
\]
\[= D\left(\rho_{AB}\|u_A \otimes \rho_B\right) + \log(t) \tag{77}
\]
\[= D\left(\rho_{AB}\|u_A \otimes \rho_B\right) - \log(t) \tag{79}
\]
such that upon substitution, we have
\[
H^I (A| B)_{\mathcal{V}(\rho)} = \log |A'| - D\left(\rho_{AB}\|u_A \otimes \rho_B\right) + \log(t) \tag{78}
\]
\[= H (A|B)_{\rho} \tag{80}
\]
The additive property of $H$ follows immediately from the additive property of $D$. Finally, when $u = \frac{1}{2}$, then
\[
H (A|B)_{u} = \log(2) - D(u\|u) = 1 \tag{82}
\]
and so normalisation is also satisfied. \(\square\)

It follows that the von-Neumann conditional entropy is a conditional entropy under our axiomatic approach. The von-Neumann conditional entropy is defined as
\[
H (A|B)_{\rho} \triangleq \log |A| - D(\rho_{AB}\|u_A \otimes \rho_B) \tag{83}
\]
where $D$ is the Umegaki relative entropy such that for any $\rho_{AB}$ and $\sigma_{AB}$, then
\[
D(\rho_{AB}\|\sigma_{AB}) = \text{Tr} [\rho_{AB}(\log(\rho_{AB}) - \log(\sigma_{AB}))] \tag{84}
\]
Finally, we define the dual of a conditional entropy function and demonstrate that the conditional von-Neumann entropy is self-dual.

**Definition 5.** Let $H$ be a conditional entropy. For any density matrix $\rho \in \mathcal{D}(AB)$ with a purification $\varphi \in \mathcal{D}(ABC)$ where $\rho_{AB} = \text{Tr}_C(\varphi_{ABC})$ the dual of $H$ is defined as
\[
H^\text{dual} (A \mid B)_{\rho} \triangleq -H (A \mid C)_{\rho} \tag{85}
\]
Let $H$ be the conditional von-Neumann entropy. Then it follows from the relation $H (A|B)_{\varphi} + H (A|C)_{\varphi} = 0$ that
\[
H^\text{dual} (A|B)_{\rho} = -H (A|C)_{\varphi} \tag{86}
\]
\[= H (A|B) \tag{87}
\]
and therefore the conditional von-Neumann entropy is self-dual.

**V. Quantum Conditional Majorization Reduces to Classical Conditional Majorization**

In [11], a pre-order $\succ_C$ between two joint probability distributions (i.e. classical states) denotes classical conditional majorization. In this section, we demonstrate that the axiomatic approach to quantum conditional majorization defined in the previous section is a generalization of classical conditional majorization. To this aim, we demonstrate that when we restrict the underlying bipartite quantum states to classical states, the quantum conditional majorization is equivalent to the classical conditional majorization.

Suppose the system $A := X$ is classical, we call a channel $M \in \text{CPTP}(X \rightarrow Y)$ conditionally doubly stochastic (CDS in short) if it has the following form
\[
\mathcal{M}_{Y \rightarrow Y'} = \sum_j D^{(j)}_{X \rightarrow X} \otimes F^{(j)}_{B \rightarrow B'} \tag{88}
\]
where each $D^{(j)}$ is a classical doubly stochastic channel and each $F^{(j)}_{B \rightarrow B'}$ is a completely positive (CP) map such that $\sum_j F^{(j)}_{B \rightarrow B'}$ is CPTP.

**Lemma 6.** Suppose all the systems involved are classical with $A := X, B := Y$, and $B' := Y'$. Then we have
\[
\text{CDS}(XY \rightarrow XY') \subseteq \text{CUSC}(XY \rightarrow XY') \tag{89}
\]
**Proof.** Recall that if $M \in \text{CDS}(XY \rightarrow XY')$, there must exist a set of classical doubly stochastic channels $\{D^{(j)}\}$ and a quantum instrument $\{F^{(j)}_{Y \rightarrow Y'}\}$ s.t.
\[
\mathcal{M}_{AB \rightarrow AB'} = \sum_j D^{(j)}_{A \rightarrow A} \otimes F^{(j)}_{B \rightarrow B'} \tag{90}
\]
Then for any input of the form $u_A \otimes \sigma_B$, it follows that
\[
\mathcal{M}_{AB \rightarrow AB'}(u_A \otimes \sigma_B) = \sum_j D^{(j)}_{A \rightarrow A}(u_A) \otimes F^{(j)}_{B \rightarrow B'}(\sigma_B) \tag{91}
\]
\[= u_A \otimes \sum_j F^{(j)}_{B \rightarrow B'}(\sigma_B) \tag{92}
\]
and therefore $\mathcal{M}$ is conditionally unital. (Note that the third line of the above follows from the fact that $D^{(j)}$ is doubly stochastic, so $D^{(j)}(u_A) = u_A$ for all $j$.)
Finally, we show that \( M \) is semi-causal as follows:

\[
\text{Tr}_A [M \circ \mathcal{T}_{A \rightarrow A}] = \text{Tr}_A \left[ \sum_j D_{A \rightarrow A}^{(j)} \circ \mathcal{T}_{A \rightarrow A} \otimes F_{B' \rightarrow B'}^{(j)} \right]
\]

\[
= \sum_j \text{Tr}_A \left[ D_{A \rightarrow A}^{(j)} \circ \mathcal{T}_{A \rightarrow A} \otimes F_{B' \rightarrow B'}^{(j)} \right] \tag{93}
\]

\[
= \sum_j \text{Tr}_A \left[ D_{A \rightarrow A}^{(j)} \otimes F_{B' \rightarrow B'}^{(j)} \right] \tag{94}
\]

\[
= \text{Tr}_A [M_{AB \rightarrow AB'}] \tag{96}
\]

Given that \( M \) is conditionally unital and semi-causal, then the theorem statement follows. \( \square \)

**Theorem 3.** Let \( \rho_{XY} \) and \( \sigma_{XY'} \), be classical bipartite states such that \( A = X, B = Y, \) and \( B' = Y' \) are all classical systems. Then \( \sigma_{XY'} \preceq \rho_{XY} \) if and only if \( \sigma_{XY'} \preceq_c \rho_{XY} \) where \( \preceq_c \) is the classical conditional majorization defined in [11].

**Proof.** It then immediately follows from lemma 6 that \( \rho_{XY} \succeq \sigma_{XY'} \), as in definition 4.

To show the other direction, we suppose that \( \rho_{XY} \succeq \sigma_{XY'} \), as in definition 4. That is, there exists \( N \in \text{CUSC}(XY \rightarrow XY') \) such that

\[
\sigma_{XY'} = N_{XY \rightarrow XY'}(\rho_{XY}) . \tag{97}
\]

Since \( N \) is semi-causal, it has the form

\[
N_{XY \rightarrow XY'} = \sum_{j=1}^k \mathcal{E}_{X \rightarrow X}^{(j)} \otimes \mathcal{F}_{Y \rightarrow Y'}^{(j)} , \tag{98}
\]

where for each \( j \in [K] \), \( \mathcal{E}_{X \rightarrow X}^{(j)} \in \text{CPTP}(X \rightarrow X) \) and \( \mathcal{F}_{Y \rightarrow Y'}^{(j)} \in \text{CP}(Y \rightarrow Y') \) with \( \sum_j \mathcal{F}_j \in \text{CPTP}(Y \rightarrow Y') \). Hence,

\[
\sigma_{XY'} = \sum_{j=1}^k \mathcal{E}_{X \rightarrow X}^{(j)} \otimes \mathcal{F}_{Y \rightarrow Y'}^{(j)}(\rho_{XY}) . \tag{99}
\]

Since \( \rho, \sigma \) are classical, we have the identification \( \rho_{XY} \cong P \) and \( \sigma_{XY'} \cong Q \). Thus, the above equation can be expressed as

\[
Q = \sum_{j=1}^k E^{(j)}_Q R^{(j)}_P , \tag{100}
\]

where for each \( j \in [k] \), \( E^{(j)}_Q \) is the transition matrix of \( \mathcal{E}_{X \rightarrow X}^{(j)} \), and where \( R^{(j)}_P \) is the transpose of transition matrix of \( \mathcal{F}_{Y \rightarrow Y'}^{(j)} \). Hence, each \( E^{(j)}_Q \) is column stochastic and \( R := \sum_{j=1}^k R^{(j)}_P \) is row stochastic. Let \( \{ p_y \}_{y \in [n]} \) be the columns of \( P \), and observe that in components form, the above relation can be expressed as

\[
q_{xw} = \sum_{j=1}^k \sum_{y=1}^n (E^{(j)}_Q)^{x,y} (R^{(j)}_P)^{w,y} . \tag{101}
\]

So far we only used the constraint that \( N \) is semi-causal. The other constraint that \( N \) is conditionally unital implies that if \( p_{xy} = \frac{1}{m} q_{yw} \) then \( q_{yw} = \frac{1}{m} q_{uw} \). Therefore, for simplicity we take \( p_{xy} = \frac{1}{m} \delta_{y,y_0} \) for some fixed \( y_0 \in [n] \).

Then the above equation becomes

\[
\frac{1}{m} q_{uw} = \sum_j (E^{(j)} u)_{x,y_0} . \tag{102}
\]

Summing over \( x \) on both sides of the equation, and using the fact that each \( E^{(j)} \) is column stochastic implying that \( E^{(j)} u \) is a probability vector, we have

\[
q_w = \sum_j r^{(j)}_{yw} . \tag{103}
\]

Note that we assumed that \( y_0 \) was fixed, and therefore as a result the equation above does not imply that \( q_w \) is independent of \( y_0 \) if we allow \( y_0 \) to vary. In other words, each value of \( y_0 \) corresponds to a different \( P \) matrix and therefore possibly resulting in a different \( Q \) matrix. For simplicity, we rename \( y_0 \) as \( y \) and denote

\[
q_w = t_{yw} := \sum_j r^{(j)}_{yw} . \tag{104}
\]

By definition, the matrix \( T = (t_{yw}) \) is row-stochastic, and with the above notations, we can express (102) as

\[
u = \sum_{j} t^{(j)}_{yw} E^{(j)}_Q u \quad y \in [n], \ w \in [\ell] . \tag{105}\]

Therefore, the matrices

\[
D^{(y,w)} := \sum_j t^{(j)}_{yw} E^{(j)}_Q y \in [n], \ w \in [\ell] \tag{106}
\]

are all doubly stochastic. To see this, observe that for each \( y \) and \( w \), the set \( \{ r^{(j)}_{yw} \} \) forms a probability distribution over \( j \), and the fact that each \( E^{(j)}_Q \) is column stochastic implying that \( D^{(y,w)} \) is column stochastic. Moreover, the relation \( u = D^{(y,w)} u \) implies that \( D^{(y,w)} \) are row stochastic. Finally, for general probability matrices \( P = [p_1, \ldots, p_n] \) and \( Q = [q_1, \ldots, q_n'] \), we conclude that (101) is equivalent to

\[
q_w = \sum_{y=1}^n t_{yw} D^{(y,w)} P_y \quad \forall w \in [n'] . \tag{107}
\]

The above relation is precisely the condition given in lemma 3 of [11]. Therefore, we conclude that \( \rho_{XY} \preceq \sigma_{XY'} \) implies classical conditional majorization. \( \square \)

VI. CONNECTION TO GAMES OF CHANCE

A. Games of Chance for Classical Bipartite States

In [20], games of chance were utilized to find an ordering between classical bipartite states, and it has been
demonstrated that such games of chance are an alternate, operational approach for characterising classical majorisation. We aim to likewise construct a set of quantum games of chance which provide an operational meaning for majorisation of quantum bipartite states.

