



## Comparison of Quantum Channels with Superchannels

Journal:	<i>IEEE Transactions on Information Theory</i>
Manuscript ID	IT-18-0571
Manuscript Type:	Regular Manuscript
Date Submitted by the Author:	17-Aug-2018
Complete List of Authors:	Gour, Gilad; University of Calgary, Mathematics and Statistics
Keywords:	Quantum mechanics, Quantum computing, Entropy, Communication channels

SCHOLARONE™  
Manuscripts

# Comparison of Quantum Channels with Superchannels

Gilad Gour<sup>1,\*</sup>

<sup>1</sup> *Department of Mathematics and Statistics, and Institute for Quantum Science and Technology, University of Calgary, AB, Canada T2N 1N4* <sup>†</sup>

(Dated: August 17, 2018)

We extend the definition of the conditional min-entropy from bipartite quantum states to bipartite quantum channels. We show that many of the properties of the conditional min-entropy carry over to the extended version, including an operational interpretation as a guessing probability when one of the subsystems is classical. We then show that the extended conditional min-entropy can be used to fully characterize when two bipartite quantum channels are related to each other via a superchannel (also known as supermap or a comb) that is acting on one of the subsystems. This relation is a pre-order that extends the definition of “quantum majorization” from bipartite states to bipartite channels, and can also be characterized with semidefinite programming. As a special case, our characterization provides necessary and sufficient conditions for when a set of quantum channels is related to another set of channels via a single superchannel. We discuss the applications of our results to channel discrimination, and to resource theories of quantum processes. Along the way we study channel divergences, entropy functions of quantum channels, and noise models of superchannels, including random unitary superchannels, and doubly-stochastic superchannels. For the latter we give a physical meaning as being completely-uniformity preserving.

## Contents

<b>I. Introduction</b>	2
<b>II. Notations and Preliminaries</b>	3
A. The space of linear maps	3
B. The space of supermaps	4
C. Completely Positive Preserving Maps	6
D. Entropies	10
1. The conditional min-entropy	11
2. Entropies and support functions of convex sets	12
<b>III. Noisy Superchannels</b>	12
A. Random Unitary Superchannels	13
B. Doubly Stochastic Superchannels	14
C. Completely Uniformity Preserving Superchannel	15
D. Completely Unital-Channel Preserving Superchannels	18
<b>IV. The Entropy of a Quantum Channel</b>	19
A. The extended conditional min-entropy	21
B. Properties of the extended conditional min entropy	22
C. Operational interpretation as a guessing probability	25
<b>V. Comparison of Quantum Channels</b>	27
A. Channel Divergences	28
B. Characterization of Quantum Majorization for Channels	30
C. Comparison of channels with the extended conditional min-entropy	33
D. Characterization with semidefinite programming	34
E. An application to thermodynamics	35

<sup>†</sup> This paper was presented in part at the Rocky Mountain Summit on Quantum Information JILA, University of Colorado Boulder, June 25-29 (2018), and at 4th Seefeld workshop on Quantum Information Seefeld, July 1-6 (2018).

\*Electronic address: [gour@ucalgary.ca](mailto:gour@ucalgary.ca)

<b>VI. Summary and Conclusions</b>	35
<b>Acknowledgments</b>	36
<b>References</b>	36
<b>A. Strong Duality in Conic Linear programming</b>	38
1. Proof of the Equivalence of (60) and (61)	38
<b>B. Operational interpretation of <math>H_{\min}^{\text{ext}}(B A)_{\Omega}</math> in the case that only system <math>B_1</math> is classical</b>	39

## I. INTRODUCTION

One of the most fundamental properties of quantum mechanics is that it is impossible to perfectly distinguish between distinct non-orthogonal states of a physical system. This property which lies at the heart of the uncertainty principle, may be perceived at first as a hindrance to the theory. In recent years, however, it was shown that it can be harnessed to the success of many interesting quantum information processing tasks such as quantum cryptography [1–3]. Consequently, different techniques have been developed to quantify the distinguishability of quantum states. For example, in quantum hypothesis testing (see e.g. [4–6] and references therein), one uses as a measure of distinguishability the maximized success probability of guessing correctly the quantum state by a quantum measurement. This can be done either in the single shot regime or asymptotically, and either in a symmetric or asymmetric ways. Alternatively, one may choose other figure of merits, such as the accessible information, to measure the distinguishability of quantum states.

The variety of figures of merits indicates that the distinguishability of quantum states cannot be fully captured with a single function, but instead can be described by a pre-order. To understand this pre-order, suppose Alice holds one out of two quantum states  $\rho_1$  and  $\rho_2$ , and not knowing which state it is. Clearly, her ability to determine which state she holds cannot increase if she sends her state  $\rho_x$  ( $x = 1, 2$ ) through a quantum channel  $\Phi$ . Therefore, the pair  $(\rho_1, \rho_2)$  is always more distinguishable than  $(\Phi(\rho_1), \Phi(\rho_2))$  and any measure of distinguishability  $D(\rho_1 \parallel \rho_2)$  must behave monotonically under such a transformation; i.e.

$$D(\Phi(\rho_1) \parallel \Phi(\rho_2)) \leq D(\rho_1 \parallel \rho_2).$$

Given two pairs of quantum states  $(\rho_1, \rho_2)$  and  $(\sigma_1, \sigma_2)$ , how can we determine if there exists a channel  $\Phi$  such that  $\Phi(\rho_x) = \sigma_x$  for both  $x = 1, 2$ ? This question was answered already in 1953 by Blackwell [7] for the classical case, and in 1980 by Alberti and Uhlmann [8] for the qubit case. More recently, it was solved for pure states in [9], characterized in [10], and finally, very recently it was shown in [11] that it can be solved with a semidefinite programming for any finite dimensions. In [11, 12] (see also references therein) it was also shown that this pre-order can be characterized completely in terms of a family of distinguishability measures that are given in terms of the conditional min-entropy [13, 14].

Quantum phenomena however are not static in general. They correspond to dynamical processes which characterize the evolution of a physical system, and described mathematically with completely positive and trace preserving (CPTP) maps (also known as quantum channels). A quantum state can be viewed as a special type of a quantum channel with one dimensional input. In this operational view, a quantum state is a preparation-process, and consequently, quantum channels can be viewed as the fundamental objects of quantum mechanics describing both static and dynamical behaviours of physical systems. Much like quantum states, quantum channels can evolve and change over a period of time. The most general evolution of a quantum channel is described with a superchannel (introduced in [15] under the name of supermaps, or 2-comb in [16]; see also [17]). A superchannel is a linear map that maps (even when act on subsystems) quantum channels to quantum channels, and has a physical realization with a pre and post processing on the quantum channel upon which it acts (see Sec. II, particularly, Fig. 1).

Discrimination of quantum channels can be defined similarly to its state (static) analog, although the theory is richer due to a variety of possible schemes (such as adaptive vs non-adaptive ones, single-shot vs asymptotic, etc) and recently, there has been an active research on the subject [16, 18–25]. Here, however, we will focus on the *comparison* of channels in general, avoiding the optimization of the success probability of a specific scheme. More precisely, suppose Alice holds at her disposal one out of two quantum channels  $\Phi_1$  and  $\Phi_2$ , but she don't know which one. By “sending” her channel  $\Phi_x$  ( $x = 1, 2$ ) through a superchannel  $\Theta$ , she ends up with the channel  $\Theta[\Phi_x]$ . Therefore, the pair  $(\Phi_1, \Phi_2)$  must be more distinguishable than the pair of channels  $(\Theta[\Phi_1], \Theta[\Phi_2])$ , and any measure of distinguishability of quantum channels must behave monotonically under such a transformation.

In this paper we provide necessary and sufficient conditions for the existence of a superchannel  $\Theta$  under which two sets of  $n$  channels  $(\Phi_1, \dots, \Phi_n)$  and  $(\Psi_1, \dots, \Psi_n)$  are related by a superchannel  $\Theta$  via  $\Psi_x = \Theta[\Phi_x]$  for all  $x = 1, \dots, n$ . Our conditions are given in terms of an SDP and therefore can be solved efficiently and algorithmically. Furthermore, we show that the conditions can be expressed in terms of a complete family of distinguishability measures given in terms of a function that we call the *extended conditional min-entropy*. The extended conditional min entropy is an extension of the conditional min-entropy from bipartite states to bipartite channels. We show that the extended conditional min-entropy satisfies many properties similar to those satisfied by the conditional min-entropy. In addition, we develop an axiomatic approach for the entropy of a quantum channel, and discuss the properties that the entropy of a quantum channel must satisfy. Particularly, we argue that entropy functions should behave monotonically under completely uniformity preserving superchannels (see Sec. III B). We show that the doubly stochastic superchannels (i.e. those for which both  $\Theta$  and  $\Theta^*$  are superchannels) have this property.

This paper is organized as follows. In Sec. II we introduce our notations, and discuss the properties of superchannels. In Sec. III we introduce and study four different noise models for superchannels, and discuss the relationships among them. In Sec. IV we define the entropy of a quantum channel, and as an example define the extended min-entropy. In addition, we introduce the extended conditional min-entropy and study its properties and physical meaning. In section V we introduce the main result about discrimination of quantum channels in terms of both an SDP and the extended conditional min-entropy. In addition, we compare our work with the analog work on quantum states as given in [11], and characterize the pre-order that extends quantum majorization (as defined in [11]) from bipartite states to bipartite channels. Finally, we end in Sec. VI with summary and conclusions.

## II. NOTATIONS AND PRELIMINARIES

In this section we introduce our notations, and cover several topics that will be used in the subsequent sections. While a significant part of the material presented in this section can be found (in some form) somewhere else, there are also new key observations that we will use extensively later on.

### A. The space of linear maps

Let  $\mathcal{B}(\mathcal{H})$  denote the space of all (bounded) operators acting on a finite dimensional Hilbert space  $\mathcal{H}$ ,  $\mathcal{B}_h(\mathcal{H})$  the subset consisting of all Hermitian matrices in  $\mathcal{B}(\mathcal{H})$ ,  $\mathcal{B}_+(\mathcal{H})$  the subset of positive semidefinite matrices in  $\mathcal{B}(\mathcal{H})$ , and  $\mathcal{D}(\mathcal{H})$  the subset of all density matrices in  $\mathcal{B}(\mathcal{H})$ . We view  $\mathcal{B}(\mathcal{H})$  as an inner product space quipped with the Hilbert-Schmidt inner product:  $\langle X, Y \rangle := \text{Tr}[X^*Y]$  for all  $X, Y \in \mathcal{B}(\mathcal{H})$ . In this paper, quite often we will consider quantum channels that both their input and output systems are accessible to a single party. In this case, we will denote a channel in Alice's system by  $\Psi^{A_0 \rightarrow A_1} : \mathcal{B}(\mathcal{H}^{A_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1})$ , where  $A_0$  and  $A_1$  are the input and output systems. Since we assume that both the input and output systems are held by Alice, we will denote by  $A$  the joint system  $A_0A_1$  and use the notation

$$\Psi^A := \Psi^{A_0 \rightarrow A_1}.$$

More generally, a multipartite quantum channel is a channel whose input and output spaces are composite quantum systems shared by several parties. If a party  $B$  holds only an output subsystem (not an input one) we will assign to it the trivial 1-dimensional system  $B_0$ . In this way, any party holds two subsystems, the input and output subsystems. For example, a channel shared by two parties will be denoted by

$$\Psi^{AB} := \Psi^{A_0B_0 \rightarrow A_1B_1} : \mathcal{B}(\mathcal{H}^{A_0B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1B_1}). \quad (1)$$

The space of all linear operators from the input space  $\mathcal{B}(\mathcal{H}^{A_0})$  to the output space  $\mathcal{B}(\mathcal{H}^{A_1})$  will be denoted by

$$\mathcal{L}^A := \left\{ \Psi^A : \mathcal{B}(\mathcal{H}^{A_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1}) \mid \Psi^A \text{ is a linear map} \right\}$$

Similarly, we will denote by  $\mathcal{L}^{AB}$  the space of all linear maps as in (1). We will view  $\mathcal{L}^A$  (and  $\mathcal{L}^{AB}$ ) as a vector space equipped with the following inner product. Let  $\{X_a\}$  be an orthonormal basis of  $\mathcal{B}(\mathcal{H}^{A_0})$  (i.e.  $\text{Tr}[X_a^*X_{a'}] = \delta_{aa'}$ ). Then, the inner product between two elements  $\Phi, \Psi \in \mathcal{L}^A$  is defined by:

$$\begin{aligned} \langle \Phi, \Psi \rangle &:= \sum_a \langle \Phi(X_a), \Psi(X_a) \rangle \\ &= \sum_a \text{Tr}[(\Phi(X_a))^* \Psi(X_a)] \end{aligned} \quad (2)$$

Note that we used the symbol  $\langle \cdot, \cdot \rangle$  to denote on the LHS the inner product in  $\mathcal{L}^A$ , while on the RHS the Hilbert Schmidt inner product in  $\mathcal{B}(\mathcal{H}^B)$ .

The above inner product is independent of the choice of the orthonormal basis  $\{X_a\}$  of  $\mathcal{B}(\mathcal{H}^{A_0})$ , and can be expressed in terms of the Choi matrices. The Choi matrix of  $\Psi^A$  is given by

$$J_\Psi^A \equiv J_\Psi^{A_0 A_1} := \left( \text{id}^{A_0} \otimes \Psi^{\tilde{A}_0 \rightarrow A_1} \right) \left( \phi_+^{A_0 \tilde{A}_0} \right) \quad (3)$$

where the tilde symbol will always indicates an identical copy of the system under it, and  $\phi_+^{A_0 \tilde{A}_0} := |\phi_+\rangle\langle\phi_+|^{A_0 \tilde{A}_0}$  is an unnormalized maximally entangled state  $|\phi_+\rangle^{A_0 \tilde{A}_0} := \sum_{i=1}^{d_{A_0}} |i\rangle^{A_0} |i\rangle^{\tilde{A}_0}$ . It is straightforward to show that by taking the canonical basis  $X_{a \equiv (i,j)} = |i\rangle\langle j|^{A_0}$  in (2), the inner product between linear maps is equivalent to inner product between their corresponding Choi matrices:

$$\langle \Phi, \Psi \rangle = \langle J_\Phi^A, J_\Psi^A \rangle = \text{Tr} \left[ (J_\Phi^A)^* J_\Psi^A \right]. \quad (4)$$

We now define the canonical orthonormal basis of  $\mathcal{L}^A$ . Let  $\{X_a\}$  and  $\{Y_b\}$  be two orthonormal bases of  $\mathcal{B}(\mathcal{H}^{A_0})$  and  $\mathcal{B}(\mathcal{H}^{A_1})$ , respectively. Then, one can construct an orthonormal basis,  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$  of  $\mathcal{L}^A$ :

$$\mathcal{E}_{ab}^{A_0 \rightarrow A_1}(\rho) := \text{Tr}[X_a^\dagger \rho] Y_b \quad \forall \rho \in \mathcal{B}(\mathcal{H}^{A_0}). \quad (5)$$

It is straight forward to check that  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$  is an orthonormal basis of  $\mathcal{L}^A$ . The canonical orthonormal basis of  $\mathcal{L}^A$ , is the one obtained from the above basis by taking  $X_{a \equiv (i,j)} = |i\rangle\langle j|^{A_0}$  and  $Y_{b \equiv (k,\ell)} = |k\rangle\langle \ell|^{A_1}$ ; i.e.

$$\mathcal{E}_{ab}^{A_0 \rightarrow A_1}(\rho) := {}^{A_0} \langle i | \rho | j \rangle {}^{A_0} | k \rangle \langle \ell |^{A_1}. \quad (6)$$

## B. The space of supermaps

We denote by  $\mathbb{L}^{AB}$  (with  $A \equiv A_0 A_1$  and  $B \equiv B_0 B_1$ ) the space of all linear maps  $\Theta : \mathcal{L}^A \rightarrow \mathcal{L}^B$ . This space is also a vector space, equipped with the following inner product: for all  $\Theta_1, \Theta_2 \in \mathbb{L}^{AB}$  and an orthonormal basis  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$  of  $\mathcal{L}^A$ :

$$\langle \Theta_1, \Theta_2 \rangle := \sum_{a,b} \left\langle \Theta_1 \left[ \mathcal{E}_{ab}^{A_0 \rightarrow A_1} \right], \Theta_2 \left[ \mathcal{E}_{ab}^{A_0 \rightarrow A_1} \right] \right\rangle, \quad (7)$$

where we used the symbol  $\langle \cdot, \cdot \rangle$  to denote on the LHS the inner product in  $\mathbb{L}^{AB}$ , while on the RHS the inner product in  $\mathcal{L}^B$  (with a definition of inner product as in (2)). As before, this definition is also independent of the choice of the orthonormal basis  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$  of  $\mathcal{L}^A$ .

In analogy with the space  $\mathcal{L}^A$ , for any  $\Theta \in \mathbb{L}^{AB}$  we associate a Choi matrix  $\mathbf{J}_\Theta^{AB} \in \mathcal{B}(\mathcal{H}^{AB})$  which is defined as follows. Let  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$  be the canonical orthonormal basis of  $\mathcal{L}^A$  given in (6), with indices  $a \equiv (i,j)$  and  $b \equiv (k,\ell)$ . Then,

$$\mathbf{J}_\Theta^{AB} := \sum_{a,b} J_{\mathcal{E}_{ab}}^A \otimes J_{\Theta[\mathcal{E}_{ab}]}^B, \quad (8)$$

where  $J_{\mathcal{E}_{ab}}^A$  and  $J_{\Theta[\mathcal{E}_{ab}]}^B$  are the Choi matrices of the maps  $\mathcal{E}_{ab}^{A_0 \rightarrow A_1} \in \mathcal{L}^A$  and  $\Theta[\mathcal{E}_{ab}^{A_0 \rightarrow A_1}] \in \mathcal{L}^B$ , respectively (recall that  $\Theta$  transforms linear maps from  $A_0$  to  $A_1$  to linear maps from  $B_0$  to  $B_1$ ). In the following we provide several motivations for this definition. We start with the preservation of inner products.

Consider two maps  $\Theta_1, \Theta_2 \in \mathbb{L}^{AB}$  and their corresponding Choi matrices  $\mathbf{J}_{\Theta_1}^{AB}$  and  $\mathbf{J}_{\Theta_2}^{AB}$ . The Hilbert Schmidt inner

product between  $\mathbf{J}_{\Theta_1}^{AB}$  and  $\mathbf{J}_{\Theta_2}^{AB}$  can be expressed as follows:

$$\begin{aligned}
 \langle \mathbf{J}_{\Theta_1}^{AB}, \mathbf{J}_{\Theta_2}^{AB} \rangle &= \sum_{a,a',b,b'} \left\langle J_{\mathcal{E}_{ab}}^A \otimes J_{\Theta_1[\mathcal{E}_{ab}]}^B, J_{\mathcal{E}_{a'b'}}^A \otimes J_{\Theta_2[\mathcal{E}_{a'b'}]}^B \right\rangle \\
 &= \sum_{a,a',b,b'} \left\langle J_{\mathcal{E}_{ab}}^A, J_{\mathcal{E}_{a'b'}}^A \right\rangle \left\langle J_{\Theta_1[\mathcal{E}_{ab}]}^B, J_{\Theta_2[\mathcal{E}_{a'b'}]}^B \right\rangle \\
 &= \sum_{a,a',b,b'} \langle \mathcal{E}_{ab}, \mathcal{E}_{a'b'} \rangle \langle \Theta_1[\mathcal{E}_{ab}], \Theta_2[\mathcal{E}_{a'b'}] \rangle \\
 &= \sum_{a,a',b,b'} \delta_{aa'} \delta_{bb'} \langle \Theta_1[\mathcal{E}_{ab}], \Theta_2[\mathcal{E}_{a'b'}] \rangle \\
 &= \sum_{a,b} \langle \Theta_1[\mathcal{E}_{ab}], \Theta_2[\mathcal{E}_{ab}] \rangle = \langle \Theta_1, \Theta_2 \rangle,
 \end{aligned} \tag{9}$$

where in the third equality we used (4).

The second motivation for the definition (8) is the following. Consider a linear map  $\Theta \in \mathbb{L}^{AB}$ , a linear map  $\Psi^A \in \mathcal{L}^A$ , and define  $\Phi^B \equiv \Theta[\Psi^A] \in \mathcal{L}^B$ . Then, the Choi matrices of  $J_\Psi^A$  and  $J_\Phi^B$  of  $\Psi^A$  and  $\Phi^B$ , respectively, are related via

$$J_\Phi^B = \text{Tr}_A \left[ \mathbf{J}_\Theta^{AB} \left( (J_\Psi^A)^T \otimes I^B \right) \right]. \tag{10}$$

That is,  $\mathbf{J}_\Theta^{AB}$  can be interpreted as the Choi matrix of the linear map  $\Delta_\Theta^{A \rightarrow B}$  that converts  $J_\Psi^A$  to  $J_\Phi^B$ . It is defined by

$$\Delta_\Theta^{A \rightarrow B} (J_\Psi^A) := J_{\Theta[\Psi]}^B \tag{11}$$

To see why (10) holds, note that the Choi matrix  $J_\Psi^A$  is linear in  $\Psi$ , and from (6)  $J_{\mathcal{E}_{ab}}^A$  is a real matrix for the canonical basis  $\{\mathcal{E}_{ab}^{A_0 \rightarrow A_1}\}$ , so that

$$\begin{aligned}
 \text{Tr}_A \left[ \mathbf{J}_\Theta^{AB} \left( (J_\Psi^A)^T \otimes I^B \right) \right] &= \\
 \sum_{ab} \text{Tr} \left[ J_{\mathcal{E}_{ab}}^A (J_\Psi^A)^T \right] J_{\Theta[\mathcal{E}_{ab}]}^B &= \sum_{ab} \text{Tr} \left[ (J_{\mathcal{E}_{ab}}^A)^* J_\Psi^A \right] J_{\Theta[\mathcal{E}_{ab}]}^B \\
 = \sum_{ab} \langle \mathcal{E}_{ab}, \Psi \rangle J_{\Theta[\mathcal{E}_{ab}]}^B &= J_{\Theta[\sum_{ab} \langle \mathcal{E}_{ab}, \Psi \rangle \mathcal{E}_{ab}]}^B = J_{\Theta[\Psi]}^B = J_\Phi^B.
 \end{aligned}$$

The matrix  $\mathbf{J}_\Theta^{AB}$  can also be expressed as the Choi matrix of the linear map  $\Lambda_\Theta^{AB} : \mathcal{B}(\mathcal{H}^{A_0 B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1})$  defined by the relation:

$$\mathbf{J}_\Theta^{AB} = \text{id}^{A_0 B_0} \otimes \Lambda_\Theta^{AB} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{B_0 \tilde{B}_0} \right). \tag{12}$$

From (8) it follows that  $\Lambda_\Theta^{AB}$  can be expressed as:

$$\begin{aligned}
 \Lambda_\Theta^{AB} &= \sum_{a,b} \mathcal{E}_{ab}^{A_0 \rightarrow A_1} \otimes \Theta \left[ \mathcal{E}_{ab}^{A_0 \rightarrow A_1} \right] \\
 &\equiv (\mathbb{1}^A \otimes \Theta) \left[ \Phi_+^{A \tilde{A}} \right],
 \end{aligned} \tag{13}$$

where we denote by  $\mathbb{1}^A : \mathcal{L}^A \rightarrow \mathcal{L}^A$  the identity map (whereas the symbol  $\text{id}$  is reserved for the identity map from  $\mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ ; see for example its use in Eq. (3)), and

$$\Phi_+^{A \tilde{A}} := \sum_{a,b} \mathcal{E}_{ab}^{A_0 \rightarrow A_1} \otimes \mathcal{E}_{ab}^{\tilde{A}_0 \rightarrow \tilde{A}_1}.$$

Note that the mapping  $\Theta \mapsto \Lambda_\Theta^{AB} = (\mathbb{1}^A \otimes \Theta) [\Phi_+^{A \tilde{A}}]$  defines an isomorphism between  $\Theta$  and  $\Lambda_\Theta^{AB}$ . The map  $\Phi_+^{A \tilde{A}}$  is completely positive, and it is the CP map analog of the maximally entangled state. Its action on matrices  $\rho^{A_0 \tilde{A}_0} \in \mathcal{B}(\mathcal{H}^{A_0 \tilde{A}_0})$  is given by:

$$\Phi_+^{A \tilde{A}} \left( \rho^{A_0 \tilde{A}_0} \right) = \text{Tr} \left[ \phi_+^{A_0 \tilde{A}_0} \rho^{A_0 \tilde{A}_0} \right] \phi_+^{A_1 \tilde{A}_1}.$$

Similar to the property of the maximally entangled state,  $\Phi_+^{A\tilde{A}}$  satisfies for any  $\Theta \in \mathbb{L}^{AB}$  the relation

$$\mathbb{1}^A \otimes \Theta[\Phi_+^{A\tilde{A}}] = \Theta^T \otimes \mathbb{1}^B[\Phi_+^{\tilde{B}B}] \quad (14)$$

where  $\Theta^T : \mathcal{L}^{\tilde{B}} \rightarrow \mathcal{L}^A$  is the transposition of  $\Theta$  which is defined by its components

$$\langle \mathcal{E}_{ab}^A, \Theta^T[\mathcal{E}_{a'b'}^B] \rangle := \langle \mathcal{E}_{a'b'}^B, \Theta[\mathcal{E}_{ab}^A] \rangle \quad \forall a, b, a', b'$$

where  $\{\mathcal{E}_{ab}^A\}$  and  $\{\mathcal{E}_{a'b'}^B\}$  are the canonical orthonormal bases of  $\mathcal{L}^A$  and  $\mathcal{L}^B$ .

