Appendix to Uniform Additivity in Classical and Quantum Information

Andrew Cross,[1] Ke Li,[1,2] and Graeme Smith[1]

[1]*IBM TJ Watson Research Center, Yorktown Heights, NY 10598, USA*

[2]*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

(Dated: July 6, 2016)

I. NOTATION AND BACKGROUND

For any collection of systems $X_1 \dots X_n$, let $\mathcal{P}(X_1 \dots X_n)$ be the power set of this collection

$$\mathcal{P}(X_1 \dots X_n) = \{X_1^{u_1} \dots X_n^{u_n} | (u_1, \dots, u_n) \in \{0,1\}^n\}. \tag{1}$$

We study channels $U_{\mathcal{N}} : A \to BE$ and are interested in formulas $f_\alpha(U_{\mathcal{N}})$ that are maximizations of linear combinations of entropies involving auxiliary variables $V_1 \dots V_n$

$$f_\alpha(U_{\mathcal{N}}) = \max_{\phi_{V_1 \dots V_n A}} f_\alpha(U_{\mathcal{N}}, \phi_{V_1 \dots V_n A}), \tag{2}$$

where the linear entropic quantity $f_\alpha(U_{\mathcal{N}}, \phi_{V_1 \dots V_n A})$ is given by

$$f_\alpha(U_{\mathcal{N}}, \phi_{V_1 \dots V_n A}) = \sum_{s \in \mathcal{P}(V_1 \dots V_n BE)} \alpha_s H\left(s, \rho_{V_1 \dots V_n BE}\right) \tag{3}$$

where $\rho_{V_1 \dots V_n BE} = (I \otimes U_{\mathcal{N}}) \phi_{V_1 \dots V_n A} (I \otimes U_{\mathcal{N}}^\dagger)$ is the channel output state and $H\left(s, \rho_{V_1 \dots V_n BE}\right)$ is the entropy of the reduced state corresponding to systems s.

II. GENERAL CONSIDERATIONS

We are interested in understanding when

$$f_\alpha(U_{\mathcal{N}} \otimes U_{\mathcal{M}}) = f_\alpha(U_{\mathcal{N}}) + f_\alpha(U_{\mathcal{M}}). \tag{4}$$

In order to do this, we study mappings from a state $\phi_{V_1 \dots V_n A_1 A_2}$ that can be acted on by $U_{\mathcal{N}} \otimes U_{\mathcal{M}}$ to two states: $\tilde{\phi}_{\tilde{V}_1 \dots \tilde{V}_n A_1}$, which can be acted on by $U_{\mathcal{N}}$ and $\hat{\phi}_{\hat{V}_1 \dots \hat{V}_n A_2}$ which can

be acted on by $U_{\mathcal{M}}$. We call such a mapping, $\mathcal{D} : \phi_{V_1...V_nA_1A_2} \to (\tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2})$ a *decoupling*.

There are two important types of decouplings that we consider: standard decouplings, and consistent decouplings. Both types of decouplings construct $\tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}$ from relabling the systems of $I \otimes U_{\mathcal{M}}\phi_{V_1...V_nA_1A_2}I \otimes U_{\mathcal{M}}^\dagger$ and construct $\hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}$ by relabling the systems of $I \otimes U_{\mathcal{N}}\phi_{V_1...V_nA_1A_2}I \otimes U_{\mathcal{N}}^\dagger$. For a *standard decoupling*, we have $\tilde{V}_i = \tilde{M}_2V_i$ and $\hat{V}_i = \hat{M}_1V_i$ with $\tilde{M}_2 \in \mathcal{P}(B_2E_2)$ and $\hat{M}_1 \in \mathcal{P}(B_1E_1)$. Each V_i will have a different $\tilde{M}_2$ and $\hat{M}_1$, but we suppress this dependence where the meaning is obvious from context. For a *consistent decoupling*, we require less: $\tilde{V}_i \in \mathcal{P}(V_1...V_nB_2E_2)$ with $\tilde{V}_i \cap \tilde{V}_j = \varnothing$ and $\hat{V}_i \in \mathcal{P}(V_1...V_nB_1E_1)$ with $\hat{V}_i \cap \hat{V}_j = \varnothing$.

We say that $f_\alpha(U_{\mathcal{N}}, \phi_{V_1...V_nA})$ is *uniformly* subadditive with respect to decoupling $\mathcal{D}$ if for all $\mathcal{N}_1$, $\mathcal{N}_2$, and $\phi_{V_1...V_nA_1A_2}$ we have

$$f_\alpha(U_{\mathcal{N}_1} \otimes U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}) \leq f_\alpha(U_{\mathcal{N}_1}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}) + f_\alpha(U_{\mathcal{N}_2}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}). \tag{5}$$

The following quantity will be useful:

$$\begin{aligned}&\Delta(f_\alpha, U_{\mathcal{N}_1}, U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}) \\ &= f_\alpha(U_{\mathcal{N}_1}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}) + f_\alpha(U_{\mathcal{N}_2}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}) - f_\alpha(U_{\mathcal{N}_1} \otimes U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}).\end{aligned} \tag{6}$$

Defined in this way, Δ is linear in f_α, so if we have

$$\Delta\left(f_{\alpha_1}, U_{\mathcal{N}_1}, U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}\right) \geq 0$$
$$\Delta\left(f_{\alpha_2}, U_{\mathcal{N}_1}, U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}\right) \geq 0$$

then for $\lambda_1, \lambda_2 \geq 0$ we also have

$$\Delta\left(\lambda_1 f_{\alpha_1} + \lambda_2 f_{\alpha_2}, U_{\mathcal{N}_1}, U_{\mathcal{N}_2}, \phi_{V_1...V_nA_1A_2}, \tilde{\phi}_{\tilde{V}_1...\tilde{V}_nA_1}, \hat{\phi}_{\hat{V}_1...\hat{V}_nA_2}\right) \geq 0. \tag{7}$$

For the standard or consistent decouplings, the Δ function defined in Eq. (6) depends only on the decoupling $\mathcal{D}$, the entropy formula f_α and the state

$$\rho_{V_1...V_nB_1E_1B_2E_2} = (I \otimes U_{\mathcal{N}_1} \otimes U_{\mathcal{N}_2})\phi_{V_1...V_nA_1A_2}(I \otimes U_{\mathcal{N}_1}^\dagger \otimes U_{\mathcal{N}_2}^\dagger). \tag{8}$$

So we abbreviate it as $\Delta^{\mathcal{D}}(\alpha, \rho_{V_1...V_nB_1E_1B_2E_2})$. It is easy to see that any state $\rho_{V_1...V_nB_1E_1B_2E_2}$ can be written the form of Eq. (8), with appropriate $U_{\mathcal{N}_1}$, $U_{\mathcal{N}_2}$ and $\phi_{V_1...V_nA_1A_2}$. Thus $f_\alpha(U_{\mathcal{N}}, \phi_{V_1...V_nA})$ is uniformly subadditive with respect to the decoupling $\mathcal{D}$ if and only if

$$\forall\, \rho_{V_1...V_nB_1E_1B_2E_2}, \quad \Delta^{\mathcal{D}}(\alpha, \rho_{V_1...V_nB_1E_1B_2E_2}) \geq 0.$$

III. NON-INFINITE FUNCTIONS THAT ARE UNIFORMLY SUBADDITIVE

We will restrict our attention to entropic formulas f_α that are not always infinite: there is at least one $U_\mathcal{N}$ such that $f_\alpha(U_\mathcal{N}) < \infty$. This requirement leads to a particularly nice structure on the α's of a uniformly additive function.

Lemma III.1. *Let $f_\alpha(U_\mathcal{N}, \phi_{V_1 \ldots V_n A})$ satisfy*

$$f_\alpha(U_\mathcal{N}) = \max_{\phi_{V_1 \ldots V_n A}} f_\alpha(U_\mathcal{N}, \phi_{V_1 \ldots V_n A}) < \infty$$

for some *$U_\mathcal{N}$ and*

$$\min_\rho \Delta^\mathcal{D}(\alpha, \rho_{V_1 \ldots V_n B_1 E_1 B_2 E_2}) \geq 0,$$

for a standard decoupling $\mathcal{D}$. In words, f_α is bounded and uniformly subadditive with respect to the standard decoupling $\mathcal{D}$. Then for all non-empty $t \in \mathcal{P}(V_1 \ldots V_n)$,

$$\eta_t := \sum_{s \in \mathcal{P}(BE)} \alpha_{s \cup t} = 0.$$

Proof. For a channel $\mathcal{N}$ such that $f_\alpha(U_\mathcal{N}) < \infty$, considering a state of the form $\rho_{V_1 \ldots V_n}^{\otimes k} \otimes \rho_A$, we have

$$\begin{aligned} & f_\alpha(U_\mathcal{N}, \rho_{V_1 \ldots V_n}^{\otimes k} \otimes \rho_A) \\ =& f_\alpha(U_\mathcal{N}, |0\rangle\langle 0|_{V_1} \otimes \ldots \otimes |0\rangle\langle 0|_{V_n} \otimes \rho_A) + k \sum_{t \in \mathcal{P}(V_1 \ldots V_n)} \eta_t H(t, \rho_{V_1 \ldots V_n}). \end{aligned}$$

So we must have

$$\sum_{t \in \mathcal{P}(V_1 \ldots V_n)} \eta_t H(t, \rho_{V_1 \ldots V_n}) \leq 0, \tag{9}$$

because otherwise, the quantity $f_\alpha(U_\mathcal{N}, \rho_{V_1 \ldots V_n}^{\otimes k} \otimes \rho_A)$ would go to ∞ as $k \to \infty$. Now, in order for f_α to be uniformly subadditive with respect to the standard decoupling $\mathcal{D}$, we need

$$\Delta^\mathcal{D}(\alpha, \rho_{V_1 \ldots V_n B_1 E_1 B_2 E_2}) \geq 0$$

for all $\rho_{V_1 \ldots V_n B_1 E_1 B_2 E_2}$. This implies

$$\Delta^\mathcal{D}(\alpha, \rho_{V_1 \ldots V_n} \otimes |0000\rangle\langle 0000|_{B_1 E_1 B_2 E_2}) = \sum_{t \in \mathcal{P}(V_1 \ldots V_n)} \eta_t H(t, \rho_{V_1 \ldots V_n}) \geq 0, \tag{10}$$

where we have used the fact that $H(s_1\tilde{X}) + H(s_2\hat{X}) - H(s_1 s_2 X) = H(X)$ for this state and any subset X of systems $V_1 \dots V_n$. Eq. (9) and Eq. (10) together imply that

$$\sum_{t \in \mathcal{P}(V_1 \dots V_n)} \eta_t H(t, \rho_{V_1 \dots V_n}) = 0 \tag{11}$$

for all $\rho_{V_1 \dots V_n}$. This implies that each $\eta_t = 0$, by uniqueness results from the classical literature (Theorem 1 of [1]). □

We let

$$\mathcal{F} = \{\alpha \,|\, \exists\, U_{\mathcal{N}},\ f_\alpha(U_{\mathcal{N}}) < +\infty\} \tag{12}$$

be the set of non-infinite entropy formulas .