Firstly, we briefly review the games of chance introduced in [20] for classical bipartite states. We first note that any classical bipartite state may be represented as a probability distribution of the form $p = \{p_{yz}\}$ where the corresponding state in quantum notation is $\rho_{AB} = \sum_{y,z} p_{yz} |y\rangle\langle y| \otimes |z\rangle\langle z|$. Each game is fixed by a classical channel $T$ with transition probabilities $\{t_{w|z'}\}$, as well as the bipartite state $\rho$ which is being used for the game.

The player is only allowed to access subsystem $B$ (corresponding to the value of $z$), and the goal is for the player to correctly guess the value of $y$. The family of gambling games includes all possible gambling games that incorporate a correlated source. Given access to $z$, the player chooses $z' = f(z)$ and the host subsequently selects $w$ from the conditional distribution $T$. In general, the player will choose $f(z)$ based on their knowledge of $z$, as well as the fixed distributions $\{p_{yz}\}$ and $\{t_{w|z'}\}$. Finally, once $w$ is selected, the player communicates a set $S$ of $w$ guesses to the host and wins if and only if $y \in S$. The game is depicted in Fig. 3.

![FIG. 3. A classical gambling game with a correlated source.](image)

The player is provided with the value $z$. Based on this value, the player chooses $z$ (or the function $f$) and sends it to the host. The host then chooses the $w$ game based on a (possibly incomplete) distribution matrix $T = (t_{w|z})$. The player will then form a set $S$ containing $w$ guesses based on $w$ and $z$, and will win the game if $y \in S$.

It has previously been demonstrated [20] that the conditional majorisation relation $\prec_C$ is equivalent to the partial ordering induced by classical games of chance, namely $\sigma_{AB} \preceq \rho_{AB'}$ if and only if $\text{Prob}_T(\sigma) \leq \text{Prob}_T(\rho)$ for all conditional distributions $T$ where $\text{Prob}_T(\rho)$ is the expected success probability for playing the classical gambling game characterised by $T$ with state $\rho$.

**B. Extension of gambling games to quantum bipartite states**

In this section, we build on the framework for classical games of chance to develop the most general games of chance for quantum bipartite states (and respectively quantum channels). The only restriction on such games is that the ordering induced by these families of games of chance must reasonably correspond to a measure of uncertainty. For example, states which are equivalent up to a unitary must have equivalent performance in these games of chance.

First, we provide a straightforward extension of the classical gambling games, and demonstrate that this direct extension is insufficient to capture necessary quantum effects such quantum entanglement being a resource. Next, we provide a broader set of games which still reduces to the same ordering on classical states, but which leads to negative entropy for maximally entangled states.

To generalize the classical gambling game, we first transform Bob’s observation of the classical output $z$ into a POVM implemented by Bob on system $B$. Evidently, $y$ can be written in quantum notation as $|y\rangle\langle y|$ such that a measurement in the computational basis allows perfect observation of $y$.

Likewise, Bob’s choice of $w$ guesses is equivalent to sending an ordered list of all potential outcomes and winning if $y$ is in the first $w$ elements of the list. In turn, this can be transformed into performing an ordered measurement on system $A$ with the output being transmitted to the host. For example, if their first guess is 2, the second guess is 1, and the third guess is 3, then they would implement the ordered measurement $\{0|1|2|3\}$ on Alice’s system. This game is depicted in Fig. 4. However, this direct extension is insufficient to distinguish between the un-entangled product state $\sigma_{AB} = |0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B$ and the entangled product state $\rho_{AB} = \phi_{AB}$. More specifically, any game can be won with certainty when played with $|0\rangle\langle 0|_A \otimes |0\rangle\langle 0|_B$. Regardless of the player’s choice of $\Pi$ or $f$, and regardless of $w$, the player can simply implement the measurement $\Pi^{*}_{z,w} = \{0\}$ on the first system. The first outcome (corresponding to measurement element $|0\rangle\langle 0|$) will be obtained with certainty, and so $y = 1$ will be less than or equal to any $w \in \{1, 2, 3, 4\}$ leading to a win for the player.

To break this equivalency, we now introduce a more
sophisticated game which allows an adversarial player to “scramble” the state of system $A$ with some probability $p$. We define a conditional quantum gambling game $G = (T, p)$ by the classical channel $T$ with transition matrix is $T = (t_{w|z'})_{w,z'}$, and the probability $p$ of adversarial involvement. The specific steps of the game are as follows:

1. Alice and Bob are given a classical description of the state they share as the entries of the corresponding density matrix $\rho_{AB}$.

2. The host generates $x \in \{0, 1\}$ from the probability distribution $p = (p, 1 - p)$, where $p \in [0,1]$. If $x = 0$, then the adversary is allowed to interfere by implementing a measurement $\tilde{\Pi}^*$ on system $A$.

3. Bob is told which measurement the adversary implemented (if any), but is not told the result. He is then allowed to perform any measurement $\tilde{\Pi}' = \{\tilde{\Pi}^{'(z)}\}$ on system $B$ and is told the measurement result $z$.

4. Bob can choose any $z' = f(z)$ to input to the classical channel $T$ with corresponding transition matrix $T = (t_{w|z'})_{w,z'}$. The host draws $w$ from $\{t_{w|z'}\}$.

5. Given $z$ and $w$, Alice implements a final rank-one projective measurement $\tilde{\Pi}(z, w) = \{|\Pi^{(k)}(z, w)\rangle\rangle_k$ on system $A$. The rank-one requirement follows from noting that if Alice were allowed to choose a trivial measurement such as $\{0, 0, \ldots, 0\}$, she would win any game with certainty.

6. Alice and Bob win if the outcome $y$ is less than $w$.

![Diagram](image.png)

FIG. 5. Alternative extension of classical gambling games for bipartite states. With probability $p$, $x = 0$ and the adversarial player is allowed to implement an adversarial measurement $\tilde{\Pi}^*$ on Alice’s system at the beginning of the game. The player selects any rank-one measurement $\tilde{\Pi}$ to implement on system $B$ and observes the outcome $z$. Given $z$, the player selects input $z' = f(z)$ to $T$, and the hosts draws $w$ according to $\{t_{w|z'}\}$. The player then implements any rank one measurement $\tilde{\Pi}(z, w)$ given knowledge of $z$ and $w$, and wins if the output $y \leq w$.

In Appendix A, we demonstrate that the ordering induced by this extended quantum gambling game on classical bipartite states reduces to classical conditional majorisation. In the following, we will describe the reward function for the above games of chance and demonstrate that the induced partial ordering between quantum bipartite states distinguishes between entangled and unentangled pure states. Additionally, in the case where system $B$ is trivial, the ordering reduces to standard majorisation.

First, we denote by $R_{T, p}$ the total expected reward for playing the game corresponding to $T$ and probability distribution $p = \{p, 1 - p\}$. Evidently, since $X$ is drawn before the players are allowed to make any moves, then

$$R_{T, p} = p \times R_{T, 0} + (1 - p) \times R_{T, 1}$$

In other words, the total expected reward is a convex combination of the total expected reward when the adversary is allowed a measurement and the total expected reward when the adversary is not allowed a measurement.

Suppose that Alice and Bob have a fixed strategy—namely, that Bob implements measurement $\tilde{\Pi}'$, selects $z' = f(z)$, and Alice implements the optimal measurement $\hat{\Pi}(z, w)$. We first consider the case where $p = 0$, i.e. there is no adversary. Then no action is taken prior to Bob’s first measurement $\tilde{\Pi}'$ on subsystem $B$, and the probability of observing a measurement result corresponding to $\hat{\Pi}'(z)$ is

$$P(z|\tilde{\Pi}') = \text{Tr} \left[ \rho_{AB} \hat{\Pi}'(z) \right]$$

The corresponding post-measurement state is

$$\rho_z(\hat{\Pi}', \rho_{AB}) = \frac{\hat{\Pi}'(z) \rho_{AB} \hat{\Pi}'(z)}{P(z|\tilde{\Pi}')}$$

In any game, Alice will always choose $\hat{\Pi}_A(z, w)$ to maximize the reward. Namely, she will measure in the eigenbasis of $\rho_z(\hat{\Pi}', \rho_{AB})$ with the first measurement element corresponding to the eigenvector of $\rho_z(\hat{\Pi}', \rho_{AB})$ with the largest eigenvalue and so on. Then the success probability for this fixed strategy is

$$\sum_z P(z|\tilde{\Pi}') \sum_w t_{w|f(z)} \left\| \hat{\Pi}_A^{(k)}(z, w) \text{Tr}_B [\rho_z(\hat{\Pi}', \rho_{AB})] \right\|$$

$$= \sum_z P(z|\tilde{\Pi}') \sum_w t_{w|f(z)} \left\| \text{Tr}_B \left[ \hat{\Pi}_B^{(z)} \rho_{AB} \hat{\Pi}_B^{(z)} \right] \right\|$$

$$= \sum_z \sum_w t_{w|f(z)} \left\| \text{Tr}_B \left[ \hat{\Pi}_B^{(z)} \rho_{AB} \hat{\Pi}_B^{(z)} \right] \right\|$$

Naturally, Bob will wish to choose the remaining game components (that is, $\hat{\Pi}_B$ and $f$) to maximize their reward, s.t. the final reward function for the game $G =$...
\begin{align}
\{ T, p = 0 \} & \text{ may then be expressed as} \\
R_G(\rho_{AB}) & \triangleq \max_{\Pi^* A} \left( \sum_{j, w} t_{w|j(z)} \left\| \text{Tr}_B \left[ \Pi^*_{B} \rho_{AB} \Pi^*_{B} \right] \right\|_{(w)} \right) \\
\tag{115}
\end{align}

Now suppose the adversarial player is allowed to interfere with probability \( p \). If the adversary chooses measurement \( \hat{\Pi}^* A = \{ \Pi^* A(j) \}_j \), then the starting state \( \rho_{AB} \) will be transformed to \( \sum_j \Pi^* A(j) \rho_{AB} \Pi^* A(j) \). Thus, following an adversarial measurement, Alice and Bob will effectively be playing the quantum gambling game with the transformed state \( \sum_j \Pi^* A(j) \rho_{AB} \Pi^* A(j) \) and will have an expected reward of \( R_{T,0} \left( \sum_j \Pi^* A(j) \rho_{AB} \Pi^* A(j) \right) \).