Both the maps  $\Lambda_{\Theta}^{AB}$  and  $\Delta_{\Theta}^{A \rightarrow B}$  correspond to the same map  $\Theta \in \mathbb{L}^{AB}$  and have the *same* Choi matrix  $\mathbf{J}_{\Theta}^{AB}$ . That is,

$$\begin{aligned} \mathbf{J}_{\Theta}^{AB} &= \text{id}^{A_0 B_0} \otimes \Lambda_{\Theta}^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1 B_1} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{B_0 \tilde{B}_0} \right) \\ &= \text{id}^{A_0 A_1} \otimes \Delta_{\Theta}^{\tilde{A}_0 \tilde{A}_1 \rightarrow B_0 B_1} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{A_1 \tilde{A}_1} \right) \end{aligned}$$

Note that (14) implies that  $\Lambda_{\Theta^T}^{BA} = \mathbb{1}^B \otimes \Theta^T[\Phi_+^{B\tilde{B}}] = \Theta \otimes \mathbb{1}^A[\Phi_+^{\tilde{A}A}]$ . Therefore, the Choi matrix

$$\mathbf{J}_{\Theta^T}^{BA} = \text{swap}(\mathbf{J}_{\Theta}^{AB}) \quad (15)$$

where the swap operator is between  $A$  and  $B$ . Moreover, the mapping  $\Theta \mapsto \Lambda_{\Theta}^{AB}$  is an isomorphism map between  $\mathbb{L}^{AB}$  and the space of bipartite maps  $\mathcal{L}^{AB}$ , and similarly the mapping  $\Theta \mapsto \Delta_{\Theta}^{A \rightarrow B}$  is an isomorphism as well.

The dual of a linear map  $\Theta \in \mathbb{L}^{AB}$  is a linear map  $\Theta^* \in \mathbb{L}^{BA}$  with the property that for all  $\Phi^A \in \mathcal{L}^A$  and for all  $\Psi^B \in \mathcal{L}^B$

$$\langle \Phi^A, \Theta^*[\Psi^B] \rangle = \langle \Theta[\Phi^A], \Psi^B \rangle. \quad (16)$$

Note that the LHS of the equation above can be expressed as:

$$\begin{aligned} \langle \Phi^A, \Theta^*[\Psi^B] \rangle &= \langle J_{\Phi}^A, J_{\Theta^*[\Psi]}^A \rangle \\ &= \langle J_{\Phi}^A, \Delta_{\Theta^*}^{B \rightarrow A}(J_{\Psi}^B) \rangle \end{aligned} \quad (17)$$

and the RHS of (16) can be expressed as:

$$\begin{aligned} \langle \Theta[\Phi^A], \Psi^B \rangle &= \langle J_{\Theta[\Phi]}^B, J_{\Psi}^B \rangle \\ &= \langle \Delta_{\Theta}^{A \rightarrow B}(J_{\Phi}^A), J_{\Psi}^B \rangle \\ &= \langle J_{\Phi}^A, \Delta_{\Theta}^{*B \rightarrow A}(J_{\Psi}^B) \rangle \end{aligned} \quad (18)$$

Comparing (17) and (18) we conclude that

$$\Delta_{\Theta^*}^{B \rightarrow A} = \Delta_{\Theta}^{*B \rightarrow A}.$$

Consequently, this also implies that

$$\mathbf{J}_{\Theta^*}^{BA} = \overline{\mathbf{J}_{\Theta}^{AB}}.$$

Finally, we point out that for two supermaps  $\Theta_1 \in \mathbb{L}^{AB}$  and  $\Theta_2 \in \mathbb{L}^{CD}$  we have

$$\begin{aligned} \Delta_{\Theta_2 \circ \Theta_1}^{A \rightarrow D} &= \Delta_{\Theta_2}^{C \rightarrow D} \circ \Delta_{\Theta_1}^{A \rightarrow B} \quad (\text{here } B \cong C) \\ \Delta_{\Theta_1 \otimes \Theta_2}^{AC \rightarrow BD} &= \Delta_{\Theta_1}^{A \rightarrow B} \otimes \Delta_{\Theta_2}^{C \rightarrow D} \end{aligned}$$

### C. Completely Positive Preserving Maps

We will denote by  $\mathcal{L}_+^A$  (and similarly  $\mathcal{L}_+^B$ ) the convex subset of  $\mathcal{L}^A$  consisting of all completely positive (CP) maps in  $\mathcal{L}^A$ . We also denote by  $\mathcal{C}^A$  all the elements in  $\mathcal{L}_+^A$  that are also trace-preserving (TP); i.e.  $\mathcal{C}^A$  is the convex set of all quantum channels in  $\mathcal{L}^A$ .

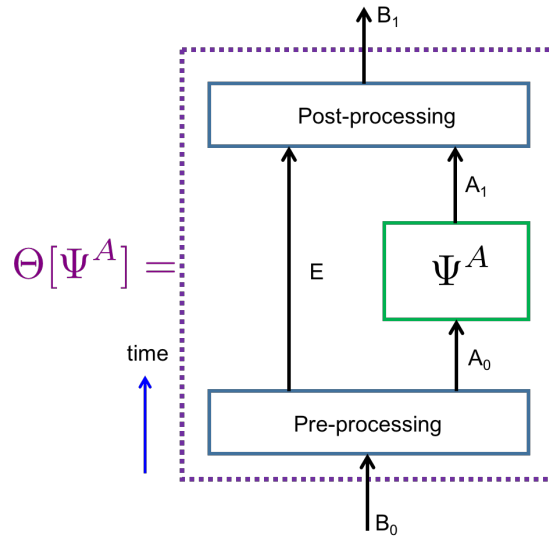


FIG. 1: Realization of a Superchannel

**Definition 1.** Let  $\Theta \in \mathbb{L}^{AB}$  be a linear map. We say that:

1.  $\Theta$  is **CP preserving (CPP)** if  $\Theta[\Psi^A] \in \mathcal{L}_+^B$  for any CP map  $\Psi^A \in \mathcal{L}_+^A$ .
2.  $\Theta$  is **completely CPP** if  $\mathbf{1}^C \otimes \Theta$  is positive for all dimensions of system  $C \equiv (C_0, C_1)$ .
3.  $\Theta$  is **TP preserving (TPP)** if  $\Theta[\Psi^A]$  is a TP map for any TP map  $\Psi^A \in \mathcal{L}^A$ .
4.  $\Theta$  is a **superchannel** if it is completely CPP and TPP.

The following theorem provides the characterization and realization of a superchannel.

**Theorem 1.** [15] Let  $\Theta \in \mathbb{L}^{AB}$ . The following are equivalent.

1.  $\Theta$  is a superchannel.
2. The Choi matrix  $\mathbf{J}_\Theta^{AB} \geq 0$  with marginals

$$\mathbf{J}_\Theta^{A_1 B_0} = I^{A_1 B_0} \quad ; \quad \mathbf{J}_\Theta^{AB_0} = \mathbf{J}_\Theta^{A_0 B_0} \otimes u^{A_1} \quad (19)$$

where  $u^{A_1} := \frac{1}{d_{A_1}} I^{A_1}$  is the maximally mixed state (i.e. the uniform state) on system  $A_1$ .

3. The map  $\Delta_\Theta^{A \rightarrow B}$  is CP, and there exist a unital CP map  $\Delta_\Theta^{A_0 \rightarrow B_0}$  such that the map  $\Delta_\Theta^{A \rightarrow B_0} := \text{Tr}_{B_1} \circ \Delta_\Theta^{A \rightarrow B}$  satisfies

$$\Delta_\Theta^{A \rightarrow B_0} = \Delta_\Theta^{A_0 \rightarrow B_0} \circ \text{Tr}_{A_1} \quad (20)$$

4. There exists a Hilbert space  $\mathcal{H}^E$ , with  $d_E \leq d_{A_0} d_{B_0}$ , and two CPTP maps  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E} : \mathcal{B}(\mathcal{H}^{B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_0 E})$  and  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1} : \mathcal{B}(\mathcal{H}^{A_1 E}) \rightarrow \mathcal{B}(\mathcal{H}^{B_1})$  such that for all  $\Psi^A \in \mathcal{L}^A$

$$\Theta[\Psi^A] = \Gamma_{\text{post}}^{A_1 E \rightarrow B_1} \circ \left( \Psi^{A_0 \rightarrow A_1} \otimes \text{id}^E \right) \circ \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E} \quad (21)$$

(see Fig. 1).

The original proof of this theorem can be found in [15]. However, for the purpose of being self contained, we provide here an alternative proof that is based on similar ideas as given in [26] for the characterization of semi-causal maps [26–29].



*Proof.* Multiplying both sides of (10) by  $Z^{B_0} \otimes I^{B_1}$  (here  $Z^{B_0} \in \mathcal{B}(\mathcal{H}^{B_0})$ ) and taking the trace we get that

$$\text{Tr} \left[ \mathbf{J}_{\Theta}^{AB_0} (J_{\Psi}^A \otimes Z^{B_0})^T \right] = \text{Tr}[Z^{B_0}] . \quad (22)$$

The above equation holds for all  $Z^{B_0} \in \mathcal{B}(\mathcal{H}^{B_0})$  and all trace preserving maps  $\Psi^A \in \mathcal{L}^A$  for which  $J_{\Psi}^{A_0} = I^{A_0}$ . Since the above equation holds for all such  $J_{\Psi}^A$ , it also holds for  $\frac{1}{d_{A_1}} I^A$ , so that

$$\text{Tr} \left[ \mathbf{J}_{\Theta}^{AB_0} (Y^A \otimes Z^{B_0})^T \right] = 0 \quad (23)$$

for any matrix  $Y^A \equiv J_{\Psi}^A - \frac{1}{d_{A_1}} I^A$  whose marginal  $Y^{A_0} = 0$  and any  $Z^{B_0}$ . Therefore,  $\mathbf{J}_{\Theta}^{AB_0}$  must have the form  $\mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{A_1}$ . Finally, substituting this form into (22) and taking  $J_{\Psi}^A = \frac{1}{d_{A_1}} I^A$  gives for all  $Z^{B_0} \in \mathcal{B}(\mathcal{H}^{B_0})$

$$\text{Tr} \left[ \mathbf{J}_{\Theta}^{B_0} (Z^{B_0})^T \right] = d_{A_1} \text{Tr}[Z^{B_0}] . \quad (24)$$

Therefore,  $\mathbf{J}_{\Theta}^{B_0} = d_{A_1} I^{B_0}$  so that  $\mathbf{J}_{\Theta}^{A_1 B_0} = I^{A_1 B_0}$ . This completes the proof of the equivalence between 1 and 2.

We now prove that 2 and 4 are equivalent. Let  $\phi^{ABC}$  be a purification of  $\mathbf{J}_{\Theta}^{AB}$ , and let  $\psi^{A_0 B_0 E}$  be a purification  $\frac{1}{d_{A_1}} \mathbf{J}_{\Theta}^{A_0 B_0}$ . The latter always exists with  $d_E \leq d_{A_0} d_{B_0}$ . Then, from the relation  $\mathbf{J}_{\Theta}^{AB_0} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{A_1}$  we conclude that  $\psi^{A_0 B_0 E} \otimes \phi_+^{A_1 \tilde{A}_1}$  is a purification of  $\mathbf{J}_{\Theta}^{AB_0}$ . Therefore, since  $\phi^{ABC}$  is also a purification of  $\mathbf{J}_{\Theta}^{AB_0}$  there exists an isometry channel  $\mathcal{V}^{\tilde{A}_1 E \rightarrow B_1 C}$  such that

$$\phi^{ABC} = \text{id}^{AB_0} \otimes \mathcal{V}^{\tilde{A}_1 E \rightarrow B_1 C} \left( \psi^{A_0 B_0 E} \otimes \phi_+^{A_1 \tilde{A}_1} \right)$$

Tracing out system  $C$  on both sides, and denoting  $\Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} := \text{Tr}_C \circ \mathcal{V}^{\tilde{A}_1 E \rightarrow B_1 C}$  gives

$$\mathbf{J}_{\Theta}^{AB} = \text{id}^{AB_0} \otimes \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \left( \psi^{A_0 B_0 E} \otimes \phi_+^{A_1 \tilde{A}_1} \right) \quad (25)$$

From its definition,  $\text{Tr}_{A_0 E}[\psi^{A_0 B_0 E}] = I^{B_0}$ . Hence, there exists a CPTP map (in fact an isometry)  $\Gamma_{\text{pre}}^{\tilde{B}_0 \rightarrow A_0 E}$  such that  $\psi^{A_0 B_0 E} = \text{id}^{B_0} \otimes \Gamma_{\text{pre}}^{\tilde{B}_0 \rightarrow A_0 E}(\phi_+^{B_0 \tilde{B}_0})$ . Hence,

$$\begin{aligned} \mathbf{J}_{\Theta}^{AB} &= \left( \text{id}^{A_1 B_0} \otimes \Gamma_{\Theta}^{\tilde{A}_1 \tilde{B}_0 \rightarrow A_0 B_1} \right) \left( \phi_+^{A_1 \tilde{A}_1} \otimes \phi_+^{B_0 \tilde{B}_0} \right) \\ \Gamma_{\Theta}^{\tilde{A}_1 \tilde{B}_0 \rightarrow A_0 B_1} &:= \\ &\left( \text{id}^{A_0} \otimes \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \right) \circ \left( \text{id}^{\tilde{A}_1} \otimes \Gamma_{\text{pre}}^{\tilde{B}_0 \rightarrow A_0 E} \right) \end{aligned} \quad (26)$$

To finish the proof, denote  $\Phi^B := \Theta[\Psi^A]$  and from (10) we get for arbitrary  $\rho \in \mathcal{B}(\mathcal{H}^{B_0})$

$$\begin{aligned} \Phi(\rho) &= \text{Tr}_{B_0} \left[ J_{\Phi}^B (\rho^T \otimes I^{B_1}) \right] \\ &= \text{Tr}_{AB_0} \left[ \mathbf{J}_{\Theta}^{AB} \left( (J_{\Psi}^A)^T \otimes \rho^T \otimes I^{B_1} \right) \right] . \end{aligned} \quad (27)$$

Next, we substitute the above expression for  $\mathbf{J}_{\Theta}^{AB}$  to get

$$\begin{aligned} \Phi(\rho) &= \text{Tr}_A \left[ \left( (J_{\Psi}^A)^T \otimes I^{B_1} \right) \times \right. \\ &\quad \left. \left( \text{id}^A \otimes \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \left( \phi_+^{A_1 \tilde{A}_1} \otimes \Gamma_{\text{pre}}^{\tilde{B}_0 \rightarrow A_0 E}(\rho) \right) \right) \right] \end{aligned} \quad (28)$$

where we traced out system  $B_0$  after using the relation

$$\left( \rho^T \otimes I^{\tilde{B}_0} \right) \phi_+^{B_0 \tilde{B}_0} = (I^{B_0} \otimes \rho) \phi_+^{B_0 \tilde{B}_0} .$$

Finally, note that

$$(J_{\Psi}^A)^T = \sum_{i,j} |j\rangle\langle i|^{A_0} \otimes \left( \Psi^{\tilde{A}_0 \rightarrow A_1} \left[ |i\rangle\langle j|^{\tilde{A}_0} \right] \right)^T ,$$

so that

$$\begin{aligned} & \text{Tr}_{A_1} \left[ \left( (J_\Psi^A)^T \otimes I^{\tilde{A}_1} \right) \left( I^{A_0} \otimes \phi_+^{A_1 \tilde{A}_1} \right) \right] \\ &= \sum_{i,j} |j\rangle\langle i|^{A_0} \otimes \Psi^{\tilde{A}_0 \rightarrow \tilde{A}_1} \left[ |i\rangle\langle j|^{\tilde{A}_0} \right]. \end{aligned}$$

Substituting this into (28) we conclude that

$$\begin{aligned} \Phi(\rho) &= \sum_{i,j} \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \\ & \left( \Psi^{\tilde{A}_0 \rightarrow \tilde{A}_1} \left[ |i\rangle\langle j|^{\tilde{A}_0} \right] \otimes \left\langle i^{A_0} \left| \Gamma_{\text{pre}}^{\tilde{B}_0 \rightarrow A_0 E}(\rho) \right| j^{A_0} \right\rangle \right) \\ &= \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \left[ \left( \Psi^{A_0 \rightarrow A_1} \otimes \text{id}^E \right) \left( \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E}(\rho) \right) \right]. \end{aligned}$$

This completes the proof that 2 and 4 are equivalent. Finally, we show now that 2 and 3 are equivalent.

Suppose  $\Delta_\Theta^{A \rightarrow B}$  has the form (20). Then, since  $\mathbf{J}_\Theta^{AB}$  is its Choi matrix we get

$$\begin{aligned} \mathbf{J}_\Theta^{AB_0} &= \Delta_\Theta^{\tilde{A} \rightarrow B_0} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{A_1 \tilde{A}_1} \right) \\ &= \Delta_\Theta^{\tilde{A} \rightarrow B_0} \left( \phi_+^{A_0 \tilde{A}_0} \right) \otimes I^{A_1}, \end{aligned}$$

where we used the form (20). Hence,  $\mathbf{J}_\Theta^{AB_0} = \mathbf{J}_\Theta^{A_0 B_0} \otimes u^{A_1}$ , and since  $\Delta_\Theta^{\tilde{A} \rightarrow B_0}$  is a unital CP map we conclude that  $\mathbf{J}_\Theta^{B_0} = d_{A_1} I^{B_0}$  so that  $\mathbf{J}_\Theta^{A_1 B_0} = I^{A_1 B_0}$ .

Conversely, suppose the Choi matrix of  $\Theta$  is positive semidefinite and has marginals as in (19). Define the map  $\Delta_\Theta^{A_0 \rightarrow B_0}$  to be the (unique) map satisfying

$$\frac{1}{d_{A_1}} \mathbf{J}_\Theta^{A_0 B_0} = \Delta_\Theta^{\tilde{A}_0 \rightarrow B_0} \left( \phi_+^{A_0 \tilde{A}_0} \right).$$

Therefore,  $\Delta_\Theta^{A_0 \rightarrow B_0}$  is a unital CP map since  $\mathbf{J}_\Theta^{B_0} = d_{A_1} I^{B_0}$ . Moreover, from the relation  $\mathbf{J}_\Theta^{AB_0} = \mathbf{J}_\Theta^{A_0 B_0} \otimes u^{A_1}$  we get that the two maps  $\Delta_\Theta^{A \rightarrow B_0}$  and  $\Delta_\Theta^{A_0 \rightarrow B_0} \circ \text{Tr}_{A_1}$  have the same Choi matrix and therefore they must be the same. This completes the proof of the equivalence between 2 and 3.  $\blacksquare$

From the theorem above, it follows that  $\Theta$  is a superchannel if and only if the CPTP map  $\Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1}$  which corresponds to the Choi matrix

$$\mathbf{J}_\Theta^{AB} = \text{id}^{A_1 B_0} \otimes \Gamma_\Theta^{\tilde{A}_1 \tilde{B}_0 \rightarrow A_0 B_1} \left( \phi_+^{A_1 \tilde{A}_1} \otimes \phi_+^{B_0 \tilde{B}_0} \right) \quad (29)$$

can be expressed as

$$\Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1} = \left( \text{id}^{A_0} \otimes \Gamma_{\text{post}}^{A_1 E \rightarrow B_1} \right) \circ \left( \text{id}^{A_1} \otimes \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E} \right). \quad (30)$$

Therefore,  $\Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1}$  is the CPTP map obtained from  $\Theta$  by taking  $A_0$  and  $B_1$  to be the outputs and  $A_1$  and  $B_0$  to be the inputs as described in Fig. 2.

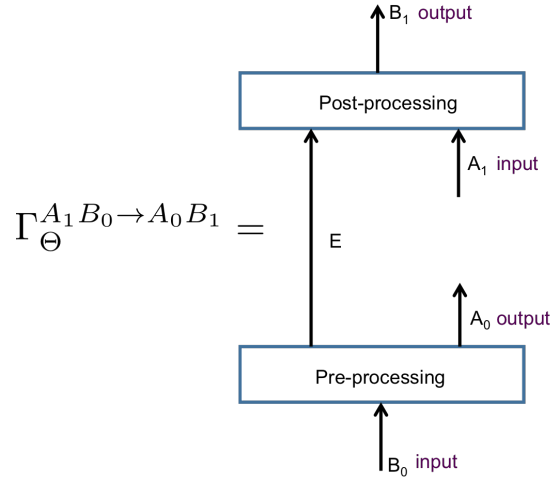
Note that the CPTP map  $\Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1}$  has the form (30) if and only if its marginal map  $\Gamma_\Theta^{A_1 B_0 \rightarrow A_0} \equiv \text{Tr}_{B_1} \circ \Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1}$  has the form:

$$\Gamma_\Theta^{A_1 B_0 \rightarrow A_0} = \Gamma_\Theta^{B_0 \rightarrow A_0} \circ \text{Tr}_{A_1}, \quad (31)$$

where  $\Gamma_\Theta^{B_0 \rightarrow A_0} \equiv \text{Tr}_E \circ \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E}$  is some CPTP map. This condition is somewhat similar to the condition in (20).

The maps  $\Delta_\Theta^{A \rightarrow B}$  and  $\Gamma_\Theta^{A_1 B_0 \rightarrow A_0 B_1}$  are closely related as they corresponds to the same Choi matrix  $\mathbf{J}_\Theta^{AB}$ . In particular, the CPTP map  $\Gamma_\Theta^{B_0 \rightarrow A_0}$  of (31) is related to the unital CP map  $\Delta_\Theta^{A_0 \rightarrow B_0}$  of (20) via

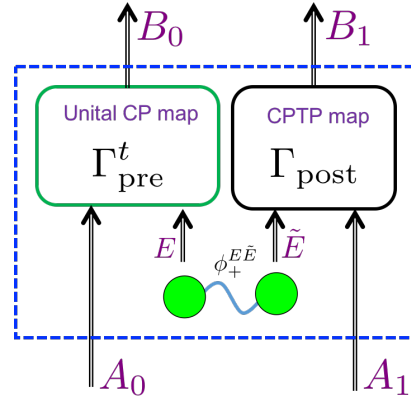
$$\Gamma_\Theta^{B_0 \rightarrow A_0}(\rho^{B_0}) = \left[ \Delta_\Theta^{*B_0 \rightarrow A_0} \left( (\rho^{B_0})^T \right) \right]^T \quad (32)$$

FIG. 2: Realization of the channel  $\Gamma_{\Theta}^{A_1 B_0 \rightarrow A_0 B_1}$ 

for all  $\rho^{B_0} \in \mathcal{B}(\mathcal{H}^{B_0})$ . More generally, it can be shown that (see also Fig. 3)

$$\Delta_{\Theta}^{A \rightarrow B}(\sigma^A) = \left( \Gamma_{\text{pre}}^{t A_0 E \rightarrow B_0} \otimes \text{id}^{B_1} \right) \circ \left( \text{id}^{A_0 E} \otimes \Gamma_{\text{post}}^{A_1 \tilde{E} \rightarrow B_1} \right) \left( \sigma^A \otimes \phi_+^{E \tilde{E}} \right),$$

where  $\Gamma_{\text{pre}}^{t A_0 E \rightarrow B_0}$  is the unital CP map obtained from the CPTP map  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E}$  by replacing all of its Kraus operators (in the operator sum representation) with their transpose.

