This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

A general theory of comparison of quantum channels (and beyond)

Anna Jenčová

Abstract—**We present a general theory of comparison of quantum channels, concerning with the question of simulability or approximate simulability of a given quantum channel by allowed transformations of another given channel. We introduce a modification of conditional min-entropies, with respect to the set F of allowed transformations, and show that under some conditions on F, these quantities characterize approximate simulability. If F is the set of free superchannels in a quantum resource theory of processes, the modified conditional min-entropies form a complete set of resource monotones. If the transformations in F consist of a preprocessing and a postprocessing of specified forms, approximate simulability is also characterized in terms of success probabilities in certain guessing games, where a preprocessing of a given form can be chosen and the measurements are restricted. These results are applied to several specific cases of simulability of quantum channels, including postprocessings, preprocessings and processing of bipartite channels by LOCC superchannels and by partial superchannels, as well as simulability of sets of quantum measurements.**

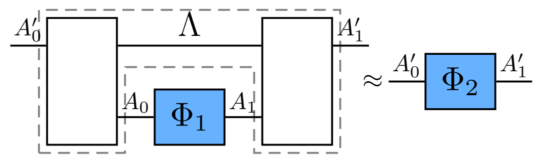
These questions are first studied in a general setting that is an extension of the framework of general probabilistic theories (GPT), suitable for dealing with channels. Here we prove a general theorem that shows that approximate simulability can be characterized by comparing outcome probabilities in certain tests. This result is inspired by the classical Le Cam randomization criterion for statistical experiments and contains its finite dimensional version as a special case.

Index Terms—**superchannels, quantum channel simulation, modified min-entropy, success probabilities**

I. Introduction

For a pair of quantum channels Φ_1 and Φ_2, we consider the following problem: is it possible to simulate Φ_2 by transforming Φ_1 by a quantum network of a specified type? Since quantum channels are the fundamental objects in quantum information theory, this question subsumes a variety of special cases already studied extensively in the literature: comparison of statistical experiments [1], [2], simulability of measurements [3], [4] or more general comparison of channels [5], [6], [7], [8], [9], [10], [11], [12]. In fact, this kind of questions goes back to the classical theory of comparison of classical statistical experiments [13], [14] (see also [15]). This problem can be also put into the setting of resource theories of processes [16], [17], [18], [19] by choosing the allowed maps to be the free operations in the theory, whence it becomes the important question of convertibility of one device to another using free operations.

Transformations between quantum channels are given by superchannels, consisting of a preprocessing and a postprocessing channel connected by an ancilla, [20], [9]. So the question is the equality or approximate equality



where Λ is a superchannel from a given family.

Two types of characterizations of simulability are mostly discussed: either by inequalities in (some modification of) conditional min-entropy (e.g. [8], [9]), or in terms of success probabilities in some discrimination tasks [5], [12]. These two characterizations are closely related, in fact, the latter can be seen as an operational interpretation of the former. These conditions provide a complete set of monotones in the given resource theory.

In the more general situation where the target channel is simulated only approximately, there are different possible approaches as to e.g. the distance measures used to assess the accuracy of the approximation. A common choice is the diamond norm, which is natural since it is well known as the distinguishability norm for channels [21]. With this choice the problem becomes a direct extension of the problem of the classical theory of comparison of statistical experiments. This framework also includes some more specific cases such as quantum dichotomies [22] (pairs of states), where we can restrict to channels that simulate one of the states exactly, but the other may differ from the target.

Similarly as in the case of exact simulations, the aim of this paper is to characterize the diamond norm accuracy of the approximation in two ways: by inequalities in terms of quantities that can be seen as modifications of the conditional min entropy and by comparison of success probabilities in some types of guessing games. The crucial observation for the first type of characterisation is the fact that the conditional min entropy is related to a norm: for any state ρ on a composite system AB, we have

$$H_{\min}(B|A)_\rho = -\log \|\rho\|^\diamond_{B|A}.$$

Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia, jenca@mat.savba.sk

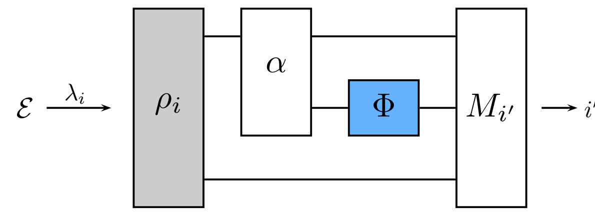
A part of this paper was presented at the Algebraic and Statistical ways into Quantum Resource Theories workshop at BIRS in July 2019

Using the Choi isomorphism, the set of operators on the bipartite Hilbert space $\mathcal{H}_{AB}$ can be seen as the dual space of the set of linear maps $B(\mathcal{H}_A) \to B(\mathcal{H}_B)$ and with this identification, $\|\cdot\|^\diamond_{A|B}$ is the dual norm to the diamond norm. A corresponding fact for semifinite von Neumann algebras was also used in [23] to extend the conditional min entropy characterization of the majorization ordering of bipartite states to the infinite dimensional setting.

The above duality relation of $H_{\min}$ and the diamond norm is based on affine duality of convex sets, this was observed also in [24] and used to extend $H_{\min}$ to quantum networks and their SDP optimization. While our primary interest lies in simulability of quantum channels, in the first part we will work in a broader setting that in a sense is an extension of general probabilistic theories (GPTs), suitable for discussion of various types of quantum channels and networks. The setting is introduced as a category called BS and the allowed transformations are specified as a convex subcategory F. The operational theories of [25] and higher-order theories of [26] can be identified with special objects in BS, in particular, this includes the sets of channels, superchannels and networks of different types. Moreover, the distinguishability norms and the duality described above are naturally defined here. The proof of the main results, in the GPT setting and for quantum channels, is based on the properties of these norms summarized in Proposition 1, together with the minimax theorem. The advantage of this approach is that it captures the basic mathematical structure that lies behind the result and, more importantly, it is applicable to a variety of specific cases. We also remark that it can be applied e.g. to the case when multiple copies of the channels are studied, under various parallel or sequential schemes.

For two elements b_1, b_2 of objects of F, the (one-way) F-conversion distance $\delta_{\mathsf{F}}(b_1\|b_2)$ is defined as the minimal distance we can get to b_2 by allowed transformations of b_1. The main result in the GPT setting (Theorem 2) can be interpreted as the fact that δ_{F} can be characterized by comparing the outcome probabilities of certain tests applied to b_1 and b_2. It is also noted that our setting includes comparison of classical statistical experiments (in the simple finite dimensional case), here Theorem 2 becomes the Le Cam randomization criterion [14] of the classical statistical decision theory. The result is then applied to the case of quantum channels, where we obtain a characterization of δ_{F} in terms of quantities that can be interpreted as modifications of the conditional min entropy (Theorem 3). In the setting of resource theories, where morphisms in F coincide with free superchannels, we show that under some mild assumptions these quantities form a complete set of resource monotones (see e.g. [16], [17], [18], [19]).

We then turn to the characterization of δ_{F} in terms of guessing games. Here we use a connection between the conditional min entropy and success probabilities that is specific to the quantum case, so this is studied only for quantum channels (apart from examples 5 and 7 where statistical experiments and measurements are treated within the GPT framework). This connection is obtained from an isomorphism between quantum channels and bipartite measurements which is close to the Choi isomorphism. We assume that the superchannels in F consist of preprocessings and postprocessings belonging to given sets $\mathcal{C}_{\mathrm{pre}}$ and $\mathcal{C}_{\mathrm{post}}$. In this case, we find some sufficient conditions (Theorem 4) under which δ_{F} is characterized by success probabilities in guessing games of the following general form. Given an ensemble $\mathcal{E} = \{\lambda_i, \rho_i\}$ of states ρ_i with prior probabilities λ_i, the guessing game is depicted in the diagram:



Here Φ is the channel in question (Φ_1 or Φ_2), we may choose the preprocessing α and the measurement M from some allowed sets $\mathcal{C}_{\mathrm{pre}}$ and $\mathcal{M}_{\mathrm{post}}$. We also permit an ancilla between α and M, but this might be restricted by the allowed sets.

These results are applied to some special cases: postprocessing, preprocessing, processing of bipartite channels by LOCC superchannels and by partial superchannels. We remark that, similarly as LOCC, Theorem 4 can be applied to other cases of restricted resource theories, such as PPT or SEP. In these cases, $\mathcal{C}_{\mathrm{pre}}$, $\mathcal{C}_{\mathrm{post}}$ and $\mathcal{M}_{\mathrm{post}}$ consist of LOCC (PPT/SEP) channels resp. measurements, see Section III-E.

In the case of postprocessings, we obtain previously known results [10]: δ_{F} is characterized by comparing the success probabilities of ensembles $(\Phi_i \otimes id)(\mathcal{E})$, see Section III-C. For preprocessings, there is only one fixed measurement M in the guessing game but the preprocessings can be chosen freely. We also characterize the related preprocessing pseudodistance as a Hausdorff distance of ranges of the two channels tensored with identity, Section III-D. The description of the guessing game for partial processings can be found in Section III-F. As another example, we treat classical simulability for sets of quantum measurements and show that it can be formulated by processing of a certain bipartite channel by a classical-to-classical partial superchannel. We show that approximate simulability is characterized by success probabilities, where no ancilla is needed in the guessing game, see Section III-G.

The outline of the paper is as follows. We start with the general GPT formulation in Section II. Here the category BS is introduced and the properties of the corresponding norms and their duality are discussed. We then treat the comparison in GPT, the main result is formulated in Theorem 2. In Section III we specialize to quantum channels. We first introduce the basic notions and show that the sets of channels and superchannels are objects in the category BS. We also recall the connections between the diamond norm, $H_{\min}$ and success probabilities of guessing games (Section III-A7) that are crucial in further sections. The main results of this part are Theorem 3, characterizing δ_{F} in terms of modified conditional min entropies, and Theorem 4, characterizing δ_{F} by success probabilities. Applications of these results are contained in Sections III-C - III-G.

II. The GPT formulation

General probabilistic theories (GPT) form a framework for description of a large class of physical theories involving probabilistic processes, see [27] for an introduction and background. This framework is built upon basic notions of states and effects and under some general assumptions on the theories, it can be put into the setting of the theory of (finite dimensional) ordered vector spaces. The classical and quantum theories are special cases of a GPT (see Example 2 below), which allows to study some well known quantum phenomena in a broader context. This is especially useful in the investigation of the mathematical foundations of quantum theory. For us, it is important that this setting describes the basic mathematical structure underlying the problem of approximate channel simulability and can be applied in many different situations.

The basic object in GPT is the set of states of a physical system in the theory, represented as a compact convex subset of an Euclidean space. Such a set can be always seen as a base of a closed convex cone in a finite dimensional real vector space. It is clear that the set of channels, or physical transformations of the systems in the theory, has a convex structure as well and can be, at least formally, treated as a "state space" in the convenient framework of GPT. For example, the set of quantum channels was considered in this way in [28]. However, observe that the set of quantum channels is, by definition, a special subset of the cone of completely positive maps, but it no longer forms a base of this cone. We will therefore need a somewhat more general representation of compact convex sets, described in the next paragraph.

A. Base sections and corresponding norms

Let $\mathcal{V}$ be a finite dimensional real vector space and let $\mathcal{V}^+\subset\mathcal{V}$ be a closed convex cone which is pointed ($\mathcal{V}^+\cap(-\mathcal{V}^+)=\{0\}$) and generating ($\mathcal{V}=\mathcal{V}^+-\mathcal{V}^+$). Below, we will say that $\mathcal{V}^+$ is a proper cone. The cone $\mathcal{V}^+$ defines a partial order in $\mathcal{V}$, defined by $x\le y$ if $y-x\in\mathcal{V}^+$. Thus the pair $(\mathcal{V},\mathcal{V}^+)$ is an ordered vector space.

A convex subset $B\subset\mathcal{V}^+$ is called a base section in $(\mathcal{V},\mathcal{V}^+)$ if

(i) B is the base of the cone $\mathcal{V}^+\cap\operatorname{span}(B)$;
(ii) $B\cap\operatorname{int}(\mathcal{V}^+)\neq\emptyset$.

Base sections were studied in [11, Appendix] and in [29] in the special case of the space of hermitian linear operators with the cone of positive operators. In this paragraph, we summarize some of the results.

For the ease of the presentation, it will be convenient to introduce the category BS whose objects are finite dimensional real vector spaces $\mathcal{V}$ endowed with a fixed proper cone $\mathcal{V}^+\subset\mathcal{V}$ and a base section $B(\mathcal{V})$ in $(\mathcal{V},\mathcal{V}^+)$. The morphisms $\Lambda:\mathcal{V}\to\mathcal{W}$ in BS are linear maps such that $\Lambda(\mathcal{V}^+)\subseteq\mathcal{W}^+$ and $\Lambda(B(\mathcal{V}))\subseteq B(\mathcal{W})$, [30].

For $\mathcal{V}\in\mathsf{BS}$, we define the dual object $\mathcal{V}^*\in\mathsf{BS}$ as the dual vector space with the dual cone

$$\mathcal{V}^{*+}:=\{\varphi\in\mathcal{V}^*,\langle\varphi,c\rangle\ge 0,\ \forall c\in\mathcal{V}^+\}$$

and the dual base section

$$B(\mathcal{V}^*)=\tilde{B}(\mathcal{V}):=\{\varphi\in\mathcal{V}^{*+},\ \langle\varphi,b\rangle=1,\ \forall b\in B(\mathcal{V})\}.$$

Note that indeed, $\mathcal{V}^{*+}$ is a proper cone in $\mathcal{V}^*$ and $B(\mathcal{V}^*)$ is a base section in $(\mathcal{V}^*,\mathcal{V}^{*+})$. Moreover, $\mathcal{V}^{**}\simeq\mathcal{V}$ in BS, where the isomorphism is given by the natural vector space isomorphism $\mathcal{V}\to\mathcal{V}^{**}$.

Let us denote

$$\begin{aligned}[-B(\mathcal{V}),B(\mathcal{V})]&:=\{x\in\mathcal{V},\ \exists b\in B(\mathcal{V}),-b\le x\le b\}\\&=\{c_1-c_2,\ c_1,c_2\in\mathcal{V}^+,\ c_1+c_2\in B(\mathcal{V})\}.\end{aligned}$$

Then we have

$$\begin{aligned}[0,B(\mathcal{V})]:=&\{x\in\mathcal{V}^+,\ \exists b\in B(\mathcal{V}),\ x\le b\}\\=&\mathcal{V}^+\cap[-B(\mathcal{V}),B(\mathcal{V})].\end{aligned}\tag{1}$$

We now define a norm in $\mathcal{V}$ as

$$\|x\|_{\mathcal{V}}:=\max_{\psi\in[-B(\mathcal{V}^*),B(\mathcal{V}^*)]}\langle\psi,x\rangle.$$

The following proposition summarizes some properties of these norms. Note that since $B(\mathcal{V})$ and $B(\mathcal{V}^*)$ are given by positivity and linear constraints, part (iii) below can be formulated as the primal and dual conic program for computing the norm $\|\cdot\|_{\mathcal{V}}$.

Proposition 1. *Let $\mathcal{V}\in\mathsf{BS}$. Then*

(i) *The unit ball of $\|\cdot\|_{\mathcal{V}}$ is $[-B(\mathcal{V}),B(\mathcal{V})]$, so that*
$$\|x\|_{\mathcal{V}}=\min_{b\in B(\mathcal{V})}\min\{\lambda>0,\ -\lambda b\le x\le\lambda b\};$$
(ii) *the norms $\|\cdot\|_{\mathcal{V}}$ and $\|\cdot\|_{\mathcal{V}^*}$ are mutually dual;*
(iii) *if $c\in\mathcal{V}^+$, then*
$$\begin{aligned}\|c\|_{\mathcal{V}}&=\max_{\varphi\in B(\mathcal{V}^*)}\langle\varphi,c\rangle\\&=\min_{b\in B(\mathcal{V})}\min\{\lambda>0,\ c\le\lambda b\};\end{aligned}$$
(iv) *if $b_1,b_2\in B(\mathcal{V})$, then*
$$\frac{1}{2}\|b_1-b_2\|_{\mathcal{V}}=\max_{\varphi\in[0,B(\mathcal{V})]}\langle\varphi,b_1-b_2\rangle.$$
(v) *Let $\Lambda:\mathcal{V}\to\mathcal{W}$ be a morphism in BS. Then Λ is a contraction with respect to the corresponding norms:*
$$\|\Lambda(x)\|_{\mathcal{W}}\le\|x\|_{\mathcal{V}},\qquad\forall x\in\mathcal{V}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

We next discuss some basic examples of objects in BS.

Example 1 (GPT state spaces). The state spaces of GPT can be also seen as objects in BS. Indeed, let K be any compact convex subset in $\mathbb{R}^N$, then there is a finite dimensional real vector space $\mathcal{V}$ with a proper cone $\mathcal{V}^+$, such that K is a base of $\mathcal{V}^+$. This means that there is some functional $u \in int((\mathcal{V}^+)^*)$ such that

$$K = \{c \in \mathcal{V}^+,\ \langle u, c\rangle = 1\}.$$

The space $\mathcal{V}$ with $\mathcal{V}^+$ and $B(\mathcal{V}) = K$ is clearly an object in BS. For the dual object, we have $B(\mathcal{V}^*) = \{u\}$. The norm $\|\cdot\|_{\mathcal{V}}$ is the base norm with respect to K and $\|\cdot\|_{\mathcal{V}^*}$ is the order unit norm with respect to u. In this way, the category BS is a common generalization of order unit and base normed spaces.

Example 2 (Classical and quantum state spaces). The prototypical examples in GPT are the classical state space of probability distributions over a finite set, and the quantum state space of all density density operators on a finite dimensional Hilbert space. In the classical GPT, we have $\mathcal{V} = \mathbb{R}^n$, with the simplicial cone $\mathcal{V}^+ = (\mathbb{R}^+)^n$ and $K = \{(p_1, \ldots, p_n) \in \mathcal{V}^+,\ \sum_i p_i = 1\}$ is the probability simplex. The morphisms in BS between such spaces are precisely the stochastic maps, given by stochastic matrices. The norm $\|\cdot\|_{\mathcal{V}} = \|\cdot\|_1$ is the L_1-norm. The dual ordered vector space is affinely isomorphic to $(\mathcal{V}, \mathcal{V}^+)$ and the unit functional $u = 1_n := (1, \ldots, 1)$. The dual norm is the L_∞-norm $\|\cdot\|_{\mathcal{V}^*} = \|\cdot\|_\infty$. In the quantum case, $(\mathcal{V}, \mathcal{V}^+)$ is the space of hermitian operators on a finite dimensional Hilbert space with the cone of positive operators and the morphisms are positive trace preserving maps. The dual ordered vector space is again affinely isomorphic to $(\mathcal{V}, \mathcal{V}^+)$ and the unit functional is the trace, $u = \mathrm{Tr}$. The two dual norms are the trace norm and the operator norm, respectively.

Example 3 (Quantum channels). The prototypical example of an object in BS is the set of quantum channels which is a base section in the vector space of hermitian maps with the cone of completely positive maps. More details will be given in Section III-A5.

Example 4 (BS morphisms). Let $\mathcal{V}$ and $\mathcal{W}$ be two objects in BS and let $\mathcal{L} = \mathcal{L}(\mathcal{V}, \mathcal{W})$ be the space of linear maps $\mathcal{V} \to \mathcal{W}$. With $\mathcal{L}^+$ the cone of positive maps and $B(\mathcal{L}) = \mathsf{BS}(\mathcal{V}, \mathcal{W})$ the set of all morphisms $\mathcal{V} \to \mathcal{W}$, it is easily seen that $\mathcal{L}$ is again an object in BS. It is this property that makes this category especially useful for description of channels and higher order theories.

In general, an object in BS can be interpreted as a set of special states or devices of a physical theory, so we can consider the problem of testing or performing measurements over such sets. A test (or yes-no measurement) is identified with an affine map $B(\mathcal{V}) \to [0, 1]$, assigning to each element the probability of the "yes" outcome. Such maps correspond to elements $\varphi \in [0, B(\mathcal{V}^*)]$, similarly to the GPT setting, these elements will be called effects. Any measurement with k outcomes is given by a k-tuple of effects $\{\varphi_i\}_{i=1}^k$ such that $\sum_i \varphi_i \in B(\mathcal{V}^*)$. Note that in the case of quantum channels this corresponds precisely to the quantum testers of [20].