IV. QUANTUM ENTROPY INEQUALITIES

All known inequalities that constrain entropy allocations in multipartite quantum states can be derived from *strong subadditivity* [2], given by

$$I(A; B|C) := H(AC) + H(BC) - H(ABC) - H(C) \geq 0. \tag{13}$$

Here A, B, and C are arbitrary systems. Pippenger distinguished an independent set of basic inequalities on n systems from which all other known inequalities arise as positive linear combinations [3]. These are (1) nonnegativity of entropy $H(A) \geq 0$, (2) strong subadditivity as stated above, (3)weak monotonicity $H(C|A) + H(C|B) \geq 0$, (4) subadditivity $I(A; B) := H(A) + H(B) - H(AB) \geq 0$ and (5) Araki-Lieb inequality $H(AB) + H(A) - H(B) \geq 0$.

V. NO AUXILIARY VARIABLES

There is only one standard decoupling, $\tilde{\phi}_{A_1} = \mathrm{tr}_{A_2}(\phi_{A_1 A_2})$ and $\hat{\phi}_{A_2} = \mathrm{tr}_{A_1}(\phi_{A_1 A_2})$, when there are no auxiliary variables. We now characterize the cone of uniformly additive linear entropic quantities. By the Minkowski-Weyl theorem, every polyhedron P has a half-space or *H-representation* $P = \{x : Ax \leq b\}$ for some real matrix A and vector b, and a vertex or *V-representation* $P = \mathrm{conv}(v_1, v_2, \dots, v_n) + \mathrm{nonneg}(r_1, r_2, \dots, r_s)$ where $v_1, v_2, \dots, v_n, r_1, r_2, \dots, r_s$ are real vectors, conv denotes the convex hull, and nonneg denotes non-negative linear combinations.

Sufficient conditions: The quantity

$$f_\alpha(U_\mathcal{N}, \phi_A) = \lambda_1 H(B) + \lambda_2 H(E) + \lambda_3 H(B|E) + \lambda_4 H(E|B) \tag{14}$$

is uniformly subadditive for all $\lambda_i \geq 0$. To see this, note first that Eq. (13) implies that $H(B_1B_2) \leq H(B_1) + H(B_2)$ and $H(B_1B_2|E_1E_2) \leq H(B_1|E_1) + H(B_2|E_2)$. The other terms $H(E)$ and $H(E|B)$ are handled similarly. We can then use Eq. (7) to show f_α is uniformly subadditive for $\lambda_i \geq 0$. This characterization of the uniform additivity cone is a V-representation where the quantities $H(B)$, $H(E)$, $H(B|E)$, and $H(E|B)$ are a set of extreme rays and the cone contains the origin.

Necessary conditions: First we express f_α in a slightly different way

$$\begin{aligned} f_\alpha(U_\mathcal{N}, \phi_A) &= \lambda_1 H(B) + \lambda_2 H(E) + \lambda_3 H(B|E) + \lambda_4 H(E|B) \\ &= (\lambda_1 - \lambda_4) H(B) + (\lambda_2 - \lambda_3) H(E) + (\lambda_3 + \lambda_4) H(BE), \end{aligned}$$

so that we have

$$f_\alpha(U_\mathcal{N}, \phi_A) = \alpha_B H(B) + \alpha_E H(E) + \alpha_{BE} H(BE)$$

with $\alpha_B = \lambda_1 - \lambda_4$, $\alpha_E = \lambda_2 - \lambda_3$, and $\alpha_{BE} = \lambda_3 + \lambda_4$. The requirement that $\lambda_i \geq 0$ translates to the conditions

$$\begin{aligned} \alpha_B + \alpha_{BE} &\geq 0 \\ \alpha_E + \alpha_{BE} &\geq 0 \\ \alpha_B + \alpha_E + \alpha_{BE} &\geq 0 \\ \alpha_{BE} &\geq 0. \end{aligned} \tag{15}$$

This characterization of the uniform additivity cone is an H-representation where each inequality corresponds to a face of the cone.

Now we show that these are necessary for uniform subadditivity. To see this, compute

$$\Delta(f_\alpha, p) = \alpha_B I(B_1; B_2) + \alpha_E I(E_1; E_2) + \alpha_{BE} I(B_1E_1; B_2E_2) \tag{16}$$

where p denotes a classical distribution on $B_1B_2E_1E_2$ corresponding to the channel output state. We will show that Eq. (15) are necessary by exhibiting distributions p that lead to a negative value of $\Delta(f_\alpha, p)$ when any of the inequalities is violated.

case	(a,b)	$\hat{M}_1$	$\tilde{M}_2$	equivalents
1.	(3,3)	B_1E_1	B_2E_2	none
2.	(3,1)	B_1E_1	B_2	(1,3), (3,2), (2,3)
3.	(3,0)	B_1E_1	$\varnothing$	(0,3)
4.	(1,1)	B_1	B_2	(2,2)
5.	(1,2)	B_1	E_2	(2,1)
6.	(1,0)	B_1	$\varnothing$	(2,0), (0,1),(0,2)
7.	(0,0)	$\varnothing$	$\varnothing$	none

TABLE I: The inequivalent standard decouplings for one auxiliary variable.

First, suppose $\alpha_B + \alpha_{BE} < 0$. Then, by choosing classical probability distribution p such that $E_1 = E_2 = 0$ and $B_1 = B_2 = R_1$, with R_1 a uniform random bit, we find $\Delta(f_\alpha, p) = \alpha_B + \alpha_{BE} < 0$. We can show $\alpha_E + \alpha_{BE} \geq 0$ is necessary for uniform subadditivity in a similar way. Now, supposing $\alpha_B + \alpha_E + \alpha_{BE} < 0$, we let $B_1 = B_2 = E_1 = E_2 = R_1$ with R_1 a random uniform bit and find $\Delta(f_\alpha, p) = \alpha_B + \alpha_E + \alpha_{BE} < 0$. Finally, if $\alpha_{BE} < 0$, we can let $B_1 = R_1$, $B_2 = R_2$, $E_1 = R_1 \oplus R_3$, $E_2 = R_2 \oplus R_3$ with R_i independent random uniform bits. In this case we find $\Delta(f_\alpha, p) = \alpha_{BE} < 0$.

VI. ONE AUXILIARY VARIABLE

For one auxiliary variable V, there are several choices of standard decouplings taking a state $\phi_{VA_1A_2}$ to states $\tilde{\phi}_{\tilde{V}A_1}$ and $\hat{\phi}_{\hat{V}A_2}$. We define standard decouplings to have $\tilde{V} = \tilde{M}_2 V$ and $\hat{V} = \hat{M}_1 V$ where $\hat{M}_1$ is a collection of output systems from $\mathcal{N}_1$ and $\tilde{M}_2$ is a collection of output systems from $\mathcal{N}_2$. Associate integer labels to each collection according to $0, 1, 2, 3 \leftrightarrow \varnothing, B, E, BE$. The standard decouplings are given by an ordered pair of integers (a, b) where a gives $\hat{M}_1$ and b gives $\tilde{M}_2$. Table I lists the inequivalent standard decouplings.

For one auxiliary variable,

$$\begin{aligned}\Delta^{(a,b)}(f_\alpha,\rho) &= \alpha_B I(B_1;B_2) + \alpha_E I(E_1;E_2) + \alpha_{BE} I(B_1E_1;B_2E_2)\\ &+ \alpha_V(H(\tilde{V}) + H(\hat{V}) - H(V))\\ &+ \alpha_{BV}(H(B_1\tilde{V}) + H(B_2\hat{V}) - H(B_1B_2V))\\ &+ \alpha_{EV}(H(E_1\tilde{V}) + H(E_2\hat{V}) - H(E_1E_2V))\\ &+ \alpha_{BEV}(H(B_1E_1\tilde{V}) + H(B_2E_2\hat{V}) - H(B_1B_2E_1E_2V))\end{aligned}$$

can be rewritten as

$$\Delta^{(a,b)}(\alpha,\rho) = \Delta^{\varnothing}(\alpha^{\varnothing},\rho) + \Delta^{V,(a,b)}(\alpha^V,\rho) \tag{17}$$

where we have replaced f_α by the simpler notation α and

$$\Delta^{\varnothing}(\alpha^{\varnothing},\rho) = \alpha_B I(B_1;B_2) + \alpha_E I(E_1;E_2) + \alpha_{BE} I(B_1E_1;B_2E_2) \tag{18}$$

$$\Delta^{V,(a,b)}(\alpha^V,\rho) = \left(\sum_{s\in\mathcal{P}(BE)} \alpha_{sV}\right) H(\hat{M}_1\tilde{M}_2V) + \sum_{s\in\mathcal{P}(BE)} \alpha_{sV}E_{sV}. \tag{19}$$

In these expressions ρ is the state at the channel outputs on which we evaluate the entropic quantities. The (a,b) index labels the different decouplings we may choose, $\alpha^{\varnothing} = (\alpha_B,\alpha_E,\alpha_{BE})$, $\alpha^V = (\alpha_V,\alpha_{BV},\alpha_{EV},\alpha_{BEV})$, and $\alpha = (\alpha^{\varnothing},\alpha^V)$. The first expression $\Delta^{\varnothing}$ is the same as Eq. (16) in the zero auxiliary case. For each s, the term corresponding to α_{sV} in the second expression has the entropic multiple

$$E_{sV} = H(s_1\tilde{M}_2V) + H(\hat{M}_1s_2V) - H(s_1s_2V) - H(\hat{M}_1\tilde{M}_2V). \tag{20}$$

If $s = \varnothing$, then Eq. (20) takes the value $I(\hat{M}_1;\tilde{M}_2|V)$. If $s = BE$, it takes the value $I(\hat{M}_1^c;\tilde{M}_2^c|\hat{M}_1\tilde{M}_2V)$ where superscript c denotes the complement in $\{B_j,E_j\}$. The expression is more complicated for other values of s. If $s = B$, it evaluates to expressions given in Table II, and if $s = E$ it evaluates to expressions in Table III.

We now show that the variables $\alpha^{\varnothing}$ and α^V can be separated and then prove that Figure 5 in the main text characterizes the uniformly additive formulas obtained using standard decouplings.