FIG. 3: Realization of the CP map  $\Delta_{\Theta}^{A \rightarrow B}$ 

#### D. Entropies

Entropy functions, such as the family of Rényi entropies, measure how noisy a quantum state is. Therefore, they do not change under unitary channels, meaning that the entropy of a quantum state  $\rho \in \mathcal{B}(\mathcal{H})$  is the same as that of  $\mathcal{U}(\rho) := U\rho U^*$ , for any unitary matrix  $U$  acting on  $\mathcal{H}$ . Now, suppose that  $U$  is chosen at random from some ensemble  $\{p_x, U_x\}_{x=1}^m$  of unitary matrices. Still the entropy of  $\mathcal{U}_x(\rho) := U_x \rho U_x^*$  and  $\rho$  are the same for each  $x$ . If the information about  $x$  is lost, the state of the system becomes:

$$\sum_{x=1}^m p_x \mathcal{U}_x(\rho).$$

Since “losing information” cannot decrease noise, the entropy of the state above can only be larger than the entropy of  $\rho$ . That is, entropy functions must behave monotonically under random unitary channels.

Since we identify the maximally mixed state as the state with the most noise, quantum channels with the same input and output dimensions, and that do not decrease noise, must preserve the maximally mixed state. Indeed, random unitary channels have this property. The set of all channels with the same input and output dimensions that preserve the maximally mixed state are called unital CPTP maps or doubly stochastic channels. While they form a strictly larger set of channels than the set of random unitary channels, if a state  $\rho$  can be converted into  $\sigma$  via a doubly stochastic channel, then this transformation can also be achieved by random unitary channels.

From the discussion above it follows that all entropy functions cannot decrease under doubly stochastic channels. In addition, entropy functions are also additive under tensor product. To summarize, an entropy function  $f : \mathcal{D}(\mathcal{H}) \rightarrow \mathbb{R}$  satisfies the following 3 conditions:

1. *Monotonicity*: For any random unitary channel  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ ,

$$f(\Phi(\rho)) \geq f(\rho) .$$

2. *Additivity*: For any two quantum states  $\rho \in \mathcal{B}(\mathcal{H}^A)$  and  $\sigma \in \mathcal{B}(\mathcal{H}^B)$  we have

$$f(\rho \otimes \sigma) = f(\rho) + f(\sigma) .$$

3. *Normalization*: on maximally mixed state  $f(\frac{1}{d}I) = \log(d)$ , and on pure state  $f(|\psi\rangle\langle\psi|) = 0$ .

One may add other conditions such as concavity, or subadditivity, but they are not as fundamental as the above three (for example, not all the Rényi entropies satisfy them).

### 1. The conditional min-entropy

The min-entropy plays an important role in quantum information. It is the smallest entropy in the family of Rényi entropies, and in this sense provides the most conservative way to quantify uncertainty. The min-entropy of a density matrix  $\rho \in \mathcal{B}(\mathcal{H}^{A_1})$  is defined by:

$$H_{\min}(A_1)_\rho := -\log \min\{t \in \mathbb{R} : tI \geq \rho\} = -\log \lambda_{\max}(\rho)$$

where  $\lambda_{\max}(\rho)$  is the maximum eigenvalue of  $\rho$ . Like the Rényi entropies, the min entropy has many interesting properties including additivity under tensor product and monotonicity under unital CPTP maps.

The conditional min-entropy of a density matrix  $\rho^A \in \mathcal{B}(\mathcal{H}^{A_0 A_1})$  is defined by [13]:

$$H_{\min}(A_1|A_0)_\rho := -\log \min \{ \text{Tr}[\sigma^{A_0}] : \sigma^{A_0} \otimes I^{A_1} \geq \rho^A \} \quad (33)$$

Note that the condition  $\sigma^{A_0} \otimes I^{A_1} \geq \rho^A$  implies that  $\sigma^{A_0} \geq 0$ . The conditional min entropy can be calculated using semi-definite programming (SDP). Any SDP optimization problem has a dual. For the conditional min-entropy, the dual is given by [14]:

$$2^{-H_{\min}(A_1|A_0)_\rho} = \max_{\Lambda \in \mathcal{C}^A} \text{Tr} \left[ \phi_+^{\tilde{A}_1 A_1} \left( \Lambda \otimes \text{id}^{A_1} (\rho^{A_0 A_1}) \right) \right] \quad (34)$$

where the maximum is over all CPTP maps  $\Lambda : \mathcal{B}(\mathcal{H}^{A_0}) \rightarrow \mathcal{B}(\mathcal{H}^{\tilde{A}_1})$ . We use the notation  $\phi_+^{\tilde{A}_1 A_1} := |\phi_+^{\tilde{A}_1 A_1}\rangle\langle\phi_+^{\tilde{A}_1 A_1}|$ , where  $|\phi_+^{\tilde{A}_1 A_1}\rangle = \sum_{x=1}^{d_{A_1}} |x\rangle^{\tilde{A}_1} |x\rangle^{A_1}$  is an unnormalized maximally entangled state. If system  $A_1$  is classical then the above expression reduces to the optimal guessing probability (that is, the optimal probability to guess the classical value of  $A_1$  after measuring the quantum system  $A_0$ ).

The conditional min-entropy has many properties reminiscent to the conditional von-Neumann entropy. First, it is indeed a generalization of the min-entropy. Particularly, if  $\dim(\mathcal{H}^{A_0}) = 1$  we get that  $H_{\min}(A_1|A_0) = H_{\min}(A_1)$ . Second, if  $\rho^A = \rho^{A_0} \otimes \rho^{A_1}$  then  $H_{\min}(A_1|A_0) = H_{\min}(A_1)$ . In addition, conditioning can only reduce the min-entropy; i.e. for any density matrix  $\rho^{AB} \in \mathcal{B}(\mathcal{H}^{AB})$  we have  $H_{\min}(A_1|A_0B) \leq H_{\min}(A_1|A_0)$ . Finally, its smoothed version satisfies the fully quantum asymptotic equipartition property, which states that in the limit of many copies of a state, the smoothed conditional min-entropy approaches the conditional von-Neuman entropy [30].

## 2. Entropies and support functions of convex sets

The support function of a convex set is one of the most central and basic concepts in convex geometry. The support function  $f_A : \mathbb{R}^n \rightarrow \mathbb{R}$  of a non-empty closed convex set  $\mathcal{C}$  in  $\mathbb{R}^n$  is defined by:

$$f_{\mathcal{C}}(\mathbf{v}) = \sup \{ \mathbf{u} \cdot \mathbf{v} : \mathbf{u} \in \mathcal{C} \} \quad \forall \mathbf{v} \in \mathbb{R}^n .$$

Consider the real vector space,  $\mathcal{B}_h(\mathcal{H})$ , of all Hermitian matrices in  $\mathcal{B}(\mathcal{H})$ . Note that the set of all density matrices in this space,  $\mathcal{D}(\mathcal{H})$ , is closed and convex. The support function of  $\mathcal{D}(\mathcal{H})$  is given by:

$$f_{\mathcal{D}(\mathcal{H})}(\rho) = \sup \{ \text{Tr}[\rho\sigma] : \sigma \in \mathcal{D}(\mathcal{H}) \} \quad \forall \rho \in \mathcal{B}_h(\mathcal{H}) .$$

This support function can be expressed as:

$$f_{\mathcal{D}(\mathcal{H})}(\rho) = \lambda_{\max}(\rho) = 2^{-H_{\min}(\rho)} . \quad (35)$$

That is, the min-entropy is simply  $-\log_2$  of the support function of the set of density matrices  $\mathcal{D}(\mathcal{H})$  in  $\mathcal{B}_h(\mathcal{H})$ .

Similarly, consider the space of all linear maps from  $\mathcal{B}(\mathcal{H}^{A_0})$  to  $\mathcal{B}(\mathcal{H}^{A_1})$ , which we denoted by  $\mathcal{L}^A$ . Since this is a complex vector space, consider its subspace  $\mathcal{L}_{\text{HP}}^A$  of all Hermitian preserving linear maps in  $\mathcal{L}^A$ . Clearly, this is a real vector space equipped with the inner product (2). Recall our notation  $\mathcal{C}^A \subset \mathcal{L}_{\text{HP}}^A$  for the set of all CPTP maps in the space  $\mathcal{L}_{\text{HP}}^A$ . Since  $\mathcal{C}^A$  is a closed convex subset of  $\mathcal{L}_{\text{HP}}^A$ , its support function is well defined: for all  $\Psi^{A_0 \rightarrow A_1} \in \mathcal{L}_{\text{HP}}^A$

$$f_{\mathcal{C}^A}(\Psi^{A_0 \rightarrow A_1}) = \sup \{ \langle \Lambda, \Psi \rangle : \Lambda \in \mathcal{C}^A \} .$$

As we show now, this expression is closely related to the conditional min entropy. Using the relation (4) we get that

$$f_{\mathcal{C}^A}(\Psi^{A_0 \rightarrow A_1}) = \sup_{\Lambda \in \mathcal{C}^A} \text{Tr} [J_{\Lambda}^A J_{\Psi}^A] \quad (36)$$

Recall that

$$J_{\Lambda}^A = \text{id}^{A_0} \otimes \Lambda \left( \phi_{+}^{A_0 \tilde{A}_0} \right) = \Lambda^T \otimes \text{id}^{A_1} \left( \phi_{+}^{\tilde{A}_1 A_1} \right)$$

where  $\Lambda^T$  is a CP unital map obtained from  $\Lambda^{A_0 \rightarrow A_1}$  by taking the transpose on the Kraus operators in an operator sum representation of  $\Lambda^{A_0 \rightarrow A_1}$ . Therefore, its dual  $(\Lambda^T)^* \equiv \bar{\Lambda}$  is a CPTP map. With these notations we get:

$$\begin{aligned} f_{\mathcal{C}^A}(\Psi^{A_0 \rightarrow A_1}) &= \sup_{\Lambda \in \mathcal{C}^A} \text{Tr} \left[ \Lambda^T \otimes \text{id}^{A_1} \left( \phi_{+}^{\tilde{A}_1 A_1} \right) J_{\Psi}^A \right] \\ &= \sup_{\Lambda \in \mathcal{C}^A} \text{Tr} \left[ \phi_{+}^{\tilde{A}_1 A_1} \left( \bar{\Lambda} \otimes \text{id}^{A_1} (J_{\Psi}^A) \right) \right] \\ &= \sup_{\Lambda \in \mathcal{C}^A} \text{Tr} \left[ \phi_{+}^{\tilde{A}_1 A_1} \left( \Lambda \otimes \text{id}^{A_1} (J_{\Psi}^A) \right) \right] \\ &= 2^{-H_{\min}(A_1|A_0)_{J_{\Psi}}} , \end{aligned} \quad (37)$$

where the third equality follows from the fact that optimization over all  $\bar{\Lambda}^{A_0 \rightarrow \tilde{A}_1}$  is equivalent to optimization over all  $\Lambda^{A_0 \rightarrow \tilde{A}_1}$ , and the last equality follows from (34). We therefore conclude that the conditional min-entropy can be viewed as the support function of the set of quantum channels, while the min-entropy can be viewed as the support function of the set of quantum states. Note that  $J_{\Psi}^A$  in the above equation is not normalized since its marginal  $J_{\Psi}^{A_0} = I^{A_0}$ .

## III. NOISY SUPERCHANNELS

We study here different types of superchannels that correspond to noisy processes. We will use these models in the next section for the definition of the entropy of a quantum channel, and particularly for the definition of the extended conditional min-entropy of a bipartite channel. Similar to noisy channels, for which we defined in the previous section both random unitaries and doubly stochastic (unital) channels, we define here different types of noisy superchannels, including random-unitary superchannels, completely uniformity preserving superchannels, completely unital-channel preserving, and doubly stochastic superchannels.

### A. Random Unitary Superchannels

If  $\Theta \in \mathbb{L}^{AB}$  is a reversible map, with  $A \cong B$ , i.e.  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ , then the entropy of a quantum channel must be defined in such a way that any channel  $\Phi^A$  has the same entropy as  $\Theta[\Phi^A]$ . Such a reversible transformation  $\Theta$  has the form

$$\Theta[\Phi^A] = \mathcal{U}_{\text{post}}^{A_1 \rightarrow B_1} \circ \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre}}^{B_0 \rightarrow A_0}, \quad (38)$$

where  $\mathcal{U}_{\text{pre}}^{\tilde{A}_0 \rightarrow A_0}$  and  $\mathcal{U}_{\text{post}}^{A_1 \rightarrow \tilde{A}_1}$  are unitary CPTP maps (i.e. acting by a conjugation with a unitary matrix). If the dimensions of systems  $A$  and  $B$  are not the same, then one can replace the pre and post unitaries above with isometries. Therefore, similar to the arguments given in the previous section, a convex combination of reversible superchannels is an entropy non-decreasing map since it corresponds to implementing a reversible transformation and then “forgetting” the information about which reversible transformation has been applied. We call such a convex combination of reversible superchannels, a *random unitary superchannel*. It can be expressed as a linear map  $\Theta \in \mathbb{L}^{AB}$  (with  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ ) given by

$$\Theta[\Phi^A] = \sum_{x=1}^m p_x \mathcal{U}_{\text{post},x}^{A_1 \rightarrow B_1} \circ \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}, \quad (39)$$

where  $\{p_x\}_{x=1}^m$  is a probability distribution.

From its definition in (8), the Choi matrix of the above map is given by:

$$\mathbf{J}_{\Theta}^{AB} = \sum_{x=1}^m p_x \sum_{a,b} J_{\mathcal{E}_{ab}}^A \otimes J_{\mathcal{U}_{\text{post},x} \circ \mathcal{E}_{ab} \circ \mathcal{U}_{\text{pre},x}}^B$$

where  $a = (i, j)$ ,  $b = (k, \ell)$ ,  $J_{\mathcal{E}_{ab}}^A = |i\rangle\langle j|^{A_0} \otimes |k\rangle\langle \ell|^{A_1}$  and

$$\begin{aligned} J_{\mathcal{U}_{\text{post},x} \circ \mathcal{E}_{ab} \circ \mathcal{U}_{\text{pre},x}}^B &= \\ \sum_{i',j'} \langle i | \mathcal{U}_{\text{pre},x} (|i'\rangle\langle j'|^{B_0}) | j \rangle & |i'\rangle\langle j'|^{B_0} \otimes \mathcal{U}_{\text{post},x} (|k\rangle\langle \ell|^{A_1}) \\ &= \mathcal{U}_{\text{pre},x}^t (|i\rangle\langle j|^{B_0}) \otimes \mathcal{U}_{\text{post},x} (|k\rangle\langle \ell|^{A_1}), \end{aligned} \quad (40)$$

where  $\mathcal{U}_{\text{pre},x}^t(\rho) := (\mathcal{U}_{\text{pre},x}^*(\rho^T))^T = U_{\text{pre},x}^T \rho (U_{\text{pre},x}^T)^*$ . Combining all this, we conclude that

$$\mathbf{J}_{\Theta}^{AB} = \sum_{x=1}^m p_x |\alpha_x\rangle\langle \alpha_x|^{A_0 B_0} \otimes |\beta_x\rangle\langle \beta_x|^{A_1 B_1} \quad (41)$$

where

$$\begin{aligned} |\alpha_x\rangle^{A_0 B_0} &:= I^{A_1} \otimes U_{\text{pre},x}^T |\phi_+\rangle^{A_0 B_0} \\ |\beta_x\rangle^{A_1 B_1} &:= I^{A_1} \otimes U_{\text{post},x} |\phi_+\rangle^{A_1 B_1} \end{aligned} \quad (42)$$

are all (unnormalized) maximally entangled states. In particular, the marginals of the above Choi matrix satisfies:

$$\begin{aligned} \mathbf{J}_{\Theta}^{A_1 B} &= u^{B_0} \otimes \mathbf{J}_{\Theta}^{A_1 B_1} \\ \mathbf{J}_{\Theta}^{A B_1} &= u^{A_0} \otimes \mathbf{J}_{\Theta}^{A_1 B_1} \\ \mathbf{J}_{\Theta}^{A_0 B} &= \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1} \\ \mathbf{J}_{\Theta}^{A B_0} &= \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{A_1} \end{aligned} \quad (43)$$

along with

$$\mathbf{J}_{\Theta}^{A_1 B_0} = I^{A_1 B_0} \quad \text{and} \quad \mathbf{J}_{\Theta}^{A_0 B_1} = I^{A_0 B_1}.$$

Note that the last condition in (43) is satisfied for *any* superchannel. Moreover, even if a Choi matrix satisfies the 4 conditions above, it is still not sufficient to guarantee the random unitary form of (41). For example, if we replace the unitary maps  $\mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}$  and  $\mathcal{U}_{\text{post},x}^{A_1 \rightarrow B_1}$  with unital CPTP maps, the conditions in (43) will still hold! In the following subsections we give physical meanings to the different conditions in (43).

## B. Doubly Stochastic Superchannels

A CPTP map  $\Phi : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H}')$  is doubly stochastic if both  $\Phi$  and  $\Phi^*$  are CPTP maps. This is possible only if  $\dim(\mathcal{H}) = \dim(\mathcal{H}')$ . For superchannels, we introduce a similar definition:

**Definition 2.** A linear map  $\Theta \in \mathbb{L}^{AB}$  is said to be doubly stochastic if both  $\Theta$  and its dual  $\Theta^* \in \mathbb{L}^{BA}$  are superchannels.

*Remark 1.* The input and output dimensions of a doubly stochastic superchannel must satisfy

$$\frac{d_{A_0}}{d_{A_1}} = \frac{d_{B_0}}{d_{B_1}}. \quad (44)$$

This follows from the fact that the traces of the Choi matrices of both  $\Theta$  and  $\Theta^*$  are the same (see the theorem below).

Random unitary superchannels are doubly stochastic and in particular satisfy trivially the above condition with  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ . However, doubly stochastic superchannels form a much larger set of operations as can be seen in the following theorem.

**Theorem 2.** Let  $\Theta$  be a superchannel in  $\mathbb{L}^{AB}$  with dimensions as in (44). The following are equivalent.

1. The dual map  $\Theta^* \in \mathbb{L}^{BA}$  is also a superchannel (i.e.  $\Theta$  is doubly stochastic).
2. The Choi matrix  $\mathbf{J}_{\Theta}^{AB} \geq 0$  with marginals:

$$\mathbf{J}_{\Theta}^{A_1 B_0} = I^{A_1 B_0} \quad ; \quad \mathbf{J}_{\Theta}^{A B_0} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{A_1} \quad (45)$$

$$\mathbf{J}_{\Theta}^{A_0 B_1} = I^{A_0 B_1} \quad ; \quad \mathbf{J}_{\Theta}^{A_0 B} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1}. \quad (46)$$

3. The superchannel  $\Theta$  can be realized as in (21) with the quantum channels  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E}$  and  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1}$  satisfying:

$$\Gamma_{\text{post}}^{A_1 E \rightarrow B_1} (u^{A_1} \otimes \rho^E) = \text{Tr} [\rho^E] u^{B_1} \quad (47)$$

$$\Gamma_{\text{pre}}^{B_0 \rightarrow A_0} (u^{B_0}) = u^{A_0} \quad (48)$$

where  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0}$  is the map  $\text{Tr}_E \circ \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E}$ .

*Remark 2.* The conditions in (45) are equivalent to those in (19), as they ensure that  $\Theta$  is a superchannel. The additional conditions are those in (46). Moreover, from the first condition in (45) we have  $\text{Tr}[\mathbf{J}_{\Theta}^{AB}] = d_{A_1} d_{B_0}$  while from the first condition in (46) we have  $\text{Tr}[\mathbf{J}_{\Theta}^{AB}] = d_{A_0} d_{B_1}$ . Hence, the dimensions of doubly stochastic superchannels must satisfy (44).

*Proof.* We first prove the equivalence of 1 and 2. From Theorem 1 the map  $\Theta^*$  is a superchannel if and only if  $\mathbf{J}_{\Theta^*}^{BA} \geq 0$  and it has marginals

$$\mathbf{J}_{\Theta^*}^{B_1 A_0} = I^{B_1 A_0} \quad ; \quad \mathbf{J}_{\Theta^*}^{B A_0} = \mathbf{J}_{\Theta^*}^{B_0 A_0} \otimes u^{B_1} \quad (49)$$

Now, since  $\Delta_{\Theta^*}^{\tilde{B} \rightarrow A} = \Delta_{\Theta}^{* \tilde{B} \rightarrow A}$  (see the discussion in the preliminary section) we conclude that the Choi matrix of  $\Theta^* \in \mathbb{L}^{BA}$  is given by

$$\begin{aligned} \mathbf{J}_{\Theta^*}^{BA} &= \text{id}^B \otimes \Delta_{\Theta^*}^{\tilde{B} \rightarrow A} \left( \phi_+^{B_0 \tilde{B}_0} \otimes \phi_+^{B_1 \tilde{B}_1} \right) \\ &= \text{id}^B \otimes \Delta_{\Theta}^{* \tilde{B} \rightarrow A} \left( \phi_+^{B_0 \tilde{B}_0} \otimes \phi_+^{B_1 \tilde{B}_1} \right) \\ &= \overline{\Delta_{\Theta}^{\tilde{A} \rightarrow B} \otimes \text{id}^A \left( \phi_+^{\tilde{A}_0 A_0} \otimes \phi_+^{\tilde{A}_1 A_1} \right)} = \overline{\mathbf{J}_{\Theta}^{BA}} \end{aligned} \quad (50)$$

where the bar above matrices indicates complex conjugation on the components of the matrix (without performing the transpose). Here  $\mathbf{J}_{\Theta^*}^{BA} = \text{swap}(\mathbf{J}_{\Theta}^{AB})$  is obtained from  $\mathbf{J}_{\Theta}^{AB}$  by swapping between systems  $A$  and  $B$ . Combining this with (49) we get the marginals in (46). This completes the proof of the equivalence between 1 and 2.

To prove part 3, note that if  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1}$  has the form (47) then from (25) it follows that

$$\mathbf{J}_{\Theta}^{A_0 B} = \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} \left( \psi^{A_0 B_0 E} \otimes I^{\tilde{A}_1} \right) = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1}$$

since  $\psi^{A_0 B_0 E}$  is a purification of  $\frac{1}{d_{A_1}} \mathbf{J}_{\Theta}^{A_0 B_0}$ . Conversely, suppose  $\mathbf{J}_{\Theta}^{A_0 B} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1}$ , and let  $d_E = \text{Rank}(\mathbf{J}_{\Theta}^{A_0 B_0})$ . Then, we can write  $|\psi\rangle^{A_0 B_0 E} = \sum_{j=1}^{d_E} r_j |u_j\rangle^{A_0 B_0} |j\rangle^E$  with  $r_j > 0$  and  $\{|u_j\rangle^{A_0 B_0}\}$  being an orthonormal set of vectors in  $\mathcal{H}^{A_0 B_0}$ . From its definition, the marginal state  $\sum_{j=1}^{d_E} r_j^2 |u_j\rangle\langle u_j|^{A_0 B_0} = \frac{1}{d_{A_1}} \mathbf{J}_{\Theta}^{A_0 B_0}$ . Therefore, from our assumption on the form of  $\mathbf{J}_{\Theta}^{A_0 B}$ , it follows that (recall that  $d_{A_1} = d_{B_1}$ )

$$\begin{aligned} & \sum_{j=1}^{d_E} r_j^2 |u_j\rangle\langle u_j|^{A_0 B_0} \otimes I^{B_1} \\ &= \sum_{j,k=1}^{d_E} r_j r_k |u_j\rangle\langle u_k|^{A_0 B_0} \otimes \Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} (|j\rangle\langle k|^E \otimes I^{\tilde{A}_1}) . \end{aligned}$$

Finally, from the independence of  $\{|u_j\rangle\langle u_k|^{A_0 B_0}\}$  it follows that

$$\Gamma_{\text{post}}^{\tilde{A}_1 E \rightarrow B_1} (|j\rangle\langle k|^E \otimes u^{\tilde{A}_1}) = \delta_{jk} u^{B_1} .$$

This proves the equivalence of  $\mathbf{J}_{\Theta}^{A_0 B} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1}$  with the condition that  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1} (u^{A_1} \otimes \rho^E) = \text{Tr}[\rho^E] u^{B_1}$  for all density matrices  $\rho^E$ . To complete the proof of the equivalence between 2 and 3, we need to show the equivalence between the condition  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0}(u^{B_0}) = u^{A_0}$  and  $\mathbf{J}_{\Theta^*}^{B_1 A_0} = I^{B_1 A_0}$ . The latter condition is equivalent to  $\Gamma_{\Theta}^{A_1 B_0 \rightarrow A_0 B_1}$  being a unital channel. Hence, (recall,  $d_{A_0} d_{B_1} = d_{A_1} d_{B_0}$ )

$$\begin{aligned} u^{A_0 B_1} &= \Gamma_{\Theta}^{A_1 B_0 \rightarrow A_0 B_1} (u^{A_1 B_0}) \\ &= \left( \text{id}^{A_0} \otimes \Gamma_{\text{post}}^{A_1 E \rightarrow B_1} \right) \left( u^{A_1} \otimes \Gamma_{\text{pre}}^{B_0 \rightarrow A_0 E} (u^{B_0}) \right) \\ &= u^{B_1} \otimes \Gamma_{\text{pre}}^{B_0 \rightarrow A_0} (u^{B_0}) , \end{aligned} \tag{51}$$

where in the last equality we used the property that  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1} (u^{A_1} \otimes \rho^E) = \text{Tr}[\rho^E] u^{B_1}$ . Hence,  $\Gamma_{\text{pre}}^{B_0 \rightarrow A_0} (u^{B_0}) = u^{A_0}$ . This completes the proof.  $\blacksquare$

Note that a doubly stochastic superchannels satisfies the last two conditions in (43) but not necessarily the first two conditions nor does it necessarily have the form (41). We now discuss how the condition (46) is related to the fact that doubly stochastic superchannels do not decrease noise.