It is not difficult to see that the restriction of $\|\cdot\|_{\mathcal{V}}$ to $\mathrm{span}(B(\mathcal{V}))$ coincides with the base norm with respect to the base $B(\mathcal{V})$ of the cone $\mathcal{V}^+ \cap \mathrm{span}(B(\mathcal{V}))$. The advantage of our extended definition is that the dual space is now an object of the same category. As the base norm, $\|\cdot\|_{\mathcal{V}}$ has an operational interpretation as a distinguishability norm for elements in $B(\mathcal{V})$. It can be seen that this holds also if the discrimination procedures are given by the effects as above, indeed, using Proposition 1 (iv) we see that that the optimal probability of correctly distinguishing two elements $b, b' \in B(\mathcal{V})$ each of which has a prior probability 1/2 is $\frac{1}{2}(1 + \frac{1}{2}\|b - b'\|_{\mathcal{V}})$.

More generally, an ensemble on $\mathcal{V}$ is a finite sequence $\mathcal{E} = \{\lambda_i, x_i\}_{i=1}^k$ of elements $x_i \in B(\mathcal{V})$ and prior probabilities λ_i. The interpretation is that x_i is prepared with probability λ_i and the task is to guess which element was prepared. Any guessing procedure is described by a k-outcome measurement $\psi = \{\psi_i\}$, the value $\langle \psi_i, x_j\rangle$ is interpreted as the probability of guessing i if the true state was x_j. The average success probability using $\psi = (\psi_i)$ is

$$P_{\mathrm{succ}}(\mathcal{E}, \psi) := \sum_i \lambda_i \langle \psi_i, x_i\rangle$$

and the optimal success probability for $\mathcal{E}$ is

$$P_{\mathrm{succ}}(\mathcal{E}) := \max_{\psi} P_{\mathrm{succ}}(\mathcal{E}, \psi).$$

B. Comparison in GPT

We now formulate the comparison problem in the above setting and prove a general theorem which in later sections will be applied to quantum channels. Assume that a subcategory F in BS is given, such that for any $\mathcal{V}, \mathcal{W} \in \mathsf{F}$, the set $\mathsf{F}(\mathcal{V}, \mathcal{W})$ of all morphisms $\mathcal{V} \to \mathcal{W}$ in F is convex (we will say in this case that F is a convex subcategory). For two objects $\mathcal{V}_1, \mathcal{V}_2 \in \mathsf{F}$, let $b_1 \in B(\mathcal{V}_1)$, $b_2 \in B(\mathcal{V}_2)$. The (one-way) F-conversion distance $\delta_{\mathsf{F}}(b_1\|b_2)$ is defined as the minimum distance we can get to b_2 by images of b_1 under all morphisms $\mathcal{V}_1 \to \mathcal{V}_2$ in F:

$$\delta_{\mathsf{F}}(b_1\|b_2) := \inf_{\Lambda \in \mathsf{F}(\mathcal{V}_1, \mathcal{V}_2)} \|\Lambda(b_1) - b_2\|_{\mathcal{V}_2}.$$

We also define the F-distance of b_1 and b_2 as

$$\Delta_{\mathsf{F}}(b_1, b_2) := \max\{\delta_{\mathsf{F}}(b_1\|b_2), \delta_{\mathsf{F}}(b_2\|b_1)\}.$$

Proposition 2. *Δ_{F} is a pseudometric on the set $\{b \in B(\mathcal{V}),\ \mathcal{V} \in \mathsf{F}\}$.*

Proof. The only thing to prove is the triangle inequality. So let $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \in \mathsf{F}$ and let $b_i \in \mathcal{V}_i$, $i = 1, 2, 3$. Let

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

$\Theta \in \mathsf{F}(\mathcal{V}_2, \mathcal{V}_3)$ and let for $\mu > 0$, $\Lambda_\mu \in \mathsf{F}(\mathcal{V}_1, \mathcal{V}_2)$ be such that $\delta_{\mathsf{F}}(b_1 \| b_2) + \mu \geq \|\Lambda_\mu(b_1) - b_2\|_{\mathcal{V}_2}$. Then

$$\begin{aligned}\delta_{\mathsf{F}}(b_1\|b_3) &\leq \|\Theta \circ \Lambda_\mu(b_1) - b_3\|_{\mathcal{V}_3} \\ &\leq \|\Theta(\Lambda_\mu(b_1) - b_2)\|_{\mathcal{V}_3} + \|\Theta(b_2) - b_3\|_{\mathcal{V}_3} \\ &\leq \mu + \delta_{\mathsf{F}}(b_1\|b_2) + \|\Theta(b_2) - b_3\|_{\mathcal{V}_3}.\end{aligned}$$

Since this holds for all $\Theta \in \mathsf{F}(\mathcal{V}_2, \mathcal{V}_3)$ and all $\mu > 0$, we get the result.

□

The following data processing inequalities for δ_{F} follow easily from Prop. 1 (v) (cf. [10, Prop. 3]).

Proposition 3. *Let $\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3 \in \mathsf{F}$, $b_i \in B(\mathcal{V}_i)$, $i = 1, 2, 3$. Then*

(i) *For any $\Lambda \in \mathsf{F}(\mathcal{V}_1, \mathcal{V}_2)$, $\delta_{\mathsf{F}}(\Lambda(b_1)\|b_3) \geq \delta_{\mathsf{F}}(b_1\|b_3)$;*
(ii) *For any $\Theta \in \mathsf{F}(\mathcal{V}_2, \mathcal{V}_3)$, $\delta_{\mathsf{F}}(b_1\|\Theta(b_2)) \leq \delta_{\mathsf{F}}(b_1\|b_2)$;*

We now turn to the main result of this section, contained in Theorem 2 below. Note that since the elements $\varphi \in \mathcal{V}^{*+}$, $\psi \in \mathcal{W}^{*+}$ in (ii) and (iii) of this theorem can always be normalized and the positive part of the unit ball of $\|\cdot\|_{\mathcal{V}^*}$ coincides with the set of all tests on $\mathcal{B}(\mathcal{V})$, this theorem says that the conversion distance can be characterized by comparing the probabilities of the "yes" outcome for certain tests applied to b_1 and b_2.

For the proof of Theorem 2, we will need the following minimax theorem (cf. [15, Thm.48.5]).

Theorem 1 (Minimax theorem)**.** *Let T be a convex and compact subset of a locally convex space and let Y be a convex subset of a vector space. Let $f : T \times Y \to \mathbb{R}$ be convex in y and continuous and concave in t. Then*

$$\inf_{y\in Y} \sup_{t\in T} f(t, y) = \sup_{t\in T} \inf_{y\in Y} f(t, y).$$

Theorem 2. *Let $\mathcal{V}_1, \mathcal{V}_2 \in \mathsf{F}$ and let $b_1 \in B(\mathcal{V}_1)$, $b_2 \in B(\mathcal{V}_2)$. Let $\epsilon \geq 0$. Then the following are equivalent.*

(i) *$\delta_{\mathsf{F}}(b_1\|b_2) \leq \epsilon$;*
(ii) *for all $\varphi \in \mathcal{V}_2^{*+}$, there is some $\Theta \in \bar{\mathsf{F}}(\mathcal{V}_1, \mathcal{V}_2)$ such that*

$$\langle \varphi, b_2\rangle \leq \langle \varphi, \Theta(b_1)\rangle + \frac{\epsilon}{2}\|\varphi\|_{\mathcal{V}^*},$$

here $\bar{\mathsf{F}}(\mathcal{V}_1, \mathcal{V}_2)$ is the closure of $\mathsf{F}(\mathcal{V}_1, \mathcal{V}_2)$ in $\mathsf{BS}(\mathcal{V}_1, \mathcal{V}_2)$;
(iii) *for all $\mathcal{W} \in \mathsf{F}$ and all $\psi \in \mathcal{W}^{*+}$, we have*

$$\begin{aligned}&\sup_{\Theta\in\mathsf{F}(\mathcal{V}_2,\mathcal{W})} \langle \psi, \Theta(b_2)\rangle \\ &\qquad \leq \sup_{\Theta'\in\mathsf{F}(\mathcal{V}_1,\mathcal{W})} \langle \psi, \Theta'(b_1)\rangle + \frac{\epsilon}{2}\|\psi\|_{\mathcal{W}^*}.\end{aligned}$$

Proof. Let $\epsilon' > \epsilon$ and let $\Lambda_0 \in \mathsf{F}(\mathcal{V}_1, \mathcal{V}_2)$ be such that $\|\Lambda_0(b_1) - b_2\|_{\mathcal{V}_2} \leq \epsilon'$. Let $\mathcal{W} \in \mathsf{F}$ and let $\psi \in \mathcal{W}^{*+}$. For any $\Theta \in \mathsf{F}(\mathcal{V}_2, \mathcal{W})$, we have

$$\begin{aligned}\langle \psi, \Theta(b_2)\rangle &= \langle \psi, \Theta \circ \Lambda_0(b_1)\rangle + \langle \psi, \Theta(b_2 - \Lambda_0(b_1))\rangle \\ &\leq \sup_{\Theta'\in\mathsf{F}(\mathcal{V}_1,\mathcal{W})} \langle \psi, \Theta'(b_1)\rangle \\ &\quad + \frac{1}{2}\|\Theta(b_2 - \Lambda_0(b_1))\|_{\mathcal{W}}\|\psi\|_{\mathcal{V}^*} \\ &\leq \sup_{\Theta'\in\mathsf{F}(\mathcal{V}_1,\mathcal{W})} \langle \psi, \Theta'(b_1)\rangle + \frac{\epsilon'}{2}\|\psi\|_{\mathcal{V}^*},\end{aligned}$$

where we have used Prop. 1 (iv) and (v). Since this holds for all $\Theta \in \mathsf{F}(\mathcal{V}_2, \mathcal{W})$ and $\epsilon' > \epsilon$, this proves that (i) implies (iii). Since (iii) obviously implies (ii), it is enough to prove (ii) $\implies$ (i).

By Prop. 1 (iv), we have

$$\frac{1}{2}\delta_{\mathsf{F}}(b_1\|b_2) = \inf_{\Lambda\in\mathsf{F}(\mathcal{V}_1,\mathcal{V}_2)} \sup_{\varphi\in[0,B(\mathcal{V}_2^*)]} \langle \varphi, b_2 - \Lambda(b_1)\rangle.$$

Note that by Prop. 1 (i), the set $[0, B(\mathcal{V}_2^*)] = \mathcal{V}_2^{*+} \cap [-B(\mathcal{V}_2^*), B(\mathcal{V}_2^*)]$ is convex and compact and the map $(\varphi, \Lambda) \mapsto \langle \varphi, b_2 - \Lambda(b_1)\rangle$ is linear in both components, so that we may apply the minimax theorem (Thm. 1). Assume that (ii) holds, then

$$\begin{aligned}\frac{1}{2}\delta_{\mathsf{F}}(b_1\|b_2) &= \sup_{\varphi\in[0,B(\mathcal{V}_2^*)]} \inf_{\Lambda\in\mathsf{F}(\mathcal{V}_1,\mathcal{V}_2)} \langle \varphi, b_2 - \Lambda(b_1)\rangle \\ &= \sup_{\varphi\in[0,B(\mathcal{V}_2^*)]} (\langle \varphi, b_2\rangle - \sup_{\Lambda\in\mathsf{F}(\mathcal{V}_1,\mathcal{V}_2)} \langle \varphi, \Lambda(b_1)\rangle) \\ &\leq \frac{\epsilon}{2}.\end{aligned}$$

□

We next some examples of an application of Theorem 2, the first two of which show the relation to the theory of comparison of statistical experiments. In these examples, $\mathcal{V} \in \mathsf{BS}$ is such that $K = B(\mathcal{V})$ is a base of $\mathcal{V}^+$. The dual object $\mathcal{V}^*$ is the order unit space $(\mathcal{V}^*, \mathcal{V}^{*+}, u)$, where $u \in \mathcal{V}^{*+}$ is the functional determined by $\langle u, x\rangle = 1$ for all $x \in K$.

Example 5 (Statistical experiments)*.* A (finite) statistical experiment in $\mathcal{V}$ is a finite set $x_1, \ldots, x_k \in K$. The set of all experiments with fixed k and $\mathcal{V}$ is an object in BS. Indeed, let $\mathcal{V}^k = \oplus_{i=1}^k \mathcal{V}$ and $\mathcal{V}^{k+} = \oplus_{i=1}^k \mathcal{V}^+$, then $K^k = \oplus_{i=1}^k K$ is a base section in the ordered vector space $(\mathcal{V}^k, \mathcal{V}^{k+})$. As for the dual object, $\mathcal{V}^{k*} = \oplus_{i=1}^k \mathcal{V}^*$ and $\mathcal{V}^{k*+} = \oplus_{i=1}^k \mathcal{V}^{*+}$. Moreover, it is easily checked that the dual section is

$$\mathcal{B}(\mathcal{V}^{k*}) = \tilde{K}^k = \{(p_1 u, \ldots, p_k u),\ p_i \in [0, 1],\ \sum_i p_i = 1\}$$

and

$$\|v\|_{\mathcal{V}^k} = \max_i \|v_i\|_{\mathcal{V}}, \qquad v = (v_i) \in \mathcal{V}^k$$
$$\|\psi\|_{\mathcal{V}^{k*}} = \sum_i \|\psi_i\|_{\mathcal{V}^*}, \qquad \psi = (\psi_i) \in \mathcal{V}^{k*}.$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

Let F be the subcategory whose objects are statistical experiments with fixed k and morphisms in $\mathsf{F}(\mathcal{V}^k, \mathcal{W}^k)$ are given by Φ^k with $\Phi \in \mathcal{F}(\mathcal{V}, \mathcal{W}) \subseteq \mathsf{BS}(\mathcal{V}, \mathcal{W})$ for some convex subset $\mathcal{F}(\mathcal{V}, \mathcal{W})$.

Let $x = (x_i) \in B(\mathcal{V}^k)$, $y = (y_i) \in B(\mathcal{W}^k)$, then the F-conversion distance has the form

$$\delta_{\mathsf{F}}(x\|y) = \inf_{\Phi\in\mathcal{F}} \max_i \|\Phi(x_i) - y_i\|_{\mathcal{V}}.$$

Under an obvious normalization, Theorem 2 (iii) says that $\delta_{\mathsf{F}}(x\|y) \leq \epsilon$ if and only if for any $(\psi_i) \in \mathcal{V}^{k*+}$ with $\sum_i \|\psi_i\|_{\mathcal{V}^*} \leq 1$, we have

$$\sup_{\Phi\in\mathcal{F}} \sum_i \langle \psi_i, \Phi(y_i)\rangle \leq \sup_{\Phi'\in\mathcal{F}} \sum_i \langle \psi_i, \Phi'(y_i)\rangle + \frac{\epsilon}{2}.$$

Since the norm $\|\cdot\|_{\mathcal{V}^*}$ is the order unit norm with respect to u, any (ψ_i) of the above form satisfies $0 \leq \psi_i \leq \|\psi_i\|_{\mathcal{V}^*} u$ and $\sum_i \psi_i \leq \sum_i \|\psi_i\|_{\mathcal{V}^*} u \leq u$. Adding a positive element $\psi_{k+1} = u - \sum_i \psi_i$, the collection $(\psi_i)_{i=1}^{k+1}$ becomes a measurement with $k+1$ outcomes and the value $\frac{1}{k}\sum_i \langle \psi_i, \Phi(y_i)\rangle =: \tilde{P}_{\mathrm{succ}}(\mathcal{E}, (\psi_i))$ becomes the success probability for the ensemble $\mathcal{E} = \{\frac{1}{k}, \Phi(y_i)\}$ in the inconclusive discrimination with the measurement (ψ_i), here ψ_{k+1} represents the inconclusive outcome. Now we see that $\delta_{\mathsf{F}}(x\|y) \leq \epsilon$ if and only if for any measurement $(\psi_i)_{i=1}^{k+1}$,

$$\sup_{\Phi\in\mathcal{F}} \tilde{P}_{\mathrm{succ}}(\{\frac{1}{k}, \Phi(y_i)\}, (\psi_i)) \leq \sup_{\Phi'\in\mathcal{F}} \tilde{P}_{\mathrm{succ}}(\{\frac{1}{k}, \Phi'(x_i)\}, (\psi_i)) + \frac{\epsilon}{2k},$$

extending the result [12, Cor. 15].

Example 6 (Le Cam randomization criterion for classical statistical experiments). In the setting of the previous example, let us further restrict F to $\mathcal{V}^k$ for classical state spaces $\mathcal{V}$ (see Example 2) and let $\mathcal{F}$ be the set of all stochastic maps. Here the objects of F are sets of classical statistical experiments with k elements. If (p^i) is such an experiment and $q^i = \Phi(p^i)$ for some stochastic map Φ, we say that (q^i) is a randomization of (p^i), so that F is the category of (finite dimensional) classical statistical experiments with randomizations. Moreover, $\delta_{\mathsf{F}}((p^i)\|(q^i))$ is the Le Cam deficiency of (p^i) with respect to (q^i) and Δ_{F} becomes the the Le Cam distance [14].

The basic idea of comparison of statistical experiments that goes back to Blackwell [13] is to compare classical statistical experiments by the performance of decision rules. Here the task is to choose a decision d from a finite set $\{1, \ldots, D\}$ using data that is known to be drawn according to one of the distributions in $\{p^1, \ldots, p^k\} \subset \Delta_m$. The decision rules are given by stochastic maps $\Theta : \Delta_m \to \Delta_D$, where $\Theta(p)(d)$ is the probability that d is chosen if the data was sampled from the distribution $p \in \Delta_m$. To each pair $i = 1, \ldots, k$ and $d = 1, \ldots, D$ a value $g(i,d) \geq 0$ is assigned expressing the gain obtained if d was chosen while the true distribution was p^i. Under a prior distribution $\lambda = (\lambda_1, \ldots, \lambda_k)$, the average gain of the decision rule Θ is given by

$$G(\{p^i\}, \lambda, g, \Theta) = \sum_i \sum_d \lambda_i \Theta(p^i)(d) g(i,d) = \langle \psi, \Theta(\{p^i\})\rangle,$$

where $\psi \in (\mathrm{R}^D_+)^k$ is given by $\psi^i_d = \lambda_i g(i,d)$, note also that we have $\|\psi\|_{\mathcal{V}^{k*}} = \sum_i \lambda_i \max_d g(i,d)$. The celebrated Le Cam randomization criterion [14] says that the deficiency of the experiment (p^i) with respect to (q^i) can be obtained by comparing the optimal values of G achievable by the two experiments, more precisely that $\delta_{\mathsf{F}}((p^i)\|(q^i)) \leq \epsilon$ if and only if for all gain functions g and prior distributions λ, we have

$$\sup_\Theta G(\{q^i\}, \lambda, g, \Theta) \leq \sup_{\Theta'} G(\{p^i\}, \lambda, g, \Theta') + \frac{\epsilon}{2} \sum_i \lambda_i \max_d g(i,d).$$

In the finite dimensional setting, this is precisely the statement of Theorem 2. Therefore, Theorem 2 can be seen as the most general GPT form of the Le Cam randomization theorem.

Similarly, restricting the objects $\mathcal{V}$ to quantum state spaces and letting $\mathcal{F}$ be the set of all quantum channels, we obtain the quantum version of the randomization criterion, cf. [1].

Example 7 (Measurements). A measurement (with k outcomes) is a collection $M = (M_i) \in \mathcal{V}^{k*+}$, $\sum_i M_i = u$. It easy to see that the set $\mathcal{M}_k(\mathcal{V})$ of all such measurements is a base section in $(\mathcal{V}^{k*}, \mathcal{V}^{k*+})$ we will denote this object of BS by $\mathcal{V}^k_{\mathcal{M}}$. The dual object is $\mathcal{V}^k, \mathcal{V}^{k+}$ with the base section

$$\tilde{\mathcal{M}}_k(\mathcal{V}) = \{(x, \ldots, x),\ x \in K\}$$

Note that for $v \in \mathcal{V}^{k+}$, the dual norm is

$$\|v\|_{\mathcal{V}^{k*}_{\mathcal{M}}} = \max_{M\in\mathcal{M}_k(\mathcal{V})} \sum_i \langle M_i, v_i\rangle.$$

Dividing v by $c := \sum_i \langle u, v_i\rangle$ and noting that $c^{-1} v_i = \lambda_i x_i$ with $x_i \in K$ and some probabilities λ_i, we obtain an ensemble: $\mathcal{E} := \{\lambda_i, x_i\}$ such that

$$\|v\|_{\mathcal{V}^{k*}_{\mathcal{M}}} = c P_{\mathrm{succ}}(\mathcal{E}).$$

Again, let F be the subcategory with objects $\mathcal{V}^k_{\mathcal{M}}$ (with k fixed) and morphisms $\Phi^{*k} : \mathcal{V}^k_{\mathcal{M}} \to \mathcal{W}^k_{\mathcal{M}}$ for $\Phi \in \mathcal{F}(\mathcal{W}, \mathcal{V}) \subseteq \mathsf{BS}(\mathcal{W}, \mathcal{V})$. The Theorem 2 tells us that $\delta_{\mathsf{F}}(M\|N) \leq \epsilon$ if and only if, for any ensemble $\mathcal{E}$ on $\mathcal{V}$,

$$P_{\mathrm{succ}}(\mathcal{E}, N) \leq \sup_{\Phi\in\mathcal{F}} P_{\mathrm{succ}}(\Phi(\mathcal{E}), N) + \frac{\epsilon}{2} P_{\mathrm{succ}}(\mathcal{E}),$$

extending the result of [12, Thm. 14].