The adversary will wish to choose the “worst case” \( \hat{\Pi}^* A \) to minimize Alice and Bob’s reward, leading to the following reward function for a general quantum bipartite gambling game

\begin{align}
R_{T,p}(\rho_{AB}) & \triangleq p \times \min_{\Pi^* A} \left( R_{T,0} \left( \sum_j \Pi^* A(j) \rho_{AB} \Pi^* A(j) \right) \right) \\
& \quad + (1-p)R_{T,0}(\rho_{AB}) \\
\tag{116}
\end{align}

Finally, we are ready to introduce conditional majorisation for quantum bipartite states based on games of chance.

**Definition 6.** The state \( \rho_{AB} \) majorizes \( \sigma_{AB'} \) with respect to subsystem \( A \), based on quantum conditional gambling games, denoted as \( \sigma_{AB'} \prec_{(A)} \rho_{AB} \), if

\[ R_{T,p}(\sigma_{AB'}) \leq R_{T,p}(\rho_{AB}, U) \]

\tag{118}

for all \( 0 \leq p \leq 1 \), \( T = \{ w | z' \} \), \( w' \in \{ 1, \ldots, |A| \} \), and \( z' \in \{ 1, \ldots, \max(|B|, |B'|) \} \).

It follows immediately from the definition of the reward function that it is sufficient to restrict to the set of games for which \( p \in \{ 0, 1 \} \), namely, games in which the adversary either always interferes or never interferes. In the remainder of this work we restrict to this set unless otherwise specified.

One key property of this partial ordering is that maximally entangled pure states outperform tensor product pure states.

**Lemma 7.** Consider the case in which \( d = |A| = |A'| = |B| \) and \( d \geq 2 \). Then any tensor product state \( \sigma_{AB'} = \sigma_A \otimes \sigma_{B'} \) is majorized by \( \phi_{AB}^{(+)} \), namely,

\[ \sigma_{AB'} \prec_{(A)} \phi_{AB}^{(+)} \]

\tag{119}

**Proof.** First, we demonstrate that \( R_{T,p}(\phi_{AB}^{(+)}|A) = 1 \) for all \( T \) and all \( p \). To do this, it is sufficient to show that

\[ R_{T,0} \left( \sum_j \Pi^* A j j A \rho_{AB} \Pi^* A j j A \right) = 1 \]

\tag{120}

for all \( T \) and all potential adversarial projective measurements \( \Pi^* A \). Since the adversarial measurement is projective, there exists some basis \( \{ |j \rangle \}_{j=1}^n \) and some partition of that basis \( \{ S_k \} \) s.t.

\[ \Pi^* A (k) = \sum_{j \in S_k} |j \rangle \langle j| \]

\tag{121}

Note that

\[ \sum_j \Pi^* A j j A \phi_{AB}^{(+)} \Pi^* A j j A = \frac{1}{d} \sum_j \Pi^* A j j A \phi_{AB}^{(+)} \Pi^* A j j A = \frac{1}{d} \sum_j |j \rangle \langle j| \otimes |j \rangle \langle j| \]

\tag{122}

Suppose that Bob’s strategy is then to implement the measurement \( \hat{\Pi} B = \{ |j \rangle \langle z| \}_{j} \) on system \( B \). Then it follows that

\[ \rho z \left( \hat{\Pi} B, \sum_j \Pi^* A j j A \phi_{AB}^{(+)} \Pi^* A j j A \right) = \sum_{z_1, z_2 \in S_j} |j \rangle \langle j| \otimes |j \rangle \langle j| \]

\tag{123}

Finally, we have

\[ R_{T,p} \left( \sum_j \Pi^* A j j A \phi_{AB}^{(+)} \Pi^* A j j A \right) \leq R_{T,0} \left( \sum_j \Pi^* A j j A \phi_{AB}^{(+)} \Pi^* A j j A \right) \]

\tag{130}

To prove the theorem statement, we now need to demonstrate that there exists some game \( G \) such that \( R_G(\sigma_{AB'}) < 1 \). Consider the game for which \( p = 1 \) and set \( T = \{ \delta_{w=1} \} \) such that \( w = 1 \) with certainty. If \( \Pi^* A \)
is a rank-one measurement in the basis which is mutually unbiased with respect to the eigenbasis of \( \sigma_A \), then 
\[ \sum_j \Pi_A^{+(j)} \sigma_{AB} \Pi_A^{+(j)} = u_A \otimes \sigma_B \]
and we have

\[ R_{T,1} \left( \sum_j \Pi_A^{+(j)} \sigma_{AB} \Pi_A^{+(j)} \right) \geq R_{T,0} \left( \sum_j \Pi_A^{+(j)} \sigma_{AB} \Pi_A^{+(j)} \right) = \max_{\Pi_B} \sum_z P(z) \| \rho_{z} \left( \Pi_B^t, u_A \otimes \sigma_B \right) \|_1 \]

where in the above, \( \| \cdot \|_1 \) represents standard majorisation.

\section{Quantum Channel Gambling Games}

In previous work involving gambling games for classical channels \cite{20}, the output of the classical channel was observed directly by the host. In the quantum case, where outputs are no longer necessarily orthogonal, the observation process is now represented through the projective quantum measurement \{\Pi_x\}^{n}_{x=1}. One natural extension of the classical gambling game introduced in \cite{20} is to determine each game with a tripartite state \( \rho_{ABX} \).

The corresponding game setup is then as follows:

1. The player is given knowledge of state \( \rho_{ABX} \) and chooses a pre-processing channel \( E \) on subsystem \( A \).

2. Channel \( N \) is implemented on subsystem \( A \). From here on, we denote this as \( N(E(\rho_{ABX})) \) with the understanding that the channel action is only on system \( A \).

3. The player chooses a rank-one measurement on system \( AB \). The player’s optimal choice will always be to measure in the ordered eigenbasis of \( N(E) \).

4. The measurement outcome \( y \) is obtained and the player wins with a reward of 1 if \( y \leq x \) and loses with a reward of 0 else.

\[ R_{\Pi_B} \left( \sum_j \Pi_A^{+(j)} \sigma_{AB} \Pi_A^{+(j)} \right) \]

and by considering games where \( w \) is fixed, the “only if” direction of the theorem statement follows. It remains to show that \( L(\sigma_A) \preceq \tilde{\lambda}(\rho_A) \) is a sufficient condition when games of chance with \( p = 1 \) are included. In this case, the adversarial measurement \( \Pi_B^t \) will be in a basis which is mutually unbiased with respect to \( \sigma_A \) (or respectively \( \rho_A \)). Thus, for both games, the state after the adversarial measurement is \( u_A \). Then

\[ R_{T,1}(\rho_A) = R_{T,0}(u_A) = R_{T,1}(\sigma_A) \]

and so the ordering is unchanged by including games where \( p > 0 \). The theorem statement then follows.

Finally, we demonstrate that when system \( B \) is trivial, the ordering reduces to standard majorisation between quantum states.

\begin{lemma}
Let \( \tilde{\lambda}(\rho) \) denote the vector whose elements consist of the eigenvalues of \( \rho \). Then for any states \( \rho_A \) and \( \sigma_A \), \( \sigma_A \preceq (\rho_A) \) if and only if

\[ \tilde{\lambda}(\sigma_A) \preceq \tilde{\lambda}(\rho_A) \]

where in the above, \( \preceq \) represents standard majorisation.
\end{lemma}

\begin{proof}
Both \( \rho_A \) and \( \sigma_A \) can be viewed as bipartite states with a trivial \( B \) subsystem (i.e. \( |B| = 1 \)). Evidently, the expected reward for a given \( T \) and games where \( p = 0 \) is then

\[ R_{T,0}(\rho_A) = \max_{x'} \left( \sum_w \alpha_w \| u_A \|_{(w)} \right) \]

and by considering games where \( w \) is fixed, the “only if” direction of the theorem statement follows. It remains to show that \( \tilde{\lambda}(\sigma_A) \preceq \tilde{\lambda}(\rho_A) \) is a sufficient condition when games of chance with \( p = 1 \) are included. In this case, the adversarial measurement \( \Pi_B^t \) will be in a basis which is mutually unbiased with respect to \( \sigma_A \) (or respectively \( \rho_A \)). Thus, for both games, the state after the adversarial measurement is \( u_A \). Then

\[ R_{T,1}(\rho_A) = R_{T,0}(u_A) = R_{T,1}(\sigma_A) \]

and so the ordering is unchanged by including games where \( p > 0 \). The theorem statement then follows.
\end{proof}
Definition 7. We denote the case where channel $\mathcal{M}_{C\rightarrow D}$ is majorized by channel $\mathcal{N}_{C\rightarrow D'}$ according to quantum games of chance as $\mathcal{M} \preceq_q \mathcal{N}$. Equivalently,

$$\mathcal{M} \preceq_q \mathcal{N} \leftrightarrow R_{\rho_{ABX}} (\mathcal{M}) \leq R_{\rho_{ABX}} (\mathcal{N}) \quad \text{for all } \rho_{ABX}$$  

(144)

where it is sufficient to bound the dimensions of systems $A, B$ by $\max (|C|, |C'|)$. Likewise, the dimension of system $X$ can be bounded by $\max (|C|, |C'|) \times \min (|D|, |D'|)$.

The bound for the dimension of $X$ follows by noting that $\mathcal{N} \circ \mathcal{E} (\rho_{AB})$ has at most $|D'| \times |B|$ nonzero eigenvalues and likewise $\mathcal{M} \circ \mathcal{E} (\rho_{AB})$ has at most $|D| \times |B|$ nonzero eigenvalues. Since $|B| \leq \max (|C|, |C'|)$, the bound on $X$ immediately follows.