### C. Completely Uniformity Preserving Superchannel

The noisiest quantum channel in  $\mathcal{L}^A$ , which we denote by  $\mathcal{N}^A$  and call the uniform channel, is given by

$$\mathcal{N}^A(X) := \text{Tr}[X] u^{A_1} \quad \forall X \in \mathcal{B}(\mathcal{H}^{A_0}) . \tag{52}$$

That is, irrespective of the input state, the output state of a uniform channel is always maximally mixed (i.e. uniform).

With this in mind, a superchannel  $\Theta \in \mathbb{L}^{AB}$  with dimensions  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$  (i.e.  $A \cong B$ ), that does not decrease noise must satisfy

$$\Theta[\mathcal{N}^A] = \mathcal{N}^B .$$

That is,  $\Theta$  is a uniformity preserving superchannel. The above condition is analogous to the condition that noisy channel  $\mathcal{E} \in \mathcal{C}^A$  must preserve the uniform (maximally mixed) state  $u^A \equiv \frac{1}{d_A} I^A$ , since otherwise, the output state  $\mathcal{E}(u^A)$  will be less noisy than the input state (the maximally mixed state).

The Choi matrix of  $\mathcal{N}^A$  and  $\mathcal{N}^B$  are given by  $I^{A_0} \otimes u^{A_1}$  and  $I^{B_0} \otimes u^{B_1}$ , respectively. Hence, from (10) the equation above becomes equivalent to

$$\mathbf{J}_{\Theta}^B = I^B .$$

Unlike the parallel discussion on noisy channels, for noisy superchannels, there is a stronger condition than the one above, that one can expect from any superchannel that does not decrease noise.



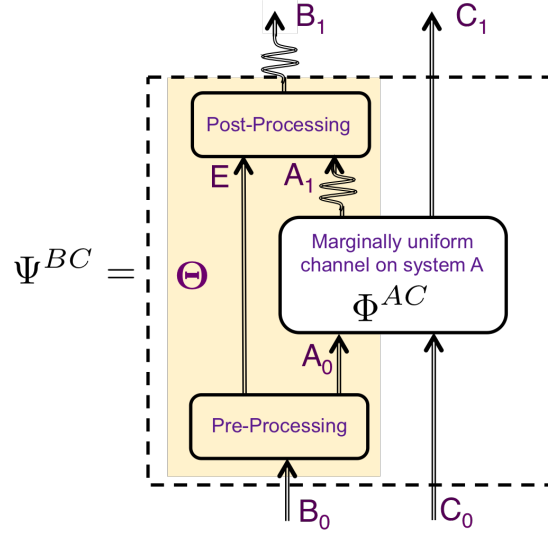


FIG. 4: The action of a superchannel  $\Theta \in \mathbb{L}^{AB}$  on a marginally uniform channel  $\Phi^{AC}$ . If  $\Theta$  is completely uniformity preserving then  $\Psi^{BC}$  is marginally uniform on  $B$ . The curly arrows indicate maximally mixed states.

Consider a bipartite channel  $\Phi^{AC}$  shared between two parties, and suppose that it has the property that the output state on subsystem  $A_1$  is always the maximally mixed state. That is,

$$\Phi^{AC}(\rho^{A_0 C_0}) = u^{A_1} \otimes \sigma^{C_1} \quad \forall \rho^{A_0 C_0} \in \mathcal{B}(\mathcal{H}^{A_0 C_0}),$$

where  $\sigma^{C_1} := \text{Tr}_{A_1} [\Phi^{AC}(\rho^{A_0 C_0})]$  is some matrix in  $\mathcal{B}(\mathcal{H}^{C_1})$ . We say that such a bipartite channel is *marginally uniform* on  $A$ . Therefore, if  $\Theta \in \mathbb{L}^{AB}$  (with dimensions  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ ) does not decrease noise, and  $\Phi^{AC}$  is marginally uniform on  $A$ , then  $\Psi^{BC} := \Theta \otimes \mathbb{1}^C [\Phi^{AC}]$  should also be marginally uniform (on  $B$ ); see Fig. (4). We call such a superchannel a *completely uniformity preserving* superchannel. Note that this condition is somewhat similar to the condition that physical operations are not only positive but completely positive.

**Theorem 3.** *Let  $\Theta$  be a superchannel in  $\mathbb{L}^{AB}$  with dimensions  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ . The following are equivalent.*

1.  $\Theta$  is a completely uniformity preserving superchannel.
2. In addition to (19), the Choi matrix of  $\Theta$  satisfies:

$$\mathbf{J}_{\Theta}^{A_0 B} = \mathbf{J}_{\Theta}^{A_0 B_0} \otimes u^{B_1}. \quad (53)$$

3. The map  $\Theta$  can be realized as in (21) with the additional condition that the quantum channel  $\Gamma_{\text{post}}^{A_1 E \rightarrow B_1}$  satisfies: for all  $\rho^E \in \mathcal{B}(\mathcal{H}^E)$

$$\Gamma_{\text{post}}^{A_1 E \rightarrow B_1} (u^{A_1} \otimes \rho^E) = \text{Tr}[\rho^E] u^{B_1}. \quad (54)$$

*Proof.* Suppose  $\Theta$  is a completely uniformity preserving superchannel. Let  $\Phi^{AC}$  be a marginally uniform channel on  $A$  and define

$$\Psi^{BC} := \Theta \otimes \mathbb{1}^C [\Phi^{AC}].$$

From our assumptions both  $\Phi^{AC}$  and  $\Psi^{BC}$  are marginally uniform, so that their Choi matrices are given by

$$J_{\Phi}^{AC} = J_{\Phi}^{A_0 C} \otimes u^{A_1} \quad \text{and} \quad J_{\Psi}^{BC} = J_{\Psi}^{B_0 C} \otimes u^{B_1},$$

respectively. Taking  $\Theta$  in (10) to be  $\Theta \otimes \mathbb{1}^C$  we have

$$J_{\Psi}^{B_0 C} \otimes u^{B_1} = \text{Tr}_{A\tilde{C}} \left[ \mathbf{J}_{\Theta \otimes \mathbb{1}^C}^{ABC\tilde{C}} \left( \left( J_{\Phi}^{A_0 \tilde{C}} \right)^T \otimes u^{A_1} \otimes I^B \right) \right]$$

From its definition,

$$\mathbf{J}_{\Theta \otimes \mathbb{1}^C}^{ABC\tilde{C}} = \mathbf{J}_{\Theta}^{AB} \otimes \phi_+^{C\tilde{C}}.$$

Moreover, note that

$$\text{Tr}_{\tilde{C}} \left[ \left( I^{A_0} \otimes \phi_+^{C\tilde{C}} \right) \left( \left( J_{\Phi}^{A_0\tilde{C}} \right)^T \otimes I^C \right) \right] = \left( J_{\Phi}^{A_0C} \right)^{T_{A_0}}$$

where  $T_{A_0}$  is the partial transpose on system  $A_0$ . We therefore get that

$$\begin{aligned} & J_{\Psi}^{B_0C} \otimes u^{B_1} \\ &= \text{Tr}_A \left[ \left( \mathbf{J}_{\Theta}^{AB} \otimes I^C \right) \left( \left( J_{\Phi}^{A_0C} \right)^{T_{A_0}} \otimes u^{A_1} \otimes I^B \right) \right] \\ &= \frac{1}{d_{A_1}} \text{Tr}_{A_0} \left[ \left( \mathbf{J}_{\Theta}^{A_0B} \otimes I^C \right) \left( \left( J_{\Phi}^{A_0C} \right)^{T_{A_0}} \otimes I^B \right) \right]. \end{aligned}$$

By multiplying both sides of the equation above by a traceless matrix  $Z^{B_1}$  and taking the trace over  $B_1$  we get that

$$\text{Tr}_{A_0B_1} \left[ \left( \mathbf{J}_{\Theta}^{A_0B} \otimes I^C \right) \left( \left( J_{\Phi}^{A_0C} \right)^{T_{A_0}} \otimes I^{B_0} \otimes Z^{B_1} \right) \right] = 0. \quad (55)$$

We show now that if the condition above holds for all traceless matrices  $Z^{B_1}$  and all positive semidefinite matrices  $M^{A_0C} \equiv J_{\Phi}^{A_0C}$  with the property that  $M^{R_0B_0} = I^{R_0B_0}$  then  $\mathbf{J}_{\Theta}^{A_0B} = \mathbf{J}_{\Theta}^{A_0B_0} \otimes u^{B_1}$ . To do that, let  $\{X_j^{B_1}\}_{j=0}^{d_{B_1}^2-1}$  be an orthogonal basis of  $\mathcal{B}(\mathcal{H}^{B_1})$  such that  $X_0^{B_1} = u^{B_1}$  and  $\text{Tr}[X_j^{B_1}] = 0$  for  $j > 0$ . With this basis we can express  $\mathbf{J}_{\Theta}^{A_0B}$  as

$$\mathbf{J}_{\Theta}^{A_0B} = \mathbf{J}_{\Theta}^{A_0B_0} \otimes u^{B_1} + \sum_{j=1}^{d_{B_1}^2-1} N_j^{A_0B_0} \otimes X_j^{B_1},$$

where  $N_j^{A_0B_0}$  are some matrices in  $\mathcal{B}(\mathcal{H}^{A_0B_0})$ . Substituting this expression into (55) and taking  $Z^{B_1} = X_j^{B_1}$  for some  $j > 0$  we conclude that

$$\text{Tr}_{A_0} \left[ \left( N_j^{A_0B_0} \otimes I^C \right) \left( \left( M^{A_0C} \right)^{T_{A_0}} \otimes I^{B_0} \right) \right] = 0$$

for all positive semidefinite matrices with  $M^{A_0C}$  with marginal  $M^{R_0B_0} = I^{R_0B_0}$ . Note also that the above equation has to hold for all dimensions of system  $C$ . If we take the dimensions of system  $C$  to be  $d_{C_0} = 1$  and  $d_C = d_{C_1} = d_{A_0}$ , then we can choose  $M^{A_0C} = \phi_+^{A_0C}$ . With this choice the equation above become:

$$0 = \text{Tr}_{A_0} \left[ \left( N_j^{A_0B_0} \otimes I^C \right) \left( \left( \phi_+^{A_0C} \right)^{T_{A_0}} \otimes I^{B_0} \right) \right] = N_j^{CB_0}$$

where  $N_j^{CB_0}$  is a copy of  $N_j^{A_0B_0}$  in  $\mathcal{B}(\mathcal{H}^{CB_0})$  (recall that  $d_C = d_{A_0}$ ). We therefore conclude that  $N_j^{A_0B_0} = 0$  for all  $j > 0$  so that  $\mathbf{J}_{\Theta}^{A_0B} = \mathbf{J}_{\Theta}^{A_0B_0} \otimes u^{B_1}$ . Note that the converse of this argument also holds. That is, following the above lines backwards we conclude that if  $\mathbf{J}_{\Theta}^{A_0B} = \mathbf{J}_{\Theta}^{A_0B_0} \otimes u^{B_1}$  then  $\Theta$  is a completely uniformity preserving superchannel.

In Theorem 2 we proved the equivalence of  $\mathbf{J}_{\Theta}^{A_0B} = \mathbf{J}_{\Theta}^{A_0B_0} \otimes u^{B_1}$  with the condition that  $\Gamma_{\text{post}}^{A_1E \rightarrow B_1}(u^{A_1} \otimes \rho^E) = u^{B_1}$  for all density matrices  $\rho^E$ . This also provides the proof of Part 3 here. ■

A comparison between Theorem 2 and Theorem 3 demonstrates that doubly stochastic superchannels are completely uniformity preserving. However, the converse is not true in general since completely uniformity preserving superchannels are not necessarily having  $\mathbf{J}_{\Theta}^{A_0B_1} = I^{A_0B_1}$ . The latter is equivalent to the condition that  $\Gamma_{\Theta}^{A_1B_0 \rightarrow A_0B_1}$  is unital, and is related to another physical condition that we discuss in the following subsection.

### D. Completely Unital-Channel Preserving Superchannels

In the preliminary section we discussed that random unitary channels can be viewed as noisy channels; i.e. channels that always increase noise no matter what the input state is. On the other hand, in any noise model, a superchannel that increase noise, should not generate a non-noisy channel from a noisy one. Therefore, in addition for being completely uniformity preserving, noisy superchannels should at least not convert random unitary channels to non-unital channels. Here we study superchannels with this property, and particularly those that are completely unital-channel preserving. Since we will consider unital channels we will assume here that  $d_{A_0} = d_{A_1}$  and  $d_{B_0} = d_{B_1}$ .

We first show that if a superchannel takes random unitary channels to unital channels then it also takes unital channels to unital channels. To see why, note that from (10), the superchannel  $\Theta$  converts random unitary channels to unital channels if and only if

$$I^{B_1} = \text{Tr}_{AB_0} \left[ \mathbf{J}_{\Theta}^{AB} \left( (J_{\Psi}^A)^T \otimes I^B \right) \right] \quad (56)$$

for any random unitary channel  $\Psi^A$ . Now, suppose  $\Psi^A$  is a unital channel. In [31] it was shown that it can be expressed as an affine linear combination of unitary channels; i.e.  $\Psi^A = \sum_j r_j \mathcal{U}_j^A$ , where  $r_j \in \mathbb{R}$  and  $\sum_j r_j = 1$ . Since for each  $\mathcal{U}_j$  the above equation holds, from its linearity it also holds for  $\Psi^A$ . We therefore conclude that the above equation holds for any unital channel which implies that  $\Theta$  is unital-channel preserving.

The condition in (56) holds in particular for the completely dephasing channel whose Choi matrix is given by  $I^{A_0} \otimes u^{A_1}$ . From the linearity, any matrix of the form  $X^A \equiv J_{\Psi}^A - I^{A_0} \otimes u^{A_1}$  with  $\Psi^A$  being unital satisfies

$$\text{Tr}_A \left[ \mathbf{J}_{\Theta}^{AB_1} \left( (X^A)^T \otimes I^{B_1} \right) \right] = 0.$$

Note that  $X^A$  is any Hermitian matrix in  $\mathcal{B}(\mathcal{H}^{A_0 A_1})$  with zero marginals  $X^{A_0} = 0$  and  $X^{A_1} = 0$ . This means that  $\mathbf{J}_{\Theta}^{AB_1}$  is orthogonal (in the Hilbert Schmidt inner product) to any Hermitian matrix in  $\mathcal{B}(\mathcal{H}^{AB_1})$  of the form  $T_1^{A_0} \otimes T_2^{A_1} \otimes Y^{B_1}$ , where  $T_1^{A_0}$  and  $T_2^{A_1}$  are arbitrary traceless Hermitian matrices and  $Y^{B_1}$  is an arbitrary Hermitian matrix. Therefore,  $\mathbf{J}_{\Theta}^{AB_1}$  must be a linear combination of matrices of the form  $I^{A_0} \otimes \alpha^{A_1 B_1}$  and  $I^{A_1} \otimes \beta^{A_0 B_1}$ , where  $\alpha^{A_1 B_1}$  and  $\beta^{A_0 B_1}$  are Hermitian matrices. We therefore conclude that the superchannel  $\Theta$  is a unital-channel preserving if and only if it's Choi matrix  $\mathbf{J}_{\Theta}^{AB}$  satisfies the superchannel condition (19) and in addition

$$\mathbf{J}_{\Theta}^{AB_1} = I^{A_0} \otimes \alpha^{A_1 B_1} + I^{A_1} \otimes \beta^{A_0 B_1}.$$

This last condition is somewhat cumbersome, but in the following theorem we show that  $\beta^{A_0 B_1}$  must be zero if  $\Theta$  is a completely unital-channel preserving superchannel. That is,  $\Theta$  is a superchannel such that for any system  $C$ , the superchannel  $\Theta \otimes \mathbb{1}^C$  is unital-channel preserving.

**Theorem 4.** *Let  $\Theta \in \mathbb{L}^{AB}$  be a superchannel with  $d_{A_0} = d_{A_1}$  and  $d_{B_0} = d_{B_1}$ . Then, the following are equivalent:*

1.  $\Theta$  is completely unital-channel preserving.
2. The Choi matrix  $\mathbf{J}_{\Theta}^{AB}$  has marginals (in addition to those in Eq. (19))

$$\mathbf{J}_{\Theta}^{AB_1} = u^{A_0} \otimes \mathbf{J}_{\Theta}^{A_1 B_1} \quad \text{and} \quad \mathbf{J}_{\Theta}^{A_0 B_1} = I^{A_0 B_1}.$$

3. The CPTP map  $\Gamma_{\Theta}^{A_1 B_0 \rightarrow A_0 B_1}$  is unital, and in addition to (31) it satisfies for any density matrix  $\rho^{A_1} \in \mathcal{D}(\mathcal{H}^{A_1})$

$$\begin{aligned} & \Gamma_{\Theta}^{A_1 B_0 \rightarrow A_0 B_1} (\rho^{A_1} \otimes u^{B_0}) \\ &= u^{A_0} \otimes \Gamma_{\Theta}^{A_1 B_0 \rightarrow B_1} (\rho^{A_1} \otimes u^{B_0}). \end{aligned} \quad (57)$$

*Proof.* We first prove the equivalence of 1 and 2. Suppose  $\Theta$  is completely unital-channel preserving. Then, from (10) it follows that the relation  $\Phi^{BC} = \Theta \otimes \mathbb{1}^C [\Psi^{AC}]$  can be expressed in the Choi form as:

$$J_{\Phi}^{BC} = \text{Tr}_A \left[ \left( \mathbf{J}_{\Theta}^{AB} \otimes I^C \right) \left( (J_{\Psi}^{AC})^{T_A} \otimes I^B \right) \right] \quad (58)$$

Suppose now that  $\Psi^{AC}$  is a bipartite unital channel. Then, from our assumption,  $\Phi^{BC}$  is unital as well. Hence,

$$I^{B_1 C_1} = \text{Tr}_A \left[ \left( \mathbf{J}_{\Theta}^{AB_1} \otimes I^{C_1} \right) \left( (J_{\Psi}^{AC_1})^{T_A} \otimes I^{B_1} \right) \right]$$

Note that the equation above holds for  $J_{\Psi}^{AC_1} = u^{A_0} \otimes I^{A_1 C_1}$ , which corresponds to the marginal of the Choi matrix of the completely dephasing map (which is unital). Hence, for this choice of  $J_{\Psi}^{AC_1}$  we get the condition that  $\mathbf{J}_{\Theta}^{B_1} = d_{A_0} I^{B_1}$ .

Next, note that for any unital bipartite channel  $\Psi^{AC}$ , the matrix  $X^{AC_1} = \left(J_{\Psi}^{AC_1}\right)^{T_A} - u^{A_0} \otimes I^{A_1 C_1}$  has the property that  $X^{A_1 C_1} = 0$  and  $X^{A_0} = 0$ . From the linearity of the equation above we conclude that for any such matrix

$$\text{Tr}_A \left[ \left( \mathbf{J}_{\Theta}^{AB_1} \otimes I^{C_1} \right) (X^{AC_1} \otimes I^{B_1}) \right] = 0$$

Let  $Z^{C_1}$  be some fixed normalized (in the Hilbert-Schmidt inner product) *traceless* Hermitian matrix in  $\mathcal{B}_h(\mathcal{H}^{C_1})$ , and let  $Y^{B_1}$  be arbitrary Hermitian matrix in  $\mathcal{B}_h(\mathcal{H}^{B_1})$ . Then, the equation above implies that

$$\left\langle \mathbf{J}_{\Theta}^{AB_1} \otimes Z^{C_1}, X^{AC_1} \otimes Y^{B_1} \right\rangle = 0$$

for all such  $Y^{B_1}$  and all  $X^{AC_1}$  with zero marginals as above. This implies that  $\mathbf{J}_{\Theta}^{AB_1} \otimes Z^{C_1}$  is orthogonal to any matrix of the form  $T_1^{A_0} \otimes T_2^{A_1 C_1} \otimes Y^{B_1}$  where  $T_1^{A_0}$  and  $T_2^{A_1 C_1}$  are arbitrary Hermitian traceless matrices. Therefore, there must exist matrices  $\alpha^{A_0 B_1}$  and  $\beta^{A_1 B_1 C_1}$  such that

$$\mathbf{J}_{\Theta}^{AB_1} \otimes Z^{C_1} = \alpha^{A_0 B_1} \otimes I^{A_1 C_1} + I^{A_0} \otimes \beta^{A_1 B_1 C_1}.$$

Finally, by multiplying both sides of the equation above by  $I^{AB_1} \otimes Z^{C_1}$  and taking the partial trace over system  $C_1$  we conclude that

$$\mathbf{J}_{\Theta}^{AB_1} = I^{A_0} \otimes \text{Tr}_{C_1} \left[ (I^{AB_1} \otimes Z^{C_1}) \beta^{A_1 B_1 C_1} \right],$$

where we used the fact that  $Z^{C_1}$  is a normalized traceless matrix. The equation above implies that

$$\mathbf{J}_{\Theta}^{AB_1} = u^{A_0} \otimes \mathbf{J}_{\Theta}^{A_1 B_1}.$$

Conversely, suppose the equation above holds, and in addition  $\mathbf{J}_{\Theta}^{B_1} = d_{A_0} I^{B_1}$ . Then, for any bipartite unital channel  $\Psi^{AC}$  we get

$$\begin{aligned} & \text{Tr}_A \left[ \left( \mathbf{J}_{\Theta}^{AB_1} \otimes I^{C_1} \right) \left( \left( J_{\Psi}^{AC_1} \right)^{T_A} \otimes I^{B_1} \right) \right] \\ &= \frac{1}{d_{A_0}} \text{Tr}_{A_1} \left[ \left( \mathbf{J}_{\Theta}^{A_1 B_1} \otimes I^{C_1} \right) \left( \left( J_{\Psi}^{A_1 C_1} \right)^{T_A} \otimes I^{B_1} \right) \right] \\ &= \frac{1}{d_{A_0}} \mathbf{J}_{\Theta}^{B_1} \otimes I^{C_1} = I^{B_1 C_1}. \end{aligned}$$

This completes the proof of the equivalence between 1 and 2. We now prove the equivalence between 2 and 3.

From (29) it follows that

$$\mathbf{J}_{\Theta}^{AB_1} = \Gamma_{\Theta}^{\tilde{A}_1 \tilde{B}_0 \rightarrow A_0 B_1} \left( \phi_{+}^{A_1 \tilde{A}_1} \otimes I^{\tilde{B}_0} \right).$$

In particular, the marginal is given by

$$\mathbf{J}_{\Theta}^{A_1 B_1} = \Gamma_{\Theta}^{\tilde{A}_1 \tilde{B}_0 \rightarrow B_1} \left( \phi_{+}^{A_1 \tilde{A}_1} \otimes I^{\tilde{B}_0} \right).$$

Hence,  $\mathbf{J}_{\Theta}^{AB_1} = u^{A_0} \otimes \mathbf{J}_{\Theta}^{A_1 B_1}$  if and only if (57) holds for all density matrices  $\rho^{A_1}$ . This completes the proof. ■

#### IV. THE ENTROPY OF A QUANTUM CHANNEL

We now extend the definition of entropies, and particularly the conditional min-entropy, from states to channels. Since the entropy of a quantum channel measures how noisy the channel is, they must behave monotonically under noisy operations. Below we give an axiomatic and minimalistic approach for the definition of entropy.

We call a function  $f : \mathcal{L}_{+}^A \rightarrow \mathbb{R}$  an entropy if it satisfies the following conditions (see the analogous conditions on entropy of states in the preliminary section)

1. *Monotonicity*: For any random unitary superchannel  $\Theta : \mathcal{L}^A \rightarrow \mathcal{L}^A$ ,

$$f(\Theta[\Phi]) \geq f(\Phi) ,$$

for all channels  $\Phi \in \mathcal{C}^A$ .