More examples will be treated in the next section.

III. Comparison of quantum channels

In this section we will present the sets of quantum channels and superchannels as objects in BS and show how Theorem 2 applies, under some conditions on the subcategory F. We need some preparation first.

A. Basic ingredients

In what follows, $\mathcal{H}_A, \mathcal{H}_B, \ldots$ will always denote a finite dimensional Hilbert space, labelled by the system it represents. The Hilbert space will often be referred to by its label, so we denote by $\mathcal{B}(A)$ the set of bounded operators on $\mathcal{H}_A$, similarly, $\mathcal{B}_h(A)$ denotes the set of self-adjoint operators, $\mathcal{B}_+(A)$ the set of positive operators and $\mathfrak{S}(A)$ the set of states on $\mathcal{H}_A$. We will also put $d_A := \dim(\mathcal{H}_A)$ and I_A denotes the identity operator on $\mathcal{H}_A$. The trivial Hilbert space $\mathbb{C}$ will be labeled by 1.

For $W \in \mathcal{B}(A_0)$, we will use the notation (cf. [20])

$$|W\rangle\rangle := \sum_i W|i\rangle_{A_0} \otimes |i\rangle_{A_0} = \sum_i |i\rangle_{A_0} \otimes W^{\mathsf{T}}|i\rangle_{A_0}, \quad (2)$$

here W^{T} denotes the transpose of W in the standard basis $\{|i\rangle\}$.

1) Linear maps and Choi representation: Let $\mathcal{L}(A_0, A_1)$ denote the set of hermitian linear maps $\mathcal{B}(A_0) \to \mathcal{B}(A_1)$, that is, linear maps satisfying

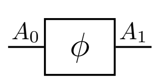
$$\Phi(X^*) = \Phi(X)^*, \qquad X \in \mathcal{B}(A_0).$$

Let $\mathcal{L}_+(A_0, A_1)$ denote the subset of completely positive maps in $\mathcal{L}(A_0, A_1)$ and $\mathcal{C}(A_0, A_1)$ the set of quantum channels, that is, trace preserving maps in $\mathcal{L}_+(A_0, A_1)$.

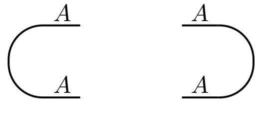
The Choi matrix of $\Phi \in \mathcal{L}(A_0, A_1)$ is defined as $C_\Phi := (\Phi \otimes id_{A_0})(|I_{A_0}\rangle\rangle\langle\langle I_{A_0}|)$. The map $\Phi \mapsto C_\Phi$ establishes a linear isomorphism between $\mathcal{L}(A_0, A_1)$ and $\mathcal{B}_h(A_1A_0)$ that maps $\mathcal{L}_+(A_0, A_1)$ onto $\mathcal{B}_+(A_1A_0)$ and $\mathcal{C}(A_0, A_1)$ onto the set

$$\{X \in \mathcal{B}_+(A_1A_0),\ \mathrm{Tr}_{A_1}[X] = I_{A_0}\}.$$

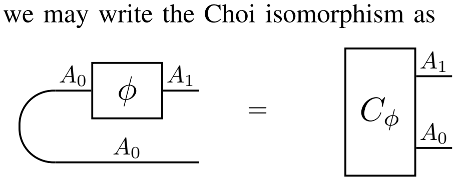
2) Diagrams: We will make use of the common diagrammatic representation of maps in $\mathcal{L}(A_0, A_1)$ as



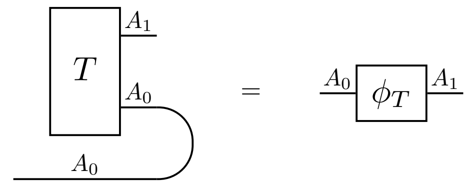
If some of the systems is trivial, the corresponding wire will be omitted. The special symbols



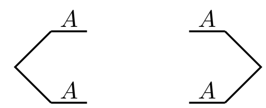
will represent $|I_A\rangle\rangle\langle\langle I_A|$ as a preparation (a map $1 \to AA$) and as an effect (a map $AA \to 1$) respectively. In this way, we may write the Choi isomorphism as



and its inverse as



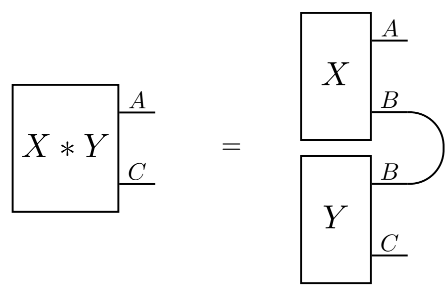
We will use similar symbols for the maximally entangled state $\psi^A := d_A^{-1}|I_A\rangle\rangle\langle\langle I_A|$:



3) The link product: The Choi matrix of a composition of maps is given by the link product of the respective Choi matrices, [20]. For general multipartite matrices $X \in \mathcal{B}(AB)$ and $Y \in \mathcal{B}(BC)$, the link product is defined as

$$X * Y = \mathrm{Tr}_B[(X \otimes I_C)(I_A \otimes Y^{\mathsf{T}_B})]$$

here $(\cdot)^{\mathsf{T}_B}$ denotes the partial transpose on the system B. Diagrammatically:



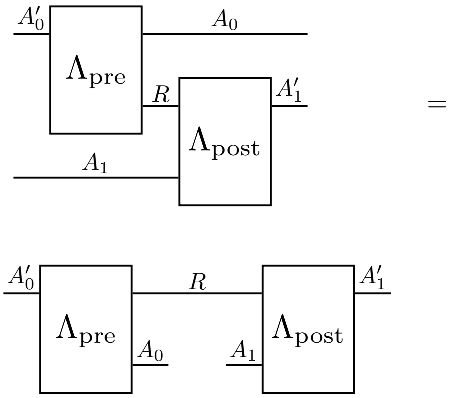
The link product is commutative (up to the order of the spaces) and associative provided that the three matrices have no labels in common. The order of the spaces is not taken into account, applying an appropriate unitary conjugation swapping the spaces in the tensor products if necessary, so, for example, if $X \in \mathcal{B}(AB_1B_2)$ and $Y \in \mathcal{B}(B_2B_1C)$, then

$$X * Y \equiv X * \mathcal{U}_{B_1,B_2}(Y), \quad (3)$$

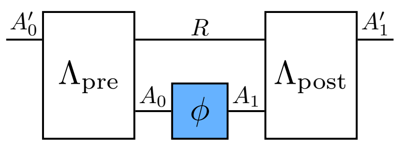
where $\mathcal{U}_{B_1,B_2}$ is the conjugation by the unitary swap $U_{B_1,B_2} : \mathcal{H}_{B_1B_2} \to \mathcal{H}_{B_2B_1}$.

4) Superchannels and 2-combs: A quantum superchannel is a special type of causal quantum network that transforms channels into channels, with possibly different input and output systems. Any superchannel Λ that maps $\mathcal{C}(A_0, A_1)$ into $\mathcal{C}(A_0', A_1')$ is a channel in $\mathcal{C}(A_0'A_1, A_1'A_0)$, consisting of a pre-processing channel $\Lambda_{\mathrm{pre}} \in \mathcal{C}(A_0', RA_0)$ and a post-processing channel $\Lambda_{\mathrm{post}} \in \mathcal{C}(RA_1, A_1')$, where R is some ancilla [20]. We will write $\Lambda = \Lambda_{\mathrm{pre}} * \Lambda_{\mathrm{post}}$ for this concatenation of channels. The set of all such superchannels will be denoted by $\mathcal{C}_2(A, A')$, where we used the abbreviation $A = A_0A_1$, $A' = A_0'A_1'$. Diagrammatically, Λ can be represented as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory



and acts on a map ϕ as



The Choi matrices of superchannels are called 2-combs in [20]. Using the link product and its properties, we have $C_\Lambda = C_{\Lambda_{\text{pre}}} * C_{\Lambda_{\text{post}}}$ and $C_{\Lambda(\Phi)} = C_\Lambda * C_\Phi = C_{\Lambda_{\text{pre}}} * C_\Phi * C_{\Lambda_{\text{post}}}$. An element $C \in \mathcal{B}_+(A_1'A_0A_0'A_1)$ is a 2-comb if and only if

$$\mathrm{Tr}_{A_1'}[C] = I_{A_1} \otimes C_2, \qquad \mathrm{Tr}_{A_0}[C_2] = I_{A_0'}, \tag{4}$$

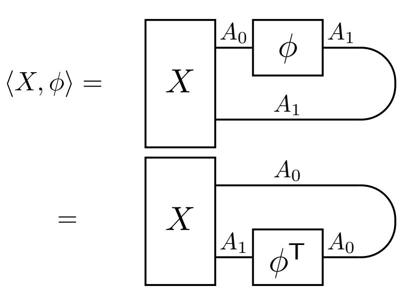
which means that C_2 is the Choi matrix of some channel in $\mathcal{C}(A_0', A_0)$.

5) Diamond norm and the conditional min-entropy: It is not difficult to see that $\mathcal{L}(A_0, A_1)$ with the cone $\mathcal{L}_+(A_0, A_1)$ and $B(\mathcal{L}(A_0, A_1)) = \mathcal{C}(A_0, A_1)$ is an object in BS, similarly for the set of superchannels. It was observed in [29] that in these cases the structures described in Section II-A yield some well known quantities. This will be discussed in the present and the next section, see [29] for more details.

We will use the identification of the dual space $\mathcal{L}^*(A_0, A_1)$ with $\mathcal{B}_h(A_0A_1)$, with duality for $X \in \mathcal{B}_h(A_0A_1)$ and $\phi \in \mathcal{L}(A_0, A_1)$ given by

$$\begin{aligned}\langle X, \phi\rangle &:= \langle\langle I_{A_1}|(\phi \otimes id)(X)|I_{A_1}\rangle\rangle = \mathrm{Tr}\,[XC_{\phi^*}] \\ &= \mathrm{Tr}\,[XC_{\phi^\mathsf{T}}^\mathsf{T}] = X * C_{\phi^\mathsf{T}} = X * C_\phi. \end{aligned}\tag{5}$$

Diagrammatically, this can be expressed as



Here $\phi^*, \phi^\mathsf{T} \in \mathcal{L}(A_1, A_0)$ are maps determined by

$$\mathrm{Tr}\,[X\phi^*(Y)] = \mathrm{Tr}\,[\phi(X)Y], \qquad \phi^\mathsf{T}(Y) = (\phi^*(Y^\mathsf{T}))^\mathsf{T}$$

for $X \in \mathcal{B}_h(A_0)$ and $Y \in \mathcal{B}_h(A_1)$. The last equality in (5) follows from (3) and $C_{\phi^\mathsf{T}} = \mathcal{U}_{A_1,A_0}(C_\phi)$. Note that by (5) we also have

$$\begin{aligned}\langle X, \phi\rangle &= \mathrm{Tr}\,[X\mathcal{U}_{A_1,A_0}(C_\phi)^\mathsf{T}] \\ &= \mathrm{Tr}\,[\mathcal{U}^*_{A_1,A_0}(X^\mathsf{T})C_\phi]. \end{aligned}\tag{6}$$

With these identifications, the dual cone is $\mathcal{L}_+^*(A_0, A_1) \simeq \mathcal{B}_+(A_0A_1)$ and the dual section

$$\begin{aligned}B(\mathcal{L}^*(A_0, A_1)) &= \tilde{\mathcal{C}}(A_0, A_1) \\ &= \{\sigma_{A_0} \otimes I_{A_1},\ \sigma_{A_0} \in \mathfrak{S}(A_0)\}.\end{aligned}$$

The corresponding base section norm is the diamond norm

$$\|\Phi\|_{\mathcal{L}(A_0,A_1)} = \|\Phi\|_\diamond := \max_{\rho \in \mathfrak{S}(A_0A_0)} \|(\Phi \otimes id)(\rho)\|_1,$$

for $\Phi \in \mathcal{L}(A_0, A_1)$, well known as the distinguishability norm for quantum channels, [21], [31].

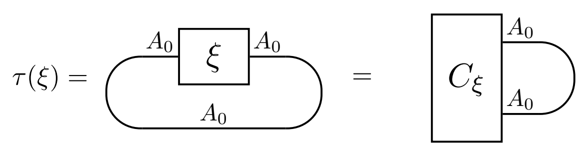
Remark 1. Using the Choi representation, we may also identify $\mathcal{L}^*(A_0, A_1) \simeq \mathcal{L}(A_1, A_0)$, with duality $\langle\cdot,\cdot\rangle_*$ given as

$$\langle\psi, \phi\rangle_* := \langle C_\psi, \phi\rangle = \tau(\phi \circ \psi),$$

where the functional $\tau : \mathcal{L}(A_0, A_0) \to \mathbb{R}$ is given by

$$\tau(\xi) = \sum_{i,j} \mathrm{Tr}\,[|i\rangle\langle j|\xi(|i\rangle\langle j|)] = \langle\langle I_{A_0}|C_\phi|I_{A_0}\rangle\rangle,$$

in diagram



Choosing the Hilbert-Schmidt inner product and the basis $\{|i\rangle\langle j|\}$ in $\mathcal{B}(A_0)$, we see that τ is the usual trace of elements in $\mathcal{L}(A_0, A_0)$ as linear maps. In this identification, the dual section becomes the set of replacement channels in $\mathcal{C}(A_1, A_0)$, mapping all states in $\mathfrak{S}(A_1)$ to a fixed state $\sigma_{A_0} \in \mathfrak{S}(A_0)$.

Let us introduce the notation

$$\|\cdot\|^\diamond_{A_1|A_0} := \|\cdot\|_{\mathcal{L}^*(A_0,A_1)}$$

for the dual norm. By Prop. 1 (iii), we have for $\rho \in \mathcal{B}_+(A_0A_1)$:

$$\begin{aligned}\|\rho\|^\diamond_{A_1|A_0} &= \min_{\sigma_{A_0} \in \mathfrak{S}(A_0)} \min\{\lambda > 0,\ \rho \le \lambda\sigma_{A_0} \otimes I_{A_1}\} \\ &= 2^{-H_{\min}(A_1|A_0)_\rho}\end{aligned}\tag{7}$$

where $H_{\min}$ denotes the conditional min-entropy [32], [33]. We also have

$$\begin{aligned}\|\rho\|^\diamond_{A_1|A_0} &= \max_{\alpha \in \mathcal{C}(A_0,A_1)} \langle\rho, \alpha\rangle \\ &= \max_{\alpha \in \mathcal{C}(A_0,A_1)} \langle\langle I_{A_1}|(\alpha \otimes id)(\rho)|I_{A_1}\rangle\rangle.\end{aligned}\tag{8}$$

Note that the last equality corresponds to the operational interpretation of the conditional min-entropy as (up to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

multiplication by d_A) the maximum fidelity with the maximally entangled state ψ^{A_1} that can be obtained by applying a quantum channel to part A_0 of the state $\rho_{A_0A_1}$ [33].

The following result follows easily from the first equality in (7).

Lemma 1. *Let $\rho \in \mathcal{B}_+(A_0A_1)$. Then there is some $V \in \mathcal{B}(A_0)$, $\mathrm{Tr}\,[VV^*] = 1$, and $G \in \mathcal{B}_+(A_0A_1)$ such that*

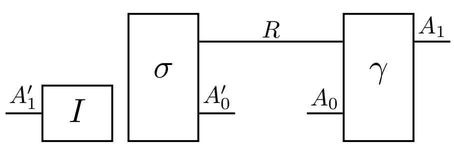
$$\rho = (\chi_V \otimes id)(G), \qquad \|G\| = \|\rho\|^{\diamond}_{A_1|A_0}.$$

Here $\chi_V := V \cdot V^* \in \mathcal{L}_+(A_0, A_0)$ and $\|\cdot\|$ denotes the operator norm.

6) Diamond 2-norm and conditional 2-min entropy: As we have seen, the set of superchannels $\mathcal{C}_2(A, A')$ is a subset of $\mathcal{C}(A'_0A_1, A'_1A_0)$. In fact, it is itself a base section. More precisely, put $\mathcal{L}_2(A, A') := \mathcal{L}(A'_0A_1, A'_1A_0)$ with $\mathcal{L}_2^+(A, A') := \mathcal{L}_+(A'_0A_1, A'_1A_0)$ and $B(\mathcal{L}_2(A, A')) = \mathcal{C}_2(A, A')$, then $\mathcal{L}_2(A, A') \in \mathsf{BS}$. Using the same identification of the dual space as before, we have $\mathcal{L}^*(A, A') = B_h(A'_1A_0A'_0A_1)$ and the dual section is

$$B(\mathcal{L}_2(A, A')) = \{\sigma * C_\gamma \otimes I_{A'_1},\\ \sigma \in \mathfrak{S}(A'_0R),\ \gamma \in \mathcal{C}(A_0R, A_1),\ R \text{ an ancilla}\}.$$

Using the identification of the dual space as in Remark 1, this corresponds to a set of superchannels where the preprocessing is a replacement channel, of the form



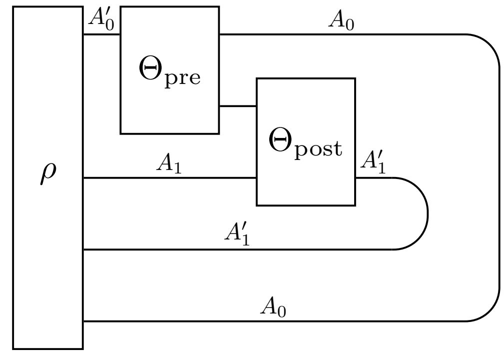
The norm $\|\cdot\|_{\mathcal{C}_2(A,A')}$ is the distinguishability norm $\|\cdot\|_{2\diamond}$ for quantum networks, see [34], [35] for the definition. Let us denote

$$\|\cdot\|^{2\diamond}_{A'|A} := \|\cdot\|_{\mathcal{L}_2^*(A,A')},$$

then for $\rho \in \mathcal{B}_+(A'_0A_1A'_1A_0)$, we have

$$\begin{aligned}\|\rho\|^{2\diamond}_{A'|A} &= \min_{\sigma,\gamma} \min\{\lambda > 0,\ \rho \le \lambda(\sigma * C_\gamma) \otimes I_{A'_1}\}\\ &=: 2^{-H^{(2)}_{\min}(A'|A)_\rho} = \max_{\Theta \in \mathcal{C}_2(A,A')} \langle \rho, \Theta\rangle\\ &= \max_{\Theta \in \mathcal{C}_2(A,A')} \langle\langle I_{A_0A'_1}|(\Theta \otimes id)(\rho)|I_{A_0A'_1}\rangle\rangle.\end{aligned}$$

Here $H^{(2)}_{\min}$ will be called the conditional 2-min entropy. Note that this quantity coincides with the extended conditional min-entropy of [9] but we prefer the present notation since it can be extended to any $N \in \mathbb{N}$ in an obvious way using the set of N-combs, see also [24]. The last equality shows an operational interpretation as the maximum fidelity (again up to multiplication by the dimension) with the maximally entangled state $\psi^{A_0B_1}$ that can be obtained by applying a structured quantum channel to the part A'_0A_1 of $\rho_{A'_0A_1A'_1A_0}$ as depicted in the diagram



7) Guessing games: Let $\mathcal{E} = \{\lambda_i, \rho_i\}_{i=1}^k$ be an ensemble of states $\rho_i \in \mathfrak{S}(A)$ and prior probabilities λ_i. Quantum measurements with k outcomes are given by operators $M = \{M_1, \dots, M_k\}$, where $M_j \in \mathcal{B}_+(A)$, $\sum_j M_j = I_A$, the set of all such measurements for the system A will be denoted by $\mathcal{M}_k(A)$. It is well known that the optimal success probability $P_{\mathrm{succ}}(\mathcal{E})$ is related to the conditional min entropy as follows, [33]. Let us define the quantum-classical state $\rho_{\mathcal{E}} = \sum_i \lambda_i\rho_i \otimes |i\rangle\langle i| \in \mathfrak{S}(AR)$, where $d_R = k$. Then for any channel $\alpha \in \mathcal{C}(A, R)$, we have

$$\langle \rho_{\mathcal{E}}, \alpha\rangle = P_{\mathrm{succ}}(\mathcal{E}, M) = \langle \rho_{\mathcal{E}}, \Phi_M\rangle, \tag{9}$$

where $M \in \mathcal{M}_k(A)$, $M_i = \alpha^*(|i\rangle\langle i|)$ and $\Phi_M \in \mathcal{C}(A, R)$ is the quantum-to-classical (q-c) channel given by $\Phi_M(\sigma) = \sum_i \mathrm{Tr}\,[\sigma M_i]|i\rangle\langle i|$. It follows that

$$\|\rho_{\mathcal{E}}\|^{\diamond}_{R|A} = P_{\mathrm{succ}}(\mathcal{E}), \tag{10}$$

see also Example 7. It was proved in [10, Prop. 2] that the dual norm $\|\cdot\|^{\diamond}_{R|A}$ can be interpreted as a success probability not only for quantum-classical states. Since this result and the related constructions will be repeatedly used below, we give the proof here.

