Conditions for the compatibility of channels in general probabilistic theory and their connection to steering and Bell nonlocality

Martin Plávala[*]

Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, Bratislava, Slovakia

We derive general conditions for the compatibility of channels in general probabilistic theory. We introduce formalism that allows us to easily formulate steering by channels and Bell nonlocality of channels as generalizations of the well-known concepts of steering by measurements and Bell nonlocality of measurements. The generalization does not follow the standard line of thinking stemming from the Einstein-Podolsky-Rosen paradox, but introduces steering and Bell nonlocality as entanglement-assisted incompatibility tests. We show that all of the proposed definitions are, in the special case of measurements, the same as the standard definitions, but not all of the known results for measurements generalize to channels. For example, we show that for quantum channels, steering is not a necessary condition for Bell nonlocality. We further investigate the introduced conditions and concepts in the special case of quantum theory and we provide many examples to demonstrate these concepts and their implications.

I. INTRODUCTION

Incompatibility of measurements is the well-known quantum phenomenon that gives rise to steering and Bell nonlocality. Historically, the idea of measurement incompatibility dates back to Bohr's principle of complementarity. Steering was first described by Schrödinger [1] and Bell nonlocality was first introduced by Bell [2], both as a reply to the paradox of Einstein, Podolsky, and Rosen (EPR) [3]. It is known that incompatibility of measurements is necessary and in some cases sufficient for both steering and Bell nonlocality, but the operational connection between incompatibility, steering, and Bell nonlocality was so far not described in general terms that would also fit channels, not only measurements.

There was extensive research into properties of quantum incompatibility of measurements [4,5], quantum incompatibility of measurements and its noise robustness, or degree of compatibility [6,7], connection of quantum incompatibility of measurements and steering [8–12], connection of quantum incompatibility of measurements and Bell nonlocality [13–16], and connection between steering and Bell nonlocality [17,18] (for a recent review, see [19]). In recent years, the problems of incompatibility of measurements on channels [20], compatibility of channels [21], the connection of channel steering to measurement incompatibility [22], and incompatibility in general probabilistic theory [23–26] were all studied.

The aim of this paper is to heavily generalize the recent results of [26], where compatibility, steering, and Bell nonlocality of measurements were formulated using convex analysis and the geometry of tensor products. In this paper, we will generalize the ideas and results of [26] for the case of two channels in general probabilistic theory. The generalizations are not straightforward and we will have to introduce several operational ideas and definitions, e.g., we introduce the operational interpretation of direct products of state spaces and we define steering and Bell nonlocality as very simple entanglement-assisted incompatibility test, that

boil down to the problem whether there exists a multipartite state with given marginal states.

During all of our calculations we will restrict ourselves to finite-dimensional general probabilistic theory and to only the case of two channels. We will restrict to only two channels just for simplicity, as one may easily formulate many of our results for more than two channels using the same operational ideas as we will present.

The paper is organized as follows: In Sec. II we describe our motivation for using general probabilistic theory. We provide several references to known applications and their connections to each other. In Sec. III we introduce general probabilistic theory. Note that in Sec. III D we introduce the operational interpretations of direct products in general probabilistic theory. In Sec. IV we define compatibility of channels and we derive a condition for compatibility of channels. In Sec. V we show that our condition for compatibility of channels yields the condition for compatibility of measurements that was presented in Ref. [26]. In Sec. VI we derive specific conditions for the compatibility of quantum channels. In Sec. VII we propose an idea for a test of incompatibility of channels, that will not work at first, but will eventually lead to both steering and Bell nonlocality. In Sec. VIII we define steering by channels as one-sided entanglement assisted incompatibility test and we derive some basic results. In Sec. IX we show that for the special case of measurements our definition of steering leads to the standard definition of steering [27] in the formalism of [26]. In Sec. X we derive the specific conditions for steering by quantum channels, we show that every pair of incompatible channels may be used for steering of maximally entangled state, and that there are entangled states that are not steerable by any pair of channels, among other results. In Sec. XI we define Bell nonlocality of channels as a two-sided entanglement assisted incompatibility test and we derive some basic results. Then, in Sec. XII we show that, when applied to measurement, the general definition of Bell nonlocality yields the standard definition of Bell nonlocality [27] in the formalism of [26], and we also show that for measurements steering is a necessary condition for Bell nonlocality. In Sec. XIII we derive conditions for the Bell nonlocality of quantum channels, we

---

[*]martin.plavala@mat.savba.sk

formulate a generalized version of the Clauser-Horne-Shimony-Holt (CHSH) inequality, we show that for such inequality Tsirelson bound [28] both holds and is reached, we show an example of violation of the generalized version of CHSH inequality, and we build on the example from Sec. X of an entangled state not steerable by any pair of channels to show that, even though the state is not steerable by any pair of channels, it leads to Bell nonlocality, which shows that steering is not a necessary condition for Bell nonlocality for quantum channels. In Sec. XIV we conclude the paper by presenting the many open questions and possible areas of research opened by our paper.

II. MOTIVATIONS FOR USING GENERAL PROBABILISTIC THEORY

There are few motivations to using general probabilistic theory. The first motivation is mathematical as general probabilistic theory is a unified framework capable of describing both classical and quantum theory, as well as other theories. In this paper, the mathematical motivation is (according to the personal opinion of the author) even stronger as some of the formulations of the presented ideas and some of the proofs of the presented theorems turn out to be clearer in the framework of general probabilistic theory. The second motivation comes from foundations of quantum theory as general probabilistic theory provides insight into the structure of entanglement and incompatibility. The third and most promising motivation comes from information theory. There were developed several models [29–31] that have very interesting information-theoretic properties and that can be described by general probabilistic theory, albeit sometimes it needs to be extended even more [32]. Apart from the well-known results on the properties of Popescu-Rohrlich boxes [33,34], it was showed that there are theories in which one can search an N-item database in $O(\sqrt[3]{N})$ queries [35] and that there is a general probabilistic theory that can be simulated by a probabilistic classical computer that can perform Deutsch-Jozsa and Simon's algorithm [36].