2. *Additivity*: For any two quantum channels  $\Phi^A \in \mathcal{C}^A$  and  $\Psi^B \in \mathcal{C}^B$  we have

$$f(\Phi^A \otimes \Psi^B) = f(\Phi^A) + f(\Psi^B) .$$

3. *Normalization*: on a uniform channel,  $\mathcal{N}^A$ , as in (52),  $f(\mathcal{N}^A) = \log(d_{A_1})$ , and on any replacement map,  $\Phi \in \mathcal{L}^A$ , of the form  $\Phi(X) = \text{Tr}[X]|\psi\rangle\langle\psi|$  with  $|\psi\rangle$  being some fixed pure state,  $f(\Phi^A) = 0$ .

The last condition is motivated by the fact that replacement maps can be viewed as quantum states and consequently the entropy of these channels should reduce to the entropy of states. For the monotonicity, we only require monotonicity under random unitary superchannels, however, we expect many entropy functions to be monotonic under a larger set of superchannels such as, for example, doubly stochastic superchannels. Regarding the additivity requirement, while it is a natural condition (since entropies of states are required to be additive), some natural candidates, as we discuss now, fail to satisfy this requirement.

The von-Neumann entropy of states is defined by  $S(\rho) := -\text{Tr}[\rho \log \rho]$  for any density matrix  $\rho \in \mathcal{B}(\mathcal{H})$ . For quantum channels, one can propose a natural generalization,  $\tilde{S}$ , given by the minimum entropy output:

$$\tilde{S}[\Phi^A] := \min_{\rho^{A_0} \in \mathcal{D}(\mathcal{H}^{A_0})} S(\Phi^A(\rho^{A_0}))$$

where the minimum is over all input density matrices  $\rho^{A_0}$ . This candidate for an entropy of channel is monotonic under random unitary superchannels. Indeed, let  $\Theta$  be a random unitary superchannel as in (39). Then,

$$\begin{aligned} & \tilde{S}[\Theta[\Phi^A]] \\ &= \min_{\rho^{A_0}} S \left( \sum_{x=1}^m p_x \mathcal{U}_{\text{post},x}^{A_1 \rightarrow B_1} \circ \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}(\rho^{A_0}) \right) \\ &\geq \min_{\rho^{A_0}} \sum_{x=1}^m p_x S \left( \mathcal{U}_{\text{post},x}^{A_1 \rightarrow B_1} \circ \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}(\rho^{A_0}) \right) \\ &= \min_{\rho^{A_0}} \sum_{x=1}^m p_x S \left( \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}(\rho^{A_0}) \right) \\ &\geq \sum_{x=1}^m p_x \min_{\rho^{A_0}} S \left( \Phi^{A_0 \rightarrow A_1} \circ \mathcal{U}_{\text{pre},x}^{B_0 \rightarrow A_0}(\rho^{A_0}) \right) = \tilde{S}[\Phi^A] , \end{aligned}$$

where in the first inequality we used the concavity of the von-Neumann entropy. Since the Rényi entropies with parameter  $\alpha \in [0, 1]$  are also concave, their extension to channels as above will also be monotonic under random unitary superchannels.

While the minimum entropy output of a quantum channel is monotonic under random unitary superchannels, it fails to satisfy the additivity property [32], and therefore, according to our definition above it is not an entropy function. One may choose to replace the additivity condition with a weaker one, in which  $f$  is only required to be additive under tensor product of two replacement maps (i.e. states). With this modification, the minimum entropy output of a quantum channel is an entropy function. However, we include the full additivity property in the definition of an entropy function, as the functions that we will consider here will be fully additive. We now give an example of such an entropy function that is based on the min-entropy.

**Definition 3.** Let  $\Phi^A \in \mathcal{L}^A$  be a quantum channel. The extended min-entropy of  $\Phi^A$  is the function

$$H_{\min}^{\text{ext}}(A)_{\Phi} := H_{\min}(A_1|A_0)_{J_{\Phi}/d_{A_0}} ,$$

where  $J_{\Phi}^{A_0 A_1}/d_{A_0}$  is the normalized Choi matrix of  $\Phi^{A_0 \rightarrow A_1}$ .

The *extended min-entropy* is an entropy function that satisfies all the 3 conditions above of monotonicity, additivity, and normalization. The additivity and the normalization follows immediately from the properties of the conditional

min entropy. To show monotonicity, we show now that the extended min entropy behaves monotonically not only under random unitary superchannels, but also under the much larger set of doubly stochastic superchannels.

Let  $\Theta \in \mathbb{L}^{AB}$  be a doubly stochastic superchannel and suppose  $d_{A_0} = d_{B_0}$  and  $d_{A_1} = d_{B_1}$ . Then, for any quantum channel  $\Phi^A \in \mathcal{L}^A$  we have

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B)_{\Theta[\Phi]}} &= 2^{-H_{\min}(B_1|B_0)_{J_{\Theta[\Phi]}/d_{B_0}}} \\ &= \frac{1}{d_{B_0}} \max_{\Psi \in \mathcal{C}^B} \langle \Psi^B, \Theta[\Phi^A] \rangle \\ &= \frac{1}{d_{A_0}} \max_{\Psi \in \mathcal{C}^B} \langle \Theta^*[\Psi^B], \Phi^A \rangle \\ &\leq \frac{1}{d_{A_0}} \max_{\Lambda \in \mathcal{C}^A} \langle \Lambda^A, \Phi^A \rangle \\ &= 2^{-H_{\min}^{\text{ext}}(A)_{\Phi}}. \end{aligned} \quad (59)$$

That is, for any doubly stochastic superchannel  $\Theta \in \mathbb{L}^{AB}$  as above (with  $A \cong B$ ), and any CPTP map  $\Phi^A \in \mathcal{C}^A$ ,

$$H_{\min}^{\text{ext}}(B)_{\Theta[\Phi]} \geq H_{\min}^{\text{ext}}(A)_{\Phi}.$$

This completes the proof that the extended min-entropy is an entropy function.

### A. The extended conditional min-entropy

We extend here the definition of the conditional min-entropy to quantum channels. This function will play a key role in our results on the comparison of quantum channels and we therefore devote the rest of this section to study it along with its properties, and its operational interpretations.

We consider here a bipartite quantum channel  $\Omega^{AB} \equiv \Omega^{A_0 B_0 \rightarrow A_1 B_1} : \mathcal{B}(\mathcal{H}^{A_0 B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1})$  and denote by

$$\omega^{AB} := \frac{1}{d_{A_0} d_{B_0}} \text{id}^{A_0 B_0} \otimes \Omega^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1 B_1} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{B_0 \tilde{B}_0} \right)$$

its (normalized) Choi matrix, where  $\phi_+^{A_0 \tilde{A}_0}$  and  $\phi_+^{B_0 \tilde{B}_0}$  are unnormalized maximally entangled states.

**Definition 4.** *The extended conditional min entropy of a bipartite channel  $\Omega^{AB}$  as above, is the function:*

$$\begin{aligned} H_{\min}^{\text{ext}}(B|A)_{\Omega} &:= -\log_2 \min \text{Tr}[\gamma^{AB_0}] \\ \text{subject to: } &\mathbf{1.} \quad \gamma^{AB_0} \otimes I^{B_1} \geq \omega^{AB} \\ &\mathbf{2.} \quad \gamma^{A_0 B_0} = u^{A_0} \otimes \gamma^{B_0} \end{aligned} \quad (60)$$

The above optimization problem is SDP, and therefore can be solved efficiently and algorithmically using standard techniques. Furthermore, since any SDP problem has a dual problem, the extended conditional min entropy can be expressed as (see SM for details):

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} &= d_{A_0} \max \text{Tr}[\alpha^{AB} \omega^{AB}] \\ \text{subject to: } &\mathbf{1.} \quad \alpha^{AB_0} = \alpha^{A_0 B_0} \otimes u^{A_1} \\ &\mathbf{2.} \quad \alpha^{A_1 B_0} = I^{A_1 B_0} \\ &\mathbf{3.} \quad \alpha^{AB} \geq 0. \end{aligned} \quad (61)$$

Note that the matrix  $\alpha^{AB}$  above satisfies precisely the conditions that a Choi matrix of a superchannel satisfies. We can therefore identify each such  $\alpha^{AB}$  with a Choi matrix  $\mathbf{J}_{\Theta}^{AB}$  of some superchannel  $\Theta$ . Denoting the set of superchannels by  $\mathbb{S}^{AB} \subset \mathbb{L}^{AB}$  we get that

$$d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} = \max_{\Theta \in \mathbb{S}^{AB}} \text{Tr}[\mathbf{J}_{\Theta}^{AB} J_{\Omega}^{AB}] \quad (62)$$

Note that the trace on the RHS can be expressed as  $\text{Tr}[\mathbf{J}_{\Theta}^{AB} J_{\Omega}^{AB}] = \langle \Theta, \Theta' \rangle$ , where  $\Theta'$  is a CPP map in  $\mathbb{L}^{AB}$  (but not necessarily a superchannel) that corresponds to the bipartite channel  $\Omega^{AB}$  via the relation  $\Omega^{AB} = \mathbb{1}^A \otimes \Theta'[\Phi_+^{A\tilde{A}}]$

(see (12) for a discussion on this relation). Therefore, the RHS of (62) can be viewed as the support function of superchannels.

A similar definition to (62) was studied in [33] for min/max entropies of more general objects known as quantum combs. However, in [33] the entropies were viewed as a function of general operators like the comb itself, and here we define the extended-conditional min-entropy as a function of bipartite channels, and not as a function of superchannels. Even the normalization (see the factor of  $d_{B_0}$  in (62)) is not arbitrary. This distinction will become clearer when we study the properties of the extended conditional min-entropy.

Moreover, recall that  $\mathbf{J}_{\Theta}^{AB}$  can be viewed as the Choi matrix of the bipartite channel  $\Lambda_{\Theta}^{AB}$  as defined in (12). Therefore, alternatively,

$$d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} = \max_{\Theta \in \mathbb{S}^{AB}} \langle \Lambda_{\Theta}^{AB}, \Omega^{AB} \rangle, \quad (63)$$

where we replaced the inner product between Choi matrices to inner product between the corresponding channels as defined in (2). Now, recall that  $\Theta \mapsto \Lambda_{\Theta}^{AB} = (\mathbb{1}^A \otimes \Theta) [\Phi_+^{A\bar{A}}]$  (see (13)) is an isomorphism. Substituting this expression of  $\Lambda_{\Theta}^{AB}$  into (63), we get that the extended conditional min-entropy can be expressed in the following form:

$$\begin{aligned} d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} &= \max_{\Theta \in \mathbb{S}^{AB}} \left\langle (\mathbb{1}^A \otimes \Theta) [\Phi_+^{A\bar{A}}], \Omega^{AB} \right\rangle \\ &= \max_{\Theta \in \mathbb{S}^{AB}} \left\langle (\Theta^T \otimes \mathbb{1}^B) [\Phi_+^{\bar{B}B}], \Omega^{AB} \right\rangle \\ &= \max_{\Theta \in \mathbb{S}^{AB}} \left\langle \Phi_+^{\bar{B}B}, ((\Theta^T)^* \otimes \mathbb{1}^B) [\Omega^{AB}] \right\rangle \\ &= \max_{\Theta \in \mathbb{S}^{AB}} \left\langle \Phi_+^{\bar{B}B}, (\Theta \otimes \mathbb{1}^B) [\Omega^{AB}] \right\rangle \\ &= \max_{\Theta \in \mathbb{S}^{AB}} \left\langle \phi_+^{\bar{B}_1 B_1} \middle| (\Theta \otimes \mathbb{1}^B) [\Omega^{AB}] (\phi_+^{\bar{B}_0 B_0}) \middle| \phi_+^{\bar{B}_1 B_1} \right\rangle \end{aligned} \quad (64)$$

where  $(\Theta \otimes \mathbb{1}^B) [\Omega^{AB}]$  is a quantum channel from  $\mathcal{B}(\mathcal{H}^{\bar{B}_0 B_0})$  to  $\mathcal{B}(\mathcal{H}^{\bar{B}_1 B_1})$ . We used the fact that  $\Theta$  is a superchannel iff  $(\Theta^T)^*$  is a superchannel. To see why, note that from (15) and (50) it follows that  $\mathbf{J}_{(\Theta^T)^*}^{AB} = \overline{\mathbf{J}_{\Theta}^{AB}}$ , and therefore  $\mathbf{J}_{(\Theta^T)^*}^{AB}$  satisfies the conditions (19) of a superchannel. For the last equality we used the definition (2) of an inner product between channels, and took an orthonormal basis  $\{X_a\}_{a=1}^{d_{B_0}^A}$  of  $\mathcal{B}(\mathcal{H}^{\bar{B}_0 B_0})$ , whose first element  $X_1 = \frac{1}{d_{B_0}} \phi_+^{B_0 \bar{B}_0}$  so that  $\Phi_+^{\bar{B}B}(X_a) = 0$  unless  $a = 1$ . Note that the expression above for the extended conditional min-entropy is reminiscent to the one given in (34) for the conditional min-entropy.

## B. Properties of the extended conditional min entropy

The extended conditional min entropy provides a generalization for the conditional min entropy. The following theorem demonstrates it by showing that many of the properties of the conditional min-entropy carry over to the extended conditional min-entropy.

**Theorem 5.** Consider a quantum channel  $\Omega^{AB} : \mathcal{B}(\mathcal{H}^{A_0 B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1})$ , and denote its normalized Choi matrix by  $\omega^{AB}$  as in Definition 4.

1. **Generalization of conditional min entropy:** If  $\Omega^{AB}$  is a replacement channel (i.e.  $\Omega^{AB}(\rho^{A_0 B_0}) = \omega^{A_1 B_1}$  for all density matrices  $\rho^{A_0 B_0}$ ), then

$$H_{\min}^{\text{ext}}(B|A)_{\Omega} = H_{\min}(B_1|A_1)_{\omega}.$$

2. **Independence:** If

$$\Omega^{AB} = \Psi^A \otimes \Phi^B, \quad (65)$$

where  $\Psi^A$  and  $\Phi^B$  are local channels, then

$$H_{\min}^{\text{ext}}(B|A)_{\Omega} = H_{\min}^{\text{ext}}(B)_{\Phi}$$

is independent of  $\Psi^A$ .

3. **Additivity:** Consider a second quantum channel  $\Gamma^{A'B'} : \mathcal{B}(\mathcal{H}^{A_0 B_0'}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1'})$ . Then,

$$H_{\min}^{\text{ext}}(BB'|AA')_{\Omega \otimes \Gamma} = H_{\min}^{\text{ext}}(B|A)_{\Omega} + H_{\min}^{\text{ext}}(B'|A')_{\Gamma}$$

4. **Monotonicity:** For any superchannel  $\Theta' : \mathcal{L}^A \rightarrow \mathcal{L}^C$

$$H_{\min}^{\text{ext}}(B|A)_{\Omega^{AB}} \leq H_{\min}^{\text{ext}}(B|C)_{(\Theta' \otimes \mathbb{1}^B)[\Omega^{AB}]} \quad (66)$$

5. **Conditioning:** Consider a tripartite quantum channel  $\Omega^{ABC} : \mathcal{B}(\mathcal{H}^{A_0 B_0 C_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1 C_1})$ . Then,

$$H_{\min}^{\text{ext}}(B|AC)_{\Omega} \leq H_{\min}^{\text{ext}}(B|A)_{\Omega} \quad (67)$$

6. **Bounds:** Upper bound:

$$H_{\min}^{\text{ext}}(B|A)_{\Omega} \leq H_{\min}(B_1|A_1)_{\omega} . \quad (68)$$

Lower bound:

$$H_{\min}^{\text{ext}}(B|A)_{\Omega} \geq H_{\min}(AB_1|B_0)_{\omega} - \log(d_A) \quad (69)$$

where  $d_A \equiv d_{A_0} d_{A_1}$ .

*Remark 3.* The conditioning property involves the marginal bipartite channel  $\Omega^{AB}$  on the RHS of (67). This marginal channel is obtained from the tripartite channel  $\Omega^{ABC}$  by inputting a fixed state  $\gamma^{C_0}$  into the input of system C and tracing out system  $C_1$ . That is,

$$\Omega^{AB}(\rho^{A_0 B_0}) := \text{Tr}_{C_1} [\Omega^{ABC}(\rho^{A_0 B_0} \otimes \gamma^{C_0})]$$

The theorem above states that the inequality in (67) holds for all density matrices  $\gamma^{C_0}$ .

*Proof. Part 1.* The condition  $\gamma^{AB_0} \otimes I^{B_1} \geq \omega^{AB}$  in (60) implies that  $\gamma^{A_1} \otimes I^{B_1} \geq \omega^{A_1 B_1}$ . Therefore, since  $\text{Tr}[\gamma^{AB_0}] = \text{Tr}[\gamma^{A_1}]$  we always have

$$2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} \geq 2^{-H_{\min}(B_1|A_1)_{\omega}} . \quad (70)$$

On the other hand, since  $\Omega$  is a replacement map, its normalized Choi matrix is given by  $\omega^{AB} = u^{A_0 B_0} \otimes \omega^{A_1 B_1}$ . Let  $\gamma^{A_1}$  be an optimal positive semidefinite matrix that satisfies  $\gamma^{A_1} \otimes I^{B_1} \geq \omega^{A_1 B_1}$  such that  $\text{Tr}[\gamma^{A_1}] = 2^{-H_{\min}^{\text{ext}}(B_1|A_1)_{\omega}}$ . Define  $\gamma^{AB_0} := u^{A_0 B_0} \otimes \gamma^{A_1}$ . It is easy to check that this  $\gamma^{AB_0}$  satisfies the two conditions in (60). We therefore get that

$$2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}} \leq 2^{-H_{\min}(B_1|A_1)_{\omega}} .$$

Hence, we must have  $H_{\min}^{\text{ext}}(B|A)_{\Omega} = H_{\min}(B_1|A_1)_{\omega}$ .

*Part 2.* From (65) it follows that the normalized Choi matrix of  $\Omega^{AB}$  can be decomposed as  $\omega^{AB} = \omega^A \otimes \omega^B$ , where  $\omega^A$  and  $\omega^B$  are the normalized Choi matrices of  $\Psi^A$  and  $\Phi^B$ , respectively. In this case, the two conditions of (60) take the form:

$$\gamma^{AB_0} \otimes I^{B_1} \geq \omega^A \otimes \omega^B \quad (71)$$

$$\gamma^{A_0 B_0} = u^{A_0} \otimes \gamma^{B_0} \quad (72)$$

Tracing out system  $A_1$  on both sides of (71) gives

$$\gamma^{A_0 B_0} \otimes I^{B_1} \geq u^{A_0} \otimes \omega^B ,$$

and when combined with (72) yields

$$\gamma^{B_0} \otimes I^{B_1} \geq \omega^B . \quad (73)$$

Note that  $\text{Tr}[\gamma^{AB_0}] = \text{Tr}[\gamma^{B_0}]$  so that we must have

$$2^{-H_{\min}^{\text{ext}}(B|A)} \geq 2^{-H_{\min}(B_1|B_0)} .$$



On the other hand, for the choice  $\gamma^{AB_0} = \omega^A \otimes \gamma^{B_0}$  with optimal  $\gamma^{B_0}$  (i.e.  $\gamma^{B_0}$  satisfies (73) and  $\text{Tr}[\gamma^{B_0}] = 2^{-H_{\min}(B_1|B_0)_\omega}$ ) we obtain the other side of the inequality. We therefore conclude that:

$$H_{\min}^{\text{ext}}(B|A)_\Omega = H_{\min}(B_1|B_0)_\omega \equiv H_{\min}^{\text{ext}}(B)_\Phi .$$

*Part 3.* Denote the Choi matrix of  $\Gamma^{A'B'}$  by  $\omega^{A'B'}$ . From Definition 4:

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(BB'|AA')_{\Omega \otimes \Gamma}} &= \min \text{Tr}[\gamma^{AA'B_0B'_0}] \\ \text{subject to: } &\mathbf{1.} \quad \gamma^{AA'B_0B'_0} \otimes I^{B_1B'_1} \geq \omega^{AB} \otimes \omega^{A'B'} \\ &\mathbf{2.} \quad \gamma^{A_0A'_0B_0B'_0} = u^{A_0A'_0} \otimes \gamma^{B_0B'_0} \end{aligned}$$

Note that the above two conditions follows from the following 5 conditions:

$$\begin{aligned} \gamma^{AA'B_0B'_0} &= \gamma^{AB_0} \otimes \gamma^{A'B'_0} \\ \gamma^{AB_0} \otimes I^{B_1} &\geq \omega^{AB} \quad , \quad \gamma^{A'B'_0} \otimes I^{B'_1} \geq \omega^{A'B'} \\ \gamma^{A_0B_0} &= u^{A_0} \otimes \gamma^{B_0} \quad , \quad \gamma^{A'_0B'_0} = u^{A'_0} \otimes \gamma^{B'_0} \end{aligned} \tag{74}$$

Therefore,  $2^{-H_{\min}^{\text{ext}}(BB'|AA')_\Omega} \geq \min \text{Tr}[\gamma^{AA'B_0B'_0}]$ , where the minimization is over all  $\gamma^{AA'B_0B'_0}$  that satisfies the 5 conditions in (74). That is,

$$2^{-H_{\min}^{\text{ext}}(BB'|AA')_{\Omega \otimes \Gamma}} \geq 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} 2^{-H_{\min}^{\text{ext}}(B'|A')_\Gamma} .$$

To prove the converse, consider the dual expression (61) for the extended conditional min entropy. We get that:

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(BB'|AA')_\Omega} &= d_{A_0} d_{A'_0} \max \text{Tr} \left[ \alpha^{ABA'B'} \omega^{AB} \otimes \omega^{A'B'} \right] \\ \text{subject to: } &\mathbf{1.} \quad \alpha^{AA'B_0B'_0} = \alpha^{A_0A'_0B_0B'_0} \otimes u^{A_1A'_1} \\ &\mathbf{2.} \quad \alpha^{A_1A'_1B_0B'_0} = I^{A_1A'_1B_0B'_0} \\ &\mathbf{3.} \quad \alpha^{AA'BB'} \geq 0 . \end{aligned} \tag{75}$$

Similarly to the previous argument, the 3 conditions above follows from the following conditions:

$$\begin{aligned} \alpha^{ABA'B'} &= \alpha_1^{AB} \otimes \alpha_2^{A'B'} \quad , \quad \alpha_1^{AB} \geq 0 \quad , \quad \alpha_2^{A'B'} \geq 0 \\ \alpha_1^{AB_0} &= \alpha^{A_0B_0} \otimes u^{A_1} \quad , \quad \alpha_2^{A'B'_0} = \alpha^{A'_0B'_0} \otimes u^{A'_1} \\ \alpha_1^{A_1B_0} &= I^{A_1B_0} \quad , \quad \alpha_2^{A'_1B'_0} = I^{A'_1B'_0} \end{aligned}$$

Therefore, if we replace the 3 conditions in (75) with the above conditions we get that

$$2^{-H_{\min}^{\text{ext}}(BB'|AA')_{\Omega \otimes \Gamma}} \leq 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} 2^{-H_{\min}^{\text{ext}}(B'|A')_\Gamma} .$$

This complete the proof of Part 3.

*Part 4.* From (64) we get that the extended conditional min-entropy can be expressed in the following form:

$$d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} = \max_{\Theta \in \mathbb{S}^{A\tilde{B}}} \left\langle \Phi_+^{\tilde{B}B}, (\Theta \otimes \mathbb{1}^B) [\Omega^{AB}] \right\rangle$$

where  $(\Theta \otimes \mathbb{1}^B) [\Omega^{AB}]$  is a quantum channel from  $\mathcal{B}(\mathcal{H}^{\tilde{B}_0B_0})$  to  $\mathcal{B}(\mathcal{H}^{\tilde{B}_1B_1})$ . Let  $\Theta' : \mathcal{L}^A \rightarrow \mathcal{L}^C$  be a superchannel. Then,

$$\begin{aligned} &d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|C)_{\Theta' \otimes \mathbb{1}^B} [\Omega^{AB}]} \\ &= \max_{\Theta \in \mathbb{S}^{C\tilde{B}}} \left\langle \Phi_+^{\tilde{B}B}, (\Theta \circ \Theta' \otimes \mathbb{1}^B) [\Omega^{AB}] \right\rangle \\ &\leq \max_{\Theta'' \in \mathbb{S}^{A\tilde{B}}} \left\langle \Phi_+^{\tilde{B}B}, (\Theta'' \otimes \mathbb{1}^B) [\Omega^{AB}] \right\rangle \\ &= d_{B_0} 2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega^{AB}}} . \end{aligned}$$

Therefore,

$$H_{\min}^{\text{ext}}(B|A)_{\Omega^{AB}} \leq H_{\min}^{\text{ext}}(B|C)_{(\Theta' \otimes \mathbb{1}^B)[\Omega^{AB}]} \quad (76)$$

for any superchannel  $\Theta' : \mathcal{L}^A \rightarrow \mathcal{L}^C$ .