A. Separation of Variables

Uniform additivity with one auxiliary variable requires us to consider 5 inequivalent decouplings. Fixing a decoupling (a,b) that maps $\phi_{VA_1A_2} \to (\tilde{\phi}_{\tilde{V}A_1}, \hat{\phi}_{\hat{V}A_2})$ define

$$\Delta^{(a,b)}(\alpha, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) = f_\alpha(U_{\mathcal{N}}, \tilde{\phi}_{\tilde{V}A_1}) + f_\alpha(U_{\mathcal{M}}, \hat{\phi}_{\hat{V}A_2}) - f_\alpha(U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) \quad (21)$$

so that f_α is uniformly additive with respect to (a,b) exactly when for all $U_{\mathcal{N}}$, $U_{\mathcal{M}}$, $\phi_{VA_1A_2}$ we have $\Delta_{(a,b)}(\alpha, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) \geq 0$. Finding the uniformly subadditive f_α is greatly simplified through the separation of variables: letting $\alpha = (\alpha^\varnothing, \alpha^V)$ with $\alpha^\varnothing = (\alpha_B, \alpha_E, \alpha_{BE})$ and $\alpha^V = (\alpha_V, \alpha_{BV}, \alpha_{EV}, \alpha_{BEV})$ and defining

$$\Delta^\varnothing(\alpha^\varnothing, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{A_1A_2}) = f_{\alpha^\varnothing}(U_{\mathcal{N}}, \phi_{A_1}) + f_{\alpha^\varnothing}(U_{\mathcal{M}}, \phi_{A_2}) - f_{\alpha^\varnothing}(U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{A_1A_2}) \quad (22)$$

and

$$\begin{aligned}\Delta^{V,(a,b)}(\alpha^V, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) = {} & f_{(0,\alpha^V)}(U_{\mathcal{N}}, \tilde{\phi}_{\tilde{V}A_1}) + f_{(0,\alpha^V)}(U_{\mathcal{M}}, \hat{\phi}_{\hat{V}A_2}) \\ & - f_{(0,\alpha^V)}(U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}),\end{aligned} \quad (23)$$

we have

$$\Delta^{(a,b)}(\alpha, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) = \Delta^\varnothing(\alpha^\varnothing, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{A_1A_2}) + \Delta^{V,(a,b)}(\alpha^V, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}). \quad (24)$$

We would like to show that $\Delta^{(a,b)}(\alpha, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) \geq 0$ for all $U_{\mathcal{N}}$, $U_{\mathcal{M}}$, $\phi_{VA_1A_2}$ exactly when $\Delta^\varnothing(\alpha^\varnothing, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{A_1A_2}) \geq 0$ for all $U_{\mathcal{N}}$, $U_{\mathcal{M}}$, $\phi_{VA_1A_2}$ and $\Delta^{V,(a,b)}(\alpha^V, U_{\mathcal{N}} \otimes U_{\mathcal{M}}, \phi_{VA_1A_2}) \geq 0$ for all $U_{\mathcal{N}}$, $U_{\mathcal{M}}$, and $\phi_{VA_1A_2}$.

We now show that the V-type terms and the $\varnothing$-type terms can be separated. Let

$$\begin{aligned}\Pi^\varnothing &= \{\alpha^\varnothing \mid \forall\rho,\ \Delta^\varnothing(\alpha^\varnothing, \rho) \geq 0\} \cap \mathcal{F} \\ \Pi^{V,(a,b)} &= \left\{\alpha^V \mid \forall\rho,\ \Delta^{V,(a,b)}(\alpha^V, \rho) \geq 0\right\} \cap \mathcal{F} \\ \Pi^{(a,b)} &= \left\{\alpha \mid \forall\rho,\ \Delta^\varnothing(\alpha^\varnothing, \rho) + \Delta^{V,(a,b)}(\alpha^V, \rho) \geq 0\right\} \cap \mathcal{F}.\end{aligned} \quad (25)$$

Lemma VI.1 (separation of variables)**.** *Let* $\Pi^\varnothing$, $\Pi^{V,(a,b)}$, *and* $\Pi^{(a,b)}$ *be as above. Then*

$$\Pi^{(a,b)} = \{(\alpha^\varnothing, \alpha^V) \mid \alpha^\varnothing \in \Pi^\varnothing \text{ and } \alpha^V \in \Pi^{V,(a,b)}\}.$$

Proof. It is clear that $\Pi^{(a,b)} \supset \Pi^{\varnothing} + \Pi^{V,(a,b)}$, since if $\Delta^{\varnothing}(\alpha^{\varnothing}, \rho) \geq 0$ and $\Delta^{V,(a,b)}(\alpha^V, \rho) \geq 0$ then $\Delta^{(a,b)}(\alpha, \rho) \geq 0$. We would also like to show that any $\alpha \in \Pi^{(a,b)}$ can be decomposed as $\alpha = (\alpha^{\varnothing}, \alpha^V)$ with $\alpha^{\varnothing} \in \Pi^{\varnothing}$ and $\alpha^V \in \Pi^{V,(a,b)}$. To this end, let $\alpha = (\alpha^{\varnothing}, \alpha^V) \in \Pi^{(a,b)}$. To begin with, Lemma III.1 tells us that $\alpha_V + \alpha_{BV} + \alpha_{EV} + \alpha_{BEV} = 0$. This lets us rewrite Eq. (19) as

$$\Delta^{V,(a,b)}(\alpha^V, p) = \sum_{s \in \mathcal{P}(BE)} \alpha_{sV} E_{sV}$$

with

$$\begin{aligned} E_{sV} &= H(s_1 \tilde{M}_2 V) + H(\hat{M}_1 s_2 V) - H(s_1 s_2 V) - H(\hat{M}_1 \tilde{M}_2 V) \\ &= H(s_1 \tilde{M}_2 | V) + H(\hat{M}_1 s_2 | V) - H(s_1 s_2 | V) - H(\hat{M}_1 \tilde{M}_2 | V). \end{aligned} \tag{26}$$

Now, suppose that $\alpha^{\varnothing} \notin \Pi^{\varnothing}$. In that case, as shown in Section V, there is a *classical* probability distribution p_1 on $B_1 B_2 E_1 E_2$ such that $\Delta^{\varnothing}(\alpha^{\varnothing}, p_1) < 0$. However, we can now extend p_1 to V by letting $V = (B_1, B_2, E_1, E_2)$ be a perfectly correlated copy of $B_1 B_2 E_1 E_2$. From Eq.(26) we see that $E_{sV}(p_1)$ is a sum of entropies of subsets of $B_1 B_2 E_1 E_2$ conditioned on V, so $E_{sV}(p_1) = 0$ and therefore $\Delta^{V,(a,b)}(\alpha, p_1) = 0$. But this means that

$$\Delta(\alpha, p_1) = \Delta^{\varnothing}(\alpha^{\varnothing}, p_1) + \Delta^{V,(a,b)}(\alpha^V, p_1) = \Delta^{\varnothing}(\alpha^{\varnothing}, p_1) < 0,$$

so we must have $\alpha \notin \Pi^{(a,b)}$ after all.

Now, suppose $\alpha = (\alpha^{\varnothing}, \alpha^V) \in \Pi^{(a,b)}$, but $\alpha^V \notin \Pi^{V,(a,b)}$. This means that there is some $\rho_{B_1 E_1 B_2 E_2 V}$ such that $\Delta^{V,(a,b)}(\alpha^V, \rho) < 0$. We use this ρ to define a new state,

$$\sigma_{B_1 B_2 E_1 E_2 V'} = \frac{1}{d_B^4 d_E^4} \sum_{i,j} \sum_{k,l} (P_i \otimes P_j \otimes P_k \otimes P_l \otimes I_V) \rho (P_i^\dagger \otimes P_j^\dagger \otimes P_k^\dagger \otimes P_l^\dagger \otimes I_V) \otimes |i,j,k,l\rangle\langle i,j,k,l|_{V_1},$$

where $i, j = 1 \ldots d_B^2$, $k, l = 1 \ldots d_E^2$, P_i and P_j label the Pauli matrices on B, P_k and P_l label the Pauli matrices on E, and $V' = V V_1$. This state is constructed so that

$$\begin{aligned} \Delta^{\varnothing}(\alpha^{\varnothing}, \sigma) &= 0 \\ \Delta^{V,(a,b)}(\alpha^V, \sigma) &= \Delta^{V,(a,b)}(\alpha^V, \rho). \end{aligned}$$

As a result, we also find that

$$\Delta^{(a,b)}(\alpha, \sigma) = \Delta^{\varnothing}(\alpha^{\varnothing}, \sigma) + \Delta^{V,(a,b)}(\alpha^V, \sigma) = \Delta^{V,(a,b)}(\alpha^V, \rho) < 0,$$

so that we have $\alpha \notin \Pi^{(a,b)}$ in this case too. □

$\hat{M}_1$	$\tilde{M}_2$	expression
$\varnothing$	$\varnothing$	$I(B_1;B_2\|V)$
B_1	$\varnothing$	0
E_1	$\varnothing$	$H(B_2\|E_1V)-H(B_2\|B_1V)=I(B_1;B_2\|V)-I(E_1;B_2\|V)$
B_1E_1	$\varnothing$	$-I(E_1;B_2\|B_1V)$
$\varnothing$	B_2	0
B_1	B_2	0
E_1	B_2	0
B_1E_1	B_2	0
$\varnothing$	E_2	$H(B_1\|E_2V)-H(B_1\|B_2V)=I(B_1;B_2\|V)-I(B_1;E_2\|V)$
B_1	E_2	0
E_1	E_2	$H(B_1E_2V)+H(E_1B_2V)-H(B_1B_2V)-H(E_1E_2V)$
B_1E_1	E_2	$I(E_1;E_2\|B_1V)-I(E_1;B_2\|B_1V)$
$\varnothing$	B_2E_2	$-I(B_1;E_2\|B_2V)$
B_1	B_2E_2	0
E_1	B_2E_2	$I(E_2;E_1\|B_2V)-I(E_2;B_1\|B_2V)$
B_1E_1	B_2E_2	$I(E_1;E_2\|B_1B_2V)$

TABLE II: The entropic quantity E_{sV} in Eq. (20) evaluates to these expressions when $s=B$.

For each standard decoupling, we want to identify parameters α such that $\Delta^{(a,b)}\geq 0$ for all states on systems $B_1B_2E_1E_2V$. We use Lemma VI.1, and our earlier characterization of the $\alpha^{\varnothing}$ satisfying $\Delta^{\varnothing}\geq 0$, to separate variables and focus solely on $\Delta^{V,(a,b)}$. Recall also that $\alpha_V+\alpha_{BV}+\alpha_{EV}+\alpha_{BEV}=0$ for all standard decouplings. In what follows, let R_1, R_2, and R_3 denote independent uniform 0-1 random variables.