The aforementioned results show that studying general probabilistic theory is interesting even from a practical viewpoint and that it could have potential applications in information processing.

III. INTRODUCTION TO GENERAL PROBABILISTIC THEORY

General probabilistic theory is a unified framework to describe the kinematics of different systems in a mathematically unified fashion. The idea of general probabilistic theory is an operational approach to setting the axioms and then carrying forward using convex analysis. Useful books on convex analysis are [37,38]. The beautiful aspect of general probabilistic theory is that it is only little bit more general than dealing with the different systems on their own, but we do not have to basically rewrite the same calculations over and over again for different theories.

During our calculations, we will use two recurring examples: one will be finite-dimensional classical theory and the other will be finite-dimensional quantum theory. The finite-dimensional classical theory is closely tied to the known results about incompatibility, steering, and Bell nonlocality of measurements and we will mainly use it to verify that the definitions we will propose are, in the special case of measurements, the same as the known definitions. The quantum theory is our main concern as this is the theory we are mostly interested in. Some results, that we will only prove for quantum theory, may be generalized for general probabilistic theory, but we will limit the generality of our calculations to make them more understandable to readers that are not so far familiar with general probabilistic theory.

Given that we will work with many different spaces, their duals, their tensor products, and many isomorphic sets, all isomorphisms will be omitted unless explicitly stated otherwise.

A. State space and the effect algebra of general probabilistic theory

There are two central notions in general probabilistic theory: the state space that describes all possible states of the system and the effect algebra that describes the measurements on the system. We will begin our construction from the state space and then define the effect algebra, but we will show how one can go the other way and start from an effect algebra and obtain state space afterwards. We will restrict ourselves to finite-dimensional spaces and always use the Euclidean topology.

Let V denote a real, finite-dimensional vector space and let $X \subset V$, then by $\text{conv}(X)$ we will denote the convex hull of X, by $\text{aff}(X)$ we will denote the affine hull of X. We will proceed with the definition of relative interior of a set $X \subset V$.

Definition 1. Let $X \subset V$, then the relative interior of X, denoted $\text{ri}(X)$ is the interior of X when it is considered as a subset of $\text{aff}(X)$.

For a more thorough discussion of relative interior, see [37, p. 44].

Let K be a compact convex subset of V, then K is a state space. The points $x \in K$ represent the states of some system and their convex combination is interpreted operationally, that is, for $x, y \in K$, $\lambda \in [0,1] \subset \mathbb{R}$ the state $\lambda x + (1 - \lambda)y$ corresponds to having prepared x with probability λ and y with probability $1 - \lambda$.

To define measurements, we have to be able to assign probabilities to states, that is, we have to have a map $f : K \to [0,1]$ such that, to follow the operational interpretation of convex combination, we have assigned the convex combination of probabilities to the convex combination of respective states. In other words, for $x, y \in K$, $\lambda \in [0,1]$ we have

$$f(\lambda x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda)f(y),$$

which means that f is an affine function. Such functions are called effects because they correspond to assigning probabilities of measurement outcomes to states. We will proceed with a more formal definition of effects and of effect algebra.

Let $A(K)$ denote the set of affine functions $K \to \mathbb{R}$. $A(K)$ is itself a real linear space, moreover it is ordered as follows: let $f, g \in A(K)$, then $f \geqslant g$ if $f(x) \geqslant g(x)$ for every $x \in K$. There are two special functions 0 and 1 in $A(K)$, such that $0(x) = 0$ and $1(x) = 1$ for all $x \in K$.

The set $A(K)^+ = \{f \in A(K) : f \geqslant 0\}$ is the convex, closed cone of positive functions. The cone $A(K)^+$ is

generating, that is, for every $f \in A(K)$ we have $f_+, f_- \in A(K)^+$ such that $f = f_+ - f_-$, and it is pointed, that is, if $f \geqslant 0$ and $-f \geqslant 0$, then $f = 0$.

Although we will provide a proper definition of measurement in Sec. III E, we will now introduce the concept of yes or no measurement, or two-outcome measurement, that will motivate the definition of the effect algebra. Our notion of measurement might seem different to the standard understanding and one may argue that what we will refer to as measurements should be called entanglement-breaking maps, but this way of defining measurement is standard in general probabilistic theory, hence, we will use it. A measurement is a procedure that assigns probabilities to possible outcomes based on the state that is measured. If we have only two outcomes and we know that the probability of the first outcome is $p \in [0,1]$, then, by normalization, the probability of the second outcome must be $1 - p$. This shows that a two-outcome measurement needs to assign only probability to one outcome and the other probability follows.

Since assigning probabilities to states is a function $f : K \to [0,1]$ and due to our operational interpretation of convex combination we want such function to be affine. Traditionally, the functions that assign probabilities to states are called effects and the set of all effects is called effect algebra.

Definition 2. The set $E(K) = \{ f \in A(K) : 0 \leqslant f \leqslant 1 \}$ is called the effect algebra.

In general, one may define effect algebra in more general fashion, using the partially defined operation of addition and a unary operation $\perp$, that would in our case correspond to $f^\perp = 1 - f$ (see [39] for a more thorough treatment).

Let $f \in E(K)$, then the two-outcome measurement m_f corresponding to the effect f is the procedure that for $x \in K$ assigns the probability $f(x)$ to the first outcome and the probability $1 - f(x)$ to the second outcome. Note that we did not mention any labels of the outcomes. Usually, the outcomes are labeled yes and no, or 0 and 1, or -1 and 1, but from an operational perspective this does not matter.