*Part 5.* Let  $\Theta_\gamma : \mathcal{L}^{AC} \rightarrow \mathcal{L}^A$  be the superchannel defined by: for all  $\Phi^{AC} \in \mathcal{L}^{AC}$  and  $\rho^{A_0} \in \mathcal{B}(\mathcal{H}^{A_0})$

$$\Theta_\gamma[\Phi^{AC}](\rho^{A_0}) := \text{Tr}_{C_1} [\Phi^{AC}(\rho^{A_0} \otimes \gamma^{C_0})] .$$

It is straightforward to see that  $\Theta_\gamma$  is a superchannel if  $\gamma^{C_0}$  is a density matrix. Let

$$\Omega_\gamma^{AB} := (\Theta_\gamma \otimes \mathbb{1}^B) [\Omega^{ABC}] .$$

In particular,

$$\Omega_\gamma^{AB}(\rho^{A_0 B_0}) = \text{Tr}_{C_1} [\Omega^{ABC}(\rho^{A_0 B_0} \otimes \gamma^{C_0})] .$$

Hence, from Part 4 we get that for any density matrix  $\gamma^{C_0} \in \mathcal{B}(\mathcal{H}^{C_0})$

$$\begin{aligned} H_{\min}^{\text{ext}}(B|AC)_{\Omega^{ABC}} &\leq H_{\min}^{\text{ext}}(B|A)_{(\Theta_\gamma \otimes \mathbb{1}^B)[\Omega^{ABC}]} \\ &= H_{\min}^{\text{ext}}(B|A)_{\Omega_\gamma^{AB}} . \end{aligned}$$

This completes the proof of Part 5.

*Part 6.* The upper bound follows from (70). For the lower bound, if we add to the two conditions in (60) a third condition that  $\gamma^{AB_0} = \gamma^{A_0 B_0} \otimes u^{A_1}$  we get that

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &\leq \min \{ \text{Tr} [\gamma^{B_0}] : I^{AB_1} \otimes \gamma^{B_0} \geq d_A \omega^{AB} \} \\ &= d_A 2^{-H_{\min}(AB_1|B_0)_\omega} . \end{aligned} \quad (77)$$

This completes the proof. ■

The properties above demonstrate that the extended conditional min entropy indeed quantifies the uncertainty about one dynamical system conditioned on another. Particularly, note that property 5 is consistent with the intuition that the uncertainty (i.e. entropy) about system  $B$  increases if the system one has access to (i.e. system  $A$ ) undergoes a physical evolution.

### C. Operational interpretation as a guessing probability

The conditional min entropy  $H(A_1|A_0)$  has an operational interpretation as a guessing probability when system  $A_1$  is classical. Here we show that a similar interpretation can be made for the extended conditional min entropy if system  $B$  is classical. Since system  $B_1$  is classical, for all  $\rho^{A_0 B_0} \in \mathcal{B}(\mathcal{H}^{A_0 B_0})$

$$\Omega^{A_0 B_0 \rightarrow A_1 B_1}(\rho^{A_0 B_0}) = \sum_{x=1}^{d_{B_1}} \Omega_x^{A_0 B_0 \rightarrow A_1}(\rho^{A_0 B_0}) \otimes |x\rangle\langle x|^{B_1} , \quad (78)$$

where  $\{\Omega_x^{A_0 B_0 \rightarrow A_1}\}$  form a quantum instrument. Moreover, since system  $B_0$  is classical, we denote

$$\Omega_{x|y}^{A_0 \rightarrow A_1}(\rho) := \Omega_x^{A_0 B_0 \rightarrow A_1}(\rho \otimes |y\rangle\langle y|) \quad \forall \rho \in \mathcal{B}(\mathcal{H}^{A_0}) \quad (79)$$

where for each  $y$  the set  $\{\Omega_{x|y}^{A_0 \rightarrow A_1}\}_{x=1}^{d_{B_1}}$  form a quantum instrument.

In Fig. 5 we describe a strategy for Alice to guess Bob's outcome  $x$  if Bob's input is  $y$ . In this strategy, Alice sends through her share of the channel  $\Omega^{AB}$  one part of a possibly entangled state  $|\eta^{A_0 A_2}\rangle$ . At her output of the channel she measures the joint system  $A_1 A_2$ . The outcome of the measurement is Alice's guess of Bob's output value. The maximum probability that Alice guesses Bob's outcome correctly, given that Bob's input is  $y$  can be expressed as:

$$\begin{aligned} P_{\text{guess}}^{(y)}(\Omega^{AB}) &:= \\ \max \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_x \left( \Omega_{x|y}^{A_0 \rightarrow A_1} \otimes \text{id}^{A_2} (|\eta\rangle\langle\eta|^{A_0 A_2}) \right) \right] \end{aligned} \quad (80)$$

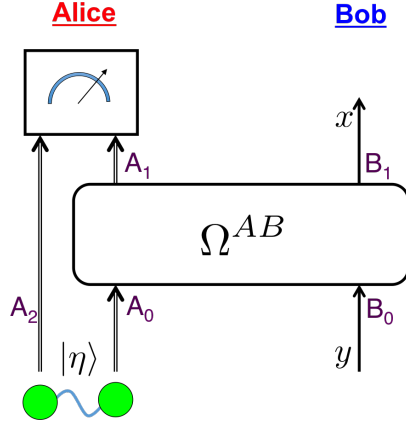


FIG. 5: A strategy for Alice to guess Bob's output.

where the maximum is over all POVMs  $\{P_x\}_{x=1}^{d_{B_1}}$  on system  $A_1A_2$ , and over all bipartite pure state  $|\eta\rangle^{A_0A_2}$  on system  $A_0A_2$ . Note that replacing  $|\eta\rangle^{A_0A_2}$  with a mixed state will not increase the optimal guessing probability, and furthermore, we can assume w.l.o.g. that  $\dim(\mathcal{H}^{A_2}) \leq \dim(\mathcal{H}^{A_0})$ . Finally, we define the guessing probability of a quantum-classical channel as:

$$P_{\text{guess}}(\Omega^{AB}) := \frac{1}{d_{B_0}} \sum_{y=1}^{d_{B_0}} P_{\text{guess}}^{(y)}(\Omega^{AB}) .$$

The above expression can be interpreted as the maximum probability that Alice can guess correctly the value  $x$  of Bob's system  $B_1$  if Bob's input  $y$  (which is known to Alice) is chosen at random according to a uniform distribution.

**Theorem 6.** Let  $\Omega^{AB}$  be a quantum channel as above with a classical system  $B$ . Then,

$$P_{\text{guess}}(\Omega^{AB}) = 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega}$$

*Proof.* Following the same notations as in (78) and (79), since  $B$  is classical, the (normalized) Choi matrix of  $\Omega^{AB}$  can be expressed as:

$$\omega^{AB} = \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \omega_{x|y}^A \otimes |y\rangle\langle y|^{B_0} \otimes |x\rangle\langle x|^{B_1} ,$$

with

$$\omega_{x|y}^A := \frac{1}{d_{A_0} d_{B_0}} \text{id}^{A_0} \otimes \Omega_{x|y}^{\tilde{A}_0 \rightarrow A_1} \left( \phi_+^{A_0 \tilde{A}_0} \right) . \quad (81)$$

Consequently, from (61), with  $\omega^{AB}$  as above and  $\alpha^{AB} \equiv \sum_{x,y} \alpha_{x|y}^A \otimes |y\rangle\langle y|^{B_0} \otimes |x\rangle\langle x|^{B_1}$ , we get

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &= d_{A_0} \max \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \text{Tr}[\alpha_{x|y}^A \omega_{x|y}^A] \\ \text{subject to: } &\sum_{x=1}^{d_{B_1}} \alpha_{x|y}^A = \eta_y^{A_0} \otimes I^{A_1} , \quad \text{Tr}[\eta_y^{A_0}] = 1 \\ &\alpha_{x|y}^A \geq 0 \\ &\text{for all } 1 \leq x \leq d_{B_1} \text{ and } 1 \leq y \leq d_{B_0} \end{aligned} \quad (82)$$

where we denoted by

$$\eta_y^{A_0} := \frac{1}{d_{A_1}} \sum_{x=1}^{d_{B_1}} \alpha_{x|y}^{A_0} .$$

Note that we can assume w.l.o.g. that  $\eta_y^{A_0}$  is full rank. Hence, we can define for each  $y$  the following POVM on system  $A$ :

$$P_{x|y}^A := \left( (\eta_y^{A_0})^{-\frac{1}{2}} \otimes I^{A_1} \right) \alpha_{x|y}^A \left( (\eta_y^{A_0})^{-\frac{1}{2}} \otimes I^{A_1} \right)$$

Note that  $P_{x|y}^A \geq 0$  and  $\sum_x P_{x|y}^A = I^A$ . With this notation

$$\begin{aligned} d_{A_0} \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \text{Tr}[\alpha_{x|y}^A \omega_{x|y}^A] &= d_{A_0} \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \\ \text{Tr} \left[ P_{x|y}^A \left( \sqrt{\eta_y^{A_0}} \otimes I^{A_1} \right) \omega_{x|y}^A \left( \sqrt{\eta_y^{A_0}} \otimes I^{A_1} \right) \right] &= \\ \frac{1}{d_{B_0}} \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_{x|y}^A \left( \text{id}^{A_0} \otimes \Omega_{x|y}^{\tilde{A}_0 \rightarrow A_1} \left( |\eta_y\rangle\langle\eta_y|^{A_0 \tilde{A}_0} \right) \right) \right] \end{aligned} \quad (83)$$

where we used (81), and the state  $|\eta_y\rangle^{A_0 \tilde{A}_0}$  is the purification of the (normalized) state  $\eta_y^{A_0}$ . That is,

$$|\eta_y\rangle^{A_0 \tilde{A}_0} := \left( \sqrt{\eta_y^{A_0}} \otimes I^{\tilde{A}_0} \right) |\phi_+^{A_0 \tilde{A}_0}\rangle.$$

We therefore conclude that

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &= \frac{1}{d_{B_0}} \max \sum_{y=1}^{d_{B_0}} \sum_{x=1}^{d_{B_1}} \\ \text{Tr} \left[ P_{x|y}^A \left( \text{id}^{A_0} \otimes \Omega_{x|y}^{\tilde{A}_0 \rightarrow A_1} \left( |\eta_y\rangle\langle\eta_y|^{A_0 \tilde{A}_0} \right) \right) \right] \\ \text{subject to: } P_{x|y}^A &\geq 0, \quad \sum_{x=1}^{d_{B_1}} P_{x|y}^A = I^A, \quad \langle\eta_y|\eta_y\rangle = 1 \\ \forall x = 1, \dots, d_{B_1} \text{ and } y = 1, \dots, d_{B_0}. \end{aligned} \quad (84)$$

This completes the proof. ■

The theorem above provides the first operational interpretation for the conditional min-entropy of a bipartite quantum channel when  $B$  is classical. Consider now the special case in which the CP and trace non-increasing maps  $\{\Omega_x^{A_0 B_0 \rightarrow A_1}\}$ , as defined above, are of the form

$$\Omega_x^{A_0 B_0 \rightarrow A_1} = p_x \Lambda_{(x)}^{A_0 B_0 \rightarrow A_1}$$

where  $\{p_x\}$  form a probability distribution, and each  $\Lambda_{(x)}^{A_0 B_0 \rightarrow A_1}$  is a CPTP. For a given fixed value  $y$  at the input system of  $B_0$ , we denote by  $\Lambda_{(x|y)}^{A_0 \rightarrow A_1}$  the corresponding CPTP map on Alice's side. In this special case, the guessing probability (80), can be interpreted as the maximum possible probability to guess which channel  $\left\{ \Lambda_{(x|y)}^{A_0 \rightarrow A_1} \right\}_x$  Alice holds. Hence, the theorem above, when applied to this case, implies that  $2^{-H_{\min}^{\text{ext}}(B|A)_\Omega}$  can be interpreted as the maximum probability to guess which channel Alice holds, out of  $\left\{ \Lambda_{(x|y)}^{A_0 \rightarrow A_1} \right\}_x$ , where  $y$  is chosen from a uniform distribution.

In Appendix B we provide another operational interpretation (for the extended conditional min-entropy) in the case that only system  $B_1$  is classical, while systems  $A_0, A_1$  and  $B_0$  are all quantum. In this case, the extended conditional min-entropy still can be expressed as the optimal probability to guess correctly the value of  $B_1$ . In this case, however, system  $B_0$  can be entangled with Alice's systems as described in Fig. 6.

## V. COMPARISON OF QUANTUM CHANNELS

Here we consider one of the main problems discussed in the introduction: given a collection of channels  $\Psi_j^A \in \mathcal{C}^A$  and  $\Phi_j^B \in \mathcal{C}^B$ , with  $j = 1, \dots, n$ , is there a superchannel  $\Theta : \mathcal{L}^A \rightarrow \mathcal{L}^B$  such that for all  $j = 1, \dots, n$

$$\Phi_j^B = \Theta[\Psi_j^A] ? \quad (85)$$

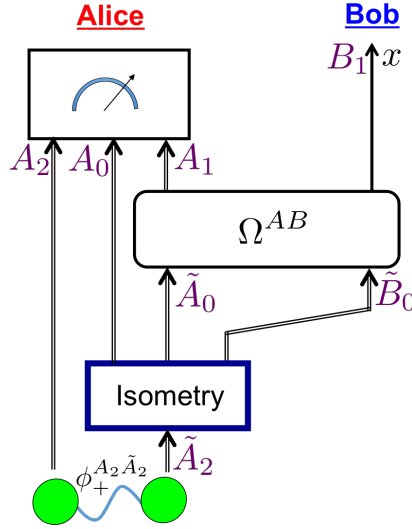


FIG. 6: A restrictive strategy for Alice to guess Bob's output. Alice has access to a maximally entangled state  $\Phi_{+}^{A_2 \tilde{A}_2}/d_{A_2}$  (with  $d_{A_2} = d_{B_0}$ ). Alice uses the maximally entangled state and send her system  $\tilde{A}_2$  through any (non-local) quantum channel with outputs  $A_0$ ,  $\tilde{A}_0$ , and  $\tilde{B}_0$  (it can be shown that an isometry channel always achieves the maximal guessing probability). Alice then performs a POVM on her systems  $A_0$ ,  $A_1$ , and  $A_2$  and guesses the output value  $x$  of  $B_1$ . The expression  $2^{-H_{\min}^{\text{ext}}(B|A)_{\Omega}}$  is the maximum probability for Alice to guess correctly Bob's output, obtained by optimizing over all possible isometries and POVMs.

These  $n$ -conditions can be expressed as a single condition given by

$$\sum_{j=1}^n |j\rangle\langle j|^R \otimes \Phi_j^B = \sum_{j=1}^n |j\rangle\langle j|^R \otimes \Theta[\Psi_j^A]$$

where  $|j\rangle\langle j|^R$  can be viewed as a channel from the 1-dimensional system  $R_0$  to the  $n$ -dimensional classical system  $R_1$ . Therefore, the problem (85) is a special case of the following problem.

Consider three physical systems  $A$ ,  $B$ , and  $R$ , and two bipartite quantum channels  $\Phi^{RA} \in \mathcal{L}^{RA}$  and  $\Psi^{RB} \in \mathcal{L}^{RB}$ . Is there a superchannel  $\Theta \in \mathbb{L}^{AB}$  such that (see Fig. 7)

$$\Psi^{RB} = \mathbb{1}^R \otimes \Theta[\Phi^{RA}] ? \quad (86)$$

If such a superchannel exists, we will say that  $\Phi^{RA}$  *quantum majorizes*  $\Psi^{RB}$  and denote this relation by:

$$\Psi^{RB} \prec_q \Phi^{RA} \quad \text{or equivalently} \quad \Phi^{RA} \succ_q \Psi^{RB}.$$

The preorder  $\prec_q$  was studied in [11] for the case of quantum states (i.e.  $d_{A_0} = d_{B_0} = d_{R_0} = 1$ ).

This section is organized as follows. We start with discussion on channel divergences and show that they can be used to provide necessary conditions for the relation (85) with  $n = 2$ . We then move to provide a full characterization of the preorder  $\prec_q$  above (and consequently of (85) as well). Our characterization is given in terms of the extended conditional min-entropy, and we also show that the problem can be completely solved with semidefinite programming. We end the section with an application of our results to the resource theory of thermodynamics.

### A. Channel Divergences

Consider (85) with  $n = 2$ . As discussed in the introduction, a measure of distinguishability  $D(\cdot||\cdot)$  of two quantum states must satisfy the following condition:

$$D(\Phi(\rho_1)||\Phi(\rho_2)) \leq D(\rho_1||\rho_2).$$

Any such function (also known as contraction) provides necessary conditions that a pair of quantum states  $(\rho_1, \rho_2)$  is related to another pair of states  $(\sigma_1, \sigma_2)$  via some quantum channel  $\Phi$ . Examples of such functions are the trace norm,

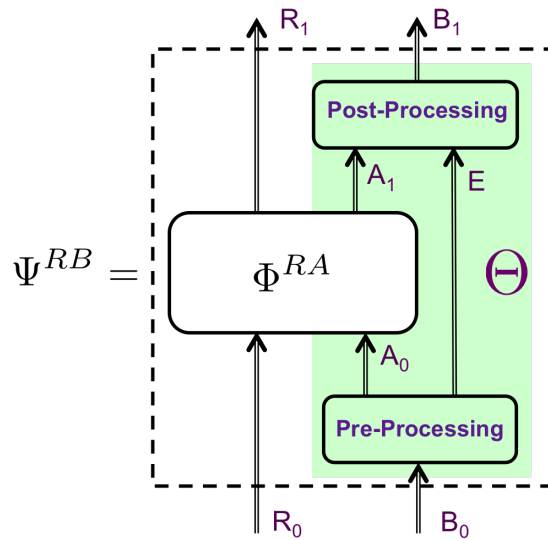


FIG. 7: The action of a superchannel  $\Theta \in \mathbb{L}^{AB}$  on a bipartite channel  $\Phi^{RA}$  yielding the bipartite state  $\Psi^{RB}$

the relative entropy, and many Rényi divergences (see [34] and reference therein for a large class of such functions). In [21, 35] a method to extend any such divergence from states to channels was introduced. This method is very similar to the extension of the trace norm into the diamond norm [36].

The completely bounded trace norm, which is commonly called the *diamond norm*, is defined as follows (we focus here only on CP maps). Let  $\Phi_1^A, \Phi_2^A : \mathcal{B}(\mathcal{H}^{A_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1})$  be two CP maps, and let  $R$  be a reference system. Then, the diamond norm distance between  $\Phi_1^A$  and  $\Phi_2^A$  is given by:

$$\|\Phi_1^A - \Phi_2^A\|_\diamond := \max_{\rho^{A_0 R}} \|\Phi_1^A \otimes \text{id}^R(\rho^{A_0 R}) - \Phi_2^A \otimes \text{id}^R(\rho^{A_0 R})\|_1$$

where the maximum can be taken over all pure bipartite states  $\rho^{A_0 R}$  with reference system  $R$  of the same dimension as  $A_0$ .

Following the above extension of the trace norm, for any divergence  $D(\|\cdot\|)$ , we define now the following induced contraction,  $C_D(\|\cdot\|)$ , that acts on CPTP maps (this induced contraction was first introduced in [21, 35]). For any two CPTP maps  $\Phi_1, \Phi_2 : \mathcal{B}(\mathcal{H}^A) \rightarrow \mathcal{B}(\mathcal{H}^B)$  define

$$\begin{aligned} C_D(\Phi_1^A \|\Phi_2^A) \\ := \sup_{\sigma^{A_0 R}} D\left(\Phi_1^A \otimes \text{id}^R(\sigma^{A_0 R}) \|\Phi_2^A \otimes \text{id}^R(\sigma^{A_0 R})\right) \end{aligned} \quad (87)$$

where the supremum is taken over all density matrices  $\sigma^{A_0 R} \in \mathcal{B}(\mathcal{H}^{A_0} \otimes \mathcal{H}^R)$  and over all dimensions of system  $R$ . Note, however, that we can assume w.l.o.g. that the supremum above is over pure states since we can always purify  $\sigma^{A_0 R}$ , and  $D$  is contractive under a partial trace. Moreover, since  $D$  is invariant under a local unitary on  $R$ , we conclude that w.l.o.g. the dimension of  $\mathcal{H}^R$  can be taken to be the same as that of  $\mathcal{H}^{A_0}$ . We now show that the above generalized divergence is indeed a contraction under superchannels.

Let  $\Theta \in \mathbb{L}^{AB}$  be a superchannel, and  $\Phi_1^A, \Phi_2^A \in \mathcal{L}^A$  be two quantum channels. First note that

$$\begin{aligned} C_D(\Theta[\Phi_1^A] \|\Theta[\Phi_2^A]) \\ = \sup_{\sigma^{B_0 R}} D\left(\Theta[\Phi_1^A] \otimes \text{id}^R(\sigma^{B_0 R}) \|\Theta[\Phi_2^A] \otimes \text{id}^R(\sigma^{B_0 R})\right) \end{aligned} \quad (88)$$

where the supremum is over all density matrices  $\sigma^{B_0 R} \in \mathcal{B}(\mathcal{H}^{B_0 R})$  and over all systems  $R$ . Although this optimization can be taken over pure states with  $R$  having the same dimension as  $B_0$ , it will be convenient not to assume this at the moment. Now, since  $\Theta$  is a superchannel, it can be realized in terms of pre and post processing as in (21). Denoting

by  $\tau^{A_0ER} \equiv \Gamma_{\text{pre}}^{B_0 \rightarrow A_0E} \otimes \text{id}^R(\sigma^{B_0R})$  we get

$$\begin{aligned}
& D \left( \Theta[\Phi_1^A] \otimes \text{id}^R(\sigma^{B_0R}) \middle\| \Theta[\Phi_2^A] \otimes \text{id}^R(\sigma^{B_0R}) \right) \\
&= D \left( \Gamma_{\text{post}}^{A_1E \rightarrow B_1} \circ \Phi_1^A(\tau^{A_0ER}) \middle\| \Gamma_{\text{post}}^{A_1E \rightarrow B_1} \circ \Phi_2^A(\tau^{A_0ER}) \right) \\
&\leq D \left( \Phi_1^A(\tau^{A_0ER}) \middle\| \Phi_2^A(\tau^{A_0ER}) \right) \\
&\leq \max_{\gamma^{A_0ER} \in \mathcal{D}(\mathcal{H}^{A_0ER})} D \left( \Phi_1^A(\gamma^{A_0ER}) \middle\| \Phi_2^A(\gamma^{A_0ER}) \right) \\
&= C_D \left( \Phi_1^A \middle\| \Phi_2^A \right), \tag{89}
\end{aligned}$$

where the first inequality follows from the contractivity of  $D$  under  $\Gamma_{\text{post}}^{A_1E \rightarrow B_1}$ , and the last equality follows from the definition. We therefore conclude that

$$C_D \left( \Theta \left[ \Phi_1^A \right] \middle\| \Theta \left[ \Phi_2^A \right] \right) \leq C_D \left( \Phi_1^A \middle\| \Phi_2^A \right). \tag{90}$$

There are a couple of simple consequences of the above data processing inequality. First, any such contraction under superchannels,  $C(\cdot \middle\| \cdot)$ , is invariant under unitary superchannels as defined in (38). For such reversible superchannel  $\Theta$  we have

$$C \left( \Theta \left[ \Phi_1^A \right] \middle\| \Theta \left[ \Phi_2^A \right] \right) = C \left( \Phi_1^A \middle\| \Phi_2^A \right).$$

Second, consider a superchannel  $\Theta_\Lambda : \mathcal{L}^A \rightarrow \mathcal{L}^{AB}$  defined by

$$\Theta_\Lambda[\Phi^A] = \Phi^A \otimes \Lambda^B$$

It is straightforward to show that it is a superchannel. Similarly, consider the linear map  $\mathcal{T} : \mathcal{L}^{AB} \rightarrow \mathcal{L}^A$  acting on a bipartite map  $\Omega^{AB} \in \mathcal{L}^{AB}$  as

$$\mathcal{T}[\Omega^{AB}] (\rho^{A_0}) = \text{Tr}_{B_1} [\Omega^{AB} (\rho^{A_0} \otimes u^{B_0})]$$

where  $u^{B_0} := \frac{1}{d_{B_0}} I^{B_0}$  is the maximally mixed (uniform) state (note that one can choose to input another state and the choice of maximally mixed state is just a convenient one). Since  $\mathcal{T}$  is also a superchannel, the contractivity of  $C(\cdot \middle\| \cdot)$  implies that

$$C \left( \Phi^A \otimes \Lambda^B \middle\| \Psi^A \otimes \Lambda^B \right) = C \left( \Phi^A \middle\| \Psi^A \right)$$

## B. Characterization of Quantum Majorization for Channels

To characterize the preorder  $\prec_q$ , we first show that we can assume w.l.o.g. that system  $R$  is classical, and that  $d_{R_0} = 1$ . For this purpose, we define two sets of CP trace non-increasing maps that we construct from the two bipartite CPTP maps  $\Phi^{RA}$  and  $\Psi^{RB}$ . Let  $\{|\varphi_x\rangle\langle\varphi_x|^{R_0}\}_{x=1}^{d_{R_0}^2}$  be a normalized rank one basis of  $\mathcal{B}_h(\mathcal{H}^{R_0})$  and let  $\{E_y^{R_1}\}_{y=0}^{d_{R_1}^2}$  be an informationally complete (basis) POVM of  $\mathcal{B}(\mathcal{H}^{R_1})$ . Then, we define for all  $\rho^{A_0} \in \mathcal{B}(\mathcal{H}^{A_0})$

$$\Phi_{y|x}^A(\rho^{A_0}) := \text{Tr}_{R_1} \left[ (E_y^{R_1} \otimes I^{A_1}) \Phi^{RA} (|\varphi_x\rangle\langle\varphi_x|^{R_0} \otimes \rho^{A_0}) \right]$$

and  $\Psi_{y|x}^B$  is defined similarly as in the above equation, by replacing  $\Phi^{RA}$  with  $\Psi^{RB}$ . By definition, for all  $x$  and  $y$ ,  $\Phi_{y|x}^A$  and  $\Psi_{y|x}^B$  are trace non-increasing CP maps, and for any  $x$  the maps  $\sum_{y=1}^{d_{R_1}^2} \Phi_{y|x}^A$  and  $\sum_{y=1}^{d_{R_1}^2} \Psi_{y|x}^B$  are CPTP maps.