Finally, we define the entropy of a quantum channel to be a function which is monotonic under channel ordering, namely

Definition 8. A non-zero function,

$$H : \bigcup_{A,B} \text{CPTP}(A \rightarrow B) \rightarrow \mathbb{R}$$

where the union is over all finite quantum systems $A$ and $B$, is a quantum channel entropy if it satisfies the following two conditions:

1. It is monotonic under channel majorization; i.e. given quantum channels $\mathcal{N}$ and $\mathcal{M}$, then

$$H(\mathcal{M}) \geq H(\mathcal{N}) \quad \text{if } \mathcal{M} \preceq_q \mathcal{N}$$  

(145)

2. It is additive under tensor products; i.e.

$$H(\mathcal{N} \otimes \mathcal{M}) = H(\mathcal{N}) + H(\mathcal{M})$$  

(146)

for all quantum channels $\mathcal{N}$ and $\mathcal{M}$.

VIII. OPERATIONAL INTERPRETATION AND RESULTS FOR SPECIAL CLASSES OF CHANNELS

We begin by showing that when $\mathcal{M}$ and $\mathcal{N}$ correspond to states or classical channels, respectively, the quantum channel majorization defined by gambling games reduces to the standard vector majorization and the classical channel majorization introduced by our previous work [20] respectively.

Lemma 9. If $\mathcal{N}$ and $\mathcal{M}$ are replacement channels which output states $\rho_{\mathcal{N}}$ and $\rho_{\mathcal{M}}$ respectively, then

$$\mathcal{M} \preceq_q \mathcal{N} \quad \text{if and only if } \tilde{\lambda} (\rho_{\mathcal{M}}) \preceq \tilde{\lambda} (\rho_{\mathcal{M}})$$  

(147)

where $\tilde{\lambda} (\rho)$ denotes the ordered eigenvalues of $\rho$.

Proof. See Appendix B. 

Previous work introduced classical gambling games parameterized by a correlated source $T = (t_{uw})$ where $t_{uw}$ represents the joint probability $\Pr (W = w, Z = z)$. The expected reward for playing a $T$-game with classical channel $\mathcal{N}$ with transition probability matrix elements $\{p_{y|x}\}$ was likewise defined as

$$\Pr_T (\mathcal{N}) \triangleq \sum_{x} \max_{y} \sum_{w \in \{1, \ldots, p\}} p_{y|x} (\rho_{AB})$$

(148)

where $p^y$ indicates that $\{p_{y|x}\}_y$ is ordered such that $p_{1|x} \geq p_{2|x} \geq \ldots p_{n|x}$ for all $x$.

We now demonstrate that when $\mathcal{M}$ and $\mathcal{N}$ are both classical, the ordering induced by quantum gambling games is equivalent to the ordering induced by classical gambling games.

Theorem 4. If $\mathcal{N}$ and $\mathcal{M}$ are both classical channels, then

$$R_{\rho_{ABX}} (\mathcal{M}) \leq R_{\rho_{ABX}} (\mathcal{N}) \quad \forall \rho_{ABX}$$

(149)

if and only if

$$\Pr_T (\mathcal{M}) \leq \Pr_T (\mathcal{N}) \quad \forall \; T = (t_{uw})$$

(150)

Proof. See Appendix C.

We now provide an operational interpretation for quantum channel majorization.

Theorem 5. If $\mathcal{M} = \sum_{z=1}^{s} p_z \mathcal{V}_z \circ \mathcal{N} \circ \mathcal{E}_z$ for some set of isometries $\mathcal{V}_z$, some set of preprocessing quantum processes $\mathcal{E}_z$, and some probability distribution $\{p_1, \ldots, p_s\}$, then:

$$\mathcal{M} \preceq_q \mathcal{N}$$

(151)

Proof. We can demonstrate the theorem statement by substituting $\mathcal{M} = \sum_{z=1}^{s} p_z \mathcal{V}_z \circ \mathcal{N} \circ \mathcal{E}_z$ into the expression for the reward function as follows:

$$R_{\rho_{ABX}} (\mathcal{M}) = \max_{\mathcal{E}_z} \left( \sum_{x} p_x \| \mathcal{M} \circ \mathcal{E}_z (\rho_{AB}) \|_x \right)$$

(152)

$$= \max_{\mathcal{E}_z} \left( \sum_{x} p_x \| \sum_{z=1}^{s} p_z \mathcal{V}_z \circ \mathcal{N} \circ \mathcal{E}_z (\mathcal{E}_z (\rho_{AB})) \|_x \right)$$

(153)

$$\leq \max_{\mathcal{E}_z} \left( \sum_{x} p_x \| \mathcal{V}_z \circ \mathcal{N} \|_x \mathcal{E}_z (\rho_{AB}) \|_x \right)$$

(154)

$$= \max_{\mathcal{E}_z} \left( \sum_{x} p_x \| \mathcal{N} \circ \mathcal{E}_z (\rho_{AB}) \|_x \right)$$

(155)

$$= R_{\rho_{ABX}} (\mathcal{N})$$

(156)

This concludes the proof.

It follows immediately from the above Theorem that two channels which are equivalent up to a unitary will have equal performance for any game of chance. We now show that for a fixed input and output system, unitaries are the “least noisy” channel.
Lemma 10. Let \( \mathcal{N}_{A \rightarrow A} \) be a quantum channel, and let \( \mathcal{U}_{A \rightarrow A} \) be any unitary channel. Then

\[
R_{\rho_{ABX}}(U) \geq R_{\rho_{ABX}}(N) \quad \forall \rho_{ABX} \tag{157}
\]

Additionally, \( \mathcal{U}_{C \rightarrow C} \prec \mathcal{U}_{C' \rightarrow C} \) if \( |C| \leq |C'| \).

Proof. We first show that unitary channels outperform any other channel of the same dimension.

\[
R_{\rho_{ABX}}(N) = \max_{\hat{E}} \left( \sum_x p_x \left\| N \circ \hat{E} \left( \rho^{(x)}_{AB} \right) \right\|_x \right) \tag{158}
\]

\[
\leq \max_{\hat{E}} \left( \sum_x p_x \left\| \hat{E} \left( \rho^{(x)}_{AB} \right) \right\|_x \right) \tag{159}
\]

\[
= \max_{\hat{E}} \left( \sum_x p_x \left\| U \hat{E} \left( \rho^{(x)}_{AB} \right) U^\dagger \right\|_x \right) \tag{160}
\]

\[
= R_{\rho_{ABX}}(U). \tag{161}
\]

The statement that higher dimensional unitary channels outperform lower dimensional unitary channels follow by setting \( \rho_{AB} = |\Phi^+_AB\rangle \langle \Phi^+_AB| \) where \( |A| = |B| = |C'| \) (namely, the dimensions of systems \( A \) and \( B \) are set to be equal to the input dimension of the higher dimensional unitary).

Thus, \( \mathcal{U}_{C' \rightarrow C'}(\rho_{AB}) \) will result in a pure state and perfect reward for any game. On the other hand, any game played with \( \mathcal{U}_{C \rightarrow C} \) will require some choice of preprocessing channel \( \mathcal{E}_{C' \rightarrow C} \), which is necessarily entanglement breaking.

\[\square\]

\section{IX. Uniqueness of the Quantum Channel Gambling Game}

In principle, one could consider more general games with additional resources. In Appendix C, we demonstrate that generalizations of the quantum gambling game do not yield any reasonable interpretation related to entropy.

First, we show that allowing the player to pass side information through a classical wire from \( \mathcal{E} \) to the final measurement would enable the player to win any game with the classical identity channel. Given that quantum entanglement is a resource, we would expect the quantum identity channel to strictly outperform the classical identity channel, and so the gambling game with an added classical communication wire does not lead to a reasonable ordering between channels.

One could additionally consider a game with quantum combs. In Appendix D, we discuss why allowing arbitrary quantum combs \( C_j \) would again prevent a reasonable channel ordering. Specifically, if \( \{C_j\} \) are allowed to be arbitrary bipartite channels, it is not possible to obtain an ordering between channels \( \mathcal{N} \) and \( \mathcal{U} \circ \mathcal{N} \) where \( \mathcal{U} \) is any unitary and \( \mathcal{N} \) is a pure state replacement channel. Given that we expect two channels which are related by a unitary to have equivalent entropy, this game must again be excluded.

\section{X. Noisy Channels}

We include a table summarizing the reward function of special qubit channels for two key games involving the maximally entangled state and a pure tensor product state respectively. More specifically, in the first game we consider, we define \( \rho_{AB} = \phi^+_AB \) and let \( X \) be drawn according to the classical state \( \rho_X \) s.t. \( \rho_{ABX} = \phi^+_AB \otimes \rho_X \). In the second game we consider, system \( B \) is set to be trivial such that \( \rho_{AB} = |0\rangle\langle 0| \) and \( \rho_{ABX} = |0\rangle\langle 0| \otimes \rho_X \). In both cases, we denote the ordered diagonal elements of \( \rho_X \) as \( \{p_j\} \).

\begin{tabular}{|c|c|c|c|}
\hline
Channel & Reward & Reward & Reward \\
\hline
\( \mathcal{N} \) & \( R_{\rho^+_AB \otimes \rho_X}(N) \) & \( R_{|0\rangle\langle 0|} \otimes \rho_X(N) \) & \\
\hline
\( \mathcal{U} \) & 1 & 1 & \\
\hline
\( \mathcal{I}_{\text{CL}} \) & \( 1 - \frac{1}{2}p_1 \) & 1 & \\
\hline
\( \mathcal{D}_\gamma \) & \( (1 - \gamma) + \frac{1}{4} \sum_x p_x \| |\sigma^{(x)}\|_2^{(x)} \) & \( 1 - \gamma p_1^{(x)} \) & \\
\hline
\( \mathcal{N}_{\hat{\Pi}} \) & \( 1 - \frac{1}{2}p_1 \) & 1 & \\
\hline
\( \mathcal{A}_\gamma \) & \( (1 - \gamma)p_1 + \frac{1}{2} \) & 1 & \\
\hline
\( \mathcal{R}_\tau \) & \( \sum_x p_x \| |\sigma^{(x)}\|_2^{(x)} \) & \( \sum_x p_x \| |\sigma^{(x)}\|_2^{(x)} \) & \\
\hline
\( \mathcal{F}_\gamma \) & \( (1 - \gamma) + \gamma (1 - \frac{1}{2}p_1) \) & 1 & \\
\hline
\end{tabular}

We now define all channels included in the above table.

1. \( \mathcal{U} \) is a unitary channel with corresponding unitary \( U \) such that \( \mathcal{U}(\rho) = U \rho U^\dagger \) for all \( \rho \in \mathcal{D}(\mathcal{H}_2) \).