We provide two standard examples of special cases of our definitions.

Example 1 (Classical theory). In classical theory, the state space K is a simplex, that is the convex hull of a set of affinely independent points $x_1, \ldots, x_n$. The special property of the simplex is that every point $x \in K$ can be uniquely expressed as convex combination of the points $x_1, \ldots, x_n$, due to their affine independence.

Example 2 (Quantum theory). Let $\mathcal{H}$ denote a finite-dimensional complex Hilbert space, let $B_h(\mathcal{H})$ denote the real linear space of self-adjoint operators on $\mathcal{H}$, for $A \in B_h(\mathcal{H})$ let $\mathrm{Tr}(A)$ denote the trace of the operator A, and let $A \geqslant 0$ denote that A is positive semidefinite. We say that $A \leqslant B$ if $0 \leqslant B - A$. Let $B_h(\mathcal{H})^+ = \{ A \in B_h(\mathcal{H}) : A \geqslant 0 \}$ denote the cone of positive-semidefinite operators.

In quantum theory the state space is given as

$$\mathfrak{D}_{\mathcal{H}} = \{ \rho \in B_h(\mathcal{H}) : \rho \geqslant 0, \mathrm{Tr}(\rho) = 1 \}$$

which is the set of density operators on $\mathcal{H}$. The effect algebra $E(\mathfrak{D}_{\mathcal{H}})$ is given as

$$E(\mathfrak{D}_{\mathcal{H}}) = \{ M \in B_h(\mathcal{H}) : 0 \leqslant M \leqslant \mathbb{1} \}.$$

The value of the effect $M \in E(\mathfrak{D}_{\mathcal{H}})$ on the state $\rho \in \mathfrak{D}_{\mathcal{H}}$ is given as

$$M(\rho) = \mathrm{Tr}(\rho M).$$

B. Structure of general probabilistic theory

This section will be rather technical, but we will introduce several mathematical results that we will use later on.

Let $x \in K$ and consider the map $\overline{x} : A(K) \to \mathbb{R}$, that to $f \in A(K)$ assigns the value $f(x)$. This is clearly a linear functional on $A(K)$. Moreover for $x, y \in K, \lambda \in [0,1]$ we have

$$\overline{\lambda x + (1 - \lambda)y} = \lambda \overline{x} + (1 - \lambda)\overline{y}$$

as the functions in $A(K)$ are affine by definition. We conclude that the state space K must be affinely isomorphic to some subset of the dual of $A(K)$. Since the aforementioned isomorphism is going to be extremely useful in later calculations, we will describe it in more detail. Let $A(K)^*$ denote the dual of $A(K)$, that is, the space of all linear functionals on $A(K)$. For $\psi \in A(K)^*$ and $f \in A(K)$ we will denote the value the functional ψ reaches on f as $\langle \psi, f \rangle$. The dual cone to $A(K)^+$ is the cone $A(K)^{*+} = \{ \psi \in A(K)^* : \langle \psi, f \rangle \geqslant 0, \forall f \in A(K)^+ \}$ that gives rise to the ordering on $A(K)^*$ given as follows: let $\psi, \varphi \in A(K)^*$, then $\psi \geqslant \varphi$ if and only if $(\psi - \varphi) \in A(K)^{*+}$, i.e., if $\psi - \varphi \geqslant 0$.

It is straightforward that the state space K is isomorphic to a subset of the cone $A(K)^{*+}$, moreover, it is straightforward to realize that the functionals isomorphic to K must map the function $1 \in A(K)$ to the value 1.

Definition 3. Let $\mathfrak{S}_K = \{ \psi \in A(K)^{*+} : \langle \psi, 1 \rangle = 1 \}$. We call $\mathfrak{S}_K$ the state space of the effect algebra $E(K)$.

It might be confusing at this point why we call $\mathfrak{S}_K$ a state space, but this will be cleared by the following.

Proposition 1. $\mathfrak{S}_K$ is affinely isomorphic to K.

Proof. It is clear that the map $x \to \overline{x}$ maps K to a convex subset of $\mathfrak{S}_K$. It is easy to show the inclusion of $\mathfrak{S}_K$ in the image of K using Hahn-Banach separation theorem (see [40, Chap. 1, Theorem 4.3] for a proof). ∎

We will omit the isomorphism between K and $\mathfrak{S}_K$, so for any $x, y \in K, \alpha \in \mathbb{R}$ we will write $\alpha x + y$ instead of $\alpha \overline{x} + \overline{y} \in A(K)^*$. Still, one must be careful when omitting this isomorphism because if $0 \in V$ denotes the zero vector and $0 \in K$, then $\overline{0} \in A(K)^*$ is not the zero functional as by construction we have $\langle \overline{0}, 1 \rangle = 1$. We will do our best to avoid such possible problems by choosing appropriate notation.

There are two more results we will heavily rely on:

Proposition 2. $\mathfrak{S}_K$ is a base of $A(K)^{*+}$, that is, for every $\psi \in A(K)^{*+}, \psi \neq 0$, there is a unique $x \in K$ and $\lambda \in \mathbb{R}, \lambda \geqslant 0$, such that $\psi = \lambda x$.

Proof. Let $\psi \in A(K)^{*+}, \psi \neq 0$, then $\langle \psi, 1 \rangle \neq 0$ as if $\langle \psi, 1 \rangle = 0$ and $\psi \geqslant 0$, then $\psi = 0$, because $1 \in ri(A(K)^+)$. Let $\psi' = \frac{1}{\langle \psi, 1 \rangle} \psi$. It is straightforward that $\psi' \in \mathfrak{S}_K$. ∎

Proposition 3. $A(K)^{*+}$ is a generating cone in $A(K)^*$, that is, for every $\psi \in A(K)$ there are $\varphi_+, \varphi_- \in A(K)^{*+}$ such that $\psi = \varphi_+ - \varphi_-$.