A key observation with these definitions is that (86) holds if and only if

$$\Psi_{y|x}^B = \Theta \left[ \Phi_{y|x}^A \right]$$

for all  $x = 1, \dots, d_{R_0}^2$  and all  $y = 1, \dots, d_{R_1}^2$ . With this at hand, we can define two classical systems  $X$  and  $Y$ , with 1-dimensional input, and with  $d_X \equiv d_{R_0}^2$  and  $d_Y \equiv d_{R_1}^2$  dimensional outputs, such that

$$\Phi^{XYA} := \frac{1}{d_X} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} |xy\rangle\langle xy|^{XY} \otimes \Phi_{y|x}^A. \tag{91}$$

Note that  $\Phi^{XYA} \in \mathcal{C}^{XYA}$ ; i.e. it is a CPTP map. Defining  $\Psi^{XYB} \in \mathcal{C}^{XYB}$  in the same way, we conclude that (86) holds if and only if

$$\Psi^{XYB} = \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}] . \quad (92)$$

Note further that this relation reduces to (85) in the special case that  $d_Y = 1$ .

We are now ready to characterize the preorder  $\prec_q$  in terms of the extended conditional min-entropy. For this purpose, we will denote by  $\mathcal{C}_\star^{XYA}$  (and similarly  $\mathcal{C}_\star^{XYB}$ ) a subset of CPTP maps in  $\mathcal{C}^{XYA}$  that has the form (91) with the property that for any  $x$  the map  $\sum_{y=1}^{d_Y} \Phi_{y|x}^A$  is CPTP.

**Theorem 7.** *Let  $\Phi^{RA} \in \mathcal{C}^{RA}$  and  $\Psi^{RB} \in \mathcal{C}^{RB}$  be two bipartite quantum channels, and let  $\Phi^{XYA} \in \mathcal{C}_\star^{XYA}$  and  $\Psi^{XYB} \in \mathcal{C}_\star^{XYB}$  be their corresponding classical-quantum channels as described above. For any quantum channel  $\Lambda^{XYB} \in \mathcal{C}_\star^{XYB}$  define the two bipartite CP maps*

$$\begin{aligned} \Lambda_\Phi^{AB} &:= \frac{1}{d_X} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \Phi_{y|x}^A \otimes \Lambda_{y|x}^B \\ \Lambda_\Psi^{\tilde{B}B} &:= \frac{1}{d_X} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \Psi_{y|x}^{\tilde{B}} \otimes \Lambda_{y|x}^B . \end{aligned}$$

Then,  $\Psi^{RB} \prec_q \Phi^{RA}$  if and only if for all  $\Lambda^{XYB} \in \mathcal{C}_\star^{XYB}$

$$H_{\min}^{\text{ext}}(B|A)_{\Lambda_\Phi} \leq H_{\min}^{\text{ext}}(B|\tilde{B})_{\Lambda_\Psi} .$$

*Remark 4.* Note that the bipartite CP maps  $\Lambda_\Phi^{AB}$  and  $\Lambda_\Psi^{\tilde{B}B}$  are in general not trace preserving. They can be expressed as:

$$\begin{aligned} \Lambda_\Phi^{AB} &= \Theta_\Lambda^{XY \rightarrow B} \otimes \mathbb{1}^A [\Phi^{XYA}] \\ \Lambda_\Psi^{\tilde{B}B} &= \Theta_\Lambda^{XY \rightarrow B} \otimes \mathbb{1}^{\tilde{B}} [\Psi^{XY\tilde{B}}] . \end{aligned}$$

where  $\Theta_\Lambda$  is a CPP map from the classical variables space  $\mathcal{L}^{XY}$  to  $\mathcal{L}^B$ . It is defined by its action on the elements of  $\mathcal{L}^{XY}$ : for any  $x \in \{1, \dots, d_X\}$  and  $y \in \{1, \dots, d_Y\}$

$$\Theta_\Lambda [|xy\rangle\langle xy|^{XY}] := \Lambda_{y|x}^B$$

(recall that  $|xy\rangle\langle xy|^{XY}$  is viewed as a preparation channel; i.e. with 1-dimensional channel input space).

*Proof.* The necessity of the condition follows trivially from the monotonicity condition of the extended conditional min-entropy (see (66) with  $\Omega^{AB}$  replaced by  $\Lambda_\Phi^{AB}$ ). To prove sufficiency consider the set

$$\mathcal{K}_\Phi := \left\{ \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}] \mid \Theta \in \mathbb{S}^{AB} \right\} \subset \mathcal{L}^{XYB} .$$

Note that the space  $\mathcal{L}^{XYB}$  consists of maps of the form  $\sum_{x,y} |xy\rangle\langle xy|^{XY} \otimes \mathcal{E}_{xy}^{B_0 \rightarrow B_1}$  with each  $\mathcal{E}_{xy}^{B_0 \rightarrow B_1} \in \mathcal{L}^B$ .  $\mathcal{K}_\Phi$  is a closed convex (compact) set in  $\mathcal{L}^{XYB}$ . Therefore, from the separation theorem it follows that  $\Psi^{XYB} \notin \mathcal{K}_\Phi$  if and only if there exists a map  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$  such that

$$\langle \Psi^{XYB}, \Lambda^{XYB} \rangle > \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle \quad (93)$$

Alternatively, this can be expressed as follows:  $\Psi^{XYB} \in \mathcal{K}_\Phi$  if and only if for all  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$

$$\langle \Psi^{XYB}, \Lambda^{XYB} \rangle \leq \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle . \quad (94)$$

We first show that (94) holds for all linear maps  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$  if and only if

$$\begin{aligned} & \max_{\Theta' \in \mathbb{S}^{\tilde{B}B}} \langle \mathbb{1}^{XY} \otimes \Theta' [\Psi^{XY\tilde{B}}], \Lambda^{XYB} \rangle \\ & \leq \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle \end{aligned} \quad (95)$$



holds for all maps  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$ . Indeed, if (95) holds, then take  $\Theta'$  to be the identity superchannel  $\mathbb{1}^B$  so that (94) holds. Conversely, suppose (94) holds for all maps  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$ , and let  $\Theta'$  be an optimal superchannel in the LHS of (95). We then get

$$\begin{aligned} \langle \mathbb{1}^{XY} \otimes \Theta' [\Psi^{XY\tilde{B}}], \Lambda^{XYB} \rangle &= \langle \Psi^{XY\tilde{B}}, \mathbb{1}^{XY} \otimes \Theta'^* [\Lambda^{XYB}] \rangle \\ &\leq \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \mathbb{1}^{XY} \otimes \Theta'^* [\Lambda^{XYB}] \rangle \\ &= \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta' \circ \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle \\ &\leq \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle, \end{aligned}$$

where for the first inequality we used (94), and for the second inequality we used the fact that  $\Theta' \circ \Theta$  is itself a superchannel in  $\mathbb{S}^{AB}$ .

We now prove that it is sufficient to consider in (95) only maps  $\Lambda^{XYB} \in \mathcal{C}_*^{XYB} \subset \mathcal{L}^{XYB}$ . From its definition  $\Lambda^{XYB} \in \mathcal{C}_*^{XYB}$  if and only if its Choi matrix is positive semidefinite and have marginal

$$J_{\Lambda}^{XB_0} = u^X \otimes I^{B_0},$$

(in particular,  $J_{\Lambda}^{B_0} = I^{B_0}$  so that  $\Lambda^{XYB}$  is trace preserving). Now, if  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$  do not have this form, then set  $T^{XB_0} \equiv u^X \otimes I^{B_0} - J_{\Lambda}^{XB_0}$  and define  $\tilde{\Lambda}^{XYB}$  by its Choi matrix

$$J_{\tilde{\Lambda}}^{XYB} := (1 - \epsilon) u^{XYB_1} \otimes I^{B_0} + \epsilon (J_{\Lambda}^{XYB} + T^{XB_0} \otimes u^{YB_1}),$$

where  $\epsilon > 0$  is small enough so that  $J_{\tilde{\Lambda}}^{XYB} \geq 0$ . Note also that  $J_{\tilde{\Lambda}}^{XB_0} = u^X \otimes I^{B_0}$ . Now, for any channel  $\Psi^{XYB} \in \mathcal{C}_*^{XYB}$  we have

$$\begin{aligned} \langle \Psi^{XYB}, \tilde{\Lambda}^{XYB} \rangle &= \\ (1 - \epsilon) \frac{d_{B_0}}{d_X d_Y d_{B_1}} + \epsilon \langle \Psi^{XYB}, \Lambda^{XYB} \rangle + \frac{\text{Tr}[T^{XB_0}]}{d_X} \end{aligned}$$

A key observation is that except for the term  $\epsilon \langle \Psi^{XYB}, \Lambda^{XYB} \rangle$  on the RHS of the equation above, all the other terms are independent on  $\Psi^{XYB}$ . Hence, equations (94) and (95) holds for all maps  $\Lambda^{XYB} \in \mathcal{L}^{XYB}$  if and only if they hold for all CPTP maps  $\tilde{\Lambda}^{XYB} \in \mathcal{C}_*^{XYB}$ .

With this form of  $\Lambda^{XYB}$  we have

$$\begin{aligned} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle &= \\ = \frac{1}{d_X^2} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \langle \Theta [\Phi_{y|x}^A], \Lambda_{y|x}^B \rangle \end{aligned}$$

Expressing the inner product in terms of the Choi matrices and using (10) we get

$$\begin{aligned} \langle \Theta [\Phi_{y|x}^A], \Lambda_{y|x}^B \rangle &= \text{Tr} \left[ \mathbf{J}_{\Theta}^{AB} \left( \left( J_{\Phi_{y|x}}^A \right)^T \otimes J_{\Lambda_{y|x}}^B \right) \right] \\ &= \text{Tr} \left[ \left( \mathbf{J}_{\Theta}^{AB} \right)^T \left( J_{\Phi_{y|x}}^A \otimes \left( J_{\Lambda_{y|x}}^B \right)^T \right) \right] \end{aligned}$$

Denoting the CPTP map  $\tilde{\Lambda}^{XYB}$  to be

$$J_{\tilde{\Lambda}}^{XYB} := \frac{1}{d_X} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} |xy\rangle\langle xy|^{XY} \otimes \left( J_{\Lambda_{y|x}}^B \right)^T,$$

we conclude that

$$\begin{aligned}
& \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle \\
&= \max_{\Theta \in \mathbb{S}^{AB}} \frac{1}{d_X^2} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \text{Tr} \left[ (\mathbf{J}_{\Theta}^{AB})^T \left( J_{\Phi_{y|x}}^A \otimes J_{\tilde{\Lambda}_{y|x}}^B \right) \right] \\
&= \max_{\Theta \in \mathbb{S}^{AB}} \frac{1}{d_X^2} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \text{Tr} \left[ \mathbf{J}_{\Theta}^{AB} J_{\Phi_{y|x}}^{AB} \tilde{\Lambda}_{y|x} \right] \\
&= \frac{1}{d_X} \max_{\Theta \in \mathbb{S}^{AB}} \text{Tr} \left[ \mathbf{J}_{\Theta}^{AB} J_{\tilde{\Lambda}_{\Phi}}^{AB} \right] \\
&= \frac{d_{B_0}}{d_X} 2^{-H_{\min}^{\text{ext}}(B|A)_{\tilde{\Lambda}_{\Phi}}}
\end{aligned} \tag{96}$$

where the last equality follows from (62), and

$$\tilde{\Lambda}_{\Phi}^{AB} \equiv \frac{1}{d_X} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \Phi_{y|x}^A \otimes \tilde{\Lambda}_{y|x}^B \tag{97}$$

Hence, Eq.(95) is equivalent to

$$H_{\min}^{\text{ext}}(B|A)_{\tilde{\Lambda}_{\Phi}} \leq H_{\min}^{\text{ext}}(B|\tilde{B})_{\tilde{\Lambda}_{\Psi}} .$$

Since  $\tilde{\Lambda}^{XYB} \in \mathcal{C}_{\star}^{XYB}$  if and only if  $\Lambda^{XYB} \in \mathcal{C}_{\star}^{XYB}$  we can remove the tilde from  $\Lambda^{XYB}$ . This completes the proof.  $\blacksquare$

### C. Comparison of channels with the extended conditional min-entropy

The family of generalized divergences discussed earlier (see also [37]), along with other methods to extend divergences [18], provide only necessary conditions for the existence of a superchannel that converts one pair of channels to another. We now focus on yet another family of channel divergences that provides both necessary and *sufficient* conditions for the existence of such a superchannel.

By taking the special case of  $d_Y = 1$  in the Theorem 7 above we get the following corollary.

**Corollary 1.** *For each  $j = 1, \dots, n$ , let  $\Psi_j^A \in \mathcal{C}^A$  and  $\Phi_j^B \in \mathcal{C}^B$  be quantum channels, and let  $R$  denote a third system of the same input and output dimensions as  $B$ . Then, there exists a superchannel satisfying (85) for all  $j = 1, \dots, n$  if and only if for any set of  $n$  quantum channels  $\Lambda_j^R \in \mathcal{C}^R$  the tripartite quantum channel*

$$\Lambda^{ABR} := \frac{1}{n} \sum_{j=1}^n \Psi_j^A \otimes \Phi_j^B \otimes \Lambda_j^R$$

*satisfies*

$$H_{\min}^{\text{ext}}(R|A)_{\Lambda} \leq H_{\min}^{\text{ext}}(R|B)_{\Lambda} .$$

*Remark 5.* In the case that  $n = 2$ , we can define for any two channels  $\Lambda_1^R$  and  $\Lambda_2^R$ , the channel contraction (i.e. divergence) as

$$C_{\Lambda}(\Psi_1^A \parallel \Psi_2^A) := H_{\min}^{\text{ext}}(R|A)_{\Lambda^{AR}} ,$$

where

$$\Lambda^{AR} := \frac{1}{2} (\Psi_1^A \otimes \Lambda_1^R + \Psi_2^A \otimes \Lambda_1^R) .$$

The corollary above states that this family of divergences  $\{C_{\Lambda}\}$  provides both necessary and sufficient conditions for (85) to hold in the case that  $n = 2$ .

#### D. Characterization with semidefinite programming

A caveat of the characterization of quantum majorization with the extended conditional min-entropy is that it involves an infinite number of conditions, and therefore it is not practical for determining whether there exists a superchannel that satisfies (86) or (85). Nonetheless, here we show that the problems in (86) and (85) can be solved with semidefinite programming (SDP).

From the proof of Theorem 7 it follows (see particularly (94)) that  $\Psi^{XYB} \prec_q \Phi^{XYA}$  if and only if the function

$$f(\Phi, \Psi) := \min_{\Lambda^{XYB} \in \mathcal{C}_*^{XYB}} \max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}] - \Psi^{XYB}, \Lambda^{XYB} \rangle \quad (98)$$

is not negative. We now show that the above optimization problem is an SDP.

Recall that in (96) we showed that

$$\max_{\Theta \in \mathbb{S}^{AB}} \langle \mathbb{1}^{XY} \otimes \Theta [\Phi^{XYA}], \Lambda^{XYB} \rangle = \frac{d_{B_0}}{d_X} 2^{-H_{\min}^{\text{ext}}(B|A)_{\tilde{\Lambda}_\Phi}}$$

where  $\tilde{\Lambda}^{AB}$  is defined in (97). Combining the expression for  $f(\Phi, \Psi)$  with the definition in (60) for the extended conditional min-entropy gives

$$f(\Phi, \Psi) = \min \{ \text{Tr}[\gamma^{AB_0}] - \text{Tr}[J_\Psi^{XYB} J_\Lambda^{XYB}] \}$$

subject to:

$$1. \quad \gamma^{AB_0} \otimes I^{B_1} \geq \frac{1}{d_{A_0} d_X^2} \sum_{x=1}^{d_X} \sum_{y=1}^{d_Y} \left( J_{\Phi_{y|x}}^A \right)^T \otimes J_{\Lambda_{y|x}}^B \quad (99)$$

$$2. \quad \gamma^{A_0 B_0} = u^{A_0} \otimes \gamma^{B_0} \quad (100)$$

$$3. \quad J_\Lambda^{XB_0} = u^X \otimes I^{B_0} \quad ; \quad J_\Lambda^{XYB} \geq 0 \quad (101)$$

where we absorbed the factor  $\frac{d_{B_0}}{d_X}$  into  $\gamma^{AB_0}$ . To bring the above optimization to a canonical SDP form define the following (real) vector spaces consisting of the direct sum of three spaces:

$$V := \mathcal{B}_h(\mathcal{H}^{AB_0}) \oplus \mathcal{B}_h(\mathcal{H}^{XYB}) \oplus \mathcal{B}_h(\mathcal{H}^{AB}) \quad (102)$$

We denote the elements of  $V$  by

$$\xi \equiv (\gamma^{AB_0}, \alpha^{XYB}, \eta^{AB}) \in V,$$

and set

$$\mu \equiv (I^{AB_0}, -J_\Psi^{XYB}, \mathbf{0}^{AB}),$$

where  $\mathbf{0}^{AB}$  is the zero matrix in  $\mathcal{B}_h(\mathcal{H}^{AB})$ . The inner product between elements in  $V$  is defined as the sum of the inner products among the three components. We also define the linear map  $\mathcal{T} : V \rightarrow \mathcal{B}_h(\mathcal{H}^{AB})$  by

$$\begin{aligned} \mathcal{T}(\xi) &= \gamma^{AB_0} \otimes I^{B_1} \\ &\quad - \text{Tr}_{XY} \left[ \left( (J_\Phi^{XYA})^T \otimes I^B \right) (\alpha^{XYB} \otimes I^A) \right] - \eta^{AB} \end{aligned}$$

Note that  $\mathcal{T}$  is indeed linear, and (99) is equivalent to  $\mathcal{T}(\xi) + \eta^{AB} \geq 0$  when we identify  $\alpha^{XYB}$  with  $J_\Lambda^{XYB}$ . Finally, set

$$\begin{aligned} G_{x,k} &:= (\mathbf{0}^{AB_0}, |x\rangle\langle x|^X \otimes |\psi_k\rangle\langle\psi_k|^{B_0} \otimes I^{YB_1}, \mathbf{0}^{AB}) \in V, \\ F_\ell &:= (M_\ell^{A_0 B_0} \otimes I^{A_1}, \mathbf{0}^{XYB}, \mathbf{0}^{AB}) \in V, \end{aligned}$$

where  $\{|\psi_k\rangle\langle\psi_k|^{B_0}\}_{k=1}^{d_{B_0}^2}$  is a pure state basis of  $\mathcal{B}_h(\mathcal{H}^{B_0})$ , and  $\{M_\ell^{A_0 B_0}\}_{\ell=1}^{(d_{A_0}^2-1)d_{B_0}^2}$  is a basis of the subspace of  $\mathcal{B}_h(\mathcal{H}^{A_0 B_0})$  consisting of all matrices with zero marginal on  $B_0$ . Note that (100) holds if and only if  $\text{Tr}[\gamma^{A_0 B_0} M_\ell^{A_0 B_0}] =$

0 for all  $\ell$ , and similarly the first equality of (101) holds if and only if  $\text{Tr} \left[ J_{\Lambda}^{XB_0} (|x\rangle\langle x|^X \otimes |\psi_k\rangle\langle\psi_k|^{B_0}) \right] = 1/d_X$  for all  $x$  and  $k$ . Hence, with these notations we get:

$$f(\Phi, \Psi) = \min \text{Tr}[\xi\mu]$$

subject to: for all  $x = 1, \dots, d_X$ ,  $k = 1, \dots, d_{B_0}^2$ , and  $\ell = 1, \dots, (d_{A_0}^2 - 1)d_{B_0}^2$

$$\mathcal{T}(\xi) = 0 \quad ; \quad \text{Tr}[\xi G_{k,x}] = \frac{1}{d_X} \quad ; \quad \text{Tr}[\xi F_{\ell}] = 0 \quad ; \quad \xi \geq 0$$

Finally, note that the condition  $\mathcal{T}(\xi) = 0$  can also be expressed in terms of inner products. That is, let  $\{E_j\}_{j=1}^d$  be a basis of  $V$ , with  $d \equiv \dim(V)$ , and for each  $j = 1, \dots, d$ , let  $F_j := \mathcal{T}^*(E_j)$ . Then, with these notations, we can replace the condition  $\mathcal{T}(\xi) = 0$  above with

$$\text{Tr}[\xi E_j] = 0 \quad \forall j = 1, \dots, d.$$

We therefore obtained a canonical form of a SDP optimization problem that can be plugged into standard packages such as CVX. Note that the number of all the constraints is polynomial in the dimensions.

### E. An application to thermodynamics

Recall that if  $d_{A_0} = d_{B_0} = 1$  then all the channels involved in Theorems 7, and the semidefinite programming above, become states and the extended conditional min-entropy reduces to the standard conditional min-entropy of states. In this case, the Theorem 7 above reduces to the state analog that was proved in [11]. As was shown in [11], the state version of the theorem above has many applications particularly in state transformations of quantum resources theories of thermodynamics and asymmetry. We expect that the theorem above will also have applications in the simulation of channels in various resource theories of quantum processes (see some very recent work on the subject [38–41]). We give here a very brief discussion on one such application in the quantum resource theory of athermality in thermodynamics.

In the resource theory of quantum thermodynamics the Gibbs state,  $\gamma$ , is known to be the only free state of the model. In this model, a replacement channel that outputs the Gibbs state irrespective of the input state is a free channel. Denote such a channel by  $\Pi_{\gamma}^A \in \mathcal{C}^A$ ; that is,

$$\Pi_{\gamma}^A(\rho^{A_0}) := \text{Tr}[\rho^{A_0}] \gamma^{A_1} \quad \forall \rho^{A_0} \in \mathcal{B}(\mathcal{H}^{A_0}).$$

Now, consider a superchannel  $\Theta : \mathcal{L}^A \rightarrow \mathcal{L}^B$  in which system  $A_1$  is associated with a Gibbs state  $\gamma^{A_1}$  and system  $B_1$  with Gibbs state  $\gamma^{B_1}$ . Then, if  $\Theta$  is a free superchannel it must take the Gibbs channel of system  $A$  to the Gibbs channel of system  $B$ . That is, it satisfies:

$$\Theta[\Pi_{\gamma}^A] = \Pi_{\gamma}^B.$$

We call such superchannels that preserve the Gibbs channel  $\Pi_{\gamma}$ , *Gibbs preserving superchannels*. In this case, Corollary 1 and Sec. VD provide the necessary and sufficient conditions that a given channel  $\Phi^B$  can be simulated by another channel  $\Psi^A$  via Gibbs preserving superchannels. While Corollary 1 provides a complete family of athermality monotones (of dynamical resources, i.e. channels) in terms of the extended conditional min-entropy, Sec. VD shows that the problem can be solved efficiently with SDP. We leave the extensions of these ideas to other, more physical models of thermodynamics (e.g. channel simulations under thermal operations, etc), and other resource theories, for future work.