2. \( \mathcal{I}_{\text{CL}} \) is a classical identity channel such that for some basis \( \{|j\rangle\} \) then

\[
\mathcal{I}_{\text{CL}}(\rho) = \sum_j \langle j | \rho | j \rangle | j \rangle \langle j | \tag{162}
\]
3. $\mathcal{D}_\gamma$ is the qubit depolarizing channel with noise parameter $\gamma$ defined as

$$\mathcal{D}_\gamma(\rho) = (1 - \gamma)\rho + \gamma u_2 \quad (163)$$

4. $\mathcal{N}_\hat{\Pi}$ is the channel implementing POVM $\hat{\Pi} = \{\Pi_j\}$ as

$$\mathcal{N}_\hat{\Pi}(\rho) = \sum_j \text{Tr}[\Pi_j \rho] |j\rangle \langle j| \quad (164)$$

5. $\mathcal{A}_\gamma$ is the amplitude damping channel with noise parameter $\gamma$ defined as

$$\mathcal{A}_\gamma \left( \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \right) \triangleq \begin{pmatrix} \rho_{00} + \gamma \rho_{11} \sqrt{1 - \gamma \rho_{01}} \\ \sqrt{1 - \gamma \rho_{10}} (1 - \gamma) \rho_{11} \end{pmatrix} \quad (165)$$

6. The replacement channel $\mathcal{R}_\sigma$ discards the input and outputs state $\sigma$ such that $\mathcal{R}_\sigma(\rho) = \sigma$ for all $\rho$.

7. The dephasing channel $\mathcal{F}_\gamma$ with parameter $\gamma$ acts as

$$\mathcal{F}_\gamma(\rho) = (1 - \gamma)\rho + \gamma I_{\text{CL}}(\rho) \quad (166)$$

Details for how the results in the table are obtained are included in Appendix E.

Finally, we note that the maximal output state purity attainable with a channel is monotonic under the ordering induced by games of chance.

**Lemma 11.** Suppose that $\mathcal{M} \preceq \mathcal{N}$. Then

$$\sup_{\rho} \left( \text{Tr} \left[ \mathcal{N}(\rho)^2 \right] \right) \geq \sup_{\rho} \left( \text{Tr} \left[ \mathcal{M}(\rho)^2 \right] \right) \quad (167)$$

**Proof.** See Appendix F

**XI. CONCLUSIONS**

In this work, we demonstrate that the set of all conditional entropy non-decreasing channels is equivalent to the set of conditionally unital, semi-causal (CUSC) channels. We then require quantum conditional entropy to be monotonic under the action of CUSC channels as well as satisfying simple axioms such as additivity and normalisation. This minimalist approach is sufficient to guarantee negativity of conditional entropy for maximally entangled states, and additionally is sufficient demonstrate that conditional entropy is non-negative for separable bipartite states. We discuss applications of this axiomatic approach to classical bipartite states and demonstrate that the resulting conditional majorization matches the classical conditional majorization first defined in [11]. Finally, we develop an operational approach for characterising entropy via games of chance, and apply this approach to both quantum bipartite states and quantum channels respectively.

*Note Added:* We recently became aware of independent work of Vempati et al. [25]

**XII. ACKNOWLEDGMENTS**

The authors would like to thank Henry Pfister for helpful discussions. 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 No. 1908730 and 1910571. 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.

[1] N. J. Cerf and C. Adami. Negative entropy and information in quantum mechanics. *Physical Review Letters*, 79(26):5194–5197, December 1997.
[2] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Partial quantum information. *Nature*, 436(7051):673–676, August 2005.
[3] Benjamin Schumacher and Michael A. Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629–2635, October 1996. arXiv:quant-ph/9604022.
[4] Yunchao Liu, Qi Zhao, and Xiao Yuan. Quantum coherence via conditional entropy. *Journal of Physics A: Mathematical and Theoretical*, 51(41):414018, September 2018.
[5] Subhasree Patro, Indranil Chakrabarty, and Nirman Ganguly. Non-negativity of conditional von neumann entropy and global unitary operations. *Physical Review A*, 96(6), December 2017.
[6] Angela Capel, Angelo Lucia, and David Perez-Garcia. Quantum conditional relative entropy and quasifactorization of the relative entropy. *Journal of Physics A: Mathematical and Theoretical*, 51(48):484001, November 2018.
[7] G. Gour. Comparison of quantum channels by superchannels. *IEEE Transactions on Information Theory*, 65(9):5880–5904. 2019.
[8] A. Kuznetsova. Quantum conditional entropy for infinite-dimensional systems. 2010. arXiv:1004.4519.
Appendix A: Quantum Gambling Games on Bipartite Classical States

In this appendix, we demonstrate that the majorisation relation \( \mathcal{Z}_A \) is equivalent to the majorisation relation \( \mathcal{Z}_C \) on classical states. We here denote the expected success probability of playing a \( T \)-game with classical bipartite state \( \rho_{AB} \) as

\[
\text{Prob}_T(\rho_{AB}) = \sum_x p_x \max_{z'} \left( \sum_w t_w |z'| \| p_x \|_w \right),
\]

where \( p_x = \{p_{y|x}\} \) is the probability that the output of \( \rho_{AB} \) on system \( B \) is \( y = y \), given that the output on system \( A \) is \( x = x \).

**Lemma 12.** For any classical bipartite states \( \rho_{AB} = \{q_{x,y}\}_{x,y} \) and \( \sigma_{AB} = \{p_{x,y}\}_{x,y} \), the following inequality holds

\[
\text{Prob}_T(\sigma_{AB}) \leq \text{Prob}_T(\rho_{AB}) \quad \forall T
\]

if and only if

\[
R_{T,p}(\sigma_{AB}) \leq R_{T,p}(\rho_{AB})
\]

for all distributions \( T \) and all \( 0 \leq p \leq 1 \).

**Proof.** First, we show that for all quantum gambling games with distribution \( T \) and \( p = 0 \), the following equality holds

\[
R_{T,0}(\sigma_{AB}) = \text{Prob}_T(\sigma_{AB}).
\]

Recall that any classical bipartite state \( \rho_{AB} \) can be written in quantum notation as

\[
\rho_{AB} = \sum_{x,y} q_{x,y} |x\rangle\langle x| \otimes |y\rangle\langle y|,
\]

\[
\sigma_{AB} = \sum_{x,y} p_{x,y} |x\rangle\langle x| \otimes |y\rangle\langle y|.
\]

First consider games where \( p = 0 \). Then
Upon relabelling variables $x$ to $k$ for consistency with the
convention used for classical gambling games, we have

$$R_{\mathcal{T},0}(\rho_{AB}) = \max_{x} \sum_{z} \max_{z'} \sum_{w} t_{w|z'} \|q_{w} \sum_{k} |k\rangle \langle k|\|_{(w)}$$

(A12)

$$= \sum_{z} \|q_{z} \max_{z'} \sum_{w} t_{w|z'} \|q_{w} \sum_{k} |k\rangle \langle k|\|_{(w)}$$

(A13)

$$= \sum_{z} \|q_{z} \max_{z'} \sum_{w} t_{w|z'} \|q_{w} \sum_{k} |k\rangle \langle k|\|_{(w)}$$

(A14)

$$= \text{Prob}_{\mathcal{T}}(\rho_{AB}).$$

(A15)

It immediately follows that $\text{Prob}_{\mathcal{T}}(\sigma_{AB}) \leq \text{Prob}_{\mathcal{T}}(\rho_{AB})$
for all $\mathcal{T}$ if $R_{\mathcal{T},p}(\sigma_{AB}) \leq R_{\mathcal{T},p}(\rho_{AB})$ for all distributions
$\mathcal{T}$ and all $p \in [0, 1]$.

Now, we demonstrate that the additional games al-
lowed in the quantum case (i.e., games where $p > 1$)
does not affect the ordering. This follows immediately upon
noting that the adversary can measure in a mutually
unbiased basis, such that the state $\rho_{AB}$ transforms to
$|u_{A} \otimes \sigma_{B}|$ for some $\sigma_{B}$.

It follows that $R_{\mathcal{T},1}(\rho_{AB}) = R_{\mathcal{T},1}(\sigma_{AB})$ for all $\mathcal{T}$ and
therefore the additional games do affect the partial
ordering.

As an aside, we note that it is sufficient to consider a
restricted set of games for ordering between two classical
states.

**Lemma 13.** If $\rho_{AB}$ and $\sigma_{AB}$ are two classical states, it
is sufficient to restrict to games where $w$ is determined.
Namely,

$$\text{Prob}_{\mathcal{G}}(\sigma_{AB}) \leq \text{Prob}_{\mathcal{G}}(\rho_{AB})$$

(A16)

for all games $G$ if and only if

$$\text{Prob}_{(w,p)}(\sigma_{AB}) \leq \text{Prob}_{(w,p)}(\rho_{AB}),$$

(A17)

for all $G$ where $w$ is fixed (i.e., $\mathcal{T} = \{\delta_{w|z'}\}_{z'}$) and where
$w = \{1, 0\}$, so that an adversarial measurement never
occurs.

**Proof.** It follows immediately from the previous proof
that it is sufficient to restrict to games where $w = \{1, 0\}$,
as for an arbitrary $w = \{p, 1 - p\}$, the reward for any
bipartite state $\rho_{AB}$ can be written as

$$R_{(\mathcal{T},\{p,1-p\})}(\rho_{AB}) = pR_{(\mathcal{T},\{1,0\})}(\rho_{AB})$$

(A18)

We demonstrated in the previous case that for a clas-
sical state $\sigma_{AB} = \sum_{x,y} p_{x,y}|x\rangle \langle y| \otimes |y\rangle \langle y|$, the following
equality holds

$$R_{(\mathcal{T},\{1,0\})}(\sigma_{AB}) = \text{Prob}_{\mathcal{T}}(\sigma_{AB})$$

(A20)

Finally, we demonstrate that this can be rewritten as a
convex combination of games where $w$ is fixed:

$$\text{Prob}_{\mathcal{T}}(\sigma_{AB}) = \sum_{x} p_{x} \max_{z} \left( \sum_{w} t_{w|z'} \|p_{x}\|_{(w)} \right)$$

(A21)

$$= \sum_{x} p_{x} \left( \sum_{w} t_{w|z'} \|p_{x}\|_{(w)} \right)$$

(A22)

$$= \sum_{x} \sum_{w} t_{w|z'} p_{x} \|p_{x}\|_{(w)}$$

(A23)

$$= \sum_{w} c_{w} \sum_{x} p_{x} \|p_{x}\|_{(w)}$$

(A24)

where for each $x$, the variable $z (x)^{*}$ is defined to be the
optimal choice and for each $w$,

$$c_{w} = \frac{\sum_{x} t_{w|z'} p_{x} \|p_{x}\|_{(w)}}{\sum_{x} p_{x} \|p_{x}\|_{(w)}}$$

(A25)

is a positive constant. Finally, if $\rho_{AB}$ and $\sigma_{AB}$ are ordered
for all games with a fixed $w$, then

$$\text{Prob}_{\mathcal{T}}(\sigma_{AB})$$

(A26)

$$\leq \sum_{w} c_{w} \text{Prob}_{w}(\sigma_{AB})$$

(A27)

$$= \sum_{w} \sum_{x} \sum_{w} t_{w|z'} p_{x} \|p_{x}\|_{(w)} \frac{q_{x} \|q_{x}\|_{(w)}}{\sum_{x} p_{x} \|p_{x}\|_{(w)}}$$

(A28)

$$\leq \sum_{x} q_{x} \sum_{w} \left( \sum_{w} t_{w|z'} \|p_{x}\|_{(w)} \right)$$

(A29)

$$\leq \sum_{x} q_{x} \max_{x'} \left( \sum_{w} t_{w|z'} \|q_{x}\|_{(w)} \right)$$

(A30)

$$= \text{Prob}_{\mathcal{T}}(\rho_{AB})$$

(A31)

where $\tilde{p}_{x} = \frac{p_{x} \|p_{x}\|_{(w)}}{\sum_{x} p_{x} \|p_{x}\|_{(w)}}$.