B. Case 1: (3,3) decoupling

Here we have $\hat{M}_1 = B_1E_1$ and $\tilde{M}_2 = B_2E_2$. We want to compute $\Delta^{V,(3,3)} = \sum_{s\in\mathcal{P}(BE)}\alpha_{sV}E_{sV}$. For $s=\varnothing$ and $s=BE$ we have $E_V=I(B_1E_1;B_2E_2|V)$ and $E_{BEV}=0$. Consulting Table II and III, we find $E_{BV}=I(E_1;E_2|B_1B_2V)$ and $E_{EV}=I(B_1;B_2|E_1E_2V)$

$\hat{M}_1$	$\tilde{M}_2$	expression
$\varnothing$	$\varnothing$	$I(E_1;E_2\|V)$
E_1	$\varnothing$	0
B_1	$\varnothing$	$H(E_2\|B_1V)-H(E_2\|E_1V)=I(E_1;E_2\|V)-I(B_1;E_2\|V)$
E_1B_1	$\varnothing$	$-I(B_1;E_2\|E_1V)$
$\varnothing$	E_2	0
E_1	E_2	0
B_1	E_2	0
E_1B_1	E_2	0
$\varnothing$	B_2	$H(E_1\|B_2V)-H(E_1\|E_2V)=I(E_1;E_2\|V)-I(E_1;B_2\|V)$
E_1	B_2	0
B_1	B_2	$H(E_1B_2V)+H(B_1E_2V)-H(E_1E_2V)-H(B_1B_2V)$
E_1B_1	B_2	$I(B_1;B_2\|E_1V)-I(B_1;E_2\|E_1V)$
$\varnothing$	E_2B_2	$-I(E_1;B_2\|E_2V)$
E_1	E_2B_2	0
B_1	E_2B_2	$I(B_2;B_1\|E_2V)-I(B_2;E_1\|E_2V)$
E_1B_1	E_2B_2	$I(B_1;B_2\|E_1E_2V)$

TABLE III: The entropic quantity E_{sV} in Eq. (20) evaluates to these expressions when $s=E$.

so that

$$\Delta^{V,(3,3)}=\alpha_V I(B_1E_1;B_2E_2|V)+\alpha_{BV}I(E_1;E_2|B_1B_2V)+\alpha_{EV}I(B_1;B_2|E_1E_2V). \tag{27}$$

We now need necessary and sufficient conditions on α for $\Delta^{V,(3,3)}\geq 0$.

Necessary conditions: The conditions

$$\alpha_V+\alpha_{BV}+\alpha_{EV}\geq 0 \tag{28}$$

$$\alpha_V+\alpha_{BV}\geq 0 \tag{29}$$

$$\alpha_V+\alpha_{EV}\geq 0 \tag{30}$$

$$\alpha_V\geq 0 \tag{31}$$

are necessary for positivity of $\Delta^{V,(3,3)}$. To see the necessity of Eq. (28), choose $B_1=R_1$, $B_2=R_2$, $E_1=R_1\oplus R_3$, $E_2=R_2\oplus R_3$ and $V=0$. This give us a distribution with

$I(B_1E_1;B_2E_2|V) = 1$, $I(B_1;B_2|E_1E_2V) = 1$, and $I(E_1;E_2|B_1B_2V) = 1$, and so we find Eq. (28). Eq. (29) can be seen by choosing $B_1 = R_1$, $B_2 = R_1 \oplus R_2$, $E_1 = 0$, $E_2 = R_2$, and $V = 0$ which results in $I(B_1E_1;B_2E_2|V) = 1$, $I(B_1;B_2|E_1E_2V) = 1$, and $I(E_1;E_2|B_1B_2V) = 0$. Eq. (30) can be seen in a similar fashion. Finally, to see Eq. (31), let $B_1 = R_1$, $B_2 = 0$, $E_1 = 0$, $E_2 = R_1$, and $V = 0$.

Sufficient conditions: We will now show that the necessary conditions for positivity of $\Delta^{V,(3,3)}$ are also sufficient. There are four cases to consider. Suppose first that $\alpha_V, \alpha_{BV}, \alpha_{EV} \geq 0$. The positivity of conditional mutual information, and thus the relevant E_{sV}'s, makes $\Delta^{V,(3,3)} \geq 0$ immediately. Suppose next that $\alpha_{BV} \geq 0$ and $\alpha_{EV} < 0$. In this case, we have

$$\begin{aligned}
\Delta^{V,(3,3)} &= \alpha_V I(B_1E_1;B_2E_2|V) + \alpha_{BV} I(E_1;E_2|B_1B_2V) + \alpha_{EV} I(B_1;B_2|E_1E_2V) \\
&= (\alpha_V + \alpha_{EV}) I(B_1E_1;B_2E_2|V) + \alpha_{BV} I(E_1;E_2|B_1B_2V) \\
&+ |\alpha_{EV}| \left[I(B_1E_1;B_2E_2|V) - I(B_1;B_2|E_1E_2V) \right] \\
&= (\alpha_V + \alpha_{EV}) I(B_1E_1;B_2E_2|V) + \alpha_{BV} I(E_1;E_2|B_1B_2V) \\
&+ |\alpha_{EV}| \left[I(E_1;B_2|E_2V) + I(B_1E_1;B_2|V) \right] \geq 0,
\end{aligned}$$

where we have used

$$I(B_1;B_2|E_1E_2V) = I(B_1E_1;B_2E_2|V) - I(E_1;B_2|E_2V) - I(B_1E_1;B_2|V)$$

and the positivity of conditional mutual information. The case $\alpha_{BV} < 0$ and $\alpha_{EV} \geq 0$ follows the same argument as the second case. Finally suppose that $\alpha_{BV} < 0$ and $\alpha_{EV} < 0$. In this case we have

$$\begin{aligned}
\Delta^{V,(3,3)} &= \alpha_V I(B_1E_1;B_2E_2|V) + \alpha_{BV} I(E_1;E_2|B_1B_2V) + \alpha_{EV} I(B_1;B_2|E_1E_2V) \\
&= (\alpha_V + \alpha_{BV} + \alpha_{EV}) I(B_1E_1;B_2E_2|V) + |\alpha_{BV}| \left[I(B_1E_1;B_2E_2|V) - I(E_1;E_2|B_1B_2V) \right] \\
&\quad + |\alpha_{EV}| \left[I(B_1E_1;B_2E_2|V) - I(B_1;B_2|E_1E_2V) \right] \\
&= (\alpha_V + \alpha_{BV} + \alpha_{EV}) I(B_1E_1;B_2E_2|V) + |\alpha_{BV}| \left[I(B_1;E_2|B_2V) + I(B_1E_1;E_2|V) \right] \\
&\quad + |\alpha_{EV}| \left[I(E_1;B_2|E_2V) + I(B_1E_1;B_2|V) \right] \geq 0.
\end{aligned}$$

C. Case 2: (3,1) decoupling

Here we have $\hat{M}_1 = B_1E_1$ and $\tilde{M}_2 = B_2$. For $s = \varnothing$ and $s = BE$ we have $E_V = I(B_1E_1;B_2|V)$ and $E_{BEV} = 0$, respectively, while for $s = B$ and $s = E$, we find $E_{BV} = 0$

and $E_{EV} = I(B_1; B_2|E_1V) - I(B_1; E_2|E_1V)$ from Table II and III. This gives us

$$\Delta^{V,(3,1)} = \alpha_V I(B_1E_1; B_2|V) + \alpha_{EV}\left[I(B_1; B_2|E_1V) - I(B_1; E_2|E_1V)\right]. \tag{32}$$

Necessary conditions: We wish to show that in order to have $\Delta^{V,(3,1)} \geq 0$ for all distributions, we need

$$\alpha_{EV} \leq 0 \tag{33}$$

$$\alpha_V + \alpha_{EV} \geq 0. \tag{34}$$

To see that Eq. (33) is true, choose $B_1 = E_2 = R_1$, and $E_1 = B_2 = V = 0$ to get $\Delta^{V,(3,1)} = -\alpha_{EV}$, so that $\alpha_{EV} \leq 0$. To see that Eq. (34) is necessary, choose $E_2 = E_1 = V = 0$ and $B_1 = B_2 = R_1$. Then $\Delta^{V,(3,1)} = \alpha_V + \alpha_{EV}$ which means we need $\alpha_V + \alpha_{EV} \geq 0$.

Sufficient conditions: Let $\alpha_{EV} \leq 0$ and $\alpha_V + \alpha_{EV} \geq 0$. Then

$$\begin{aligned}
\Delta^{V,(3,1)} &= \alpha_V I(B_1E_1; B_2|V) + \alpha_{EV}\left[I(B_1; B_2|E_1V) - I(B_1; E_2|E_1V)\right] \\
&= (\alpha_V - |\alpha_{EV}|)I(B_1E_1; B_2|V) + |\alpha_{EV}|\left[I(B_1E_1; B_2|V) - I(B_1; B_2|E_1V)\right] + |\alpha_{EV}|I(B_1; E_2|E_1V) \\
&= (\alpha_V + \alpha_{EV})I(B_1E_1; B_2|V) + |\alpha_{EV}|\left[I(E_1; B_2|V)\right] + |\alpha_{EV}|I(B_1; E_2|E_1V) \geq 0.
\end{aligned}$$

D. Case 3: (3,0) decoupling

Here we have $\hat{M}_1 = B_1E_1$ and $\tilde{M}_2 = \varnothing$ which leads to $E_V = 0$, $E_{BEV} = 0$, $E_{BV} = -I(E_1; B_2|B_1V)$, and $E_{EV} = -I(B_1; E_2|E_1V)$. Therefore,

$$\Delta^{V,(3,0)} = -\alpha_{BV} I(E_1; B_2|B_1V) - \alpha_{EV} I(B_1; E_2|E_1V). \tag{35}$$

Necessary conditions: We will need to have

$$\alpha_{BV} \leq 0 \tag{36}$$

$$\alpha_{EV} \leq 0. \tag{37}$$

To see Eq. (36), choose $E_1 = B_2 = R_1$ and $E_2 = B_1 = V = 0$ to get $\Delta^{V,(3,0)} = -\alpha_{BV} \geq 0$. Similarly, choosing $B_1 = E_2 = R_1$ and $E_1 = B_2 = V = 0$ gives $\Delta^{V,(3,0)} = -\alpha_{EV} \geq 0$ and Eq. (37).

Sufficient conditions: Eq.(35) is explicitly nonnegative when $\alpha_{BV} \leq 0$ and $\alpha_{EV} \leq 0$.