Proof. The result follows from the fact that $A(K)^+$ is a pointed cone (see [38, Sec. 2.6.1]). ∎

C. Tensor products of state spaces and effect algebras

Tensor products are a way to describe joint systems of several other systems. There are several approaches to introducing a tensor product in general probabilistic theory. There is a category theory based approach [41] that is a viable way to introduce the tensor products, but we will use a simpler, operational approach. Note that the state space of the joint system will be a compact convex subset of a real, finite-dimensional vector space as it itself must be a state space of some general probabilistic theory. Also, keep in mind that describing a tensor product of state spaces K_A, K_B is equivalent to describing the tensor product of the cones $A(K_A)^{*+}$, $A(K_B)^{*+}$. This is going to be useful as some things are easier to express in terms of the positive cones.

Let V, W be real finite-dimensional vector spaces and let $v \in V$, $w \in W$. $v \otimes w$ will refer to the element of the algebraic tensor product $V \otimes W$ (see, e.g., [42]). We will first describe the minimal and maximal tensor products of state spaces that set bounds on the real state space of the joint system. Note that when describing the joint state space of two state spaces or states of two systems, we will refer to them as bipartite state space or bipartite states.

Let K_A, K_B denote two state spaces of Alice and Bob, respectively. For every $x_A \in K_A$, $x_B \in K_B$ there must be a state of the joint system describing the situation that Alice's system is in the state x_A and Bob's system is in the state x_B. We will denote such state $x_A \otimes x_B$ and we will call it a product state. Since the state space must be convex, the state space of the joint system must contain at least the convex hull of the product states. This leads to the definition of minimal tensor product.

Definition 4. The minimal tensor product of state spaces K_A and K_B, denoted $K_A \dot{\otimes} K_B$, is the compact convex set

$$K_A \dot{\otimes} K_B = \text{conv}(\{x_A \otimes x_B : x_A \in K_A, x_B \in K_B\}).$$

The bipartite states $y \in K_A \dot{\otimes} K_B$ are also called separable states. For the positive cones we get

$$A(K_A \dot{\otimes} K_B)^{*+}$$
$$= \text{conv}(\{\psi_A \otimes \psi_B : \psi_A \in A(K_A)^{*+}, \psi_B \in A(K_B)^{*+}\}).$$

Example 3. In quantum theory, the minimal tensor product $\mathfrak{D}_\mathcal{H} \dot{\otimes} \mathfrak{D}_\mathcal{H}$ is the set of all separable states, that is, of all states of the form $\sum_{i=1}^n \lambda_i \rho_i \otimes \sigma_i$ for $n \in \mathbb{N}$ and $\rho_i \in \mathfrak{D}_\mathcal{H}$, $\sigma_i \in \mathfrak{D}_\mathcal{H}$, $0 \leqslant \lambda_i$ for $i \in \{1, \ldots, n\}$, $\sum_{i=1}^n \lambda_i = 1$.

In a similar fashion, let $f_A \in E(K_A)$, $f_B \in E(K_B)$, then we can define a function $f_A \otimes f_B$ as the unique affine function such that for $x_A \in K_A$, $x_B \in K_B$ we have

$$(f_A \otimes f_B)(x_A \otimes x_B) = f_A(x_A) f_B(x_B).$$

This function is used in the most simple measurement on the joint system, such that Alice applies the two-outcome measurement m_{f_A} and Bob applies the two-outcome measurement m_{f_B}, so $f_A \otimes f_B$ must be an effect on the joint state space. This leads to the definition of the maximal tensor product.

Definition 5. The maximal tensor product of the state spaces K_A and K_B, denoted $K_A \hat{\otimes} K_B$, is defined as

$$K_A \hat{\otimes} K_B = \{\psi \in A(K_A)^* \otimes A(K_B)^* : \forall \, f_A \in A(K_A)^+,$$
$$\forall \, f_B \in A(K_B)^+, \langle \psi, f_A \otimes f_B \rangle \geqslant 0\}.$$

States in $K_A \hat{\otimes} K_B \setminus K_A \dot{\otimes} K_B$ are called entangled states. Equivalent definition, in terms of the positive cones, would be

$$A(K_A \hat{\otimes} K_B)^{*+} = (A(K_A)^+ \dot{\otimes} A(K_B)^+)^{*+},$$

where

$$A(K_A)^+ \dot{\otimes} A(K_B)^+$$
$$= \text{conv}(\{f_A \otimes f_B : f_A \in A(K_A)^+, f_B \in A(K_B)^+\}).$$

As we see, the definition of tensor product of cones of positive functionals goes hand in hand with the definition of tensor product of cones of positive functions.

Example 4. In quantum theory, the maximal tensor product of the cones $B_h(\mathcal{H})^+ \hat{\otimes} B_h(\mathcal{H})^+$ is the cone of entanglement witnesses [43, Sec. 6.3.1], i.e., $W \in B_h(\mathcal{H})^+ \hat{\otimes} B_h(\mathcal{H})^+$ if for every $\rho \in \mathfrak{D}_\mathcal{H}$, $\sigma \in \mathfrak{D}_\mathcal{H}$ we have $\text{Tr}(W\rho \otimes \sigma) \geqslant 0$. Note that this does not imply the positivity of W.

From the constructions it is clear that the state space of the joint system has to be a subset of the maximal tensor product and it has to contain the minimal tensor product. But, there is no other specification of the state space of the joint system in general; it has to be provided by the theory we are working with.