□
Appendix B: Proof of reduction for quantum replacement channels

The “only if” part follows trivially. We now prove the “if” part of the statement. First, recall that in [20], it was demonstrated that for classical games of chance $\mathcal{M} \preceq \mathcal{N}$ if and only if $\rho_{\mathcal{M}} \preceq \rho_{\mathcal{N}}$. Additionally, classical games of chance can be represented as quantum games of chance where $\rho_{ABX} = |0\rangle|0\rangle_A \otimes \rho_X$. Thus, the “if” direction of the theorem statement may be rewritten as

$$R_{\rho_{ABX}}(\mathcal{M}) \leq R_{\rho_{ABX}}(\mathcal{N}) \quad \forall \rho_{ABX}$$ (B1)

if

$$R_{\rho_{AX}}(\mathcal{M}) \leq R_{\rho_{AX}}(\mathcal{N}) \quad \forall \rho_{ABX} = |0\rangle|0\rangle_A \otimes \rho_X$$ (B2)

Clearly, since $\mathcal{M}$ is a replacement channel, then $\mathcal{M} \circ \mathcal{E} = \mathcal{M}$ for all $\mathcal{E}$. Hence, the choice of preprocessing channel does not affect the reward, so w.l.o.g. we may set $\mathcal{E} = \mathcal{I}$. It follows that the reward then becomes

$$R_{\rho_{ABX}}(\mathcal{M}) = \max_{\mathcal{E}} \left( \sum_{x=1}^{\ell} p_x \| \mathcal{M} \circ \mathcal{E} (\rho_{AB}) \|_x \right)$$ (B3)

$$= \sum_{x=1}^{\ell} p_x \| \mathcal{M} (\rho_{AB}) \|_x$$ (B4)

$$= \sum_{x=1}^{\ell} p_x \| \rho_{\mathcal{M}} \otimes \text{Tr}_A[\rho_{AB}] \|_x$$ (B5)

We now denote $\lambda_j^i (\rho_{\mathcal{M}}) = [\lambda_1^M, \lambda_2^M, \ldots, \lambda_d^M]$ and likewise $\lambda_j^i \left( \text{Tr}_A[\rho_{AB}] \right) = [\lambda_1^B, \lambda_2^B, \ldots, \lambda_d^B]$. Then

$$\text{eig} \left( \rho_{\mathcal{M}} \otimes \text{Tr}_A[\rho_{AB}] \right) = \{ \lambda_j^{(M)} \lambda_k^{(B)} \}_{j,k}$$ (B6)

We define $f(j, k, x)$ to return the position of $\lambda_j^{(M)} \lambda_k^{(B)}$ in the full ordered set $\lambda_j^{(M)} \lambda_k^{(B)}$ (for example, $f(1, 1, x) = 1$ since $\lambda_1^M \lambda_1^B$ is the largest eigenvalue and $f(d, d, x) = d^2$ since $\lambda_d^M \lambda_d^B$ is the smallest eigenvalue). Finally, we can define the set

$$S(j) = \{ x, k \mid \text{ such that } f(j, k, x) \leq x \}$$ (B7)

In other words, $S(j)$ provides all values of $x, k$ such that the term $p_x \lambda_j^{(M)} \lambda_k^{(B)}$ appears in the sum.

We now rewrite the reward function as

$$R_{\rho_{ABX}}(\mathcal{M}) = \sum_{x=1}^{\ell} p_x \left( \sum_{j,k \text{ s.t. } f(j,k,x) \leq x} \lambda_j^{(M)} \lambda_k^{(B)} \right)$$ (B8)

$$= \sum_{j=1}^{d} \lambda_j^{(M)} \left( \sum_{x,j \in S(j)} p_j \lambda_k^{(B)} \right)$$ (B9)

In previous work, we proved that games of the form $\sum_j \lambda_j^{(M)} C_j$ for $C_j > 0$ can be rewritten in the form $\alpha R_{\tilde{p}}(\mathcal{M})$. In this case, evidently

$$C_j \triangleq \sum_{x,k \in S(j)} p_x \lambda_k^{(B)} > 0$$ (B10)

so there exists some $\alpha$ and $\tilde{p}$ such that $R_{\rho_{ABX}}(\mathcal{M}) = \alpha R_{\tilde{p}}(\mathcal{M})$.

Since we are assuming that $\mathcal{M} \preceq \mathcal{N}$, then $R_{\tilde{p}}(\mathcal{M}) \leq R_{\tilde{p}}(\mathcal{N})$. It follows that

$$R_{\rho_{ABX}}(\mathcal{M}) = \alpha R_{\tilde{p}}(\mathcal{M}) \leq \alpha R_{\tilde{p}}(\mathcal{N})$$ (B11)

$$= \max_z \left( \sum_{k=1}^{d} \lambda_k^B (\rho_{\mathcal{N}}) \left( \sum_{j=1}^{d} \lambda_j^M \sum_{\ell = f(k)} p_\ell \right) \right)$$ (B12)

$$\leq \max_z \left( \sum_{x} p_x \| \rho_{\mathcal{N}} \otimes \text{Tr}_A[\rho_{AB}] \|_z \right)$$ (B13)

$$= R_{\rho_{ABX}}(\mathcal{N})$$ (B14)

The statement then follows from recalling that classical gambling games are equivalent to quantum gambling games where $\rho_{ABX}$ has the form $|0\rangle|0\rangle_A \otimes \rho_X$.

Appendix C: Proof of reduction for Classical Channels

The “only if” part follows immediately from noting that the quantum gambling games are a strict generalisation of the classical gambling games. We now prove the “if” part of the statement. For $\mathcal{M}$ classical, $\mathcal{E}$ can be reduced to selecting the optimal classical output $z$, such that $\mathcal{E} (\rho) = |z\rangle\langle z|$ for every $\rho$. It follows that the reward then becomes

$$R_{\rho_{ABX}}(\mathcal{M}) = \max_{\mathcal{E}} \left( \sum_{x=1}^{\ell} p_x \| \mathcal{E} (\rho_{AB}) \|_x \right)$$ (C1)

$$= \max_z \left( \sum_{x=1}^{\ell} p_x \| \mathcal{M} (|x\rangle\langle x|) \otimes \text{Tr}_A[\rho_{AB}] \|_x \right)$$ (C2)

$$= \sum_{x=1}^{\ell} p_x \| \lambda_j^M (|z^*\rangle\langle z^*|) \otimes \text{Tr}_A[\rho_{AB}] \|_x$$ (C3)

where $z^*$ is the optimal value (i.e. the value of $z$ which maximises the reward). We now denote $\lambda_j^i (|z^\ast\rangle\langle z^\ast|) = [\lambda_1^M, \lambda_2^M, \ldots, \lambda_d^M]$ and likewise $\lambda_j^i \left( \text{Tr}_A[\rho_{AB}] \right) = [\lambda_1^B, \lambda_2^B, \ldots, \lambda_d^B]$. Then

$$\text{eig} \left( \mathcal{M} (|z^\ast\rangle\langle z^\ast|) \otimes \text{Tr}_A[\rho_{AB}] \right) = \{ \lambda_j^{(M)} \lambda_k^{(B)} \}_{j,k}$$ (C4)
We define \( f(j, k, x) \) to return the position of \( \lambda_j^{(M)}\lambda_{k,x}^B \) in the full ordered set \( \{\lambda_j^{(M)}\lambda_{k,x}^B\}_{j, k} \) (for example, \( f(1, 1, x) = 1 \) since \( \lambda_1^{(M)}\lambda_{1,x}^B \) is the largest eigenvalue and \( f(d, d, x) = d^2 \) since \( \lambda_d^{(M)}\lambda_{d,x}^B \) is the smallest eigenvalue). Finally, we can define the set

\[
S(j) = \{x, k \mid f(j, k, x) \leq x\} \tag{C5}
\]

In other words, \( S(j) \) provides all values of \( x, k \) such that the term \( p_x\lambda_j^{(M)}\lambda_{k,x}^B \) appears in the sum.