E. Case 4: (1,1) decoupling

Here we have $\hat{M}_1 = B_1$ and $\tilde{M}_2 = B_2$, which gives $E_V = I(B_1; B_2|V)$, $E_{BEV} = I(E_1; E_2|B_1B_2V)$, $E_{BV} = 0$, and $E_{EV} = H(E_1B_2V) + H(B_1E_2V) - H(B_1B_2V) - H(E_1E_2V)$. Therefore

$$\Delta^{V,(1,1)} = \alpha_V I(B_1; B_2|V) + \alpha_{EV}\left[H(E_1B_2V) + H(B_1E_2V) - H(B_1B_2V) - H(E_1E_2V)\right] \tag{38}$$

$$+ \alpha_{BEV} I(E_1; E_2|B_1B_2V). \tag{39}$$

Necessary conditions: We need to have

$$\alpha_{EV} = 0 \tag{40}$$

$$\alpha_V \geq 0 \tag{41}$$

$$\alpha_{BEV} \geq 0. \tag{42}$$

Choosing $B_1 = E_2 = R_1$ and $E_1 = B_2 = V = 0$, we find $-\alpha_{EV} \geq 0$ so that $\alpha_{EV} \leq 0$. Choosing $B_1 = R_1$, $B_2 = R_2$, $E_1 = E_2 = R_1 \oplus R_2$, and $V = 0$, we find $\Delta^{V,(1,1)} = \alpha_{EV} \geq 0$, so that $\alpha_{EV} = 0$, showing Eq.(40). Thus, we have

$$\Delta^{V,(1,1)} = \alpha_V I(B_1; B_2|V) + \alpha_{BEV} I(E_1; E_2|B_1B_2V),$$

from which we see Eq. (41) and Eq. (42).

Sufficient conditions: The sufficiency of $\alpha_V \geq 0$, $\alpha_{BEV} \geq 0$ and $\alpha_{EV} = 0$ is immediate from positivity of conditional mutual information.

F. Case 5: (1,2) decoupling

Here we have $\hat{M}_1 = B_1$ and $\tilde{M}_2 = E_2$, which gives $E_V = I(B_1; E_2|V)$, $E_{BEV} = I(E_1; B_2|B_1E_2V)$, $E_{BV} = 0$, and $E_{EV} = 0$. This leads to

$$\Delta^{V,(1,1)} = \alpha_V I(B_1; E_2|V) + \alpha_{BEV} I(E_1; B_2|B_1E_2V).$$

Necessary conditions: We need to have

$$\alpha_V \geq 0 \tag{43}$$

$$\alpha_{BEV} \geq 0. \tag{44}$$

Choosing $B_1 = E_2 = V = 0$, and $E_1 = B_2 = R_1$, we get Eq.(44). Letting $B_1 = B_2 = E_1 = E_2 = R_1$ and $V = 0$ we get Eq.(43).

Sufficient conditions: Sufficiency is immediate from positivity of conditional mutual information.

VII. MULTIPLE AUXILIARY VARIABLES

We now consider the general case with multiple auxiliary variables $V_1, \ldots, V_n$. We will prove that we can separate the variables, similar to the one-variable case. As a result, under a standard decoupling, the cone of uniformly additive entropic formulas is decomposed into a sum of smaller cones, each of which involves one specific subset of the auxiliary variables. Furthermore, the characterizations of these smaller cones is identical with the ones for zero and one auxiliary variable, which we have given in the previous sections. This will finish the characterization of the additive cone under standard decouplings.

Let $\rho_{V_1 \ldots V_n B_1 E_1 B_2 E_2} = (I \otimes U_{\mathcal{N}_1} \otimes U_{\mathcal{N}_2}) \phi_{V_1 \ldots V_n A_1 A_2} (I \otimes U^\dagger_{\mathcal{N}_1} \otimes U^\dagger_{\mathcal{N}_2})$ be a state generated by the channels $\mathcal{N}_1$ and $\mathcal{N}_2$. We are considering entropic quantities evaluated on systems $V_1 ... V_n B_1 E_1 B_2 E_2$. A standard decoupling is an assignment $\tilde{V}_i = N_2^i V_i$, $\hat{V}_i = N_1^i V_i$, where N_1^i is picked from $\mathcal{P}(B_1 E_1)$ and N_2^i is picked from $\mathcal{P}(B_2 E_2)$. We require the decoupling to be consistent: each of B_2 and E_2 appears in at most one N_2^i and each of B_1 and E_1 appears in at most one N_1^i, such that the new auxiliary variables have no overlaps. A consistent standard decoupling will be indexed by $(a_1, b_1) ... (a_n, b_n)$.

Let $J \subseteq [n] := \{1, 2, \ldots, n\}$ be a set of indices and V_J denote the collection of systems $V_1 \ldots V_n$ indexed by J. Likewise let N_1^J and N_2^J denote collections of systems $\{N_1^i\}$ and $\{N_2^i\}$ respectively. Note that $V_\varnothing = \varnothing$, $N_1^\varnothing = \varnothing$ and $N_2^\varnothing = \varnothing$. For α being the coefficient vector of an entropy formula f_α, we let $\alpha^{V_J} = (\alpha_{V_J}, \alpha_{BV_J}, \alpha_{EV_J}, \alpha_{BEV_J})$. In Lemma III.1 we found that if f_α is bounded and uniformly additive with respect to a standard decoupling, then for all J it must hold that

$$\alpha_{V_J} + \alpha_{BV_J} + \alpha_{EV_J} + \alpha_{BEV_J} = 0. \tag{45}$$

So in the following we assume Eq.(45). Thus, we can write

$$\Delta^{V_J,(a_J,b_J)}(\alpha^{V_J},\rho) = \sum_{s\in\mathcal{P}(BE)} \alpha_{sV_J}\Big(H(s_1N_2^J|V_J) + H(N_1^Js_2|V_J) - H(s_1s_2|V_J) - H(N_1^JN_2^J|V_J)\Big), \tag{46}$$

$$\Delta^{(a_1,b_1)...(a_n,b_n)}(\alpha,\rho) = \sum_J \Delta^{V_J,(a_J,b_J)}(\alpha^{V_J},\rho), \tag{47}$$

where (a_J, b_J) tells us which systems from $\{B_1, E_1, B_2, E_2\}$ go with V_J and is induced from $(a_1, b_1)...(a_n, b_n)$. We can now define the cones

$$\Pi^{V_J,(a_J,b_J)} = \big\{\alpha^{V_J}|\ \forall\rho,\ \Delta^{V_J,(a_J,b_J)}(\alpha^{V_J},\rho) \geq 0\big\} \cap \mathcal{F},$$
$$\Pi^{(a_1,b_1)...(a_n,b_n)} = \big\{\alpha|\ \forall\rho,\ \Delta^{(a_1,b_1)...(a_n,b_n)}(\alpha,\rho) \geq 0\big\} \cap \mathcal{F}.$$

If $V_J = \varnothing$, the characterization of $\Pi^{V_J,(a_J,b_J)}$ has been given in Section V. Note that in this case $a_J = 0$ and $b_J = 0$ correspond to empty sets and they are meaningless. If $V_J \neq \varnothing$, we can regard V_J as a single auxiliary variable and find the explicit description of $\Pi^{V_J,(a_J,b_J)}$ in Section VI. On the other hand, the cone $\Pi^{(a_1,b_1)...(a_n,b_n)}$ includes all the uniformly additive quantities f_α, under the decoupling $(a_1, b_1)...(a_n, b_n)$. Our main result in this section is the following Theorem VII.1, which gives a simple characterization of $\Pi^{(a_1,b_1)...(a_n,b_n)}$, in terms of $\Pi^{V_J,(a_J,b_J)}$.

Theorem VII.1. *Given α, we have*

$$\alpha \in \Pi^{(a_1,b_1)...(a_n,b_n)} \tag{48}$$

if and only if

$$\forall J \subseteq [n], \quad \alpha^{V_J} \in \Pi^{V_J,(a_J,b_J)}. \tag{49}$$

The proof of Theorem VII.1 uses the following two lemmas.

Lemma VII.2. *Let $\Pi^{V_J,(a_J,b_J)}$ be defined as above. Then, if $\alpha^{V_J} \notin \Pi^{V_J,(a_J,b_J)}$ there is a classical probability distribution p on $B_1E_1B_2E_2V_J$ such that $\Delta^{V_J,(a_J,b_J)}(\alpha^{V_J},p) < 0$.*

Proof. This is shown in Section V and Section VI. □

Lemma VII.3. *Fix a probability distribution p on $V_1 \dots V_n B_1 E_1 B_2 E_2$ and a consistent standard decoupling $(a_1, b_1)...(a_n, b_n)$. Let $T \subseteq [n]$ be a fixed set and (a_T, b_T) be the induced standard decoupling associated with the set of variables V_T. Then we can construct a probability distribution p' on $V_1' \dots V_n' B_1' E_1' B_2' E_2'$ such that*

$$\Delta^{(a_1,b_1)\dots(a_n,b_n)}(\alpha, p') = \Delta^{V_T,(a_T,b_T)}(\alpha^{V_T}, p). \tag{50}$$

Proof. If the systems B_1, B_2, E_1, E_2 do not have the same size, we extend them such that their sizes are the same. Denote $d = |B_1| = |B_2| = |E_1| = |E_2|$. We let $k = |T|$, the number of indices in T, and let t_i, $i = 1, \dots, k$, be the ith element of T. For each $f = 0, \dots, 3$ and $i = 1, \dots, k$, choose R_i^f to be an independent uniformly distributed variable on $\{0, 1, \dots, d-1\}$. Let $R^f = \sum_{i=1}^k R_i^f \pmod d$. To define p', we let

$$B_1' = R^0 + B_1 \pmod d \qquad E_1' = R^1 + E_1 \pmod d \tag{51}$$

$$B_2' = R^2 + B_2 \pmod d \qquad E_2' = R^3 + E_2 \pmod d. \tag{52}$$

For $i = 1, \dots, k$, we let $V_{t_i}' = V_{t_i} R_i^0 R_i^1 R_i^2 R_i^3$, and for $r \notin T$ we choose $V_r' = B_1 E_1 B_2 E_2$.