Lemma 2. *For any state $\rho \in \mathfrak{S}(AR)$ there is an ensemble $\mathcal{E}^R_\rho$ on AR such that*

$$\|\rho\|^{\diamond}_{R|A} = d_R P_{\mathrm{succ}}(\mathcal{E}^R_\rho).$$

The proof is based on a relation between quantum channels $A \to R$ and measurements on AR, close to the Choi representation of channels. Let $\{U^R_1, \dots, U^R_{d_R^2}\}$ be the group of generalized Pauli unitaries on R and let $\mathcal{U}^R_x$ denote the conjugation by U^R_x, so that we have

$$\sum_x \mathcal{U}^R_x(X) = d_R \mathrm{Tr}\,[X] I_R.$$

Let

$$\mathsf{B}^R_x := (\mathcal{U}^R_x \otimes id)(\psi^R) = d_R^{-1}|U^R_x\rangle\rangle\langle\langle U^R_x|$$

then $\mathsf{B}^R = \{\mathsf{B}^R_1, \dots, \mathsf{B}^R_{d_R^2}\}$ defines the Bell measurement on RR. For a channel $\beta \in \mathcal{C}(A, R)$, let $M^\beta \in \mathcal{M}_{d_R^2}(AR)$ be defined as

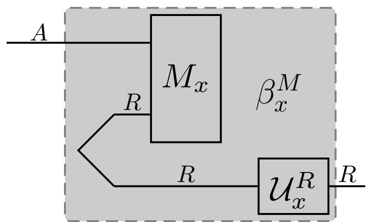
$$M^\beta_x = (\beta^* \otimes id)(\mathsf{B}^R_x), \quad x = 1, \dots, d_R^2. \tag{11}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

Conversely, for any $M \in \mathcal{M}_{d_R^2}(AR)$, let $\beta^M \in \mathcal{C}(A,R)$ be the channel obtained from the measurement M in the teleportation scheme, that is

$$\beta^M(\sigma) = \sum_x \mathcal{U}_x^R(\mathrm{Tr}_{A\tilde{R}}[(\sigma \otimes \psi^R)(M_x \otimes I_R)]) \quad (12)$$

for $\sigma \in \mathfrak{S}(A)$ (here we take $\tilde{R} \simeq R$ and view ψ^R as a state on $\tilde{R}R$ and M as a measurement on $A\tilde{R}$). It is easily seen that we have $M^{\beta^M} = M$ and $\beta^{M^\beta} = \beta$. Note that we have $\beta^M = \sum_x \beta_x^M$, where $\{\beta_x^M\}$ is an instrument whose elements are depicted as

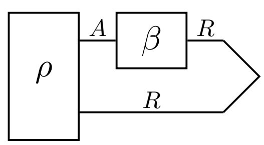


Proof of Lemma 2. For $\rho \in \mathfrak{S}(AR)$ we introduce the equiprobable ensemble

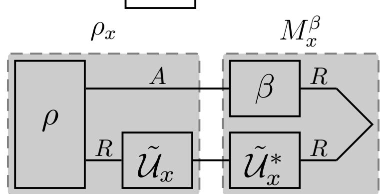
$$\mathcal{E}_\rho^R = \{d_R^{-2}, \rho_x\}_{x=1}^{d_R^2}, \quad \rho_x = (id_A \otimes \tilde{\mathcal{U}}_x^R)(\rho),$$

where $\tilde{\mathcal{U}}_x^R$ denotes the conjugation by $(U_x^R)^\mathsf{T}$. Note that for any channel $\beta \in \mathcal{C}(A,R)$ and $x = 1,\dots,d_R^2$, we have

$d_R^{-1}\langle \rho, \beta\rangle =$



$=$



Multiplying by the probability d_R^{-2} and summing up over x we obtain

$$\begin{aligned} d_R^{-1}\langle \rho, \beta\rangle &= P_{\mathrm{succ}}((\beta \otimes id_R)(\mathcal{E}_\rho^R), \mathsf{B}^R) \\ &= P_{\mathrm{succ}}(\mathcal{E}_\rho^R, M^\beta). \end{aligned} \quad (13)$$

Also conversely, it is readily checked that for any measurement $M \in \mathcal{M}_{d_R^2}(AR)$ we have

$$\begin{aligned} P_{\mathrm{succ}}(\mathcal{E}_\rho, M) &= \sum_x \mathrm{Tr}\,[\rho_x M_x] = d_R^{-1}\sum_x \langle \rho, \beta_x^M\rangle \\ &= d_R^{-1}\langle \rho, \beta^M\rangle. \end{aligned} \quad (14)$$

Using (8), this proves Lemma 2.

□

As an application, we have the following expression for the diamond norm distance of quantum channels in terms of the success probabilities in guessing games.

Corollary 1. *Let $\Phi_1, \Phi_2 \in \mathcal{C}(A_0, A_1)$. Then*

$$\frac{1}{2}\|\Phi_1 - \Phi_2\|_\diamond = \sup_{\mathcal{E},M} \frac{P_{\mathrm{succ}}((\Phi_1 \otimes id_R)(\mathcal{E}), M) - P_{\mathrm{succ}}((\Phi_2 \otimes id_R)(\mathcal{E}), M)}{P_{\mathrm{succ}}(\mathcal{E})},$$

where the supremum is taken over all ensembles $\mathcal{E}$ on A_0R, all measurements M on A_1R and any ancilla R. Moreover, the supremum is attained with $R \simeq A_1$ and the Bell measurement $M = \mathsf{B}^{A_1}$.

Proof. Let s denote the supremum on the left hand side of the equality to be proved. Let $\mathcal{E} = \{\lambda_j, \rho_j\}_{j=1}^k$ be an ensemble on A_0R and $M \in \mathcal{M}_k(A_1R)$. Put $\rho = \rho_\mathcal{E} \in \mathfrak{S}(A_0RR')$, with $d_{R'} = k$. Then by (9)

$$\begin{aligned} P_{\mathrm{succ}}((\Phi_i \otimes id)(\mathcal{E}), M) &= P_{\mathrm{succ}}(\mathcal{E}, (\Phi_i^* \otimes id)(M)) \\ &= \langle \rho, \Phi_{(\Phi_i^* \otimes id)(M)}\rangle, \end{aligned}$$

$i = 1,2$. Note that the q-c channel $\Phi_{(\Phi_i^* \otimes id)(M)} = \Phi_M \circ (\Phi_i \otimes id_R)$ and therefore

$$\begin{aligned} &P_{\mathrm{succ}}((\Phi_1 \otimes id)(\mathcal{E}), M) - P_{\mathrm{succ}}((\Phi_2 \otimes id)(\mathcal{E}), M) \\ &\quad = \langle \rho_\mathcal{E}, \Phi_M \circ (\Phi_1 \otimes id_R) - \Phi_M \circ (\Phi_2 \otimes id_R)\rangle \\ &\quad \le \frac{1}{2}\|\rho_\mathcal{E}\|_{R|A}^\diamond \|\Phi_M \circ [(\Phi_1 - \Phi_2) \otimes id_R]\|_\diamond \\ &\quad \le \frac{1}{2}P_{\mathrm{succ}}(\mathcal{E})\|\Phi_1 - \Phi_2\|_\diamond, \end{aligned}$$

here we used Proposition 1. This shows that $s \le \frac{1}{2}\|\Phi_1 - \Phi_2\|_\diamond$.

For the converse, note that by Proposition 1 (iii) and (iv), we have

$$\begin{aligned} \frac{1}{2}\|\Phi_1 - \Phi_2\|_\diamond &= \sup_{\rho \in \mathcal{B}_+(A_0A_1)} \frac{\langle \rho, \Phi_1 - \Phi_2\rangle}{\|\rho\|_{A_1|A_0}^\diamond} \\ &= \sup_{\rho \in \mathfrak{S}(A_0A_1)} \frac{\langle \rho, \Phi_1 - \Phi_2\rangle}{\|\rho\|_{A_1|A_0}^\diamond}. \end{aligned} \quad (15)$$

Let now $R \simeq A_1$ and $M = \mathsf{B}^{A_1}$. For $\rho \in \mathfrak{S}(A_0A_1)$, we take the ensemble $\mathcal{E}_\rho^{A_1}$ on A_0A_1. By (15), (13) and Lemma 2, we obtain

$$\begin{aligned} \frac{1}{2}\|\Phi_1 - \Phi_2\|_\diamond = \sup_{\rho \in \mathfrak{S}(A_0A_1)} \Big[&P_{\mathrm{succ}}((\Phi_1 \otimes id)(\mathcal{E}_\rho^{A_1}), \mathsf{B}^{A_1}) \\ &- P_{\mathrm{succ}}((\Phi_2 \otimes id)(\mathcal{E}_\rho^{A_1}), \mathsf{B}^{A_1})\Big] P_{\mathrm{succ}}(\mathcal{E}_\rho^R)^{-1} \le s. \end{aligned}$$

This finishes the proof.

□

Remark 2. The above construction implies another operational interpretation of $\|\rho\|_{R|A}^\diamond$. To see this, let $\mathcal{E} = \{\lambda_i, \Phi_i\}$ be an ensemble of quantum channels, $\Phi_i \in \mathcal{C}(R,R)$. The guessing procedures for ensembles of channels can be described by pairs (ρ, M), consisting of an input state $\rho \in \mathfrak{S}(AR)$ with some ancilla A and M is a measurement on AR, such triples are also called quantum

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

testers [34]. The average success probability for the tester (ρ, M) is then

$$P_{\rm succ}(\mathcal{E}, \rho, M) := P_{\rm succ}(\mathcal{E}(\rho), M)$$

where $\mathcal{E}(\rho) = \{\lambda_i, (\Phi_i \otimes id)(\rho)\}$.

Any state $\rho \in \mathfrak{S}(AR)$ can be seen as the input state of some tester. We claim that the norm $\|\rho\|^\diamond_{R|A}$ can be interpreted as (d_R times) the maximal success probability that can be obtained by all testers with input state ρ for equiprobable ensembles $\mathcal{E} = \{d_R^{-2}, \Phi_i\}_{x=1}^{d_R^2}$ of unital channels $\Phi_x \in \mathcal{C}(R, R)$. Indeed, let $M \in \mathcal{M}_{d_R^2}(AR)$ be any measurement, we have

$$\begin{aligned} P_{\rm succ}(\mathcal{E}, \rho, M) &= d_R^{-2} \sum_x \mathrm{Tr}\left[(id \otimes \Phi_x)(\rho) M_x\right] \\ &= d_R^{-1} \mathrm{Tr}\left[\rho C\right], \end{aligned}$$

where $C = d_R^{-1} \sum_i (id \otimes \Phi_x^*)(M_x) \in \mathcal{B}_+(AR)$. Since Φ_x^* is trace preserving, we see that $\mathrm{Tr}_R[C] = d_R^{-1}\mathrm{Tr}_R[I] = I_A$. Hence there is a channel $\alpha \in \mathcal{C}(R, A)$ such that $C = C_{\alpha^*}$. Finishing the above computation, we obtain

$$\begin{aligned} P_{\rm succ}(\mathcal{E}, \rho, M) &= d_R^{-1} \mathrm{Tr}\left[\rho C_{\alpha^*}\right] = d_R^{-1} \langle \rho, \alpha \rangle \\ &\le d_R^{-1} \|\rho\|_{R|A}. \end{aligned}$$

As we have seen, equality is attained for $\Phi_x = \tilde{\mathcal{U}}_x^R$.

B. General comparison theorems for quantum channels

We are now ready to apply the results of Section II-B to the comparison of a pair of quantum channels Φ_1 and Φ_2. For this, we consider a subcategory F in BS whose objects are some spaces of channels (as in Sec. III-A5) and morphisms between them are given by some convex subsets of superchannels, (so F is in fact a convex subcategory of the category of quantum channels with superchannels).

If $\mathcal{L}(A_0, A_1)$ is an object in F we will say that $A = A_0A_1$ is an input-output space admissible in F. We will use the notation

$$\mathsf{F}(A, A') := \mathsf{F}(\mathcal{L}(A_0, A_1), \mathcal{L}(A_0', A_1')) \subseteq \mathcal{C}_2(A, A')$$

for a pair of admissible spaces A, A'.

Let A, A' be admissible in F and let $\Phi_1 \in \mathcal{C}(A_0, A_1)$, $\Phi_2 \in \mathcal{C}(A_0', A_1')$. The F-conversion distance becomes

$$\delta_{\mathsf{F}}(\Phi_1 \| \Phi_2) = \inf_{\Lambda \in \mathsf{F}(A, A')} \|\Lambda(\Phi_1) - \Phi_2\|_\diamond.$$

For $\rho \in \mathcal{B}_+(A_0'A_1A_1'A_0)$, we define

$$\|\rho\|^{\mathsf{F}}_{A'|A} := \sup_{\Theta \in \mathsf{F}(A, A')} \langle \rho, \Theta \rangle.$$

This notation may be somewhat misleading, since for a general subcategory F there is no guarantee that $\|\cdot\|^{\mathsf{F}}_{A'|A}$ can be extended to a norm. However, there are some choices that lead to the norms introduced in Sections III-A5 and III-A6. Indeed, with the choice of a subcategory where all objects are spaces of states (that is, all admissible input spaces are $A_0 = 1$) and the morphisms are all (super)channels between them, we obtain $\|\cdot\|^{\mathsf{F}}_{A'|A} = \|\cdot\|^\diamond_{A_1'|A_1}$. If F coincides with the category of quantum channels with superchannels, we similarly get $\|\rho\|^{\mathsf{F}}_{A'|A} = \|\rho\|^{2\diamond}_{A'|A}$. Since $\mathsf{F}(A, A') \subseteq \mathcal{C}_2(A, A') \subseteq \mathcal{C}(A_0'A_1, A_1'A_0)$, the following inequalities are immediate:

$$\|\rho\|^{\mathsf{F}}_{A'|A} \le \|\rho\|^{2\diamond}_{A'|A} \le \|\rho\|^\diamond_{A_1'A_0|A_0'A_1}.$$

Moreover, the quantity

$$H^{\mathsf{F}}_{\min}(A'|A)_\rho := -\log(\|\rho\|^{\mathsf{F}}_{A'|A})$$

can be seen as a modified conditional min entropy, since it coincides with $H_{\min}$ and $H^{(2)}_{\min}$ in the above cases.

More generally, it may happen that $\mathsf{F}(A, A')$ is a base section in $\mathcal{L}_2(A, A')$, possibly with respect to some subcone $\mathcal{L}^+_{\mathsf{F}}(A, A') \subseteq \mathcal{L}^+_2(A, A')$. Then $\|\cdot\|^{\mathsf{F}}_{A'|A}$ coincides with the base section norm with respect to the dual section, see Proposition 1 (iii). Consequently, $\|\cdot\|^{\mathsf{F}}_{A'|A}$ satisfies the properties in Proposition 1, so that there is a dual expression for this norm. In particular, it can be expressed as a conic program.

Example 8. Let F be the subcategory whose objects are spaces $\mathcal{L}(A_0B_0, A_1B_1)$ of bipartite channels and the morphisms in F are PPT-superchannels defined by the property that they stay completely positive when pre and postcomposed by the partial transpose supermap, see [18] for more details. In this case, $\mathsf{F}(AB, A'B')$ is a base section in $\mathcal{L}_2$ with respect to the subcone $\mathcal{L}^+_{2,\rm PPT} \subsetneq \mathcal{L}^+_2$ of PPT-supermaps, [18, Definition V.2]. Similarly, if the morphisms are given by separable superchannels, characterized by the condition that the Choi matrix is separable, then $\mathsf{F}(AB, A'B')$ is a base section with respect to the cone $\mathcal{L}^+_{2,\rm SEP}$ of separable supermaps. Note that these sets of superchannels might be different from their restricted variants that will be considered below.

Another such example is the subcategory with spaces of channels as objects and morphisms given by no-signaling superchannels which can be easily described by properties of their Choi matrices: besides the condition (4) of a 2-comb they also satisfy an analogical condition with input and output spaces exchanged. In this case, $\mathsf{F}(A, A')$ is a base section in the cone $\mathcal{L}^+_2(A, A')$.

For any choice of F, $\|\cdot\|^{\mathsf{F}}_{A'|A}$ is obviously convex and monotone under morphisms in F, more precisely, for any $\Lambda \in \mathsf{F}(A', A'')$, we have

$$\|\rho\|^{\mathsf{F}}_{A'|A} \ge \|\rho * C_\Lambda\|^{\mathsf{F}}_{A''|A}.$$

Further properties depend on the details of the structure of F, e.g. with respect to the tensor products.

The importance of these quantities is seen from the following general comparison theorem for quantum channels, which is a straightforward reformulation of Theorem 2 from the GPT setting.

Theorem 3. *Let $\epsilon \ge 0$. The following are equivalent.*

(i) $\delta_{\mathsf{F}}(\Phi_1 \| \Phi_2) \le \epsilon$;

(ii) *for any* $\rho \in \mathfrak{S}(A_0' A_1')$, *there is some* $\Theta \in \bar{\mathsf{F}}(A, A')$ *such that*

$$\langle \rho, \Phi_2 \rangle \leq \langle \rho, \Theta(\Phi_1) \rangle + \frac{\epsilon}{2} \|\rho\|^{\diamond}_{A_1'|A_0'}$$

(here $\bar{\mathsf{F}}(A, A')$ *is the closure of the set* $\mathsf{F}(A, A')$*);*

(iii) *for any* $R = R_0 R_1$ *admissible in* F *and any* $\rho \in \mathfrak{S}(R_0 R_1)$,

$$\|\rho * C_{\Phi_2}\|^{\mathsf{F}}_{R|A'} \leq \|\rho * C_{\Phi_1}\|^{\mathsf{F}}_{R|A} + \frac{\epsilon}{2} \|\rho\|^{\diamond}_{R_1|R_0}.$$

Moreover, in (iii) it is enough to assume $R_0 \simeq A_0'$, $R_1 \simeq A_1'$.