Definition 6. We will call the joint state space of the systems described by the state spaces K_A and K_B the real tensor product of K_A and K_B and we will denote it $K_A \tilde{\otimes} K_B$. We always have

$$K_A \dot{\otimes} K_B \subseteq K_A \tilde{\otimes} K_B \subseteq K_A \hat{\otimes} K_B.$$

Example 5. In quantum theory, the real tensor product of the state spaces is defined as the set of density matrices on the tensor product of the Hilbert spaces, that is,

$$\mathfrak{D}_\mathcal{H} \tilde{\otimes} \mathfrak{D}_\mathcal{H} = \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}.$$

It is tricky to work with the tensor products in general probabilistic theory as the real tensor product is not always specified, or it may not be clear what it should be. We will always assume that every tensor product we need to be defined is defined. Moreover, when working with a tensor product of more than two state spaces, say K_A, K_B, K_C, we will always assume that

$$(K_A \tilde{\otimes} K_B) \tilde{\otimes} K_C = K_A \tilde{\otimes} (K_B \tilde{\otimes} K_C)$$

and we will simply write $K_A \tilde{\otimes} K_B \tilde{\otimes} K_C$. In the applications of general probabilistic theory to quantum and classical theory it will always be clear how to construct the needed tensor products and we consider this sufficient for us since we are mainly interested in the applications of our results.

We will state and prove a result about classical state spaces that we will use several times later on.

Proposition 4. Let S be a simplex with the extremal points $x_1, \ldots, x_n$, i.e., $S = \text{conv}(\{x_1, \ldots, x_n\})$ and let K be any state space, then we have

$$S \dot{\otimes} K = S \hat{\otimes} K.$$

Proof. Let S be a simplex and let $x_i \in A(S)^{*+}$, $i \in \{1, \ldots, n\}$, be the extreme points of S. The points $x_1, \ldots, x_n$ form a basis of $A(S)^*$. Let $\psi \in S \hat{\otimes} K$, then we have

$$\psi = \sum_{i=1}^n x_i \otimes \varphi_i$$

for some $\varphi_i \in A(K)^*$. Our aim is to prove that $\varphi_i \in A(K)^{*+}$, then $\psi \in S \dot{\otimes} K$ follows by definition.

Let $b_1, \ldots, b_n$ denote the basis of $A(S)$ dual to the basis $x_1, \ldots, x_n$ of $A(S)^*$, i.e., we have $b_i(x_j) = \delta_{ij}$, where $i, j \in \{1, \ldots, n\}$ and δ_{ij} is the Kronecker delta. We have $b_i \in E(S)$ because S is a simplex. For any $f \in E(K)$ we have

$$0 \leqslant (\psi, b_i \otimes f) = (\varphi_i, f)$$

for all $i \in \{1, \ldots, n\}$, which implies $\varphi_i \in A(K)^{*+}$. ∎

Note that tensor product of the simplexes S_1, S_2 is also a simplex, so we have

$$K \hat{\otimes} S_1 \hat{\otimes} S_2 = K \dot{\otimes} S_1 \dot{\otimes} S_2.$$

D. Direct product of state spaces and effect algebras

For certain reasons, we will need to use direct products together with tensor products. The idea of why they will be used is going to be clear in the end, but now we will present several of their properties that will be required later. As in Sec. III C, we will work mostly with the cones of the positive functionals.

Let K_A, K_B be two state spaces. Given $A(K_A)^{*+}$ and $A(K_B)^{*+}$, there are two ways to define the direct product of these cones. The first is to use the cone $A(K_A)^{*+} \times A(K_B)^{*+}$. The second is to realize that we can construct $K_A \times K_B$ that will be a compact and convex set, i.e., a state space that gives rise to the cone $A(K_{B_1} \times K_{B_2})^{*+}$.

It may seem that these cones are fairly similar, but they are not and they have different physical interpretations. Let $\psi \in A(K_A \times K_B)^{*+}$, then there are unique $\lambda \in \mathbb{R}, x_A \in K_A, x_B \in K_B$ such that $\psi = \lambda(x_A, x_B)$. Now, let $\varphi \in A(K_A)^{*+} \times A(K_B)^{*+}$, then there are $y_A \in K_A, y_B \in K_B, \alpha_A, \alpha_B \in \mathbb{R}, \alpha_A, \alpha_B \geqslant 0$ such that $\varphi = (\alpha_A y_A, \alpha_B y_B)$. In other words, the normalization may be different in every component of the product. This can be rewritten as

$$\varphi = (\alpha_A y_A, \alpha_B y_B)$$
$$= (\alpha_A + \alpha_B)\left(\frac{\alpha_A}{\alpha_A + \alpha_B} y_A, \frac{\alpha_B}{\alpha_A + \alpha_B} y_B\right)$$
$$= (\alpha_A + \alpha_B)\left(\frac{\alpha_A}{\alpha_A + \alpha_B}(y_A, 0) + \frac{\alpha_B}{\alpha_A + \alpha_B}(0, y_B)\right)$$

that shows that every element of $A(K_A)^{*+} \times A(K_B)^{*+}$ can be uniquely expressed as a multiple of a convex combination of elements of the form $(y_A, 0)$ and $(0, y_B)$. The operational interpretation of such states is that we do not even know which system we are working with, but we know that with some probability p we have the first system and with probability $1 - p$ we have the second system.

The operational interpretation of $A(K_A \times K_B)^{*+}$ is a bit harder to grasp. We may understand $\psi \in A(K_A \times K_B)^{*+}$ as a (multiple of) conditional state. That is, we will interpret the object (x_A, x_B) as a state that corresponds to making a choice in the past between the systems K_A and K_B and keeping track of both of the outcomes at once. The cone $A(K_A \times K_B)^{*+}$ will play a central role in our results on incompatibility, steering, and Bell nonlocality because in the problem of incompatibility, we wish to implement two channels at the same time, and in

steering and Bell nonlocality we are choosing between two incompatible channels.