## VI. SUMMARY AND CONCLUSIONS

We discussed in this paper several different noise models for superchannels, and used that to define entropy functions for quantum channels. Our approach was axiomatic and minimalistic in nature, requiring the entropy function to be additive and monotonic only under random unitary superchannels. As an example, we found that the extended min-entropy is an entropy function that is monotonic under a much larger set of operations than the random unitary ones. We called these operations doubly stochastic superchannels since they consist of superchannels whose dual maps are also superchannels. We gave doubly stochastic superchannels a physical interpretation by showing that they have the property that they are completely uniformity preserving (see Sec. III C).

We then introduced an extension to the conditional min-entropy from bipartite states to bipartite channels. Given our definition of an entropy of a channel, we were able to show that the extended conditional min-entropy has many similar properties to the ones of its state version (i.e. the conditional min-entropy), including an operational interpretation in terms of a guessing probability if one of the subsystems is classical.

The extended conditional min-entropy turned out to play a key role in our extension of quantum majorization from bipartite states to bipartite channels. Quantum majorization, as defined originally in [11], is a pre-order for bipartite states that can be viewed as a generalization of majorization. It has applications to quantum resource theories, degradability of channels, and quantum statistical comparisons. Here we extended this definition from bipartite states to bipartite channels. A special case of this pre-order is the problem of comparison of channels given in (85). In theorem 7 we showed that quantum majorization for channels can be fully characterized with a family of functions given in terms of the extended conditional min-entropy. In particular, for the comparison between two channels, Corollary 1 provides a complete set of channel divergences that are both necessary and sufficient to determine if (85) holds. We also showed that determining whether one bipartite channel quantum majorizes another can be solved efficiently with semidefinite programming.

We expect that the results and techniques used here will be useful particularly in resource theories of quantum processes. We gave an indication for that in Sec. V E. Moreover, some of the definitions given here can be extended further. For example, it is straightforward to define the smoothed version of the extended conditional min entropy. Let  $\Psi^{AB} : \mathcal{B}(\mathcal{H}^{A_0 B_0}) \rightarrow \mathcal{B}(\mathcal{H}^{A_1 B_1})$  be a bipartite quantum channel. The  $\epsilon$ -extended conditional min entropy is defined by:

$$H_{\min}^{\text{ext},\epsilon}(B|A)_{\Psi} := \sup_{\|\Phi - \Psi\|_{\diamond} \leq \epsilon} H_{\min}^{\text{ext}}(B|A)_{\Phi},$$

where the supremum is over all bipartite channels  $\Phi^{AB}$  that are  $\epsilon$ -close (in the diamond norm) to the channel  $\Psi^{AB}$ . With this definition we can also define

$$S^{\text{ext}}(B|A)_{\Psi} := \lim_{\epsilon \rightarrow 0} \lim_{n \rightarrow \infty} \frac{1}{n} H_{\min}^{\text{ext},\epsilon}(B^n|A^n)_{\Psi^{\otimes n}}$$

In [42] the asymptotic equipartition property was proved for states. This means that the function above becomes the von-Neumann conditional entropy whenever  $\Psi^{AB}$  is a replacement map; i.e. for any input density matrix  $\sigma^{A_0 B_0}$  we have  $\Psi^{AB}(\sigma^{A_0 B_0}) = \rho^{A_1 B_1}$ , where  $\rho^{A_1 B_1}$  is a fixed output density matrix. In this case,

$$S^{\text{ext}}(B|A)_{\Psi} = S(B_1|A_1)_{\rho},$$

where  $S(B_1|A_1)_{\rho}$  is the conditional von-Neumann entropy. We therefore expect that the above quantity will have an interesting operational interpretation and leave its investigation for future work.

### Acknowledgments

The author would like to thank Francesco Buscemi, Giulio Chiribella, Eric Chitambar, Teiko Heinosaari, Barbara Kraus, Rob Spekkens, Mark Wilde, and Andreas Winter for useful discussions related to the topic of this paper. Particularly, the author appreciates the many useful comments by Mark Wilde on the first draft. The author acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

- 
- [1] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013).
  - [2] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
  - [3] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, 2018).
  - [4] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, L. Masanes, A. Acín, and F. Verstraete, Phys. Rev. Lett. **98**, 160501 (2007), URL <https://link.aps.org/doi/10.1103/PhysRevLett.98.160501>.
  - [5] F. G. S. L. Brandão and M. B. Plenio, Communications in Mathematical Physics **295**, 791 (2010), ISSN 1432-0916, URL <https://doi.org/10.1007/s00220-010-1005-z>.
  - [6] M. Mosonyi and T. Ogawa, Communications in Mathematical Physics **334**, 1617 (2015), ISSN 1432-0916, URL <https://doi.org/10.1007/s00220-014-2248-x>.
  - [7] D. Blackwell, Ann. Math. Statist. **24**, 265 (1953), URL <https://doi.org/10.1214/aoms/1177729032>.
  - [8] P. Alberti and A. Uhlmann, Reports on Mathematical Physics **18**, 163 (1980), ISSN 0034-4877, URL <http://www.sciencedirect.com/science/article/pii/003448778090083X>.

- [9] A. Chefles, R. Jozsa, and A. Winter, International Journal of Quantum Information **02**, 11 (2004).
- [10] F. Buscemi, Commun. Math. Phys. **310**, 625 (2012).
- [11] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian (2018), arXiv:quant-ph/1708.04302.
- [12] F. Buscemi (2018), arXiv:quant-ph/1803.02945.
- [13] R. Renner (2005), PhD Thesis.
- [14] R. Konig, R. Renner, and C. Schaffner, IEEE Transactions on Information Theory **55**, 4337 (2009).
- [15] G. Chiribella, G. M. D'Ariano, and P. Perinotti, EPL (Europhysics Letters) **83**, 30004 (2008), URL <http://stacks.iop.org/0295-5075/83/i=3/a=30004>.
- [16] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. Lett. **101**, 060401 (2008), URL <https://link.aps.org/doi/10.1103/PhysRevLett.101.060401>.
- [17] K. Zyczkowski, Journal of Physics A: Mathematical and Theoretical **41**, 355302 (2008), URL <http://stacks.iop.org/1751-8121/41/i=35/a=355302>.
- [18] M. Berta, C. Hirche, E. Kaur, and M. Wilde M. (2018), arXiv:quant-ph/1808.01498.
- [19] M. Wilde M. (2018), arXiv:quant-ph/1807.11939.
- [20] D. Puzzioli and J. Watrous, Annales Henri Poincaré **18**, 1153 (2017), ISSN 1424-0661, URL <https://doi.org/10.1007/s00023-016-0537-y>.
- [21] T. Cooney, M. Mosonyi, and M. M. Wilde, Communications in Mathematical Physics **344**, 797 (2016), ISSN 1432-0916, URL <https://doi.org/10.1007/s00220-016-2645-4>.
- [22] A. W. Harrow, A. Hassidim, D. W. Leung, and J. Watrous, Phys. Rev. A **81**, 032339 (2010), URL <https://link.aps.org/doi/10.1103/PhysRevA.81.032339>.
- [23] R. Duan, Y. Feng, and M. Ying, Phys. Rev. Lett. **103**, 210501 (2009), URL <https://link.aps.org/doi/10.1103/PhysRevLett.103.210501>.
- [24] A. Jencova, Journal of Mathematical Physics **55**, 022201 (2014), <https://doi.org/10.1063/1.4863715>, URL <https://doi.org/10.1063/1.4863715>.
- [25] A. Jencova, Journal of Mathematical Physics **53**, 012201 (2012), <https://doi.org/10.1063/1.3676294>, URL <https://doi.org/10.1063/1.3676294>.
- [26] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **74**, 012305 (2006), URL <https://link.aps.org/doi/10.1103/PhysRevA.74.012305>.
- [27] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, Phys. Rev. A **64**, 052309 (2001), URL <https://link.aps.org/doi/10.1103/PhysRevA.64.052309>.
- [28] T. Eggeling, D. Schlingemann, and R. F. Werner, EPL (Europhysics Letters) **57**, 782 (2002), URL <http://stacks.iop.org/0295-5075/57/i=6/a=782>.
- [29] G. M. D'Ariano, S. Facchini, and P. Perinotti, Phys. Rev. Lett. **106**, 010501 (2011), URL <https://link.aps.org/doi/10.1103/PhysRevLett.106.010501>.
- [30] M. Tomamichel, R. Colbeck, and R. Renner, IEEE Transactions on Information Theory **55**, 5840 (2009), ISSN 0018-9448.
- [31] C. B. Mendl and M. M. Wolf, Communications in Mathematical Physics **289**, 1057 (2009), ISSN 1432-0916, URL <https://doi.org/10.1007/s00220-009-0824-2>.
- [32] M. B. Hastings, Nature Physics **5** (2009), URL <http://dx.doi.org/10.1038/nphys1224>.
- [33] G. Chiribella and D. Ebler, New Journal of Physics **18**, 093053 (2016), URL <http://stacks.iop.org/1367-2630/18/i=9/a=093053>.
- [34] K. M. R. Audenaert and N. Datta, Journal of Mathematical Physics **56**, 022202 (2015), <https://doi.org/10.1063/1.4906367>, URL <https://doi.org/10.1063/1.4906367>.
- [35] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Phys. Rev. A **97**, 012332 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevA.97.012332>.
- [36] A. Kitaev, Russian Mathematical Surveys **52**, 1191?1249 (1997), URL <http://dx.doi.org/10.1070/RM1997v052n06ABEH002155>.
- [37] K. Sharma, M. M. Wilde, S. Adhikari, and M. Takeoka, New Journal of Physics **20**, 063025 (2018), URL <http://stacks.iop.org/1367-2630/20/i=6/a=063025>.
- [38] P. Faist, M. Berta, and F. Brandt (2018), arXiv:quant-ph/1807.05610.
- [39] D. Rosset, F. Buscemi, and Y.-C. Liang, Phys. Rev. X **8**, 021033 (2018), URL <https://link.aps.org/doi/10.1103/PhysRevX.8.021033>.
- [40] M. García Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter (2018), arXiv:quant-ph/1805.04045.
- [41] K. Ben Dana, M. García Díaz, M. Mejjaty, and A. Winter, Phys. Rev. A **95**, 062327 (2017), URL <https://link.aps.org/doi/10.1103/PhysRevA.95.062327>.
- [42] M. Tomamichel, R. Colbeck, and R. Renner, IEEE Transactions on Information Theory **55**, 5840 (2009).
- [43] A. Barvinok, *A Course in Convexity*, Graduate studies in mathematics (American Mathematical Society, 2002), ISBN 9780821829684, URL <https://books.google.ca/books?id=PaieFEiPn9cC>.

## Supplementary Material

### Comparison of quantum channels with superchannels

#### Appendix A: Strong Duality in Conic Linear programming

There were several places in the paper that we were using the strong duality of SDP or conic linear programming.

We present here the strong duality relation as given in [43], and use it in the following subsections to prove the various statements made in the paper for its specific applications.

Let  $V_1$  and  $V_2$  be two (real) vector spaces (here we will assume that they consist of Hermitian matrices) and let  $\Gamma : V_1 \rightarrow V_2$  be a linear map. Let  $\mathcal{K}_1 \subset V_1$  and  $\mathcal{K}_2 \subset V_2$  be two convex cones. Moreover, let  $H_1 \in V_1$  and  $H_2 \in V_2$  be two (fixed) elements.

1. The Primal Problem:

$$\begin{aligned} \text{Find } & \alpha := \inf \text{Tr}[XH_1] \\ \text{Subject to } & \Gamma(X) - H_2 \in \mathcal{K}_2 \quad \text{and} \\ & X \in \mathcal{K}_1 \end{aligned} \tag{A1}$$

2. The Dual Problem:

$$\begin{aligned} \text{Find } & \beta := \sup \text{Tr}[YH_2] \\ \text{Subject to } & H_1 - \Gamma^*(Y) \in \mathcal{K}_1^* \quad \text{and} \\ & Y \in \mathcal{K}_2^* \end{aligned} \tag{A2}$$

Here  $\Gamma^* : V_2 \rightarrow V_1$  is the dual map of  $\Gamma$ , and  $\mathcal{K}_1^*$  and  $\mathcal{K}_2^*$  are the dual cones, respectively, of  $\mathcal{K}_1$  and  $\mathcal{K}_2$ .

#### Weak duality:

For any feasible plan  $X$  (i.e.  $X$  satisfies  $\Gamma(X) - H_2 \in \mathcal{K}_2$  and  $X \in \mathcal{K}_1$ ) and a dual feasible plan  $Y$  (i.e.  $Y$  satisfies  $H_1 - \Gamma^*(Y) \in \mathcal{K}_1^*$  and  $Y \in \mathcal{K}_2^*$ ), we have

$$\text{Tr}[XH_1] \geq \text{Tr}[YH_2] \quad \text{and, in particular} \quad \alpha \geq \beta.$$

#### Strong Duality:

1. Consider the cone  $\mathcal{K} \subset V_2 \oplus \mathbb{R}$  defined by

$$\mathcal{K} := \left\{ \left( \Gamma(X) - Y, \text{Tr}[XH_1] \right) : X \in \mathcal{K}_1, Y \in \mathcal{K}_2 \right\}.$$

If  $\mathcal{K}$  is closed in  $V_2 \oplus \mathbb{R}$  and there exists a primal feasible plan then  $\alpha = \beta$ . Moreover, if  $\alpha > -\infty$  then there exists a primal optimal plan (i.e. a feasible  $X$  such that  $\alpha = \text{Tr}[XH_1]$ ).

2. The Slater's condition: Suppose that there is a primal feasible plan  $X_0 \in \text{int}(\mathcal{K}_1)$  such that  $\Gamma(X_0) - H_2 \in \text{int}(\mathcal{K}_2)$ . Suppose also that there exists a primal optimal plan. Then, there is no duality gap; i.e.  $\alpha = \beta$ .

In our cases, the strong duality will always hold.

#### 1. Proof of the Equivalence of (60) and (61)

Consider now the primal problem in (60):

$$\begin{aligned} & \min \text{Tr}[\gamma^{AB_0}] \\ \text{subject to: } & \mathbf{1.} \quad \gamma^{AB_0} \otimes I^{B_1} \geq \omega^{AB} \\ & \mathbf{2.} \quad \gamma^{A_0B_0} = u^{A_0} \otimes \gamma^{B_0} \end{aligned} \tag{A3}$$



We can identify it with the primal problem of the above conic programming in which  $V_1 := \mathcal{B}_h(\mathcal{H}^{AB_0})$ ,  $V_2 := \mathcal{B}_h(\mathcal{H}^{AB}) \oplus \mathcal{B}_h(\mathcal{H}^{A_0B_0})$ ,  $H_1 := I^{AB_0} \in V_1$ ,  $H_2 := (\omega^{AB}, 0^{A_0B_0}) \in V_2$ ,  $\mathcal{K}_1 := V_1$  (hence,  $\mathcal{K}_1^* = \{0^{AB_0}\}$ ), and

$$\mathcal{K}_2 := \{(\eta^{AB}, 0^{A_0B_0}) : \eta^{AB} \geq 0\} \subset V_2$$

and  $\Gamma : V_1 \rightarrow V_2$  defined by: for all  $\gamma^{AB_0} \in V_1$

$$\Gamma(\gamma^{AB_0}) := (\gamma^{AB_0} \otimes I^{B_1}, \gamma^{A_0B_0} - u^{A_0} \otimes \gamma^{B_0}) .$$

With these identifications, we get that the problem is identical to (A1). Hence, to get its dual, observe that  $\mathcal{K}_2^* = \mathcal{B}_+(\mathcal{H}^{AB}) \oplus \mathcal{B}_h(\mathcal{H}^{A_0B_0})$ , and  $\Gamma^* : V_2 \rightarrow V_1$  satisfies for any  $(\eta^{AB}, \zeta^{A_0B_0}) \in V_2$

$$\Gamma^*(\eta^{AB}, \zeta^{A_0B_0}) = \eta^{AB_0} + (\zeta^{A_0B_0} - u^{A_0} \otimes \zeta^{B_0}) \otimes I^{A_1} \quad (\text{A4})$$

With these identifications at hand, we get that the dual problem in (A2) is given by:

$$\begin{aligned} & \max \text{Tr} [\eta^{AB} \omega^{AB}] \\ \text{Subject to} \quad & I^{AB_0} = \eta^{AB_0} + (\zeta^{A_0B_0} - u^{A_0} \otimes \zeta^{B_0}) \otimes I^{A_1} \\ & \eta^{AB} \geq 0 \quad ; \quad \zeta^{A_0B_0} \in \mathcal{B}_h(\mathcal{H}^{A_0B_0}) . \end{aligned}$$

Finally, note that the condition

$$\eta^{AB_0} = I^{AB_0} - (\zeta^{A_0B_0} - u^{A_0} \otimes \zeta^{B_0}) \otimes I^{A_1} = \eta^{A_0B_0} \otimes u^{A_1}$$

where  $\eta^{A_0B_0} = d_{A_1} (I^{A_0B_0} - \zeta^{A_0B_0} + u^{A_0} \otimes \zeta^{B_0})$ . Furthermore,  $\eta^{B_0} = d_{A_0} d_{A_1} I^{B_0}$ . Finally, denoting by  $\alpha^{AB} := \frac{1}{d_{A_0}} \eta^{AB}$  we conclude that the above dual problem can be expressed as:

$$\begin{aligned} & d_{A_0} \max \text{Tr} [\alpha^{AB} \omega^{AB}] \\ \text{Subject to} \quad & \alpha^{AB_0} = \alpha^{A_0B_0} \otimes u^{A_1} \\ & \alpha^{AB} \geq 0 \quad ; \quad \alpha^{B_0} = d_{A_1} I^{B_0} . \end{aligned}$$

This expression is equivalent to (61).

## Appendix B: Operational interpretation of $H_{\min}^{\text{ext}}(B|A)_\Omega$ in the case that only system $B_1$ is classical

Since system  $B_1$  is classical, for all  $\rho^{A_0B_0} \in \mathcal{B}(\mathcal{H}^{A_0B_0})$

$$\Omega^{A_0B_0 \rightarrow A_1B_1}(\rho^{A_0B_0}) = \sum_{x=1}^{d_{B_1}} \Omega_x^{A_0B_0 \rightarrow A_1}(\rho^{A_0B_0}) \otimes |x\rangle\langle x|^{B_1} ,$$

where  $\{\Omega_x^{A_0B_0 \rightarrow A_1}\}$  form a quantum instrument, and the Choi matrix of  $\Omega^{AB}$  can be expressed as:

$$\omega^{AB} = \sum_{x=1}^{d_{B_1}} \omega_x^{AB_0} \otimes |x\rangle\langle x|^{B_1} ,$$

with

$$\omega_x^{AB_0} = \frac{1}{d_{A_0} d_{B_0}} \text{id}^{A_0B_0} \otimes \Omega_x^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1} \left( \phi_+^{A_0 \tilde{A}_0} \otimes \phi_+^{B_0 \tilde{B}_0} \right) \quad (\text{B1})$$

Consequently, from (61), with  $\omega^{AB}$  as above and  $\alpha^{AB} \equiv \sum_{x=1}^{d_{B_1}} \alpha_x^{AB_0} \otimes |x\rangle\langle x|^{B_1}$ , we get

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &= d_{A_0} \max \sum_{x=1}^{d_{B_1}} \text{Tr}[\alpha_x^{AB_0} \omega_x^{AB_0}] \\ \text{subject to:} \quad & \sum_{x=1}^{d_{B_1}} \alpha_x^{AB_0} = \eta^{A_0B_0} \otimes I^{A_1} , \quad \eta^{B_0} = I^{B_0} \\ & \alpha_x^{AB_0} \geq 0 \quad \forall x \in \{1, \dots, d_{B_1}\} \end{aligned} \quad (\text{B2})$$



where we denoted by

$$\eta^{A_0 B_0} := \frac{1}{d_{A_1}} \sum_{x=1}^{d_{B_1}} \alpha_x^{A_0 B_0}.$$

Note that we can assume w.l.o.g. that  $\eta^{A_0 B_0}$  is full rank. Hence, we can define the following POVM on system  $AB_0$ :

$$P_x^{AB_0} := \left( (\eta^{A_0 B_0})^{-\frac{1}{2}} \otimes I^{A_1} \right) \alpha_x^{AB_0} \left( (\eta^{A_0 B_0})^{-\frac{1}{2}} \otimes I^{A_1} \right)$$

Note that  $P_x^{AB_0} \geq 0$  and  $\sum_x P_x^{AB_0} = I^{AB_0}$ . With this notation

$$\begin{aligned} d_{A_0} \sum_{x=1}^{d_{B_1}} \text{Tr}[\alpha_x^{AB_0} \omega_x^{AB_0}] &= d_{A_0} \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_x^{AB_0} \left( \sqrt{\eta^{A_0 B_0}} \otimes I^{A_1} \right) \omega_x^{AB_0} \left( \sqrt{\eta^{A_0 B_0}} \otimes I^{A_1} \right) \right] \\ &= \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_x^{AB_0} \left( \text{id}^{A_0 B_0} \otimes \Omega_x^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1} \left( |\eta\rangle\langle\eta|^{A_0 B_0 \tilde{A}_0 \tilde{B}_0} \right) \right) \right] \end{aligned} \quad (\text{B3})$$

where we used (B1), and the state  $|\eta\rangle^{A_0 B_0 \tilde{A}_0 \tilde{B}_0}$  is the purification of the normalized state  $\frac{1}{d_{B_0}} \eta^{A_0 B_0}$ . That is,

$$|\eta\rangle^{A_0 B_0 \tilde{A}_0 \tilde{B}_0} := \frac{1}{\sqrt{d_{B_0}}} \left( \sqrt{\eta^{A_0 B_0}} \otimes I^{\tilde{A}_0 \tilde{B}_0} \right) |\phi_+^{A_0 \tilde{A}_0}\rangle |\phi_+^{B_0 \tilde{B}_0}\rangle.$$

We therefore conclude that

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &= \max \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_x^{AB_0} \left( \text{id}^{A_0 B_0} \otimes \Omega_x^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1} \left( |\eta\rangle\langle\eta|^{A_0 B_0 \tilde{A}_0 \tilde{B}_0} \right) \right) \right] \\ \text{subject to: } &\text{Tr}_{A_0 \tilde{A}_0 \tilde{B}_0} \left[ |\eta\rangle\langle\eta|^{A_0 B_0 \tilde{A}_0 \tilde{B}_0} \right] = u^{B_0} \text{ and } P_x^{AB_0} \geq 0 \quad \forall x = 1, \dots, d_{B_1} \end{aligned} \quad (\text{B4})$$

Note that we can think of system  $B_0$  above in Alice's system. Denoting it by  $A_2$  (hence  $d_{A_2} = d_{B_0}$ ), and expressing

$$|\eta\rangle\langle\eta|^{A_0 A_2 \tilde{A}_0 \tilde{B}_0} = \text{id}^{A_2} \otimes \mathcal{V}^{\tilde{A}_2 \rightarrow A_0 \tilde{A}_0 \tilde{B}_0} \left( \frac{1}{d_{A_2}} \phi_+^{A_2 \tilde{A}_2} \right)$$

with  $\mathcal{V}^{\tilde{A}_2 \rightarrow A_0 \tilde{A}_0 \tilde{B}_0}$  is an isometry, we get

$$\begin{aligned} 2^{-H_{\min}^{\text{ext}}(B|A)_\Omega} &= \max \sum_{x=1}^{d_{B_1}} \text{Tr} \left[ P_x^{A_0 A_1 A_2} \left( \text{id}^{A_0 A_2} \otimes \Omega_x^{\tilde{A}_0 \tilde{B}_0 \rightarrow A_1} \right) \circ \left( \text{id}^{A_2} \otimes \mathcal{V}^{\tilde{A}_2 \rightarrow A_0 \tilde{A}_0 \tilde{B}_0} \right) \left( \phi_+^{A_2 \tilde{A}_2} / d_{A_2} \right) \right] \\ \text{subject to: } &\mathcal{V}^{\tilde{A}_2 \rightarrow A_0 \tilde{A}_0 \tilde{B}_0} \text{ being isometry, and } P_x^{A_0 A_1 A_2} \geq 0 \quad \forall x = 1, \dots, d_{B_1}. \end{aligned} \quad (\text{B5})$$

This optimization problem is illustrated in Fig. 6.