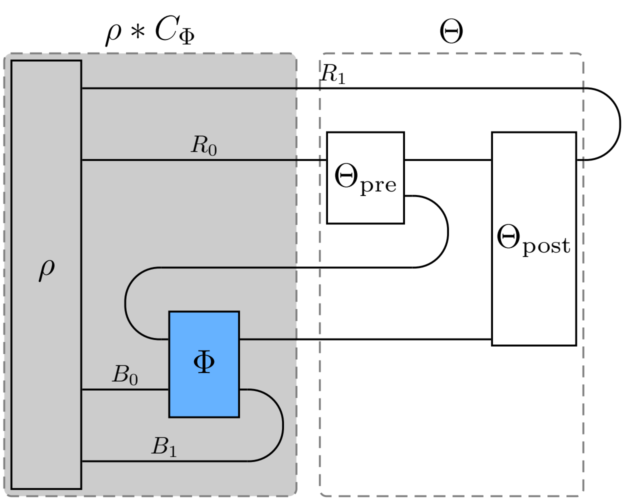
Proof. We only need to observe that for any $\Theta \in \mathsf{F}(A, R)$ and $\rho \in \mathfrak{S}(R)$, we have

$$\langle \rho, \Theta(\Phi_1) \rangle = \rho * C_{\Theta(\Phi_1)} = (\rho * C_{\Phi_1}) * C_\Theta = \langle \rho * C_{\Phi_1}, \Theta \rangle,$$

similarly for Φ_2. The proof now follows directly from Theorem 2 and the definition of $\|\cdot\|^{\mathsf{F}}_{R|A}$ ($\|\cdot\|^{\mathsf{F}}_{R|A'}$). □

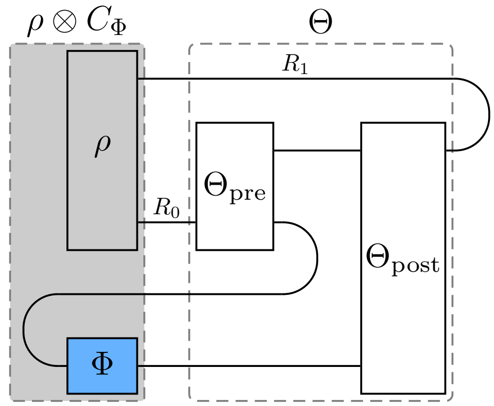
Remark 3. There is an ambiguity around $\rho * C_{\Phi_i}$ that might cause some confusion: it may happen that parts of R_0 or R_1 coincide with some parts of the input or output spaces of Φ_i, so it is unclear how to apply the link product. In some cases, such as in some of the sections below, the subcategory F does not permit any processing on some parts of the input and/or output spaces. In this case, these parts are always fixed and are viewed as the same, so that they are connected in the link product. In diagram, if the fixed input of the channels is B_0 and the fixed output is B_1, we obtain in this case

$$\langle \rho, \Theta(\Phi) \rangle = \langle \rho * C_\Phi, \Theta \rangle =$$



In all other cases all the involved spaces are treated as independent, so that the link product is in fact the tensor product, $\rho * C_{\Phi_i} = \rho \otimes C_{\Phi_i}$, in diagram

$$\langle \rho, \Theta(\Phi) \rangle = \langle \rho \otimes C_\Phi, \Theta \rangle =$$



The questions of simulability/convertibility of quantum channels naturally appear in the quantum resource theory of processes, [17], [16], [18], [19]. In this case, the subset $\mathcal{O}(A_0, A_1) := \mathsf{F}(1, A) \subseteq \mathcal{C}(A_0, A_1)$ is the set of free channels and $\mathsf{F}(A, A')$ is the set of free superchannels, that is, transformations that can be performed at no cost. The subcategory is usually assumed to have further properties, such as that it behaves well under tensor products (i.e. it is a symmetric monoidal subcategory in the category of quantum channels with superchannels). Note that a convex symmetric monoidal category is called a convex resource theory in [36].

Assume that the resource theory is such that all free channels can be converted one into another by free superchannels, then for each $\rho \in \mathfrak{S}(R_0 R_1)$ the map

$$\varphi_\rho : \Phi \mapsto \|\rho * C_\Phi\|^{\mathsf{F}}_{R|A}$$

is constant over $\Phi \in \mathcal{O}$, namely $\varphi_\rho(\Phi) = \sup_{\Psi \in \mathcal{O}} \langle \rho, \Psi \rangle$ for any $\Phi \in \mathcal{O}$. Therefore each φ_ρ can be easily normalized to be 0 on $\mathcal{O}$. Furthermore, if any channel Φ can be converted into a channel arbitrarily close to $\mathcal{O}$ by some elements in F, then Theorem 3 implies that $\varphi_\rho(\Phi) \geq \sup_{\Psi \in \mathcal{O}} \langle \rho, \Psi \rangle$ and the set of normalized φ_ρ becomes a complete family of resource monotones. In fact, the case $\epsilon = 0$ is closely related to [18, Theorem III.3].

Specific examples of the subcategory F will be studied in detail in the next sections: postprocessings and preprocessings of quantum channels, processings of bipartite channels by LOCC and by partial superchannels. In all these cases, the superchannels in F consist of pre- and postprocessings belonging to some specified families of channels $\mathcal{C}_{\text{pre}}$ and $\mathcal{C}_{\text{post}}$. More precisely, for $R = R_0 R_1$, $S = S_0 S_1$ admissible in F, any superchannel $\Theta \in \mathsf{F}(R, S)$ has the form $\Theta = \Lambda_{\text{pre}} * \Lambda_{\text{post}}$, where $\Lambda_{\text{pre}} \in \mathcal{C}_{\text{pre}}(S_0, U R_0)$ and $\Lambda_{\text{post}} \in \mathcal{C}_{\text{post}}(U R_1, S_1)$, where U is some ancilla. To ensure that F is a convex subcategory, we have to assume that for any input-output spaces R, S, T admissible in F and any ancillas U, V available in $\mathcal{C}_{\text{pre}}$ and $\mathcal{C}_{\text{post}}$, we have

(a) $id_{R_0} \in \mathcal{C}_{\text{pre}}(R_0, R_0)$, $id_{R_1} \in \mathcal{C}_{\text{post}}(R_1, R_1)$;
(b) both $\mathcal{C}_{\text{pre}}$ and $\mathcal{C}_{\text{post}}$ are closed under composition,

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

that is,

$$\alpha \in \mathcal{C}_{\rm pre}(S_0, UR_0),\ \alpha' \in \mathcal{C}_{\rm pre}(T_0, VS_0) \\ \implies (id_V \otimes \alpha) \circ \alpha' \in \mathcal{C}_{\rm pre}(T_0, VUR_0)$$

and

$$\beta \in \mathcal{C}_{\rm post}(UR_1, S_1),\ \beta' \in \mathcal{C}_{\rm post}(VS_1, T_1) \\ \implies \beta' \circ (id_V \otimes \beta) \in \mathcal{C}_{\rm post}(VUR_1, T_1);$$

(c) the sets $\mathcal{C}_{\rm pre}(S_0, UR_0)$ and $\mathcal{C}_{\rm post}(UR_1, S_1)$ are convex.

Further conditions may be needed to ensure convexity, such as

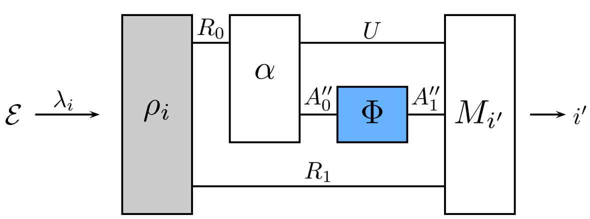
(d) $\mathcal{C}_{\rm pre/post}$ are closed under appending/discarding a classical register, more precisely, that we have $|i\rangle\langle i| \otimes (\cdot) \in \mathcal{C}_{\rm pre}(R_0, UR_0)$ and $\sum_i \langle i| \cdot |i\rangle \otimes \Phi_i \in \mathcal{C}_{\rm post}(UVR_1, S_1)$ if $\Phi_i \in \mathcal{C}_{\rm post}(VR_1, S_1)$, $\forall i$.

We will write

$$\mathsf{F} = \mathcal{C}_{\rm pre} * \mathcal{C}_{\rm post}$$

to emphasize that the morphisms in F have this form. This is a natural assumption in the framework of resource theories, where $\mathcal{C}_{\rm pre} = \mathcal{C}_{\rm post} = \mathcal{O}$ is the set of free channels [36], [18]. We will study the subcategory of LOCC superchannels, where the objects are spaces of bipartite channels and $\mathcal{O} = \mathcal{C}_{\rm LOCC}$. We may similarly consider the category of restricted PPT superchannels, where $\mathcal{O} = \mathcal{C}_{\rm PPT}$, or restricted separable superchannels where $\mathcal{O} = \mathcal{C}_{\rm SEP}$, [18], [19].

We will use the decomposition $\mathsf{F} = \mathcal{C}_{\rm pre} * \mathcal{C}_{\rm post}$ to characterize δ_{F} by success probabilities in modified guessing games of the form depicted in the diagram



Here $\mathcal{E} = \{\lambda_i, \rho_i\}$ is an ensemble on an admissible space $R = R_0R_1$, a preprocessing α can be chosen from $\mathcal{C}_{\rm pre}(R_0, UA_1'')$, M is a measurement chosen from a family $\mathcal{M}_{\rm post}(UA_1''R_1)$ of allowed measurements and A_0'', A_1'' are the input and output spaces of Φ, which is either Φ_1 or Φ_2. The optimal success probability with this scheme is

$$P_{\rm succ}^{\mathcal{C}_{\rm pre}(R_0,\cdot A_0''),\Phi,\mathcal{M}_{\rm post}}(\mathcal{E}) := \\ \sup_{\alpha\in\mathcal{C}_{\rm pre}, M\in\mathcal{M}_{\rm post}} P_{\rm succ}((\alpha \otimes id)(\mathcal{E}), (\Phi^* \otimes id)(M)),$$

here $\mathcal{C}_{\rm pre}(R_0, \cdot A_0'')$ indicates that we allow any ancilla U that is permitted by the structure of $\mathcal{C}_{\rm pre}$ and $\mathcal{M}_{\rm post}$. We will also assume that the family $\mathcal{M}_{\rm post}$ is closed under preprocessings by $\mathcal{C}_{\rm post}$, more precisely, for any R, S, T admissible in F and any ancillas U, V, we have

$$\beta \in \mathcal{C}_{\rm post}(UR_1, S_1),\ M \in \mathcal{M}_{\rm post}(VS_1T_1) \\ \implies (id_{VT_1} \otimes \beta^*)(M) \in \mathcal{M}_{\rm post}(UVR_1T_1).$$

Theorem 4. *Let $\mathsf{F} = \mathcal{C}_{\rm pre}*\mathcal{C}_{\rm post}$ and let $\mathcal{M}_{\rm post}$ be a family of measurements closed under preprocessings by $\mathcal{C}_{\rm post}$. If $\delta_{\mathsf{F}}(\Phi_1\|\Phi_2) \le \epsilon$, then for any $R = R_0R_1$ admissible in F and any ensemble $\mathcal{E}$ on R_0R_1 we have*

$$P_{\rm succ}^{\mathcal{C}_{\rm pre}(R_0,\cdot A_0'),\Phi_2,\mathcal{M}_{\rm post}}(\mathcal{E}) \\ \le P_{\rm succ}^{\mathcal{C}_{\rm pre}(R_0,\cdot A_0),\Phi_1,\mathcal{M}_{\rm post}}(\mathcal{E}) + \frac{\epsilon}{2} P_{\rm succ}(\mathcal{E}).$$

Moreover, if the Bell measurement $\mathsf{B}^{A_1'} \in \mathcal{M}_{\rm post}(A_1'A_1')$ and for any measurement $M \in \mathcal{M}_{\rm post}(\cdot A_1A_1')$ with $d_{A_1'}^2$ outcomes we have $\beta^M \in \mathcal{C}_{\rm post}$, then the converse also holds, with $R_0 \simeq A_0'$ and $R_1 \simeq A_1'$.

Proof. Assume that there is some $\Lambda \in \bar{\mathsf{F}}(A, A')$ such that $\|\Lambda(\Phi_1) - \Phi_2\|_\diamond \le \epsilon$. Let $\mathcal{E}$ be an ensemble on R_0R_1 and let $\alpha \in \mathcal{C}_{\rm pre}(R_0, UA_0')$, $M \in \mathcal{M}_{\rm post}(UA_1'R_1)$. Then

$$\|\Lambda(\Phi_1) \circ \alpha - \Phi_2 \circ \alpha\|_\diamond = \|(\Lambda(\Phi_1) - \Phi_2) \circ \alpha\|_\diamond \le \epsilon$$

and using Corollary 1, we have

$$P_{\rm succ}((\alpha \otimes id)(\mathcal{E}), (\Phi_2^* \otimes id)(M)) \\ \le P_{\rm succ}((\alpha\otimes id)(\mathcal{E}), (\Lambda(\Phi_1)^* \otimes id)(M)) + \frac{\epsilon}{2} P_{\rm succ}(\mathcal{E}).$$

Let $\Lambda = \Lambda_{\rm pre}*\Lambda_{\rm post}$, with $\Lambda_{\rm pre} \in \mathcal{C}_{\rm pre}(A_0', VA_0)$, $\Lambda_{\rm post} \in \mathcal{C}_{\rm post}(VA_1, A_1')$. Then $\alpha' = \Lambda_{\rm pre} \circ \alpha \in \mathcal{C}_{\rm pre}(R_0, UVA_0)$ and $M' = \Lambda_{\rm post}^*(M) \in \mathcal{M}_{\rm post}(UVA_1R_1)$, so that

$$P_{\rm succ}((\alpha \otimes id)(\mathcal{E}), (\Lambda(\Phi_1)^* \otimes id)(M)) \\ = P_{\rm succ}((\alpha' \otimes id)(\mathcal{E}), (\Phi_1^* \otimes id)(M')) \\ \le P_{\rm succ}^{\mathcal{C}_{\rm pre}(R_0,\cdot A_0),\Phi_1,\mathcal{M}_{\rm post}}.$$

We now prove the converse, assuming the two additional conditions. Suppose that the inequality holds with $R_0 \simeq A_0'$ and $R_1 \simeq A_1'$. Let $\rho \in \mathfrak{S}(A_0'A_1')$ and let $\mathcal{E} = \mathcal{E}_\rho^{A_1'}$ be as in Section III-A7, then by (13), we have

$$\langle \rho, \Phi_2 \rangle = d_{A_1'} P_{\rm succ}(\mathcal{E}, (\Phi_2^* \otimes id)(\mathsf{B}^{A_1'})) \\ \le d_{A_1'} P_{\rm succ}^{\mathcal{C}_{\rm pre}(A_0',\cdot A_0'),\Phi_2,\mathcal{M}_{\rm post}}(\mathcal{E}) \\ \le d_{A_1'} P_{\rm succ}^{\mathcal{C}_{\rm pre}(A_0',\cdot A_0),\Phi_1,\mathcal{M}_{\rm post}}(\mathcal{E}) + \frac{\epsilon}{2} d_{A_1'} P_{\rm succ}(\mathcal{E}),$$

here we have used that $id_{A_0'} \in \mathcal{C}_{\rm pre}(A_0', A_0')$ and $\mathsf{B}^{A_1'} \in \mathcal{M}_{\rm post}(A_1'A_1')$. Choose any $\alpha \in \mathcal{C}_{\rm pre}(A_0', UA_0)$ and $M \in \mathcal{M}_{\rm post}(UA_1A_1')$ and consider the success probability $P_{\rm succ}((\alpha \otimes id)(\mathcal{E}), (\Phi_1^* \otimes id)(M))$. Since the channel $\Phi_1 \circ \alpha$ acts only on A_0', we wee that

$$P_{\rm succ}((\alpha \otimes id)(\mathcal{E}), (\Phi_1^* \otimes id)(M)) = P_{\rm succ}(\mathcal{E}_{\tilde\rho}^{A_1'}, M),$$

with $\tilde\rho = (\Phi_1 \circ \alpha \otimes id)(\rho)$. By (14) and the assumption, we have $\beta = \beta^M \in \mathcal{C}_{\rm post}(UA_1, A_1')$ and

$$d_{A_1'} P_{\rm succ}(\mathcal{E}_{\tilde\rho}^{A_1'}, M) = \langle \tilde\rho, \beta \rangle \\ = \langle (\Phi_1 \circ \alpha \otimes id)(\rho), \beta \rangle = \langle \rho, \Theta(\Phi_1) \rangle,$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

with $\Theta = \alpha * \beta \in \mathsf{F}(A, A')$. Using the definition of $\|\cdot\|^{\mathsf{F}}_{A'|A}$ and Lemma 2 we now obtain

$$\langle \rho, \Phi_2 \rangle \leq \|\rho * C_{\Phi_1}\|^{\mathsf{F}}_{A'|A} + \frac{\epsilon}{2}\|\rho\|^{\diamond}_{A'_1|A'_0}.$$

By the condition (ii) in Theorem 3 this finishes the proof. □

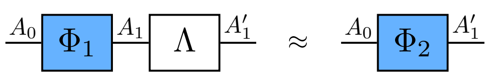
C. Comparison by postprocessings

Comparison of quantum channels by postprocessings was already considered in [10] (see also [37] for a longer version with complete proofs), where quantum versions of the randomization theorem (Example 6) were studied and the results below were obtained. We consider this case for completeness, just to show that they fit into the setting of Theorem 3.

Assume that the channels $\Phi_1 \in \mathcal{C}(A_0, A_1)$ and $\Phi_2 \in \mathcal{C}(A_0, A'_1)$ have the same input space A_0. The postprocessing deficiency of Φ_1 with respect to Φ_2 was defined in [10] as

$$\delta_{\text{post}}(\Phi_1 \| \Phi_2) := \min_{\Lambda \in \mathcal{C}(A_1, A'_1)} \|\Lambda \circ \Phi_1 - \Phi_2\|_\diamond.$$

Let F be the subcategory with admissible spaces of the form $A_0 R_1$, with a fixed input system A_0, and morphisms $\mathcal{L}(A_0, R_1) \to \mathcal{L}(A_0, S_1)$ given by postprocessings $\Phi \mapsto \Lambda \circ \Phi$ for some $\Lambda \in \mathcal{C}(R_1, S_1)$. The sought approximation now becomes



It is clear that $\delta_{\text{post}}(\Phi_1 \| \Phi_2) = \delta_{\mathsf{F}}(\Phi_1 \| \Phi_2)$ and we may apply the results of the previous section. Note also that $\mathsf{F} = \mathcal{C}_{\text{pre}} * \mathcal{C}_{\text{post}}$, where $\mathcal{C}_{\text{pre}} = \{id_{A_0}\}$ and $\mathcal{C}_{\text{post}} = \mathcal{C}$.

Theorem 5. *Let $\Phi_1 \in \mathcal{C}(A_0, A_1)$, $\Phi_2 \in \mathcal{C}(A_0, A'_1)$ and let $\epsilon \geq 0$. The following are equivalent.*

(i) $\delta_{\text{post}}(\Phi_1 \| \Phi_2) \leq \epsilon$;

(ii) *For any ancilla R_1 and $\rho \in \mathfrak{S}(A_0 R_1)$, we have*

$$\|(\Phi_2 \otimes id_{R_1})(\rho)\|^{\diamond}_{R_1|A'_1} \leq \|(\Phi_1 \otimes id_{R_1})(\rho)\|^{\diamond}_{R_1|A_1} + \frac{\epsilon}{2}\|\rho\|^{\diamond}_{R_1|A_0};$$

(iii) *For any ancilla R_1 and any ensemble $\mathcal{E}$ on $A_0 R_1$, we have*

$$P_{\text{succ}}((\Phi_2 \otimes id_{R_1})(\mathcal{E})) \leq P_{\text{succ}}((\Phi_1 \otimes id_{R_1})(\mathcal{E})) + \frac{\epsilon}{2} P_{\text{succ}}(\mathcal{E}).$$

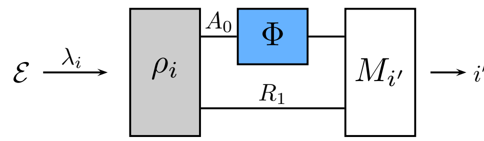
Moreover, in (ii) and (iii), it is enough to use $R_1 \simeq A'_1$.

Proof. Note that for any $\rho \in \mathfrak{S}(A_0 R_1)$ we have $\rho * C_{\Phi_i} = (\Phi_i \otimes id)(\rho)$ (see Remark 3). It follows that

$$\begin{aligned}\|\rho * C_{\Phi_2}\|^{\mathsf{F}}_{A_0 R_1 | A_0 A'_1} &= \sup_{\Lambda \in \mathcal{C}(A'_1, R_1)} \langle (\Phi_2 \otimes id)(\rho), \Lambda \rangle \\ &= \|(\Phi_2 \otimes id)(\rho)\|^{\diamond}_{R_1|A'_1}\end{aligned}$$

and similarly for Φ_1. This shows the equivalence (i) $\iff$ (ii) and the fact that we may assume $R_1 \simeq A'_1$, by a direct application of Theorem 3. The rest of the proof follows by Theorem 4 with $\mathcal{M}_{\text{post}}$ being the set of all measurements. □

As we have seen, the guessing games have a simple form here, with no preprocessings and no restrictions on the measurement M:



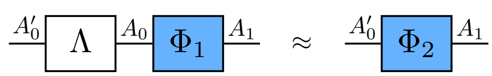
Note that classical-to-quantum channels can be identified with statistical experiments, see Example 5 in the more general GPT case. This provides another possible characterization for classical-to-quantum channels, in terms of guessing games with an inconclusive outcome, where no ancilla is needed.