At last, we will need to describe the set $A(K_A \times K_B)$ and its structure with respect to the sets $A(K_A)$ and $A(K_B)$. We will show that $A(K_A \times K_B)$ corresponds to a certain subset of $A(K_A) \times A(K_B)$ by using the following two ideas: since all of the vector spaces are finite dimensional, we have that $A(K_A) \times A(K_B)$ is the dual to $A(K_A)^* \times A(K_B)^*$ and $A(K_A \times K_B)^*$ can be identified with a subset of $A(K_A)^* \times A(K_B)^*$. Note that this identification holds only between the vector spaces and not between the corresponding state spaces.

Proposition 5. We have

$$A\left(K_{B_1} \times K_{B_2}\right)^{*+} \subset A\left(K_{B_1}\right)^{*+} \times A\left(K_{B_2}\right)^{*+}.$$

Proof. The idea of the proof is that if we have $\varphi \in A(K_A)^{*+} \times A(K_B)^{*+}$ such that $\varphi = (\alpha_A y_A, \alpha_B y_B)$, then $\varphi \in A(K_A \times K_B)^{*+}$ if and only if $\alpha_A = \alpha_B$. Therefore, we can identify $A(K_A \times K_B)^{*+}$ with the set $\{\psi \in A(K_A)^{*+} \times A(K_B)^{*+} : \langle \psi, (1, -1) \rangle = 0\}$. It is easy to verify this constraint on the positive cones and since it is linear it must hold everywhere else. ∎

The above proof shows that the function $(1, -1) \in A(K_A) \times A(K_B)$ is equal to zero when restricted to $A(K_A \times K_B)^*$ or, in other words, $(1, 0) = (0, 1)$ when restricted to $A(K_A \times K_B)^*$. We introduce a relation of equivalence on $A(K_A) \times A(K_B)$ as follows: for $f, g \in A(K_A) \times A(K_B)$ we say that f and g are equivalent and we write $f \sim g$ if $f - g = k(1, -1)$ for some $k \in \mathbb{R}$. Equivalently, $f \sim g$ if for every $\psi \in A(K_A \times K_B)^*$ we have $\langle \psi, f \rangle = \langle \psi, g \rangle$. $A(K_A \times K_B)$ corresponds to the set of equivalence classes of $A(K_A) \times A(K_B)$ with respect to the relation of equivalence $\sim$.

To demonstrate this, consider the constant function $1 \in E(K_A \times K_B)$ and let $x \in K_A, y \in K_B$, then we have

$$\langle (x, y), (1, 0) \rangle = \langle x, 1 \rangle = 1 = \langle (x, y), 1 \rangle,$$
$$\langle (x, y), (0, 1) \rangle = \langle y, 1 \rangle = 1 = \langle (x, y), 1 \rangle.$$

This is not a coincidence because $(1, 0) - (0, 1) = (1, -1)$, so we have $(1, 0) \sim (0, 1)$.

E. Channels and measurements in general probabilistic theory

It is not easy to define channels in general probabilistic theory as we would like all of the channels to be completely positive. We will use the following definition:

Definition 7. Let K_A, K_B be state spaces, then channel Φ is a linear map

$$\Phi : A(K_A)^* \to A(K_B)^*$$

that is positive, i.e., for every $\psi \in A(K_A)^{*+}$ we have $\Phi(\psi) \in A(K_B)^{*+}$ and that for $\psi \in K_A$ we have $\Phi(\psi) \in K_B$.

One may also require a channel to be completely positive, that is, if K_C is some state space such that we can define $K_C \tilde{\otimes} K_A$, then we can consider the map $\text{id} \otimes \Phi : K_C \tilde{\otimes} K_A \to K_C \hat{\otimes} K_B$ and require it to be positive. In the applications of general probabilistic theory to classical and quantum theories, we always know how to create joint systems of given two systems, so in the examples we will always require complete positivity of channels, but one still has to bear in mind that

in the general case, complete positivity is not a well-defined concept.

One can identify the channel $\Phi : A(K_A)^* \to A(K_B)^*$ with an element of $A(K_A) \otimes A(K_B)^*$ as follows: let $x \in K_A$ and $f \in A(K_B)$, then the expression $\langle \Phi(x), f \rangle$ gives rise to a linear functional on $A(K_A)^* \otimes A(K_B)$. This means that we have $\Phi \in A(K_A) \otimes A(K_B)^*$, where we omit the isomorphism between the channel and the functional. If we also consider the positivity of the channel on the elements of the form $x \otimes f \in K_A \dot{\otimes} E(K_B)$, we get

$$\Phi \in A(K_A)^+ \hat{\otimes} A(K_B)^{*+}.$$

This is a well-known construction that may be also used to define the tensor product of linear spaces [42, Chap. 1.3].