For any $X \in \mathcal{P}(B_1 E_1)$ and $Y \in \mathcal{P}(B_2 E_2)$, we let X' and Y' be the corresponding collections of systems from $B_1' E_1'$ and $B_2' E_2'$, respectively (i.e., if $X = B_1 E_1$, then $X' = B_1' E_1'$). Since V_T' includes V_T, as well as all the R_i^f variables from which we know R^0, R^1, R^2 and R^3, we have

$$H(X'Y'|V_T') = H(XY|V_T).$$

This, combined with Eq. (46), gives

$$\Delta^{V_T,(a_T,b_T)}(\alpha^{V_T}, p') = \Delta^{V_T,(a_T,b_T)}(\alpha^{V_T}, p). \tag{53}$$

Next, we show that

$$\Delta^{V_J,(a_J,b_J)}(\alpha^{V_J}, p') = 0, \qquad \text{for all } J \neq T. \tag{54}$$

For this, we consider two cases. If $T \subset J$, then B_1', E_1', B_2', E_2' are all known given V_J', because V_J' includes $B_1 E_1 B_2 E_2 R^0 R^1 R^2 R^3$. As a result,

$$H(X'Y'|V_J') = 0.$$

On the other hand, if $T \not\subseteq J$, There must exist i, such that none of $R_i^0, R_i^1, R_i^2, R_i^3$ is included in V_J'. Thus given V_J', the variables R^0, R^1, R^2, R^3 are independent and uniformly distributed,

and so are B_1', E_1', B_2', E_2'. As a result,

$$H(X'Y'|V_J') = H(X') + H(Y').$$

In both cases, using Eq. (46) we obtain Eq. (54). At last, using Eq. (47) we easily see that Eq. (53) and Eq. (54) together lead to Eq. (50). □

Proof of Theorem VII.1. It is obvious that Eq. (49) implies Eq. (48). For the other direction, we suppose that Eq. (49) is not true: there is a subset $T \subseteq [n]$, such that $\alpha^{V_T} \notin \Pi^{V_T,(a_T,b_T)}$. Then by Lemma VII.2, there is a probability distribution p on $B_1E_1B_2E_2V_T$, satisfying

$$\Delta^{V_T,(a_T,b_T)}(\alpha^{V_T}, p) < 0.$$

Due to Lemma VII.3, this further implies that we have probability distribution p' such that

$$\Delta^{(a_1,b_1)\ldots(a_n,b_n)}(\alpha, p') < 0,$$

which indicates that $\alpha \notin \Pi^{(a_1,b_1)\ldots(a_n,b_n)}$. □

VIII. NON-STANDARD DECOUPLINGS

The motivation of our consideration of standard decouplings comes from the experience in proving additivity of certain well-known quantities. However, a general treatment should consider all possible ways to generate the new auxiliary variables in the decoupling. In this section, we investigate the usefulness of non-standard decouplings. Interestingly, we find that all uniformly additive quantities $f_\alpha(U_\mathcal{N})$ derived from consistent decouplings that are non-standard (cf. definitions in Section II), can be obtained by using standard decouplings. This proves that standard decouplings are really typical.

Theorem VIII.1. *Let the linear entropy formula $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_nA})$ be bounded and uniformly subadditive with respect to a non-standard, consistent decoupling. Then there is $f_\beta(U_\mathcal{N}, \varphi_{V_1\ldots V_mA})$ defined on states with $m \leq n$ auxiliary variables, such that $f_\beta(U_\mathcal{N}, \varphi_{V_1\ldots V_mA})$ is uniformly subadditive with respect to a standard decoupling and*

$$\max_{\phi_{V_1\ldots V_nA}} f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_nA}) = \max_{\varphi_{V_1\ldots V_mA}} f_\beta(U_\mathcal{N}, \varphi_{V_1\ldots V_mA}).$$

Theorem VIII.1 guarantees that there is no need to find out the uniformly sub-additive entropy formulas $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_n A})$ under non-standard consistent decouplings. This is because our interest is in searching for uniformly additive quantities $f_\alpha(U_\mathcal{N}) := \max_\phi f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_n A})$, other than in the entropy formulas themselves. For this purpose, Theorem VIII.1 shows that our consideration of standard decouplings suffices.

Before going to the proof, we specify some of the notations. Since the linear entropy formula $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_n A})$ is defined with respect to the state $\rho_{V_1\ldots V_n BE} = (I\otimes U_\mathcal{N})\phi_{V_1\ldots V_n A}(I\otimes U_\mathcal{N}^\dagger)$, we also denote $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_n A})$ as

$$f_\alpha(\rho_{V_1\ldots V_n BE}) := \sum_{t\in\mathcal{P}(V_1\ldots V_n)} \sum_{s\in\mathcal{P}(BE)} \alpha_{s\cup t} H(st)_\rho. \tag{55}$$

When non-standard decouplings are considered, we may encounter the situation that some of the auxiliary variables are empties. Let the state $\sigma_{V_1\ldots V_n BE}$ have empty auxiliary variables, say, we suppose $V_n = \varnothing$. Then $f_\alpha(\sigma_{V_1\ldots V_n BE})$ is evaluated according to Eq. (55) by letting $H(V_n)_\sigma = 0$ and $H(MV_n)_\sigma = H(M)_\sigma$ for any $M \in \mathcal{P}(V_1\ldots V_{n-1}BE)$. Such a state $\sigma_{V_1\ldots V_n BE}$ with $V_n = \varnothing$ is not artifical: we can identify it in a natural way with $\sigma_{V_1\ldots V_{n-1}BE} \otimes |0\rangle\langle 0|_{V_n}$, that is, empty variables are actually each in a pure state and are hence isolated from the other ones.

Let $\mathcal{D} : \rho_{V_1\ldots V_n B_1 B_2 E_1 E_2} \to (\rho_{\tilde{V}_1\ldots\tilde{V}_n B_1 E_1}, \rho_{\hat{V}_1\ldots\hat{V}_n B_2 E_2})$ be the non-standard decoupling in the assumption of Theorem VIII.1. It is determined by a grouping and relabeling of the systems $V_1, \ldots, V_n, B_2, E_2$ to form $\tilde{V}_1, \ldots, \tilde{V}_n$, and another grouping and relabeling of the systems $V_1, \ldots, V_n, B_1, E_1$ to form $\hat{V}_1, \ldots, \hat{V}_n$. That is, $\tilde{V}_i \in \mathcal{P}(V_1\ldots V_n B_2 E_2)$ and $\hat{V}_i \in \mathcal{P}(V_1\ldots V_n B_1 E_1)$, and as a consistence condition we require $\tilde{V}_i \cap \tilde{V}_j = \hat{V}_i \cap \hat{V}_j = \varnothing$. We further write $\tilde{V}_i$ and $\hat{V}_i$ as the joint of the "V" part and the "BE" part: $\tilde{V}_i = V_i' N_2^i$ with $V_i' \in \mathcal{P}(V_1\ldots V_n)$ and $N_2^i \in \mathcal{P}(B_2E_2)$, $\hat{V}_i = V_i'' N_1^i$ with $V_i'' \in \mathcal{P}(V_1\ldots V_n)$ and $N_1^i \in \mathcal{P}(B_1E_1)$. In this section, the notations $\tilde{V}_i, \hat{V}_i, V_i', V_i'', N_1^i, N_2^i$ with $i = 1, \ldots, n$ are all reserved to denote the fixed sets of variables given by the decoupling $\mathcal{D}$, as described above.

Definition VIII.2. *Given the sets $T_1, \ldots, T_n \in [n]$ such that $T_i \cap T_j = \varnothing$ for $1 \le i \ne j \le n$, we define a relocation rule g of the variables $W_1, \ldots, W_n$, via*

$$g(W_1, \ldots, W_n) := (W_{T_1}, \ldots, W_{T_n}),$$

where W_{T_i} is a collection of the systems W_j such that $j \in T_i$.

According to Definition VIII.2, we now define two relocation rules g_1 and g_2, which are associated with the decoupling $\mathcal{D}$ and satisfy

$$g_1(V_1,\ldots,V_n) = (V_1',\ldots,V_n'),$$
$$g_2(V_1,\ldots,V_n) = (V_1'',\ldots,V_n'').$$

That is, g_1 is given by the sets $T_i := \{j|\ 1 \le j \le n, V_j \in V_i'\}$ with $i = 1,\ldots,n$, and g_2 is given by the sets $S_i := \{j|\ 1 \le j \le n, V_j \in V_i''\}$ with $i = 1,\ldots,n$.

The following lemma will be very useful. Note that in Eqs. (56) and (57), V_i' and V_i'' are actually collections of the variables $V_1,\ldots,V_n$, formulated by the relocation rules g_1 and g_2. So in later applications of Lemma VIII.3, we may also use g_1 and g_2 to specify the relations between the auxiliary variables.

Lemma VIII.3. *Under the same assumption of Theorem VIII.1 and using the notations described above, we have for any state $\rho_{V_1\ldots V_nBE}$,*

$$f_\alpha(\rho_{V_1\ldots V_nBE}) \le f_\alpha(\rho_{V_1'\ldots V_n'BE}), \tag{56}$$

$$f_\alpha(\rho_{V_1\ldots V_nBE}) \le f_\alpha(\rho_{V_1''\ldots V_n''BE}). \tag{57}$$

Proof. At first, it has been shown in Lemma III.1 (Eq. (9)) that $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_nA})$ being bounded implies that

$$f_\alpha(\rho_{V_1\ldots V_n} \otimes |00\rangle\langle 00|_{BE}) = \sum_{t\in\mathcal{P}(V_1\ldots V_n)} \sum_{s\in\mathcal{P}(BE)} \alpha_{s\cup t} H(t)_{\rho_{V_1\ldots V_n}} \le 0 \tag{58}$$

for any state $\rho_{V_1\ldots V_n}$. Now since $f_\alpha(U_\mathcal{N}, \phi_{V_1\ldots V_nA})$ is uniformly subadditive with respect to the decoupling $\mathcal{D}$, we have

$$\Delta(\alpha, \rho_{V_1\ldots V_nB_1B_2E_1E_2}) = f_\alpha(\rho_{\tilde{V}_1\ldots\tilde{V}_nB_1E_1}) + f_\alpha(\rho_{\hat{V}_1\ldots\hat{V}_nB_2E_2}) - f_\alpha(\rho_{V_1\ldots V_nB_1B_2E_1E_2}) \ge 0 \tag{59}$$

for any state $\rho_{V_1\ldots V_nB_1B_2E_1E_2}$. Considering a state of the form $\rho_{V_1\ldots V_nB_1E_1} \otimes \rho_{B_2E_2}$, we derive from Eq. (59) that

$$\begin{aligned}
\Delta(\alpha, \rho_{V_1\ldots V_nB_1E_1} \otimes \rho_{B_2E_2}) &= \sum_{s,t} \alpha_{s\cup t}\big(H(s_1\tilde{t}) + H(s_2\hat{t}) - H(s_1s_2t)\big) \\
&= \sum_{s,t} \alpha_{s\cup t}\big(H(s_1t') + H(\tilde{t}/t') + H(s_2) + H(\hat{t}) - H(s_2) - H(s_1t)\big) \\
&= \sum_{s,t} \alpha_{s\cup t}\big(H(s_1t') - H(s_1t)\big) + \sum_{s\cup t} \alpha_{s\cup t}\big(H(\tilde{t}/t') + H(\hat{t})\big) \\
&\ge 0,
\end{aligned} \tag{60}$$

where the sums are over all subsets $s \in \mathcal{P}(BE)$ and $t \in \mathcal{P}(V_1 \dots V_n)$, and the notation $\tilde{t}/t'$ indicates the collection of variables resulting from removing t' from $\tilde{t}$. Eq. (58) gives

$$\sum_{s,t} \alpha_{s\cup t} H(\tilde{t}/t') \leq 0 \quad \text{and} \quad \sum_{s\cup t} \alpha_{s\cup t} H(\hat{t}) \leq 0.$$

Combining this with Eq. (60) we conclude that for any state $\rho_{V_1 \dots V_n B_1 E_1}$,

$$\sum_{s,t} \alpha_{s\cup t} \big(H(s_1 t') - H(s_1 t) \big) \geq 0,$$

which proves Eq. (56). Since Eq. (57) can be proved in the same way we have finished the proof. □

Now we are ready for the proof of Theorem VIII.1. We will not construct an explicit expression for $f_\beta(U_\mathcal{N}, \varphi_{V_1 \dots V_m A})$. Instead, we prove the existence.