For $\epsilon = 0$, we obtain an ordering on the set of channels that was treated also in [6], [7], [5], [8]. It was observed in [2] that in this case we may assume that the ensembles in (iii) consist of separable states, see [37, Theorem 2] for a proof close to the present setting. Equivalently, we may restrict to separable states in (ii). This also corresponds to the results of [8].

For similar results in the infinite dimensional case, see [23].

D. Comparison by preprocessings

This time the admissible spaces for F are $R_0 A_1$ with fixed output system A_1 and the morphisms in $\mathsf{F}(R_0 A_1, S_0 A_1)$ are restricted to preprocessings $\Phi \mapsto \Phi \circ \Lambda$ for some $\Lambda \in \mathcal{C}(S_0, R_0)$, so our problem has the form



Here $\mathsf{F} = \mathcal{C}_{\text{pre}} * \mathcal{C}_{\text{post}}$, where $\mathcal{C}_{\text{pre}} = \mathcal{C}$ and $\mathcal{C}_{\text{post}} = \{id_{A_1}\}$. We have

$$\delta_{\mathsf{F}} = \delta_{\text{pre}}(\Phi_1 \| \Phi_2) := \min_{\Lambda \in \mathcal{C}(A'_0, A_0)} \|\Phi_1 \circ \Lambda - \Phi_2\|_\diamond.$$

We will also consider the corresponding F-distance, which will be denoted by

$$\Delta_{\text{pre}} := \max\{\delta_{\text{pre}}(\Phi_1 \| \Phi_2), \delta_{\text{pre}}(\Phi_2 \| \Phi_1)\}.$$

A part of the following theorem was proved in [11].

Theorem 6. *Let $\Phi_1 \in \mathcal{C}(A_0, A_1)$, $\Phi_2 \in \mathcal{C}(A'_0, A_1)$ and let $\epsilon \geq 0$. The following are equivalent.*

(i) $\delta_{\text{pre}}(\Phi_1 \| \Phi_2) \leq \epsilon$;

(ii) *For any ancilla R_0 and $\rho \in \mathfrak{S}(R_0 A_1)$, we have*

$$\|(id_{R_0} \otimes \Phi_2^{\mathsf{T}})(\rho)\|^{\diamond}_{A'_0|R_0} \leq \|(id_{R_0} \otimes \Phi_1^{\mathsf{T}})(\rho)\|^{\diamond}_{A_0|R_0} + \frac{\epsilon}{2}\|\rho\|^{\diamond}_{A_1|R_0},$$

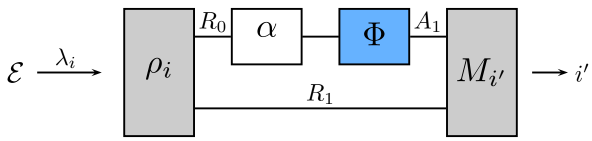
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

(iii) *For any ancilla R_0, any ensemble $\mathcal{E}$ on R_0A_1 and any fixed measurement M on A_1A_1, we have*

$$P_{\rm succ}^{\mathcal{C}(R_0,A_0'),\Phi_2,\{M\}}(\mathcal{E}) \\ \leq P_{\rm succ}^{\mathcal{C}(R_0,A_0),\Phi_1,\{M\}}(\mathcal{E}) + \frac{\epsilon}{2}P_{\rm succ}(\mathcal{E}).$$

Moreover, in (ii) and (iii), it is enough to use $R_0 \simeq A_0'$.

The guessing games in (iii) are depicted in the diagram

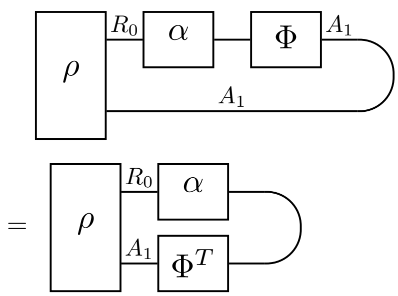


Here the preprocessing α can be chosen freely and the measurement M is fixed.

Proof. The equivalence (i) $\iff$ (ii) follow from Thm. 3 and

$$\langle \rho, \Phi\circ\alpha\rangle = (\rho * C_\Phi) * C_\alpha = \langle (id_{R_0}\otimes \Phi^{\mathsf{T}})(\rho), \alpha\rangle$$

(see Remark 3). In diagram:



The implication (i) $\implies$ (iii) follows by Theorem 4, with $\mathcal{M}_{\rm post} = \{M\}$. To finish the proof, we put $M = \mathsf{B}^{A_1}$ and observe that we have $\beta^M = id_{A_1} \in \mathcal{C}_{\rm post}$. $\square$

Let us consider the more general situation when $\mathsf{F} = \mathcal{O}*\{id_{A_1}\}$ where $\mathcal{O}(S_0,R_0)\subset\mathcal{C}(S_0,R_0)$ is some suitable subset. Then

$$\delta_{\mathsf{F}}(\Phi_1\|\Phi_2) = \inf_{\Lambda\in\mathcal{O}(A_0',A_0)} \|\Phi_1\circ\Lambda - \Phi_2\|_\diamond.$$

It can be seen as in the above proof that we have $\delta_{\mathsf{F}}(\Phi_1\|\Phi_2) \leq \epsilon$ if and only if for any ensemble $\mathcal{E}$ on $A_0'A_1$ and any fixed measurement M on A_1A_1 we have

$$P_{\rm succ}((\Phi_2\otimes id)(\mathcal{E}), M) \\ \leq \sup_{\Lambda\in\mathcal{O}} P_{\rm succ}((\Phi_1\circ\Lambda\otimes id)(\mathcal{E}), M) + \frac{\epsilon}{2}P_{\rm succ}(\mathcal{E}).$$

Let now $\Phi_1 = id_{A_1}$, $\Phi_2 = \Phi \in \mathcal{C}(A_0,A_1)$, then

$$\delta_{\mathsf{F}}(id\|\Phi) = \inf_{\Lambda\in\mathcal{O}(A_0,A_1)} \|\Lambda - \Phi\|_\diamond,$$

that is, the distance of Φ to the set $\mathcal{O}$. If $\mathcal{O}$ is the set of free channels in a resource theory for quantum channels, then this distance is a resource measure, [16]. The above considerations now give the following operational characterization of this distance.

Corollary 2. $\inf_{\Lambda\in\mathcal{F}(A_0,A_1)} \|\Lambda - \Phi\|_\diamond \leq \epsilon$ *if and only if for any ensemble $\mathcal{E}$ on A_0A_1 and any measurement M on A_1A_1 we have*

$$P_{\rm succ}((\Phi\otimes id)(\mathcal{E}), M) \\ \leq \sup_{\Lambda\in\mathcal{O}} P_{\rm succ}((\Lambda\otimes id)(\mathcal{E}), M) + \frac{\epsilon}{2}P_{\rm succ}(\mathcal{E}).$$

1) $\Delta_{\rm pre}$ as the distance of ranges: In this paragraph, we obtain a characterization of $\delta_{\rm pre}$ and the pseudo-distance $\Delta_{\rm pre}$ in terms of the ranges of channels. Recall that the range of a channel $\Phi\in\mathcal{C}(A_0,A_1)$ is defined as

$$\mathcal{R}(\Phi) = \Phi(\mathfrak{S}(A_0)).$$

Our first result in this direction is based on the following simple lemma. The proof is rather standard and is included for the convenience of the reader.

Lemma 3. *Let $\sigma\in\mathfrak{S}(AR)$ and let $Z\in\mathcal{B}(R)$ be such that $\sigma_R = {\rm Tr}_A[\sigma] = ZZ^*$. Then there is some channel $\beta\in\mathcal{C}(R,A)$ such that $\sigma = (\beta\otimes id)(|Z^{\mathsf{T}}\rangle\!\rangle\langle\!\langle Z^{\mathsf{T}}|)$.*

Proof. Let $p = {\rm supp}(\sigma)$ and let $p_R = {\rm supp}(\sigma_R)$, then $p \leq I\otimes p_R$ and we may put

$$C := (I\otimes U^*\sigma_R^{-1/2})\sigma(I\otimes\sigma_R^{-1/2}U) + d_A^{-1}I_A\otimes(I-U^*p_RU),$$

where $U\in\mathcal{B}(R)$ is a unitary such that $Z = \sigma_R^{1/2}U$ and the inverse is restricted to p_R. Since $C\geq 0$ and ${\rm Tr}_A[C] = I_R$, there is some $\beta\in\mathcal{C}(R,A)$ such that $C = C_\beta$. We have

$$\sigma = (I\otimes Z)C_\beta(I\otimes Z^*) = (\beta\otimes id)(|Z^{\mathsf{T}}\rangle\!\rangle\langle\!\langle Z^{\mathsf{T}}|).$$

$\square$

Corollary 3. *Let $\Phi_1\in\mathcal{C}(A_0,A_1)$, $\Phi_2\in\mathcal{C}(A_0',A_1)$. Then*

$$\delta_{\rm pre}(\Phi_1\|\Phi_2) \\ = \sup_{\xi\in\mathfrak{S}(A_0'R)} \inf_{\substack{\sigma\in\mathfrak{S}(A_0R)\\ \sigma_R=\xi_R}} \|(\Phi_1\otimes id)(\sigma) - (\Phi_2\otimes id)(\xi)\|_1$$

where $R\simeq A_0'$.

Proof. Let $\Lambda\in\mathcal{C}(A_0',A_0)$ be such that $\|\Phi_1\circ\Lambda - \Phi_2\|_\diamond \leq \epsilon$. Then for any $\xi\in\mathfrak{S}(A_0'R)$, we have

$$\|(\Phi_1\circ\Lambda\otimes id)(\xi) - (\Phi_2\otimes id)(\xi)\|_1 \leq \|\Phi_1\circ\Lambda - \Phi_2\|_\diamond \leq \epsilon.$$

Put $\sigma = (\Lambda\otimes id)(\xi)$, then we also have $\sigma_R = \xi_R$, so that the supremum on the right hand side is upper bounded by $\delta_{\rm pre}$. For the converse, let $\rho\in\mathfrak{S}(A_0'A_1)$. By Lemma 1, there is some $V\in\mathcal{B}(A_0')$, ${\rm Tr}\,[VV^*] = 1$ and an element $G\in\mathcal{B}_+(A_0'A_1)$ such that (recall that $\chi_V = V\cdot V^*$)

$$\rho = (\chi_V\otimes id)(G), \qquad \|\rho\|^\diamond_{A_1|A_0'} = \|G\|.$$

Using (5) and (6), we have

$$\langle\rho,\Phi_2\rangle = \langle(\chi_V\otimes id)(G),\Phi_2\rangle = \langle G, \Phi_2\circ\chi_V\rangle \\ = {\rm Tr}\,[C_{\Phi_2\circ\chi_V}\tilde{G}] = {\rm Tr}\,[(\Phi_2\otimes id)(|V\rangle\!\rangle\langle\!\langle V|)\tilde{G}],$$

where $\tilde{G} = \mathcal{U}^*_{A_1,A_0'}(G^\mathsf{T})$. Note that $\xi := |V\rangle\!\rangle\langle\!\langle V| \in \mathfrak{S}(A_0'A_0')$, with $\mathrm{Tr}_1[\xi] = V^\mathsf{T}(V^\mathsf{T})^*$. Assume that there is some $\sigma \in \mathfrak{S}(A_0A_0')$ with $\sigma_{A_0'} = \mathrm{Tr}_1[\xi]$ and

$$\|(\Phi_2 \otimes id)(\xi) - (\Phi_1 \otimes id)(\sigma)\|_1 \leq \epsilon.$$

By Lemma 3, there is some channel $\beta \in \mathcal{C}(A_0', A_0)$ such that $\sigma = (\beta \otimes id)(|V\rangle\!\rangle\langle\!\langle V|)$. We now have

$$\begin{aligned}\langle \rho, \Phi_2 - \Phi_1 \circ \beta\rangle &= \langle G, (\Phi_2 - \Phi_1 \circ \beta) \circ \chi_V\rangle \\ &= \mathrm{Tr}\,[((\Phi_2 \otimes id)(\xi) - (\Phi_1 \otimes id)(\sigma))\tilde{G}] \\ &\leq \frac{1}{2}\|(\Phi_2 \otimes id)(\xi) - (\Phi_1 \otimes id)(\sigma)\|_1\|G\| \\ &\leq \frac{\epsilon}{2}\|\rho\|^\diamond_{A_1|A_0'},\end{aligned}$$

here the inequalities follow from the fact that $\tilde{G} \in \mathcal{B}_+(A_1A_0')$, properties of the trace norm $\|\cdot\|_1$ and $\|\tilde{G}\| = \|G\| = \|\rho\|^\diamond_{A_1|A_0'}$. From Theorem 3 (ii), we obtain that $\delta_{\mathrm{pre}}(\Phi_1\|\Phi_2) \leq \epsilon$.

□

Using the above corollary, we immediately obtain that for any R with $d_R \geq d_{A_0'}$, $\delta_{\mathrm{pre}}(\Phi_1\|\Phi_2) = 0$ is equivalent to the inclusion

$$\mathcal{R}(\Phi_2 \otimes id_R) \subseteq \mathcal{R}(\Phi_1 \otimes id_R).$$

In the case of q-c channels, that is for measurements (POVMs), this result was proved in [38], where also a counterexample was given, showing that inclusion of the ranges of the channels is not enough for existence of even a positive preprocessing, so that tensoring with id_R is necessary in general.

We next show that the pseudo-distance Δ_{pre} can be expressed as a distance of ranges. Recall that for two subsets S, T of a metric space with metric m, the Hausdorff distance is defined by

$$m_H(S,T) = \max\{\sup_{s\in S}\inf_{t\in T} m(s,t), \sup_{t\in T}\inf_{s\in S} m(s,t)\}.$$

A natural choice for a metric on the set of states would be the trace distance

$$\|\sigma - \rho\|_1 = \mathrm{Tr}\,|\sigma - \rho|, \qquad \sigma, \rho \in \mathfrak{S}(A_1R).$$

As it turns out, we will have to add a term for the distance of the restrictions to R. For $\sigma_1, \sigma_2 \in \mathfrak{S}(R)$, let $p(\sigma_1, \sigma_2)$ denote the purified distance

$$\begin{aligned}p(\sigma_1,\sigma_2) &= \sqrt{1 - F(\sigma_1,\sigma_2)^2} \\ &= \inf_{\substack{V_1V_1^*=\sigma_1,\\ V_2V_2^*=\sigma_2}} \frac{1}{2}\||V_1^\mathsf{T}\rangle\!\rangle\langle\!\langle V_1^\mathsf{T}| - |V_2^\mathsf{T}\rangle\!\rangle\langle\!\langle V_2^\mathsf{T}|\|_1,\end{aligned}$$

where F denotes the fidelity $F(\sigma_1, \sigma_2) = \|\sigma_1^{1/2}\sigma_2^{1/2}\|_1$.

Corollary 4. *For $\Phi_1 \in \mathcal{C}(A_0, A_1)$ and $\Phi_2 \in \mathcal{C}(A_0', A_1)$, we have*

$$\Delta_{\mathrm{pre}}(\Phi_1,\Phi_2) = m_H(\mathcal{R}(\Phi_1 \otimes id_R), \mathcal{R}(\Phi_2 \otimes id_R)),$$

where $d_R = \max\{d_{A_0}, d_{A_0'}\}$ and m_H is the Hausdorff distance with respect to the metric m in $\mathfrak{S}(A_1R)$, given as

$$m(\xi,\sigma) = \|\xi - \sigma\|_1 + 2p(\xi_R, \sigma_R)$$

Proof. From Corollary 3, we easily obtain that $m_H(\mathcal{R}(\Phi_1 \otimes id_R), \mathcal{R}(\Phi_2 \otimes id_R)) \leq \Delta_{\mathrm{pre}}(\Phi_1,\Phi_2)$. For the converse, put

$$\epsilon := m_H(\mathcal{R}(\Phi_1 \otimes id_R), \mathcal{R}(\Phi_2 \otimes id_R)).$$

The idea of the proof is similar to the previous proof. Let $\rho \in \mathfrak{S}(RA_1)$, $\alpha \in \mathcal{C}(R, A_0')$ and let V and G be connected to ρ as in the proof of Corollary 3. Let also $\xi := (\alpha \otimes id)(|V\rangle\!\rangle\langle\!\langle V|) \in \mathfrak{S}(A_0'R)$. Then there is some $\sigma \in \mathfrak{S}(A_0R)$ such that

$$m((\Phi_2 \otimes id)(\xi), (\Phi_1 \otimes id)(\sigma)) \leq \epsilon.$$

Note that unlike the previous proof, we may now have $\xi_R \neq \sigma_R$. Let $W \in \mathcal{B}(R)$ be such that $W^\mathsf{T}(W^\mathsf{T})^* = \sigma_R$ and

$$2p(\xi_R,\sigma_R) = \||V\rangle\!\rangle\langle\!\langle V| - |W\rangle\!\rangle\langle\!\langle W|\|_1$$

(such W always exists since $\|\cdot\|_1$ is unitarily invariant). By Lemma 3, there is a channel $\gamma \in \mathcal{C}(R, A_0)$ such that

$$\sigma = (\gamma \otimes id)(|W\rangle\!\rangle\langle\!\langle W|) = C_{\gamma\circ\chi_W}.$$

We now have

$$\begin{aligned}&\langle \rho, \Phi_2 \circ \alpha - \Phi_1 \circ \gamma\rangle = \langle G, (\Phi_2 \circ \alpha - \Phi_1 \circ \gamma) \circ \chi_V\rangle \\ &= \langle G, \Phi_2 \circ \alpha \circ \chi_V - \Phi_1 \circ \gamma \circ \chi_W\rangle \\ &\quad + \langle G, \Phi_1 \circ \gamma \circ (\chi_W - \chi_V)\rangle \\ &= \mathrm{Tr}\,[((\Phi_2 \otimes id)(\xi) - (\Phi_1 \otimes id)(\sigma))\tilde{G}] \\ &\quad + \mathrm{Tr}\,[(\Phi_1 \circ \gamma \otimes id)(|W\rangle\!\rangle\langle\!\langle W| - |V\rangle\!\rangle\langle\!\langle V|)\tilde{G}] \\ &\leq \frac{1}{2}\|\rho\|^\diamond_{A_1|R}\Big(\|(\Phi_2 \otimes id)(\xi) - (\Phi_1 \otimes id)(\sigma)\|_1 \\ &\quad + \tilde{m}_B(\xi_R,\sigma_R)\Big) \leq \frac{\epsilon}{2}\|\rho\|^\diamond_{A_1|R}.\end{aligned}$$

The last inequality implies that $\delta_{\mathrm{pre}}(\Phi_1\|\Phi_2) \leq \epsilon$ and we similarly obtain that also $\delta_{\mathrm{pre}}(\Phi_2\|\Phi_1) \leq \epsilon$.

□

E. Comparison of bipartite channels by LOCC superchannels

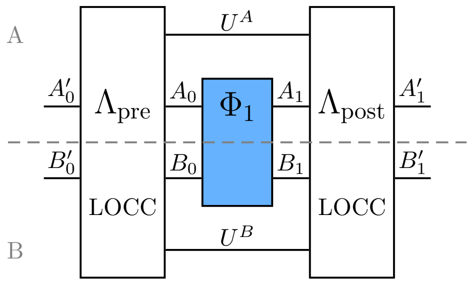
In this section, the objects in F are spaces of bipartite quantum channels $\mathcal{L}(A_0B_0, A_1B_1)$ and the morphisms are restricted to LOCC superchannels, that is,

$$\mathsf{F} = \mathcal{C}_2^{\mathrm{LOCC}} := \mathcal{C}_{\mathrm{LOCC}} * \mathcal{C}_{\mathrm{LOCC}}$$

where $\mathcal{C}_{\mathrm{LOCC}}$ is the set of $A|B$ LOCC channels. To be more precise, in this case the admissible spaces are of the form $R^AR^B = R_0^AR_0^BR_1^AR_1^B$ and the morphisms in $\mathsf{F}(R^AR^B, S^AS^B)$ have the form $\Lambda_{\mathrm{pre}} * \Lambda_{\mathrm{post}}$ with $\Lambda_{\mathrm{pre}} \in \mathcal{C}_{\mathrm{LOCC}}(S_0^A|S_0^B, U^AR_0^A|R_0^BU^B)$, $\Lambda_{\mathrm{post}} \in \mathcal{C}_{\mathrm{LOCC}}(U^AR_1^A|R_1^BU^B, S_1^A|S_1^B)$, so that also the ancilla

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

consists of two parts $U = U^A U^B$. The task becomes to simulate the channel Φ_2 as



In this case,

$$\delta_{\mathsf{F}}(\Phi_1\|\Phi_2) = \delta_{\rm LOCC}(\Phi_1\|\Phi_2) := \inf_{\Lambda\in\mathcal{C}_2^{\rm LOCC}} \|\Lambda(\Phi_1)-\Phi_2\|_\diamond$$

is the LOCC conversion distance $\Phi_1 \to \Phi_2$. We will use the notation $\|\rho\|_{S|R}^{2-{\rm LOCC}} := \|\rho\|_{S|R}^{\mathsf{F}}$ for $\rho \in \mathcal{B}_+(R^A R^B S^A S^B)$.