There is one more construction with channels that will be important in our formulation of compatibility of channels: compositions with effect. Let $\Phi : K_A \to K_B$ be a channel and let $f \in E(K_B)$, then they give rise to an effect $(f \circ \Phi) \in E(K_A)$ defined for $x_A \in K_A$ as

$$\langle x_A, (f \circ \Phi) \rangle = \langle \Phi(x_A), f \rangle.$$

By the same idea, we can define a map $f \otimes \mathrm{id} : A(K_B)^* \otimes A(K_C)^* \to A(K_C)^*$ such that for $x_B \in K_B$ and $x_C \in K_C$ we have $(f \otimes \mathrm{id})(x_B \otimes x_C) = f(x_B) x_C$ and we extend the map by linearity. Also, given a channel $\Phi : K_A \to K_B \tilde{\otimes} K_C$ we can compose the map $f \otimes \mathrm{id}$ with the channel Φ to obtain $(f \otimes \mathrm{id}) \circ \Phi' : A(K_A)^* \to A(K_C)^*$ such that the corresponding functional on $A(K_A) \otimes A(K_C)^*$ is for $x_A \in K_A$ and $g \in A(K_C)$ given as

$$\langle (f \otimes \mathrm{id}) \circ \Phi, x_A \otimes g \rangle = \langle \Phi(x_A), f \otimes g \rangle.$$

Specifically, we will be interested in the expressions $(1 \otimes \mathrm{id}) \circ \Phi$ and $(\mathrm{id} \otimes 1) \circ \Phi$. If Φ is a channel, then $(1 \otimes \mathrm{id}) \circ \Phi$ and $(\mathrm{id} \otimes 1) \circ \Phi$ are channels as well and they are called marginal channels of Φ.

A special type of channel is a measurement.

Definition 8. A channel $m : K_A \to K_B$ is called a measurement if K_B is a simplex.

The interpretation is simple: the vertices of the simplex correspond to the possible measurement outcomes and the resulting state is a probability distribution over the measurement outcomes, i.e., an assignment of probabilities to the possible outcomes. Since we require all state spaces to be finite dimensional, this implies that we consider only finite-outcome measurements. Let K_B be a simplex with vertices $\omega_1, \ldots, \omega_n$, then we can identify a measurement m with an element of $A(K_A)^+ \dot{\otimes} A(K_B)^{*+}$ of the form

$$m = \sum_{i=1}^{n} f_i \otimes \delta_{\omega_i},$$

where for $i \in \{1, \ldots, n\}$ we have $f_i \in E(K_A)$, $\sum_{i=1}^{n} f_i = 1$ and $\delta_{\omega_i} \in \mathfrak{S}(K_B)$ are the functionals corresponding to the extreme points of K_B (where we have not omitted the isomorphism this time). This expression has an operational interpretation that for $x \in K_A$ the measurement m assigns the probability $f_i(x)$ to the outcome ω_i.

Example 6. Quantum channels are completely positive, trace-preserving maps $\Phi : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$. The complete positiv-

ity means that for any $\rho \geqslant 0$ we have $(\mathrm{id} \otimes \Phi)(\rho) \geqslant 0$. We denote the set of channels $\Phi : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ as $\mathfrak{C}_{\mathcal{H} \to \mathcal{H}}$.

Let $|1\rangle, \ldots, |n\rangle$ be an orthonormal base of $\mathcal{H}$. To every quantum channel we may assign its unique Choi matrix $C(\Phi)$ defined as

$$C(\Phi) = (\Phi \otimes \mathrm{id}) \left(\sum_{i,j=1}^{n} |ii\rangle\langle jj| \right),$$

where we use the shorthand $|ii\rangle = |i\rangle \otimes |i\rangle$. Note that $C(\Phi) \geqslant 0$ and $\mathrm{Tr}_1(C(\Phi)) = \mathbb{1}$, where Tr_1 denotes the partial trace. Also, every matrix $C \in B_h(\mathcal{H} \otimes \mathcal{H})$ such that $C \geqslant 0$ and $\mathrm{Tr}_1(C) = \mathbb{1}$ is a Choi matrix of some channel (see [43, Sec. 4.4.3]).

The Choi matrix $C(\Phi)$ is isomorphic to a state $\frac{1}{\dim(\mathcal{H})} C(\Phi)$, which corresponds to the channel $\Phi \otimes \mathrm{id}$ acting on the maximally entangled state $|\psi^+\rangle\langle\psi^+|$, where

$$|\psi^+\rangle = \frac{1}{\sqrt{\dim(\mathcal{H})}} \sum_{i=1}^{n} |ii\rangle.$$

IV. COMPATIBILITY OF CHANNELS

Definition 9. Let K_A, K_{B_1}, K_{B_2} be state spaces and let Φ_1, Φ_2 be channels

$$\Phi_1 : K_A \to K_{B_1},$$
$$\Phi_2 : K_A \to K_{B_2}.$$

We say that Φ_1, Φ_2 are compatible if and only if there exists a channel

$$\Phi : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$$

such that Φ_1 and Φ_2 are the marginal channels of Φ, i.e., we have

$$\Phi_1 = (\mathrm{id} \otimes 1) \circ \Phi, \quad (1)$$
$$\Phi_2 = (1 \otimes \mathrm{id}) \circ \Phi. \quad (2)$$

The channel Φ is also called the joint channel of the channels Φ_1, Φ_2.

The operational meaning of compatibility of channels is that if the channels Φ_1, Φ_2 are compatible, then we can apply them both to the input state at once and select which one we actually want the output from latter. If the channels are incompatible, we have to choose from which one we want the output before applying anything. For a more in-depth explanation, see [19]. The important thing is that there is a choice from which channel we want to get the output so we can expect to see $A(K_{B_1} \times K_{B_2})^{*+}$ come up in the calculations.

Consider the channel $\Phi : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$. One can realize that the maps $(\mathrm{id} \otimes 1) : \Phi \mapsto (\mathrm{id} \otimes 1) \circ \Phi$ and $(1 \otimes \mathrm{id}) : \Phi \mapsto (1 \otimes \mathrm{id}) \circ \Phi$ are linear maps of channels. Moreover, Eqs. (1) and (2) both have Φ on the right-hand side in the same position. We are going to exploit this to obtain a simpler condition for compatibility of the channels Φ_1, Φ_2. To do so, we have to introduce a new map J.