Proof of Theorem VIII.1. We will use mathematical induction. Let us consider the following two cases.

Case 1: $V_i' \neq \varnothing$ and $V_i'' \neq \varnothing$ for all $1 \leq i \leq n$. In this case, $V_1', \dots, V_n'$ and $V_1'', \dots, V_n''$ are respectively permutations of $V_1, \dots, V_n$: there are permutations $\pi, \tau \in S_n$ such that $V_i' = V_{\pi(i)}$ and $V_i'' = V_{\tau(i)}$ for all $1 \leq i \leq n$. Denote the order of π and τ as a and b, respectively. That is

$$\pi^a = \tau^b = I,$$

where I is the identity of the symmetric group S_n. Now define

$$\left(\tilde{V}_1^{(a-1)}, \dots, \tilde{V}_n^{(a-1)} \right) := g_1^{a-1} \left(\tilde{V}_1, \dots, \tilde{V}_n \right), \tag{61}$$

$$\left(\hat{V}_1^{(b-1)}, \dots, \hat{V}_n^{(b-1)} \right) := g_2^{b-1} \left(\hat{V}_1, \dots, \hat{V}_n \right). \tag{62}$$

Then

$$\tilde{V}_i^{(a-1)} = \tilde{V}_{\pi^{a-1}(i)} = V'_{\pi^{a-1}(i)} N_2^{\pi^{a-1}(i)} = V_{\pi^a(i)} N_2^{\pi^{a-1}(i)} = V_i N_2^{\pi^{a-1}(i)}, \tag{63}$$

$$\hat{V}_i^{(b-1)} = \hat{V}_{\tau^{b-1}(i)} = V''_{\tau^{b-1}(i)} N_1^{\tau^{b-1}(i)} = V_{\tau^b(i)} N_1^{\tau^{b-1}(i)} = V_i N_1^{\tau^{b-1}(i)}. \tag{64}$$

To proceed, for any state $\rho_{V_1 \dots V_n B_1 B_2 E_1 E_2}$, we have

$$\begin{aligned} f_\alpha(\rho_{V_1 \dots V_n B_1 B_2 E_1 E_2}) &\leq f_\alpha(\rho_{\tilde{V}_1 \dots \tilde{V}_n B_1 E_1}) + f_\alpha(\rho_{\hat{V}_1 \dots \hat{V}_n B_2 E_2}) \\ &\leq f_\alpha\big(\rho_{\tilde{V}_1^{(a-1)} \dots \tilde{V}_n^{(a-1)} B_1 E_1}\big) + f_\alpha\big(\rho_{\hat{V}_1^{(b-1)} \dots \hat{V}_n^{(b-1)} B_2 E_2}\big), \end{aligned} \tag{65}$$

where the first inequality is by assumption, and for the second inequality we have applied Lemma VIII.3 iteratively and used the notations defined in Eqs. (61) and (62). Eq. (65) shows that $f_\alpha(\rho_{V_1\ldots V_nBE})$ itself is uniformly subadditive with respect to a *standard* decoupling given by Eqs. (63) and (64).

Case 2: at least one of V_i' for $i=1,\ldots,n$ or one of V_i'' for $i=1,\ldots,n$ is $\varnothing$. Without loss of generality, we suppose $V_i'=\varnothing$ for some values of i, and further suppose that all the empty variables are in the end. So there is $k<n$, such that $V_1'\ldots V_n'=V_1'\ldots V_k'\varnothing\ldots\varnothing$ (i.e., $V_i'=\varnothing$ for $i=k+1,\ldots,n$). Note that it is possible that $k=0$. Now Eq. (56) of Lemma VIII.3 translates to

$$f_\alpha(\rho_{V_1\ldots V_nBE})\leq f_\alpha\big(\rho_{V_1'\ldots V_k'BE}\otimes|0\rangle\langle 0|_{V_{k+1}'}\otimes\cdots\otimes|0\rangle\langle 0|_{V_n'}\big). \tag{66}$$

Define a linear entropy formula $f_\gamma(\rho_{V_1\ldots V_kBE})$ on states with k auxiliary variables, as

$$f_\gamma(\rho_{V_1\ldots V_kBE}):=f_\alpha\big(\rho_{V_1\ldots V_kBE}\otimes|0\rangle\langle 0|_{V_{k+1}}\otimes\cdots\otimes|0\rangle\langle 0|_{V_n}\big). \tag{67}$$

We now claim:

(A) It holds that

$$\max_{\phi_{V_1\ldots V_nA}} f_\alpha\big(U_{\mathcal{N}}(\phi_{V_1\ldots V_nA})\big)=\max_{\varphi_{V_1\ldots V_kA}} f_\gamma\big(U_{\mathcal{N}}(\varphi_{V_1\ldots V_kA})\big).$$

In particular, this equality implies that $f_\gamma\big(U_{\mathcal{N}},\varphi_{V_1\ldots V_kA}\big)$ is also bounded.

(B) $f_\gamma\big(\rho_{V_1\ldots V_kBE}\big)$ is uniformly subadditive with respect to a *consistent* decoupling.

Claim (A) is easy to see. The "$\leq$" part follows from Eq. (66), and the "$\geq$" part is obvious by the definition Eq. (67). To verify claim (B), for any state $\rho_{V_1\ldots V_kB_1B_2E_1E_2}$ we have

$$\begin{aligned}
f_\gamma(\rho_{V_1\ldots V_kB_1B_2E_1E_2})=&f_\alpha\big(\rho_{V_1\ldots V_kB_1B_2E_1E_2}\otimes|0\rangle\langle 0|_{V_{k+1}}\otimes\cdots\otimes|0\rangle\langle 0|_{V_n}\big)\\
\leq&f_\alpha(\rho_{\tilde V_1\ldots\tilde V_nB_1E_1})+f_\alpha(\rho_{\hat V_1\ldots\hat V_nB_2E_2})\\
\leq&f_\alpha\big(\rho_{\tilde V_1'\ldots\tilde V_k'B_1E_1}\otimes|0\rangle\langle 0|_{\tilde V_{k+1}'}\otimes\ldots\otimes|0\rangle\langle 0|_{\tilde V_n'}\big)+f_\alpha\big(\rho_{\hat V_1'\ldots\hat V_k'B_2E_2}\otimes|0\rangle\langle 0|_{\hat V_{k+1}'}\otimes\ldots\otimes|0\rangle\langle 0|_{\hat V_n'}\big)\\
=&f_\gamma\big(\rho_{\tilde V_1'\ldots\tilde V_k'B_1E_1}\big)+f_\gamma\big(\rho_{\hat V_1'\ldots\hat V_k'B_2E_2}\big),
\end{aligned} \tag{68}$$

where we have defined $(\tilde V_1',\ldots,\tilde V_n'):=g_1(\tilde V_1,\ldots,\tilde V_n)$ and $(\hat V_1',\ldots,\hat V_n'):=g_1(\hat V_1,\ldots,\hat V_n)$, and since the second line, we have set $V_{k+1}=\cdots=V_n=\varnothing$. In Eq. (68), the first line is

by definition (67), the second line is by assumption that f_α is uniformly subadditive with respect to the decoupling $\mathcal{D}$, the third line is by Lemma VIII.3 (in the form of Eq. (66)), and the last line is again by definition (67). We can check that $\tilde{V}_i' \in \mathcal{P}(V_1 \dots V_k B_2 E_2)$ and $\hat{V}_i' \in \mathcal{P}(V_1 \dots V_k B_1 E_1)$ for $1 \le i \le k$, and also $\tilde{V}_i' \cap \tilde{V}_j' = \varnothing$ and $\hat{V}_i' \cap \hat{V}_j' = \varnothing$ for $i \neq j$. Thus

$$\rho_{V_1 \dots V_k B_1 B_2 E_1 E_2} \to \left(\rho_{\tilde{V}_1' \dots \tilde{V}_k' B_1 E_1}, \rho_{\hat{V}_1' \dots \hat{V}_k' B_2 E_2}\right) \tag{69}$$

is a consistent decoupling and Eq. (68) indeed verifies that $f_\gamma\left(\rho_{V_1 \dots V_k BE}\right)$ is uniformly subadditive with respect to this decoupling.

At last, we argue that the above considerations of Case 1 and Case 2 suffice to conclude the proof, by applying the method of mathematical induction, of which the basis here is the fact that with zero auxiliary variable, the unique consistent decoupling $\rho_{B_1 B_2 E_1 E_2} \to (\rho_{B_1 E_1}, \rho_{B_2 E_2})$ is standard. Note that the proofs for the two claims in Case 2 work as well when $k = 0$ and in this case the decoupling (69) reduces to the standard decoupling $\rho_{B_1 B_2 E_1 E_2} \to (\rho_{B_1 E_1}, \rho_{B_2 E_2})$. □

IX. ENTROPIC CRITERION FOR C-Q STRUCTURE OF QUANTUM STATE

Lemma IX.1. *For a quantum state $\rho_{R_1 R_2 R_3 A}$, suppose the conditional entropies satisfy $H(R_i|R_j) = 0$ and $H(R_i|R_j R_k) = 0$ for all $i, j, k \in \{1, 2, 3\}$. Then the reduced state $\rho_{R_i A}$ is classical-quantum, e.g., $\rho_{R_1 A}$ can be written as*

$$\rho_{R_1 A} = \sum_x p_x |x\rangle\langle x|^{R_1} \otimes \rho_x^A, \tag{70}$$

with $\{|x\rangle\}$ a set of orthogonal states.