Theorem 7. *Let $\Phi_1 \in \mathcal{C}(A_0B_0, A_1B_1)$ and $\Phi_2 \in \mathcal{C}(A_0'B_0', A_1'B_1')$, $\epsilon \geq 0$. The following are equivalent.*

(i) $\delta_{\rm LOCC}(\Phi_1\|\Phi_2) \leq \epsilon$;

(ii) *for any spaces $R^A R^B$ and any $\rho \in \mathfrak{S}(R^A R^B)$, we have*

$$\begin{aligned}&\|\rho\otimes C_{\Phi_2}\|^{2-{\rm LOCC}}_{R^AR^B|A'B'}\\ &\quad\leq \|\rho\otimes C_{\Phi_1}\|^{2-{\rm LOCC}}_{R^AR^B|AB} + \frac{\epsilon}{2}\|\rho\|^\diamond_{R_1^AR_1^B|R_0^AR_0^B}\end{aligned}$$

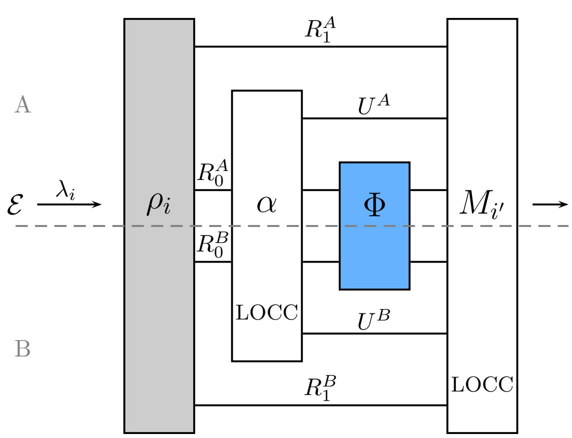
(iii) *For any spaces $R^A R^B$ and any ensemble $\mathcal{E}$ on $R^A R^B$, we have*

$$\begin{aligned}&P_{\rm succ}^{\mathcal{C}_{\rm LOCC}(R_0^AR_0^B,\cdot A_0'B_0'\cdot)\Phi_2,\mathcal{M}_{\rm LOCC}}(\mathcal{E})\\ &\leq P_{\rm succ}^{\mathcal{C}_{\rm LOCC}(R_0^AR_0^B,\cdot A_0B_0\cdot),\Phi_1,\mathcal{M}_{\rm LOCC}}(\mathcal{E})+\frac{\epsilon}{2}P_{\rm succ}(\mathcal{E})\end{aligned}$$

where $\mathcal{M}_{\rm LOCC}$ is the set of LOCC measurements.

Moreover, in (ii) and (iii) it is enough to take $R_0^A \simeq A_0'$, $R_1^A \simeq A_1'$, $R_0^B \simeq B_0'$, $R_1^B \simeq B_1'$.

The guessing games in (iii) have the form

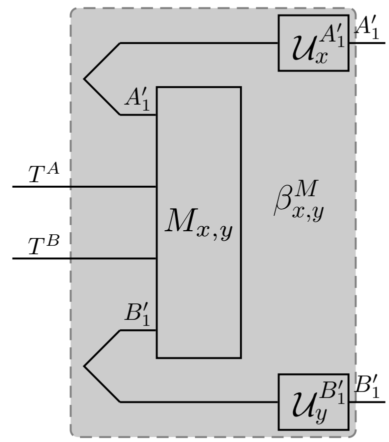


Here α can be any LOCC preprocessing channel and M any LOCC measurement.

Proof. The equivalence of (i) and (ii) is proved exactly as before from Theorem 3 and Remark 3. To prove the condition (iii), we invoke Theorem 4 with $\mathcal{M}_{\rm post} = \mathcal{M}_{\rm LOCC}$, this is obviously closed under $\mathcal{C}_{\rm LOCC}$. To prove the two additional conditions in Theorem 4, observe that the group of generalized Pauli unitaries on $A_1'B_1'$ has the form

$$\{U_{x,y}^{A_1'B_1'} = U_x^{A_1'}\otimes U_y^{B_1'},\ x=1,\dots,d_{A_1'}^2,\ y=1,\dots,d_{B_1'}^2\}$$

and we have $\mathsf{B}_{x,y}^{A_1'B_1'} = \mathsf{B}_x^{A_1'} \otimes \mathsf{B}_y^{B_1'} \in \mathcal{M}_{\rm LOCC}(A_1'A_1'|B_1'B_1')$. It is now enough to show that for any measurement $M \in \mathcal{M}_{\rm LOCC}(A_1'U^AA_1|B_1U^BB_1')$ with outcomes labeled by x, y we have $\beta^M \in \mathcal{C}_{\rm LOCC}(U^AA_1|B_1U^B, A_1'|B_1')$. Recall that β^M is a channel associated with the instrument $\{\beta_{x,y}^M\}$, where the operations $\beta_{x,y}^M$ have the form



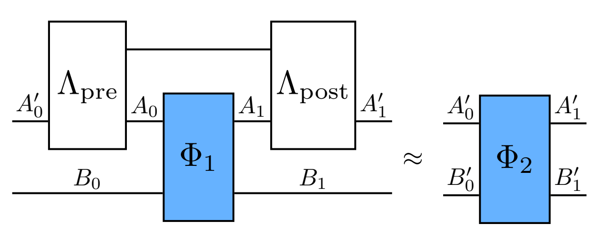
here $T^A = U^AA_1$ and $T_B = B_1U^B$. It is quite obvious that β^M is LOCC if M is.

□

Note that we have a similar situations for example for restricted PPT or SEP superchannels, where $\mathcal{C}_{\rm pre} = \mathcal{C}_{\rm post} = \mathcal{C}_{\rm PPT}$ or $\mathcal{C}_{\rm SEP}$.

F. Comparison of bipartite channels by partial superchannels

In this section, the objects of F are again spaces of bipartite channels, but this time the morphisms are given by applying arbitrary superchannels on the A part, in diagram



More precisely, the admissible spaces for F are of the form RB, where $R = R_0R_1$ are arbitrary and $B = B_0B_1$ with B_0, B_1 fixed. The morphisms $\mathsf{F}(RB, SB)$ in F are given by elements of $\mathcal{C}_2(R, S)$. Here we have $\mathsf{F} = \mathcal{C}_{\rm pre} * \mathcal{C}_{\rm post}$ with $\mathcal{C}_{\rm pre} = \mathcal{C}\otimes\{id_{B_0}\}$ and $\mathcal{C}_{\rm post} = \mathcal{C}\otimes\{id_{B_1}\}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

For $\Phi_1 \in \mathcal{C}(A_0B_0, A_1B_1)$ and $\Phi_2(A_0'B_0, A_1'B_1)$ we denote

$$\delta_{A|A'}(\Phi_1\|\Phi_2) := \delta_{\mathsf{F}}(\Phi_1\|\Phi_1) \\ = \min_{\Theta\in\mathcal{C}_2(A,A')} \|(\Theta\otimes id_B)(\Phi_1)-\Phi_2\|_\diamond.$$

Theorem 8. *Let $\Phi_1 \in \mathcal{C}(A_0B_0, A_1B_1)$, $\Phi_2 \in \mathcal{C}(A_0'B_0, A_1'B_1)$ and let $\epsilon \geq 0$. The following are equivalent.*

(i) $\delta_{A|A'}(\Phi_1\|\Phi_2) \leq \epsilon$;

(ii) *For any spaces R_0, R_1 and $\rho \in \mathfrak{S}(R_0B_0R_1B_1)$, we have*

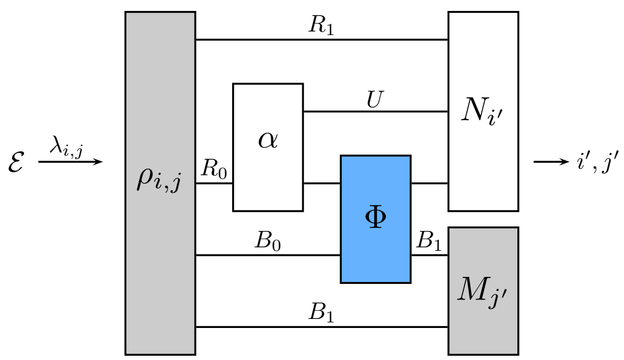
$$\|\rho * C_{\Phi_2}\|^{2\diamond}_{R|A'} \leq \|\rho * C_{\Phi_1}\|^{2\diamond}_{R|A} + \frac{\epsilon}{2}\|\rho\|^{\diamond}_{R_1B_1|R_0B_0};$$

(iii) *For any spaces R_0, R_1, $k, l \in \mathbb{N}$, any ensemble $\mathcal{E}$ with kl elements on $R_0B_0R_1B_1$, any fixed measurement $M \in \mathcal{M}_l(B_1B_1)$, we have*

$$P_{\rm succ}^{\mathcal{C}(R_0,A_0'\cdot),\Phi_2,\mathcal{M}_k\otimes M}(\mathcal{E}) \\ \leq P_{\rm succ}^{\mathcal{C}(R_0,A_0\cdot),\Phi_1,\mathcal{M}_k\otimes M}(\mathcal{E}) + \frac{\epsilon}{2}P_{\rm succ}(\mathcal{E}).$$

Moreover, in (ii) and (iii) it is enough to put $R_0 \simeq A_0'$, $R_1 \simeq A_1'$.

The guessing games have the form depicted in the diagram:



Here $\mathcal{E} = \{\lambda_{i,j}, \rho_{i,j}\}_{\substack{i=1,\dots,k\\ j=1,\dots,l}}$ is an and ensemble of states on $R_0B_0R_1B_1$ and $M \in \mathcal{M}_l(B_1B_1)$ is a fixed measurement. The preprocessing α and the measurement N with k outcomes can be chosen freely, with no restriction on the ancilla U.

Proof. As before, the equivalence (i) $\iff$ (ii) follows from Thm. 3 and the definition of the norm $\|\cdot\|^{2\diamond}$, taking into account Remark 3.

For the implication (i) $\implies$ (iii) we use Theorem 4 with $\mathcal{M}_{\rm post} = \mathcal{M}_k \otimes M$, which is clearly closed under preprocessings from $\mathcal{C}_{\rm post}$. For the converse, let $l = d_{B_1}^2$, $k = d_{a_1'}^2$ and put the fixed measurement to $M_y = \mathsf{B}_y^{B_1}$. Then $\mathsf{B}^{A_1'B_1} = \mathsf{B}^{A_1'} \otimes \mathsf{B}^{B_1} \in \mathcal{M}_k \otimes M = \mathcal{M}_{\rm post}$. Moreover, any measurement in $\mathcal{M}_k\otimes M$ on $UA_1B_1A_1'B_1$ has the form $\tilde{N} = N\otimes \mathsf{B}^{B_1}$ for some $N \in \mathcal{M}_k(UA_1A_1')$. The channel $\beta^{\tilde{N}} \in \mathcal{C}(UA_1B_1, A_1'B_1)$ clearly satisfies

$$\beta^{\tilde{N}} = \beta^N \otimes \beta^{\mathsf{B}^{B_1}} = \beta^N \otimes id_{B_1} \in \mathcal{C}_{\rm post}.$$

The proof is finished by the second part of Theorem 4. $\square$

The case $\epsilon = 0$ was also treated in [9]. By a careful comparison of the results there to the present setting, one can see that Theorem 7 in [9] corresponds to our condition (ii), but restricted to ρ of the form

$$\rho = \sum_{x=1}^{d^2 B_0}\sum_{y=1}^{d_{B_1}^2} |x\rangle\langle x| \otimes E_y \otimes \Lambda_{y|x},$$

where $|x\rangle\langle x|$ is a normalized rank 1 basis of B_0, E_y is an informationally complete POVM on B_1 and $\Lambda_{y|x} \in \mathcal{L}_+(R_0, R_1)$ is such that $\sum_y \Lambda_{y|x}$ is a channel for any x. The possibility of such a restriction seems specific to the case $\epsilon = 0$, as it is basically a consequence of the fact that to check equality of two channels it is enough to input some basis states and measure by an IC POVM. Apart from some special cases (e.g. the one treated below), this cannot be extended to $\epsilon > 0$ since in general one needs entangled states to attain the diamond norm.

G. Classical simulability of measurements

As a further application, we investigate the problem of classical simulability of measurements. In this problem, two sets of measurements $\mathbb{M} = \{M^1, \dots, M^k\}$, $M^i \in \mathcal{M}_l(C)$, $i = 1, \dots, k$ and $\mathbb{N} = \{N^1, \dots, N^m\}$, $N^y \in \mathcal{M}_n(C)$, $y = 1, \dots, m$ are given. We will say that $\mathbb{M}$ can simulate $\mathbb{N}$ if all elements in $\mathbb{N}$ can be obtained as convex combinations of postprocessings of elements in $\mathbb{M}$, [4]. It can be seen that we may exchange the order of convex combinations and postprocessings, and always obtain the same notion of simulability. Our aim is to study an approximate version with respect to some suitable norm. In particular, we will show that this problem can be put into the setting of Section III-F: we represent $\mathbb{M}$ and $\mathbb{N}$ by bipartite channels and express the simulations as applications of superchannels to one of the parts.

We will need the following type of guessing games. Let $\mathcal{E} = \{\lambda_x, \rho_x\}_{x=1}^n$ be an ensemble of states of A, but assume that only one fixed measurement $M \in \mathcal{M}_l(A)$ can be performed and the true state has to be guessed using its outcome. Any guessing procedure is described by conditional probabilities $\{p(x|j)\}$, giving the probability of guessing $x \in \{1, \dots, n\}$ if $j \in \{1, \dots, l\}$ was measured. This defines a measurement $N \in \mathcal{M}_n(A)$ given as

$$N_x = \sum_j p(x|j)M_j, \quad x = 1, \dots, n.$$

Such measurement is a postprocessing of M. The average probability of a correct guess using this procedure is

$$P_{\rm succ}(\mathcal{E}, N) = P_{\rm succ}(\mathcal{E}, M, p) := \sum_{x,j} \lambda_x p(x|j) {\rm Tr}\,[\rho_x M_j]$$

and the maximal success probability is denoted by (cf. [3])

$$P_{\rm succ}^Q(\mathcal{E}, M) := \sup_{\{p(x|j)\}} P_{\rm succ}(\mathcal{E}, M, p).$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIT.2021.3070120, IEEE Transactions on Information Theory

Any set of conditional probabilities can be identified with the classical-to-classical (c-c) channel in $\mathcal{C}(S,R)$, $d_S = l$, $d_R = n$ determined by the Choi matrix

$$C_p := \sum_{x,j} p(x|j)|x\rangle\langle x| \otimes |j\rangle\langle j|.$$

This channel will be denoted by p. We then have $\Phi_N = p \circ \Phi_M$ for the q-c channels given by the measurements M and N. Moreover, for any $\alpha \in \mathcal{C}(S,R)$, there are conditional probabilities $\{p(x|j) := \langle x|\alpha(|j\rangle\langle j|)|x\rangle\}$, such that

$$\begin{aligned}\langle (\Phi_M \otimes id)(\rho_\mathcal{E}), \alpha\rangle &= \langle \rho_\mathcal{E}, \alpha \circ \Phi_M\rangle \\ &= \langle \rho_\mathcal{E}, p \circ \Phi_M\rangle = P_{\mathrm{succ}}(\mathcal{E}, M, p), \quad (16)\end{aligned}$$

This proves the following result.

Lemma 4. *For any ensemble $\mathcal{E}$ on A and $M \in \mathcal{M}_{d_S}(A)$,*

$$P^Q_{\mathrm{succ}}(\mathcal{E}, M) = \|(\Phi_M \otimes id)(\rho_\mathcal{E})\|^\diamond_{S|A} = P_{\mathrm{succ}}(\Phi_M(\mathcal{E})).$$

Let now $\mathbb{M}$ and $\mathbb{N}$ be sets of measurements as above. By definition, $\mathbb{M}$ can simulate $\mathbb{N}$ if for each $N^y \in \mathbb{N}$ there are probabilities $q(i|y)$ such that

$$\Phi_{N^y} = \sum_i q(i|y) p_{i,y} \circ \Phi_{M^i}$$

for some c-c channels $p_{i,y}$ determined by sets of conditional probabilities $\{p_{i,y}(x|j),\ x = 1,\dots,n,\ j = 1,\dots,l\}$, $i = 1,\dots,k$, $y = 1,\dots,m$. That is,

$$N^y_x = \sum_{i,j} p(i,x|j,y) M^i_j, \qquad (17)$$

wfor $x = 1,\dots,n,$, $y = 1,\dots,m$, here $p(i,x|j,y) := q(i|y)p_{i,y}(x|j)$ are conditional probabilities. Let $\Phi_\mathbb{M}$ be a channel in $\mathcal{C}(A_0C, A_1)$, $d_{A_0} = k$, $d_{A_1} = l$ with the Choi matrix

$$C_\mathbb{M} := \sum_{i,j} |j\rangle\langle j| \otimes |i\rangle\langle i| \otimes (M^i_j)^\mathsf{T}.$$

Note that $\Phi_\mathbb{M}$ is a bipartite channel as in the setting of Theorem 8, with $B_0 = C$ and $B_1 = 1$, moreover, the first input and the output of $\Phi_\mathbb{M}$ is classical. Similarly, $\mathbb{N}$ is represented by the channel $\Phi_\mathbb{N} \in \mathcal{C}(A'_0C, A'_1)$, with $d_{A'_0} = m$, $d_{A'_1} = n$. It can be easily checked that (17) can be expressed as

$$C_p * C_\mathbb{M} = C_\mathbb{N},$$

where $p \in \mathcal{C}(A'_0A_1, A_0A'_1)$ is the c-c channel given by $\{p(i,x|j,y)\}$. Using (4), we ca see that p is a superchannel if and only if there are conditional probabilities $\{q(i|y)\}$ such that

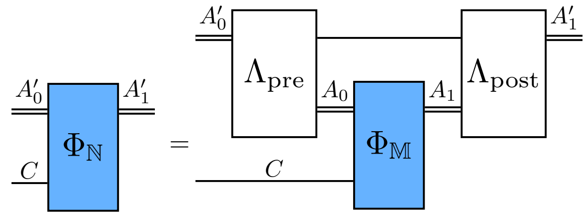
$$\sum_x p(x,i|y,j) = q(i|y), \qquad \forall i,y,\ \forall j.$$

By putting $p_{i,y}(x|j) := q(i|y)^{-1}p(x,i|y,j)$ if $q(i|y) > 0$ and choosing any conditional probabilities for $p_{i,y}$ otherwise, we obtain the following result.

Lemma 5. *A c-c channel $p \in \mathcal{C}(A'_0A_1, A_0A'_1)$ is a superchannel if and only if there are conditional probabilities $\{p_{i,y}(x|j)\}$ and $\{q(i|y)\}$, with $i = 1,\dots,d_{A_0}$, $j = 1,\dots,d_{A_1}$, $x = 1,\dots,d_{A'_1}$, $y = 1,\dots,d_{A'_0}$ such that*

$$p(x,i|y,j) = q(i|y)p_{i,y}(x|j).$$

We obtain that $\mathbb{N}$ is simulable by $\mathbb{M}$ if and only if $\Phi_\mathbb{N} = p(\Phi_\mathbb{M})$ for some c-c superchannel p, in diagram



here the double lines denote classical inputs and outputs.

We now introduce the following notion of approximate simulability: for $\epsilon \geq 0$, we say that $\mathbb{N}$ is ϵ-simulable by $\mathbb{M}$ if $\mathbb{M}$ can simulate some set of measurements $\mathbb{N}' = \{(M')^1,\dots,(M')^m\} \subset \mathcal{M}_n(C)$ such that

$$\|\Phi_\mathbb{N} - \Phi_{\mathbb{N}'}\|_\diamond \leq \epsilon.$$

The next result shows that approximate simulability is expressed by the conversion distance $\delta_{A|A'}(\Phi_\mathbb{M}\|\Phi_\mathbb{N})$.