We now rewrite the reward function as

\[
R_{\rho_{ABX}}(\mathcal{M}) = \frac{1}{\sqrt{\det(\rho_{ABX})}} \left( \sum_{x, j \in S(j)} p_j \lambda_{k,x}^B \right) \tag{C6}
\]

However, this is of the form of a classical gambling game (possibly up to a scaling). In previous work, we proved that games of the form \( \sum_j \lambda_j^{(M)}C_j \) for \( C_j > 0 \) can be rewritten in the form \( \alpha R_{\tilde{\rho}}(M) \). In this case, evidently,

\[
C_j \triangleq \sum_{x, k \in S(j)} p_x \lambda_{k,x}^B > 0 \tag{C7}
\]

so there exists some \( \alpha \) and \( \tilde{\rho} \) such that \( R_{\rho_{ABX}}(\mathcal{M}) = \alpha R_{\tilde{\rho}}(M) \). Since we are assuming that \( \mathcal{M} \lesssim \mathcal{N} \), then \( R_{\tilde{\rho}}(\mathcal{M}) \leq R_{\tilde{\rho}}(\mathcal{N}) \). It follows that

Finally, we have

\[
R_{\rho_{ABX}}(\mathcal{M}) = \alpha R_{\tilde{\rho}}(\mathcal{M}) \leq \alpha R_{\tilde{\rho}}(\mathcal{N}) \tag{C8}
\]

\[
= \max_z \left( \sum_{k=1}^d \lambda_k^2 (N(|z|z)) \sum_{j=1}^d \lambda_j^B \sum_{x, k \in S(j)} p_x \right) \tag{C9}
\]

\[
\leq \max_z \left( \sum_{x} p_x \left\| N(|z|z) \otimes \text{Tr}_A[\rho_{ABX}^x] \right\| \right) \tag{C10}
\]

\[
= R_{\rho_{ABX}}(\mathcal{N}) \tag{C11}
\]

\[
= R_{\rho_{ABX}}(\mathcal{N}) \tag{C12}
\]

Appendix D: Games with Wires and Resources

1. Game with Classical Wire

We first consider the game with an additional classical wire, and demonstrate that one can win any game of the form \( \rho_{ABX} = |\psi_{AB}|\psi_{AB}\rangle \otimes \rho_X \) with a qubit classical identity channel. Define \( \mathcal{E} \) as

\[
\mathcal{E}(\rho) = \text{Tr}[|0\rangle\langle 0| \rho] |0\rangle\langle 0| + \text{Tr}[|1\rangle\langle 1| \rho] |1\rangle\langle 1| \tag{D1}
\]

Thus, \( \mathcal{E} \) is a POVM and if outcome 0 (respectively 1) is obtained the post measurement state is,

\[
\rho_{AB}^{(1)} = \frac{|1\rangle\langle 1|_A \otimes I_B (|\psi_{AB}\rangle\langle \psi_{AB}|) |1\rangle\langle 1|_A \otimes I_B}{\text{Tr}[|1\rangle\langle 1|_A \otimes I_B (|\psi_{AB}\rangle\langle \psi_{AB}|)]} \tag{D3}
\]

The most general expression for \( |\psi_{AB}\rangle \) is

\[
|\psi_{AB}\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle + \alpha_{10} |10\rangle + \alpha_{11} |11\rangle \tag{D4}
\]

\[
\rightarrow |0\rangle|0\rangle_\rho \otimes I_B |\psi_{AB}\rangle = \alpha_{00} |00\rangle + \alpha_{01} |01\rangle \tag{D5}
\]

\[
\rightarrow |1\rangle|1\rangle_\rho \otimes I_B |\psi_{AB}\rangle = \alpha_{10} |10\rangle + \alpha_{11} |11\rangle \tag{D6}
\]

from which it follows that

\[
\rho_{AB}^{(0)} = \frac{|0\rangle\langle 0|_A \otimes |\psi^0\rangle\langle \psi^0|_B}{\text{Tr}[|0\rangle\langle 0|_A \otimes |\psi^0\rangle\langle \psi^0|_B]} \tag{D7}
\]

\[
\rho_{AB}^{(1)} = \frac{|1\rangle\langle 1|_A \otimes |\psi^1\rangle\langle \psi^1|_B}{\text{Tr}[|1\rangle\langle 1|_A \otimes |\psi^1\rangle\langle \psi^1|_B]} \tag{D8}
\]

We now show that the following measurements will allow the player to always win the game with \( \rho_{AB}^{(0)} \) and \( \rho_{AB}^{(1)} \) respectively:

\[
\tilde{\Gamma}^{(0)} = \left\{ |0\rangle|0\rangle \otimes |\psi^0\rangle\langle \psi^0|, |0\rangle|0\rangle \otimes |\psi^0\rangle\langle \psi^0|, |0\rangle|0\rangle \otimes |\psi^0\rangle\langle \psi^0| \right\} \tag{D9}
\]

\[
\tilde{\Gamma}^{(1)} = \left\{ |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1|, |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1|, |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1| \right\} \tag{D10}
\]

\[
\tilde{\Gamma}^{(1)} = \left\{ |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1|, |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1|, |1\rangle|1\rangle \otimes |\psi^1\rangle\langle \psi^1| \right\} \tag{D11}
\]

\[
|0\rangle|0\rangle \otimes |\psi^1\rangle\langle \psi^1|, |0\rangle|0\rangle \otimes |\psi^1\rangle\langle \psi^1| \tag{D12}
\]

It is sufficient to show that the above measurements allow the player to always win the game where diag \( \rho_X = [1, 0, \ldots, 0] \) (namely, \( x \) is always equal to one.) The corresponding reward function for the given choice of \( \mathcal{E} \) then becomes:

\[
R_J(\mathcal{C}_L) = \sum_{x \in \{0, 1\}} \frac{\text{Tr} [|x\rangle\langle x|_A \rho_{AB}]}{\text{Tr} [|x\rangle\langle x|_A \otimes |\psi^x\rangle\langle \psi^x|]} \rho_{AB}^{(1)} \tag{D13}
\]

\[
= \sum_{x \in \{0, 1\}} \frac{\text{Tr} [|x\rangle\langle x|_A \rho_{AB}]}{\text{Tr} [|x\rangle\langle x|_A \otimes |\psi^x\rangle\langle \psi^x|]} \rho_{AB}^{(1)} \tag{D15}
\]

\[
= 1 \tag{D16}
\]

2. Game with Quantum Combs

We now consider the quantum comb-based game depicted in Figure (reference), and demonstrate that allowing any element \( C_j \) to be an arbitrary quantum channel would result in a channel ordering that cannot correspond to channel entropy.

First, we demonstrate that \( C_3 \) cannot be an arbitrary bipartite channel. To do so, consider the following channels:

\[
\mathcal{N}_1(\rho) = |0\rangle|0\rangle \forall \rho, \quad \mathcal{N}_2(\rho) = |+\rangle|+\rangle \forall \rho \tag{D17}
\]
We now consider the performance of both channels in the game where \( \rho_{ABC}^{(x)} = |000\rangle \langle 000 | \) for all \( x \), where combs \( C_1 \) and \( C_2 \) are taken to be trivial (identity) channels, and where

\[
C_3 (\rho_{AB}) = \text{Tr}[[0][0]_A \rho_{AB} |00\rangle \langle 00 | + \text{Tr}[1][1]_A \rho_{AB} |11\rangle \langle 11 |]
\]

(D18)

Additionally, suppose that \( X \) is such that \( x = 1 \) with certainty, or equivalently, the player wins only if they always obtain the first measurement outcomes. Evidently, in the above, \( N_1 \) and \( N_2 \) are equivalent pure state replacement channels (up to a unitary). However,

\[
R \left( N_1 \right) = \left\| C_3 \left( N_1 \left( |0\rangle \langle 0|_A \otimes |00\rangle \langle 00 |_{BC} \right) \right) \right\|_{(1)} = \left\| C_3 \left( |000\rangle \langle 000 |_{ABC} \right) \right\|_{(1)} = 1
\]

(D20)

whereas

\[
R \left( N_2 \right) = \left\| C_3 \left( N_2 \left( |0\rangle \langle 0|_A \otimes |00\rangle \langle 00 |_{BC} \right) \right) \right\|_{(1)} = \left\| C_3 \left( |1+00\rangle \langle +00 |_{ABC} \right) \right\|_{(1)} = \left\| \text{Tr}[\text{Tr}[|0\rangle \langle 0 |_{A} |+\rangle \langle + |_B] |00\rangle \langle 00 |_{AB} \otimes |0\rangle \langle 0 |_C \right\|_{(1)}
\]

(D21)

\[
+ \text{Tr}[|1\rangle \langle 1 |_{A} |+\rangle \langle + |_B] |11\rangle \langle 11 |_{AB} \otimes |0\rangle \langle 0 |_C \right\|_{(1)} = \left\| \text{Tr}\left[ \frac{1}{2} |000\rangle \langle 000 |_{ABC} + \frac{1}{2} |110\rangle \langle 110 |_{ABC} \right] \right\|_{(1)} = \left\| \text{Tr}\left[ \frac{1}{2} |000\rangle \langle 000 |_{ABC} + \frac{1}{2} |110\rangle \langle 110 |_{ABC} \right] \right\|_{(1)}
\]

(D26)

Thus, there is no ordering between channels that are equivalent up to a unitary.

We now similarly demonstrate that \( C_1 \) cannot be an arbitrary bipartite channel. We consider the following channels and games:

\[
N_1 (\rho) = \text{Tr} [[0][0]_A \rho_A |0\rangle \langle 0 | + \text{Tr}[1][1]_A \rho_A |1\rangle \langle 1 |]
\]

(D32)

\[
N_2 (\rho) = \text{Tr} [[+\rangle \langle + |_A \rho_A [+\rangle \langle + | + \text{Tr} [[-\rangle \langle - |_A \rho_A [-\rangle \langle - |]
\]

(D33)

\[
\rho_{ABC}^{(x)} = |000\rangle \langle 000 | \forall x
\]

(D35)

\[
C_1 (\rho_{AB}) = |00\rangle \langle 00 |_{AB} \quad \forall \rho_{AB}
\]

(D36)

and where \( C_2 \) and \( C_3 \) are taken again to be the identity. Additionally, suppose that \( X = 1 \) so the player wins only if they always obtain the first measurement outcomes. Evidently, in the above, \( N_1 \) and \( N_2 \) are both POVMs. However,

\[
R \left( N_1 \right) = \left\| N_1 \circ C_1 \left( |00\rangle \langle 00 |_{AB} \otimes |0\rangle \langle 0 |_C \right) \right\|_{(1)} = \left\| N_1 \left( |0\rangle \langle 0 |_A \otimes |00\rangle \langle 00 |_{BC} \right) \right\|_{(1)} = \left\| |000\rangle \langle 000 |_{ABC} \right\|_{(1)} = 1
\]

(D38)

whereas

\[
R \left( N_2 \right) = \left\| N_2 \circ C_1 \left( |00\rangle \langle 00 |_{AB} \otimes |0\rangle \langle 0 |_C \right) \right\|_{(1)} = \left\| N_2 \left( |0\rangle \langle 0 |_A \otimes |00\rangle \langle 00 |_{BC} \right) \right\|_{(1)} = \left\| I_A \otimes |00\rangle \langle 00 |_{BC} \right\|_{(1)} = \frac{1}{2}
\]

(D42)

Thus, there is no ordering between channels that are equivalent up to a unitary.