Proof. Since $I(R_1; R_2|R_3) = H(R_1|R_3) - H(R_1|R_2 R_3) = 0$, by the result of [4], we know that the reduced state $\rho_{R_1 R_2}$ is separable. Thus we can write

$$\rho_{R_1 R_2} = \sum_{x=1}^{M} p_x \sigma_x^{R_1} \otimes \omega_x^{R_2},$$

and without loss of generality we assume $\sigma_{x_1} \neq \sigma_{x_2}$ and $\omega_{x_1} \neq \omega_{x_2}$ for all $1 \le x_1 \neq x_2 \le M$. Let

$$\rho_{X R_1 R_2} = \sum_{x=1}^{M} p_x |x\rangle\langle x|^X \otimes \sigma_x^{R_1} \otimes \omega_x^{R_2}$$

be an extension of $\rho_{R_1R_2}$, with $\{|x\rangle\}$ a set of orthogonal states. Then we have

$$0 = H(R_1|R_2) \geq H(R_1|R_2X) = \sum_x p_x H(\sigma_x) \geq 0. \tag{71}$$

On the one hand, Eq. (71) implies that σ_x is pure for all values of x. On the other hand, from Eq. (71) we have

$$H(R_1|R_2) = H(R_1|R_2X),$$

which implies that we can recover $\rho_{XR_1R_2}$ from $\rho_{R_1R_2}$ by a CPTP map acting on system R_2 only [5, 6]. This further implies that the set of states $\{\omega_x\}_x$ are mutually orthogonal. In similar ways, we can show that ω_x is pure for all values of x, and the set of states $\{\sigma_x\}_x$ are mutually orthogonal. These consequences all together give us that $\rho_{R_1R_2}$ has the follow form:

$$\rho_{R_1R_2} = \sum_{x=1}^{M} p_x |x\rangle\langle x|^{\otimes 2},$$

with $\{|x\rangle\}$ a set of orthogonal states. This obviously implies Eq. (70), and we are done. □

X. INFORMATIONALLY DEGRADABLE CHANNELS HAVE ADDITIVE COHERENT INFORMATION

We say that a quantum channel $\mathcal{N} : A \to B$ is informationally degradable, if

$$I(V;B) \geq I(V;E)$$

for any state $\rho_{VBE} = (I \otimes U_{\mathcal{N}})\phi_{VA}(I \otimes U_{\mathcal{N}}^{\dagger})$, where $U_{\mathcal{N}} : A \to BE$ is the unitary interaction associated with the channel. This class of channels is a generalization of degradable channels [7], and we will show that they enjoy the same property of additivity for coherent information.

Proposition X.1. *Let quantum channels $\mathcal{N}_1, \ldots, \mathcal{N}_n$ be informationally degradable. Then they have additive coherent information:*

$$I_c(\mathcal{N}_1 \otimes \ldots \otimes \mathcal{N}_n) = \sum_{i=1}^{n} I_c(\mathcal{N}_i). \tag{72}$$

Especially, $Q(\mathcal{N}) = I_c(\mathcal{N})$ for any informationally degradable channel $\mathcal{N}$.

Proof. It suffices to show subadditivity. At first, we notice that due to the informational degradability, for any i it holds that

$$H(B_iV) - H(E_iV) \leq H(B_i) - H(E_i),$$

where the entropies are evaluated on any state $\rho_{VB_iE_i} := U_{\mathcal{N}_i}\phi_{VA_i}U_{\mathcal{N}_i}^\dagger$. Using this, we can actually show uniform subadditivity,

$$\begin{aligned}
I_c(\phi_{A_1\ldots A_n}, \mathcal{N}_1 \otimes \ldots \otimes \mathcal{N}_n) =& H(B_1\ldots B_n) - H(E_1\ldots E_n) \\
=& H(B_1\ldots B_n) - H(E_1B_2\ldots B_n) \\
&+ H(E_1B_2\ldots B_n) - H(E_1E_2B_3\ldots B_n) \\
&+ \ldots \\
&+ H(E_1\ldots E_{n-1}B_n) - H(E_1\ldots E_n) \\
\leq& (H(B_1) - H(E_1)) + (H(B_2) - H(E_2)) + \ldots + (H(B_n) - H(E_n)) \\
=& \sum_{i=1}^{n} I_c(\phi_{A_i}, \mathcal{N}_i).
\end{aligned}$$

□

This implies Eq. (72). The single-letter formula for quantum capacity follows as a consequence directly.

Remark: We do not know whether product of informationally degradable channels is still informationally degradable. This is why we prove additivity for multiple uses of channels, instead of proving this for any two channels. A interesting problem we leave for future study is to find informationally degradable channels that are not degradable in the sense of Ref. [7].

XI. COMPLETELY COHERENT INFORMATION AND QUANTUM SUM RATE

Given an isometry $U_{\mathcal{N}} : A \to BE$, we say that the rate pair (R_1, R_2) is an achievable joint quantum communication rate if for all $\epsilon > 0$ there is an n_0 such that for all $n \geq n_0$ there are isometries $U_{\mathcal{E}_n} : A_1A_2 \to A^nF$ and decoders $U_{\mathcal{D}_1} : B^n \to A_1'D_1$ and $U_{\mathcal{D}_2} : E^n \to A_2'D_2$

with $\log|A_1| \geq nR_1$ and $\log|A_2| \geq nR_2$, and $A_1 \cong A_1'$ and $A_2 \cong A_2'$, such that

$$\begin{aligned}\rho_{V_1A_1V_2A_2D_1D_2F} :&= (U_{\mathcal{D}_1} \otimes U_{\mathcal{D}_2})U_{\mathcal{N}}^{\otimes n}U_{\mathcal{E}}(|\Phi_{V_1A_1}\rangle\langle\Phi_{V_1A_1}| \otimes |\Phi_{V_2A_2}\rangle\langle\Phi_{V_2A_2}|)U_{\mathcal{E}}^{\dagger}(U_{\mathcal{N}}^{\otimes n})^{\dagger}(U_{\mathcal{D}_1}^{\dagger} \otimes U_{\mathcal{D}_2}^{\dagger}) \\ &\approx_{\epsilon} |\Phi_{V_1A_1'}\rangle\langle\Phi_{V_1A_1'}| \otimes |\Phi_{V_2A_2'}\rangle\langle\Phi_{V_2A_2'}| \otimes |\sigma_{D_1D_2F}\rangle\langle\sigma_{D_1D_2F}| =: \sigma_{V_1A_1'V_2A_2'D_1D_2F}\end{aligned} \tag{73}$$

where

$$|\Phi_{V_1A_1}\rangle = \frac{1}{\sqrt{|A_1|}}\sum_{i=1}^{|A_1|}|i\rangle_{V_1}|i\rangle_{A_1} \tag{74}$$

$$|\Phi_{V_2A_2}\rangle = \frac{1}{\sqrt{|A_2|}}\sum_{i=1}^{|A_2|}|i\rangle_{V_2}|i\rangle_{A_2}, \tag{75}$$

$\rho \approx_{\epsilon} \sigma$ means $\|\rho - \sigma\|_1 \leq \epsilon$ and $|\sigma_{D_1D_2F}\rangle$ is some pure state on D_1D_2F.

We define the *joint quantum capacity* of an isometry $U_{\mathcal{N}} : A \to BE$ to be

$$Q_J(U_{\mathcal{N}}) = \max\{R_1 + R_2 | (R_1, R_2) \text{ is acheivable}\}. \tag{76}$$

Now, if (R_1, R_2) is an achievable rate pair, for any ϵ we have encoders and decoders satisfying Eq. (73). Thus, we have

$$\begin{aligned}n(R_1 + R_2) + H(D_1)_{\sigma} - H(D_1F)_{\sigma} &\leq \log|A_1'| + \log|A_2'| + H(D_1)_{\sigma} - H(D_1F)_{\sigma} \\ &= H(A_1')_{\sigma} + H(V_2)_{\sigma} + H(D_1)_{\sigma} - H(D_1F)_{\sigma} \\ &= H(A_1'V_2D_1)_{\sigma} - H(D_2)_{\sigma} \\ &= H(A_1'V_2D_1)_{\sigma} - H(A_2'V_2D_2)_{\sigma} \\ &\approx H(A_1V_2D_1)_{\rho} - H(A_2V_2D_2)_{\rho} \\ &= H(B^nV_2)_{\mu} - H(E^nV_2)_{\mu}\end{aligned} \tag{77}$$

where

$$\mu_{B^nE^nV_1V_2F} = U_{\mathcal{N}}^{\otimes n}U_{\mathcal{E}}(|\Phi_{V_1A_1}\rangle\langle\Phi_{V_1A_1}| \otimes |\Phi_{V_2A_2}\rangle\langle\Phi_{V_2A_2}|)U_{\mathcal{E}}^{\dagger}(U_{\mathcal{N}}^{\otimes n})^{\dagger}. \tag{78}$$

We also have

$$\begin{aligned} n(R_1+R_2)+H(D_1F)_\sigma - H(D_1)_\sigma &\leq \log|A_1'| + \log|A_2'| + H(D_1F)_\sigma - H(D_1)_\sigma \\ &= H\left(A_1'\right)_\sigma + H(V_2)_\sigma + H(D_1F)_\sigma - H(D_1)_\sigma \\ &= H\left(A_1'D_1V_2F\right)_\sigma - H(D_2F)_\sigma \\ &= H\left(A_1'D_1V_2F\right)_\sigma - H(A_2'V_2D_2F)_\sigma \\ &\approx H(B^nV_2F)_\mu - H(E^nV_2F)_\mu. \end{aligned} \tag{79}$$

Now, let

$$\tilde{\mu}_{B^nE^nV_2FT} = \frac{1}{2}\mu_{B^nE^nV_2F}\otimes|0\rangle\langle 0|_T + \frac{1}{2}\mu_{B^nE^nV_2}\otimes|0\rangle\langle 0|_F\otimes|1\rangle\langle 1|_T. \tag{80}$$

This is a state that can be made with n copies of $U_{\mathcal{N}}$. Taking the average of Eq. (77) and Eq. (79), and letting $W = V_2FT$ we find

$$n(R_1+R_2) \lesssim \frac{1}{2}\left(H(B^nV_2)_\mu - H(E^nV_2)_\mu + H(B^nV_2F)_\mu - H(E^nV_2F)_\mu\right) \tag{81}$$

$$= H(B^nW)_{\tilde{\mu}} - H(E^nW)_{\tilde{\mu}} \tag{82}$$

$$\leq I^{cc}(U_{\mathcal{N}}^{\otimes n}) = nI^{cc}(U_{\mathcal{N}}). \tag{83}$$

This implies $Q_J(U_{\mathcal{N}}) \leq I^{cc}(U_{\mathcal{N}})$.

---

[1] R. W. Yeung, Information Theory, IEEE Transactions on **43**, 1924 (1997).

[2] E. H. Lieb and M.-B. Ruskai, J. Math. Phys. **14**, 1938 (1973).

[3] N. Pippenger, Information Theory, IEEE Transactions on **49**, 773 (2003).

[4] P. Hayden, R. Jozsa, D. Petz, and A. Winter, Communications in mathematical physics **246**, 359 (2004).

[5] D. Petz, Communications in mathematical physics **105**, 123 (1986).

[6] D. Petz, The Quarterly Journal of Mathematics **39**, 97 (1988).

[7] I. Devetak and P. W. Shor, Comm. Math. Phys. **256**, 287 (2005), arXiv:quant-ph/0311131.