Proposition 4. *Let $\mathbb{M} = \{M^1,\dots,M^k\} \subset \mathcal{M}_l(C)$, $\mathbb{N} = \{N^1,\dots,N^m\} \subset \mathcal{M}_n(C)$, $\epsilon \geq 0$. Let A_0, A_1, A'_0, A'_1 be systems such that $d_{A_0} = k$, $d_{A_1} = l$, $d_{A'_0} = m$, $d_{A'_1} = n$. Then $\mathbb{N}$ is ϵ-simulable by $\mathbb{M}$ if and only if*

$$\delta_{A|A'}(\Phi_\mathbb{M}\|\Phi_\mathbb{N}) \leq \epsilon.$$

Proof. Assume that $\mathbb{N}$ is ϵ-simulable by $\mathbb{M}$. As shown above, there is some c-c superchannel $p \in \mathcal{C}_2(A, A')$ such that $\|p(\Phi_\mathbb{M}) - \Phi_\mathbb{N}\|_\diamond \leq \epsilon$ and hence $\delta_{A|A'}(\Phi_\mathbb{M}\|\Phi_\mathbb{N}) \leq \epsilon$. For the converse, for any system D, let $\Delta_D \in \mathcal{C}(D,D)$ denote the channel that maps any $X \in \mathcal{B}(D)$ to its diagonal elements in the basis $|i\rangle_D$:

$$\Delta_D(X) = \sum_{i=1}^{d_D} \langle i|X|i\rangle |i\rangle\langle i|.$$

For any $\Theta \in \mathcal{C}_2(A, A')$, we have

$$\begin{aligned}\|\Theta(\Phi_\mathbb{M}) - \Phi_\mathbb{N}\|_\diamond &\geq \|\Delta_{A'_1} \circ (\Theta(\Phi_\mathbb{M}) - \Phi_\mathbb{N}) \circ \Delta_{A'_0}\|_\diamond \\ &= \|\Delta_{A'_1} \circ \Theta(\Phi_\mathbb{M}) \circ \Delta_{A'_0} - \Phi_\mathbb{N}\|_\diamond,\end{aligned}$$

so that the minimum in $\delta_{A|A'}(\Phi_\mathbb{M}\|\Phi_\mathbb{N})$ is attained at some c-c superchannel p and we have seen that $p(\Phi_\mathbb{M}) = \Phi_{\mathbb{N}'}$ for some set $\mathbb{N}'$ of measurements that are simulated by $\mathbb{M}$. It follows that $\mathbb{N}$ is ϵ-simulable by $\mathbb{M}$.

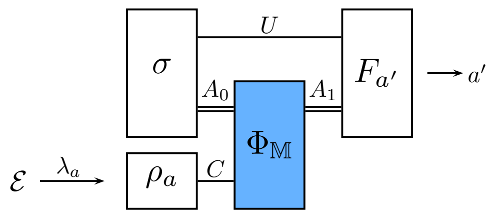
□

We are ready to apply the results of Section III-F and prove that ϵ-simulability can be characterized by guessing games. Note that in this case, it is enough to use ensembles on the system C (so $R_0 = R_1 = 1$ in Theorem 8 (iii)). For $\mathbb{M}$ consisting of a single element and $\epsilon = 0$, this result was proved in [3].

Corollary 5. *Let $\mathbb{M}$ and $\mathbb{N}$ be sets of measurements as above, $\epsilon \geq 0$. Then $\mathbb{N}$ is ϵ-simulable by $\mathbb{M}$ if and only if for any ensemble $\mathcal{E}$ on C,*

$$\max_{1\leq y\leq m} P^Q_{\rm succ}(\mathcal{E}, N^y) \leq \max_{1\leq i\leq k} P^Q_{\rm succ}(\mathcal{E}, M^i) + \frac{\epsilon}{2} P_{\rm succ}(\mathcal{E}).$$

Proof. We start by expressing the success probabilities of part (iii) of Theorem 8 for $R_0 = R_1 = 1$. Let $\mathcal{E} = \{\lambda_a, \rho_a\}$ be an ensemble on C. Since $B_0 = C$ and $B_1 = 1$, the guessing games with $\Phi_{\mathbb{M}}$ can be represented as in the diagram



Here, as a preprocessing, we pick a quantum-classical state $\sigma \in \mathfrak{S}(UA_0)$ seen as a channel in $\mathcal{C}(1, SA_0)$ and we also pick a measurement F on UA_1. The success probability is then by (9)

$$P_{\rm succ}(\sigma \otimes \mathcal{E}, \Phi^*_{\mathbb{M}}(F)) = \langle \rho_{\mathcal{E}}, \Theta(\Phi_{\mathbb{M}})\rangle,$$

with $\Theta := \sigma * \Phi_F \in \mathcal{C}_2(A, Q)$, where $Q = Q_0 Q_1$, $Q_0 = 1$, Q_1 is the output system of Φ_F. Since Θ is obviously a c-c superchannel, by Lemma 5 there are conditional probabilities $p_i(a|j)$ and probabilities $q(i)$ such that $\Theta = p$ with $p(a, i|j) = q(i)p_i(a|j)$. We obtain

$$\begin{aligned}\langle \rho_{\mathcal{E}}, \Theta(\Phi_{\mathbb{M}})\rangle &= \rho_{\mathcal{E}} * C_p * C_{\mathbb{M}} \\ &= \sum_i q(i) \sum_{j,a} \lambda_a p_i(a|j) {\rm Tr}\,[\rho_a M^i_j] \\ &= \sum_i q(i) P_{\rm succ}(\mathcal{E}, M^i, p_i).\end{aligned}$$

Since any c-c superchannel in $\mathcal{C}_2(A, Q)$ consists of a preprocessing and postprocessing of the above form, we see that

$$\begin{aligned}P^{\mathcal{C}(1,\cdot A_0),\Phi_{\mathbb{M}},\mathcal{M}}_{\rm succ}(\mathcal{E}) &= \sup_{\sigma,F} P_{\rm succ}(\sigma\otimes\mathcal{E}, \Phi^*_{\mathbb{M}}(F)) \\ &= \sup_{p\in\mathcal{C}_2(A,Q)} \langle \rho_{\mathcal{E}}, p(\Phi_{\mathbb{M}})\rangle \\ &= \sup_{q,\{p_i\}} \sum_i q(i) P_{\rm succ}(\mathcal{E}, M^i, p_i) \\ &= \sup_q \sum_i q(i) P^Q_{\rm succ}(\mathcal{E}, M^i) \\ &= \max_{1\leq i\leq m} P^Q_{\rm succ}(\mathcal{E}, M^i),\end{aligned}$$

where the second supremum is taken over all c-c superchannels. Since we have a similar equality for $\Phi_{\mathbb{N}}$, we obtain the 'only if' part.

For the converse, let $\rho \in \mathfrak{S}(A'_0 C A_1)$. It is easy to see from the shape of the Choi matrix $C_{\mathbb{N}}$ that there are probabilities $\lambda(y)$, conditional probabilities $\mu(x|y)$ and states $\rho^y_x \in \mathfrak{S}(C)$ such that

$$\begin{aligned}\langle \rho, \Phi_{\mathbb{N}}\rangle = \rho * C_{\mathbb{N}} &= \sum_y \lambda(y) \sum_x \mu(x|y) {\rm Tr}\,[\rho^y_x N^y_x] \\ &= \sum_y \lambda(y) P_{\rm succ}(\mathcal{E}_y, N^y),\end{aligned}$$

here $\mathcal{E}_y = \{\mu(x|y), \rho^y_x\}^n_{x=1}$. For each y, $P_{\rm succ}(\mathcal{E}_y, N^y) \leq P^Q_{\rm succ}(\mathcal{E}_y, N^y)$, so that by the assumption, there is some $1 \leq i_y \leq k$ and conditional probabilities $p_y(x|j)$ such that

$$P_{\rm succ}(\mathcal{E}_y, N^y) \leq P_{\rm succ}(\mathcal{E}_y, M^{i_y}, p_y) + \frac{\epsilon}{2} P_{\rm succ}(\mathcal{E}_y).$$

Put $q(i|y) = \delta_{i,i_y}$, then $q(i|y)$ are conditional probabilities. Put $p(i, x|y, j) = q(i|y)p_y(x|j)$, then p is a c-c superchannel in $\mathcal{C}_2(A, A')$. It can be easily computed that

$$\begin{aligned}\langle \rho, p(\Phi_{\mathbb{M}})\rangle &= \rho * C_p * C_{\mathbb{M}} \\ &= \sum_{i,j,x,y} \lambda(y)\mu(x|y) {\rm Tr}\,[\rho^y_x M^i_j] p(i, x|y, j) \\ &= \sum_{y,x,j} \lambda(y)\mu(x|y) {\rm Tr}\,[\rho^y_x M^{i_y}_j] p_y(x|j) \\ &= \sum_y \lambda(y) P_{\rm succ}(\mathcal{E}_y, M^{i_y}, p_y)\end{aligned}$$

It follows that

$$\langle \rho, \Phi_{\mathbb{N}}\rangle \leq \langle \rho, p(\Phi_{\mathbb{M}})\rangle + \frac{\epsilon}{2} \sum_y \lambda(y) P_{\rm succ}(\mathcal{E}_y).$$

Let now $F^y \in \mathcal{M}_n(C)$ be such that $P_{\rm succ}(\mathcal{E}_y) = P_{\rm succ}(\mathcal{E}_y, F^y)$ and let $\mathbb{F} = \{F^1, \ldots, F^m\}$. As we have seen before,

$$\begin{aligned}\sum_y \lambda(y) P_{\rm succ}(\mathcal{E}_y) &= \sum_y \lambda(y) P_{\rm succ}(\mathcal{E}_y, F^y) \\ &= \langle \rho, \Phi_{\mathbb{F}}\rangle \leq \|\rho\|^{\diamond}_{A'_1|A'_0 C}.\end{aligned}$$

By Theorem 3 (ii), this finishes the proof. □

IV. Conclusion

We have introduced a general framework for comparison of channels, in quantum information theory as well as in the broader setting of GPT. The framework is based on the category BS, which is a special category of ordered (finite dimensional) vector spaces, modelled on the set of channels. In this setting, we defined a notion of an F-conversion distance δ_{F} with respect to a convex subcategory F and proved a general result giving an operational characterization of this distance. This result was then applied to quantum channels, where we proved that the F-conversion distance can be characterized by a set of modified conditional min-entropies. In the setting of quantum resource theories of processes, these quantities form a complete set of resource monotones. Under some conditions of the subcategory F, the modified conditional min-entropies can be obtained by a conic program.

In the case when elements of F can be characterized as concatenations $\Theta_{\rm pre} * \Theta_{\rm post}$ with $\Theta_{\rm pre} \in \mathcal{C}_{\rm pre}$ and $\Theta_{\rm post} \in \mathcal{C}_{\rm post}$ for some suitable sets of channels, we also characterized the F-conversion distance in terms of success probabilities in some guessing games. We discussed several choices of such subcategories: postprocessings, preprocessings, LOCC and partial superchannels on bipartite channels. We also noted that our results hold for other classes of restricted superchannels, such as PPT or SEP. As another application, we studied the problem of approximate classical simulability for sets of measurements.

The advantage of the general formulation in the GPT setting of Section II is that our results can be applied not only to pairs of channels but also to more specialized networks. In particular, one can study the problem of converting m copies of a channel to n copies of another, using parallel or sequential schemes (see also [17]). If the m copies of a channel, or more generally an m-tuple of channels are used in parallel, we can treat their tensor product simply as a channel and some variants of the present theory may be applied. If we use a sequential scheme with respect to some fixed ordering, the tensor product is a special case of an m-comb [20]. Similarly to 2-combs, spaces of m-combs are also objects of the category BS, so convertibility in this setting can be treated using Theorem 2. On the other hand, if there is no fixed ordering in the use of the channels, then we may see the tensor product of the channels as a product element in an object obtained from a symmetric monoidal structure in BS. Note also that the choice of the subcategory F not only restricts the allowed transformations (by the choice of morphisms) but also determines the distance measure in δ_{F} (by the choice of the objects).

This work concerns only the one shot situation, when the channels in question are used only once. To go beyond the one shot setting in the general framework, we need to discuss possible symmetric monoidal structures (tensor products) in BS, their properties and the corresponding behaviour of the related norms. Another important direction is an extension to infinite dimensions. The present framework strongly depends on the finite dimensional setting, but some corresponding results for post- and preprocessings for quantum channels on semifinite von Neumann algebras were proved in [23].

Acknowledgements

I am indebted to Martin Plávala for the idea of introducing the category that is here called BS and sharing his notes, as well as, together with Gejza Jenča, for discussions on its definition and properties. I am also grateful to anonymous referees whose comments helped me greatly to improve the presentation and results of the paper. Part of the results (Section III-F) were presented at the Algebraic and Statistical ways into Quantum Resource Theories workshop at BIRS in July 2019. The research was supported by grants VEGA 2/0142/20 and APVV-16-0073.

References

[1] K. Matsumoto. (2010) A quantum version of randomization criterion. arxiv:1012.2650. [Online]. Available: https://arxiv.org/abs/1012.2650

[2] F. Buscemi, "Comparison of quantum statistical models: Equivalent conditions for sufficiency," *Communications in Mathematical Physics*, vol. 310, no. 3, p. 625647, Jan 2012. [Online]. Available: http://dx.doi.org/10.1007/s00220-012-1421-3

[3] P. Skrzypczyk and N. Linden, "Robustness of measurement, discrimination games, and accessible information," *Phys. Rev. Lett.*, vol. 122, p. 140403, Apr 2019. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.122.140403

[4] L. Guerini, J. Bavaresco, M. Terra Cunha, and A. Acín, "Operational framework for quantum measurement simulability," *Journal of Mathematical Physics*, vol. 58, no. 9, p. 092102, Sep 2017. [Online]. Available: http://dx.doi.org/10.1063/1.4994303

[5] A. Chefles. (2009) The quantum Blackwell theorem and minimum error state discrimination. arxiv:0907.0866. [Online]. Available: https://arxiv.org/abs/0907.0866

[6] F. Buscemi, "Degradable channels, less noisy channels, and quantum statistical morphisms: An equivalence relation," *Problems of Information Transmission*, vol. 52, no. 3, p. 201213, Jul 2016. [Online]. Available: http://dx.doi.org/10.1134/S0032946016030017

[7] ——, "Comparison of noisy channels and reverse data-processing theorems," in *IEEE Information Theory Workshop (ITW)*, 2017, pp. 489–493.

[8] G. Gour, D. Jennings, F. Buscemi, R. Duan, and I. Marvian, "Quantum majorization and a complete set of entropic conditions for quantum thermodynamics," *Nature Communications*, vol. 9, p. 5352, 2018.

[9] G. Gour, "Comparison of quantum channels by superchannels," *IEEE Transactions on Information Theory*, vol. 65, no. 9, pp. 5880–5904, 2019.

[10] A. Jenčová, "Comparison of quantum channels and quantum statistical experiments," in *2016 IEEE International Symposium on Information Theory (ISIT)*, ser. IEEE Conference Publications, 2016, pp. 2249 – 2253.

[11] ——. (2014) Randomization theorems for quantum channels. arxiv:1404.3900. [Online]. Available: https://arxiv.org/abs/1404.3900

[12] R. Takagi and B. Regula, "General resource theories in quantum mechanics and beyond: operational characterization via discrimination tasks," *Phys. Rev. X*, vol. 9, p. 031053, 2019.

[13] D. Blackwell, "Comparison of experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, Calif.: University of California Press, 1951, pp. 93–102. [Online]. Available: https://projecteuclid.org/euclid.bsmsp/1200500222

[14] L. Le Cam, "Sufficiency and approximate sufficiency," *Ann. Math. Statist.*, vol. 35, no. 4, pp. 1419–1455, 12 1964. [Online]. Available: https://doi.org/10.1214/aoms/1177700372

[15] H. Strasser, *Mathematical Theory of Statistics*. Berlin, New York: de Gruyter, 1985.

[16] Z.-W. Liu and A. Winter. (2019) Resource theories of quantum channels and the universal role of resource erasure. arxiv:1904.04201. [Online]. Available: https://arxiv.org/abs/1904.04201

[17] Y. Liu and X. Yuan, "Operational resource theory of quantum channels," *Physical Review Research*, vol. 2, no. 1, p. 012035, 2020.

[18] G. Gour and C. M. Scandolo. (2019) The entanglement of a bipartite channel. arxiv:1907.02552. [Online]. Available: https://arxiv.org/abs/1907.02552

[19] ——, "Dynamical entanglement," *Physical Review Letters*, vol. 125, no. 18, p. 180505, 2020.

[20] G. Chiribella, G. M. DAriano, and P. Perinotti, "Theoretical framework for quantum networks," *Physical Review A*, vol. 80, no. 2, Aug 2009. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.80.022339

[21] A. Kitaev, "Quantum computations: Algorithms and error correction," *Russian Mathematical Surveys*, vol. 52, pp. 1191–1249, 1997.

[22] F. Buscemi, D. Sutter, and M. Tomamichel, "An information-theoretic treatment of quantum dichotomies," *Quantum*, vol. 3, p. 209, Dec 2019. [Online]. Available: http://dx.doi.org/10.22331/q-2019-12-09-209

[23] P. Ganesan, L. Gao, S. K. Pandey, and S. Plosker, "Quantum majorization on semi-finite von neumann algebras," *Journal of Functional Analysis*, vol. 279, no. 7, p. 108650, 2020.

[24] G. Chiribella and D. Ebler, "Optimal quantum networks and one-shot entropies," *New Journal of Physics*, vol. 18, no. 9, p. 093053, 2016.

[25] G. Chiribella, G. M. D'Ariano, and P. Perinotti, "Informational derivation of Quantum Theory," *Phys. Rev. A*, vol. 84, p. 012311, 2011.

[26] A. Bisio and P. Perinotti, "Theoretical framework for higher-order quantum theory," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 475, no. 2225, p. 20180706, May 2019. [Online]. Available: http://dx.doi.org/10.1098/rspa.2018.0706

[27] L. Lami, "Non-classical correlations in quantum mechanics and beyond," Ph.D. dissertation, Universitat Autonoma de Barcelona, 2018. [Online]. Available: https://arxiv.org/abs/1803.02902

[28] A. Jenčová and M. Plávala, "Conditions on the existence of maximally incompatible two-outcome measurements in general probabilistic theory," *Physical Review A*, vol. 96, no. 2, Aug 2017. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.96.022113

[29] A. Jenčová, "Base norms and discrimination of generalized quantum channels," *Journal of Mathematical Physics*, vol. 55, no. 2, p. 022201, 2014.

[30] M. Plávala, "Private communication," 2019.

[31] J. Watrous, *The Theory of Quantum Information*. Cambridge University Press, 2018.

[32] R. Renner, "Security of quantum key distribution," *International Journal of Quantum Information*, vol. 6, no. 01, pp. 1–127, 2008.

[33] R. Konig, R. Renner, and C. Schaffner, "The operational meaning of min- and max-entropy," *IEEE Transactions on Information Theory*, vol. 55, no. 9, p. 43374347, Sep 2009. [Online]. Available: http://dx.doi.org/10.1109/TIT.2009.2025545

[34] G. Chiribella, G. M. DAriano, and P. Perinotti, "Memory effects in quantum channel discrimination," *Physical Review Letters*, vol. 101, no. 18, Oct 2008. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.101.180501

[35] G. Gutoski, "On a measure of distance for quantum strategies," *Journal of Mathematical Physics*, vol. 53, no. 3, p. 032202, Mar 2012. [Online]. Available: http://dx.doi.org/10.1063/1.3693621

[36] B. Coecke, T. Fritz, and R. W. Spekkens, "A mathematical theory of resources," *Information and Computation*, vol. 250, pp. 59–86, 2016.

[37] A. Jenčová. (2015) Comparison of quantum channels and statistical experiments. arXiv:1512.07016. [Online]. Available: https://arxiv.org/abs/1512.07016

[38] F. Buscemi, M. Keyl, G. M. DAriano, P. Perinotti, and R. F. Werner, "Clean positive operator valued measures," *Journal of Mathematical Physics*, vol. 46, no. 8, p. 082109, Aug 2005. [Online]. Available: http://dx.doi.org/10.1063/1.2008996