From the above, evidently \( C_1 \) and \( C_3 \) need to be restricted. If we suppose that \( C_1 \) and \( C_3 \) are restricted to be the identity (or a unitary), then the remaining non-trivial channel is \( C_2 \). However, \( C_2 \) can be absorbed into the initial state preparation, so

\[
\rho_{ABC}^{(x)} \rightarrow C_2 \left( \rho_{ABC}^{(x)} \right)
\]

(D46)

Appendix E: Examples

1. Unitary Channel

Consider any unitary channel \( U : A \rightarrow A \) defined as \( U (\rho) = U \rho U^\dagger \). Then if a game is described by \( \rho_{ABX} \) where \( \text{dim} (B) \leq \text{dim} (A) \), the reward is as follows:

\[
R_{\rho_{ABX}} (U) = \max \sum_x p_x \left\| U \mathcal{E}_A \left( \rho_{AB}^{(x)} \right) U^\dagger \right\|_x
\]

(E1)

\[
= \max \sum_x p_x \left\| \mathcal{E}_A \left( \rho_{AB}^{(x)} \right) \right\|_x
\]

(E2)

Consider the game where \( \rho_{AB}^{(x)} = \phi_{AB}^{(x)} \) for all \( x \) and where \( X \) is drawn according to distribution \( p \). Then

\[
R_{\rho_{ABX}} (U) = \max \sum_x p_x \left\| \mathcal{E}_A \left( \phi_{AB}^{(x)} \right) \right\|_x
\]

(E4)
Clearly, for the game where $\rho_{AB}^{(x)} = |0\rangle|0\rangle$ we likewise have

\begin{align*}
R_{\rho_{AB}^{(x)}}(\mathcal{U}) &= \max_{\xi} \left( \sum_{x} p_x \| E_A(|0\rangle|0\rangle) \|_x \right) \\
&\leq \sum_{x} p_x \| |0\rangle|0\rangle \|_x \\
&= \sum_{x} p_x \\
&= 1
\end{align*}

3. Amplitude Damping Channel

Recall that the amplitude damping channel is defined as

$$A_\gamma(\rho) \triangleq \left( \rho_{00} + \gamma \rho_{11} \frac{1}{\sqrt{1 - \gamma}} \rho_{01}, \frac{1}{\sqrt{1 - \gamma}} \rho_{10} (1 - \gamma) \rho_{11} \right)$$

Then for the game where $\rho_{AB}^{(x)} = |0\rangle|0\rangle$ for all $x$, we have

\begin{align*}
R_{\rho_{AB}^{(x)}}(A_\gamma) &= \max_{\xi} \left( \sum_{x} p_x \| A_\gamma \circ E(|0\rangle|0\rangle) \|_x \right) \\
&= \sum_{x} p_x \| A_\gamma \circ E(|0\rangle|0\rangle) \|_x \\
&= 1
\end{align*}

Likewise if $\rho_{AB}^{(x)} = \phi_{AB}^+$ for all $x$, we have

\begin{align*}
R_{\rho_{AB}^{(x)}}(A_\gamma) &= \max_{\xi} \left( \sum_{x} p_x \| A_\gamma \circ E(\phi_{AB}^+) \|_x \right) \\
&= \max_{\phi} \left( \sum_{x} p_x \sum_{y,z=0}^1 A_\gamma(\kappa |y\rangle\langle z|) \otimes |y\rangle\langle z| \right)_x \\
&= \max_{\phi} \left( \sum_{x} p_x \sum_{y,z=0}^1 A_\gamma(\kappa |y\rangle\langle z|) \otimes |y\rangle\langle z| \right)_x \\
&= \sum_{x} p_x \sum_{y,z=0}^1 A_\gamma(\kappa |y\rangle\langle z|) \otimes |y\rangle\langle z| \right)_x \\
&= p_1(1 - \gamma^2) + (1 - p_1) \\
&= 1 - p_1 \gamma^2
\end{align*}

4. Dephasing Channel:

The dephasing channel $\mathcal{F}_\gamma$ can be written as

$$\mathcal{F}_\gamma(\rho) = (1 - \gamma) \rho + \gamma I_{CL}(\rho)$$

where $I_{CL}$ is the classical identity channel (completely dephasing channel.) Evidently for $\rho_{AB}^{(x)} = |0\rangle|0\rangle$ for all $x$ we have

\begin{align*}
R_{\rho_{AB}^{(x)}}(\mathcal{F}_\gamma) &= \max_{\xi} \left( \sum_{x} p_x \| \mathcal{F}_\gamma \circ E(|0\rangle|0\rangle) \|_x \right) \\
&= \max_{\xi} \left( \sum_{x} p_x \| \mathcal{F}_\gamma \circ E(|0\rangle|0\rangle) \|_x \right)
\end{align*}
Immediately by setting $\gamma = 1$.

\[
R_{\rho_{AB}}(\Phi) = \max_{\xi} \left( \sum_{x} p_{x} \left\| \Phi_{x} \right\|^{2} \right)
\]

Finally, we consider the case where $\rho_{AB}^{(x)} = 0$ for all $x$.

\[
R_{\rho_{AB}}(\Phi) = \max_{\xi} \left( \sum_{x} p_{x} \left\| \Phi_{x} \right\|^{2} \right)
\]

Results for the classical identity channel follow immediately by setting $\gamma = 1$.

5. Projective measurements

Finally, we consider the case where $N_{\Phi}$ is a channel which implements the projective qubit measurement $\Pi = \{\Pi_0, \Pi_1\}$ s.t.

\[
N_{\Phi}(\rho) = \text{Tr} \left[ \Pi_0 \rho \right] \times \Pi_0 + \text{Tr} \left[ \Pi_1 \rho \right] \times \Pi_1
\]

where $\Pi_0$ and $\Pi_1$ are both rank-one measurement elements.

We now calculate the reward for the game where $\rho_{AB}^{(x)} = 0$ for all $x$

\[
R_{\rho} (N_{\Phi}) = \max_{\xi} \left( \sum_{x=1}^{2} p_{x} \left\| \left( \Pi_{i} \mathcal{E} (|0\rangle\langle 0|) \right) \Pi_{i} \right\|_{(x)} \right)
\]

Finally, we calculate the reward for the game where $\rho_{AB}^{(x)} = 0$ for all $x$

\[
R_{\rho} (N_{\Phi}) = \max_{\xi} \left( \sum_{x=1}^{2} p_{x} \left\| \left( \Pi_{i} \mathcal{E} (\phi_{AB}^{(x)}) \right) \Pi_{i} \right\|_{(x)} \right)
\]

\[
= \sum_{x=1}^{2} p_{x} \left\| \sum_{i=1}^{2} \text{Tr} \left[ \Pi_{i} \mathcal{E} (|0\rangle\langle 0|) \Pi_{i} \right] \right\|_{(x)}
\]

\[
= \sum_{x=1}^{2} p_{x} \left\| \sum_{i=1}^{2} \Pi_{i} \otimes \frac{I}{2} \right\|_{(x)}
\]

\[
= \sum_{x=1}^{2} p_{x} \left\| \Pi_{i} \otimes \frac{I}{2} \right\|_{(x)}
\]

where in the second line we note that the optimal choice of $\mathcal{E}$ is the replacement channel which always prepares output state $\Pi_0$ (or alternatively, which always prepares output state $\Pi_1$).

Appendix F: Monotonicity Proof

Suppose that $\rho^*$ is the state which satisfies

\[
\rho^* \triangleq \arg\max_{\rho} \left( \text{Tr} \left[ \mathcal{M} (\rho)^{2} \right] \right)
\]

Consider diagonal matrices $U, V$. Through Holder’s inequality, we have:

\[
\text{Tr} [UV] \leq \sqrt{\text{Tr} [U^{2}] \sqrt{\text{Tr} [V^{2}]}}
\]

It follows that if $\text{Tr} [UV] \geq \text{Tr} [U^{2}]$, then

\[
\text{Tr} [V^{2}] \geq \text{Tr} [UV] \geq \text{Tr} [U^{2}]
\]

Now consider the game where $\rho_{AB}^{(x)} = |0\rangle\langle 0|_A$ for all $x$ (namely, the system $B$ is trivial), and where $p_{x}$ satisfy

\[
\sum_{x=1}^{\ell} p_{x} = \alpha \lambda_{j}^{(x)} \left( \mathcal{M} (\rho^*) \right)
\]

where $\alpha$ is some positive normalisation constant. Then the reward for $\mathcal{M}$ is simply

\[
R (\mathcal{M}) = \max_{\xi} \sum_{x=1}^{\ell} p_{x} \sum_{j=1}^{x} \lambda_{j}^{(x)} \left( \mathcal{M} (|0\rangle\langle 0|) \right)
\]

\[
= \max_{\rho} \sum_{x=1}^{\ell} p_{x} \sum_{j=1}^{x} \lambda_{j}^{(x)} \left( \mathcal{M} (\rho) \right)
\]

\[
= \max_{\rho} \sum_{j=1}^{x} \sum_{x=1}^{\ell} p_{x} \lambda_{j}^{(x)} \left( \mathcal{M} (\rho) \right)
\]

\[
= \alpha \max_{\rho} \sum_{j} \lambda_{j}^{(x)} \lambda_{j}^{(x)} \left( \mathcal{M} (\rho^*) \right)
\]

\[
= \alpha \sum_{j} \lambda_{j}^{(x)} \left( \mathcal{M} (\rho^*) \right)
\]

\[
= \text{Tr} \left[ \mathcal{M} (\rho^*)^{2} \right]
\]
Evidently, since $\mathcal{M} \preceq \mathcal{N}$,

$$R(\mathcal{N}) = \alpha \max_{\rho} \sum_j \lambda_j^N(\mathcal{N}(\rho)) \lambda_j^M(\mathcal{M}(\rho^*)) \quad \text{(F11)}$$

$$= \alpha \sum_j \lambda_j^N(\mathcal{N}(\sigma)) \lambda_j^M(\mathcal{M}(\rho^*)) \quad \text{(F12)}$$

$$\geq \alpha \text{Tr} \left[ \mathcal{M}(\rho^*)^2 \right] \quad \text{(F13)}$$

where in the above we denote the state which attains the optimum for $\mathcal{N}$ as $\sigma$. Evidently, upon rewriting the above we have

$$\text{Tr} \left[ \text{diag}(\mathcal{N}(\sigma)) \text{diag}(\mathcal{M}(\rho^*)) \right] \geq \text{Tr} \left[ \mathcal{M}^*(\rho) \right] \quad \text{(F14)}$$

This implies

$$\text{Tr} \left[ \mathcal{N}(\sigma)^2 \right] \geq \text{Tr} \left[ \mathcal{M}(\rho^*)^2 \right] \quad \text{(F15)}$$

$$\triangleq \sup_{\rho} \left( \text{Tr} \left[ \mathcal{M}(\rho)^2 \right] \right) \quad \text{(F16)}$$

and the statement immediately follows.