Let us define a map $J : A(K_A) \otimes A(K_{B_1})^* \otimes A(K_{B_2})^* \to A(K_A) \otimes A(K_{B_1} \times K_{B_2})^*$ given for $\Xi \in A(K_A) \otimes A(K_{B_1})^* \otimes A(K_{B_2})^*$ as

$$J(\Xi) = ((\mathrm{id} \otimes 1) \circ \Xi, (1 \otimes \mathrm{id}) \circ \Xi).$$

For $\Xi = f \otimes \psi \otimes \varphi$, we have

$$J(\Xi) = f \otimes (\langle \varphi, 1 \rangle \psi, \langle \psi, 1 \rangle \varphi).$$

Proposition 6. J is a linear mapping.

Proof. Let $\Xi_1, \Xi_2 \in A(K_A) \otimes A(K_{B_1} \otimes K_{B_2})^*$ and $\lambda \in \mathbb{R}$, then we have

$$\begin{aligned}
J(\lambda \Xi_1 + \Xi_2) &= (\lambda (\mathrm{id} \otimes 1) \circ \Xi_1 + (\mathrm{id} \otimes 1) \circ \Xi_2, 0) \\
&\quad + (0, \lambda(1 \otimes \mathrm{id}) \circ \Xi_1 + (1 \otimes \mathrm{id}) \circ \Xi_2) \\
&= \lambda((\mathrm{id} \otimes 1) \circ \Xi_1, (1 \otimes \mathrm{id}) \circ \Xi_1) \\
&\quad + ((\mathrm{id} \otimes 1) \circ \Xi_2, (1 \otimes \mathrm{id}) \circ \Xi_2) \\
&= \lambda J(\Xi_1) + J(\Xi_2).
\end{aligned}$$

$\blacksquare$

Assume that the channels Φ_1, Φ_2 are compatible and that Φ is their joint channel, then we must have

$$J(\Phi) = (\Phi_1, \Phi_2)$$

which is just a more compact form of Eqs. (1) and (2).

Proposition 7. The channels Φ_1, Φ_2 are compatible if and only if there is $\Phi \in A(K_A)^+ \hat{\otimes} A(K_{B_1} \tilde{\otimes} K_{B_2})^{*+}$ such that

$$J(\Phi) = (\Phi_1, \Phi_2). \tag{3}$$

Proof. If the channels Φ_1, Φ_2 are compatible, then Eq. (3) must hold for their joint channel Φ. If Eq. (3) holds for some $\Phi \in A(K_A)^+ \hat{\otimes} A(K_{B_1} \tilde{\otimes} K_{B_2})^{*+}$, then the channels Φ_1, Φ_2 are compatible and Φ is their joint channel. $\blacksquare$

The operational interpretation is that (Φ_1, Φ_2) represents a conditional channel in the same way as the states from $A(K_{B_1} \times K_{B_2})^{*+}$ represent conditional states that keep track of some choice made in the past. If the channels are compatible, then we actually do not have to make the choice of either using Φ_1 or Φ_2, but we can use their joint channel, that has the property that its marginals reproduce the outcomes of the two channels Φ_1, Φ_2. We will investigate several of the properties of the map J.

Proposition 8. For every $(\xi_1, \xi_2) \in A(K_A) \otimes A(K_{B_1} \times K_{B_2})^*$ there is a $\Xi \in A(K_A) \otimes A(K_{B_1})^* \otimes A(K_{B_2})^*$ such that

$$J(\Xi) = (\xi_1, \xi_2).$$

Moreover, if we have

$$(1,1) \circ (\xi_1, \xi_2) = 1,$$

then

$$(1 \otimes 1) \circ \Xi = 1.$$

Proof. Let $f_1, \ldots, f_n$ be a basis of $A(K_A)$, then we have

$$\xi_1 = \sum_{i=1}^{n} f_i \otimes \psi_i, \quad \xi_2 = \sum_{i=1}^{n} f_i \otimes \varphi_i$$

for some $\psi_i \in A(K_{B_1})^*$ and $\varphi_i \in A(K_{B_2})^*$. Since we must have

$$(1,0) \circ (\xi_1, \xi_2) = (0,1) \circ (\xi_1, \xi_2),$$

we obtain

$$\sum_{i=1}^{n} \langle \psi_i, 1 \rangle f_i = \sum_{i=1}^{n} \langle \varphi_i, 1 \rangle f_i,$$

which implies

$$\langle \psi_i, 1 \rangle = \langle \varphi_i, 1 \rangle = k_i$$

for all $i \in \{1, \ldots, n\}$ as $f_1, \ldots, f_n$ is linearly independent. Let

$$\Xi = \sum_{i=1}^{n} k_i^{-1} f_i \otimes \psi_i \otimes \varphi_i,$$

then we have

$$\begin{aligned}
J(\Xi) &= \sum_{i=1}^{n} k_i^{-1} f_i \otimes \left(\langle \varphi_i, 1_{B_2} \rangle \psi_i, \langle \psi_i, 1_{B_1} \rangle \varphi_i \right) \\
&= \sum_{i=1}^{n} f_i \otimes (\psi_i, \varphi_i).
\end{aligned}$$

If we have $1 \circ (\xi_1, \xi_2) = 1$, then

$$\sum_{i=1}^{n} k_i f_i = 1$$

and we get

$$\begin{aligned}
(1 \otimes 1) \circ \Xi &= (1 \otimes 1) \circ \left(\sum_{i=1}^{n} k_i^{-1} f_i \otimes \psi_i \otimes \varphi_i \right) \\
&= \sum_{i=1}^{n} k_i^{-1} \langle \psi_i, 1 \rangle \langle \varphi_i, 1 \rangle f_i = 1.
\end{aligned}$$

$\blacksquare$

Proposition 9. We have

$$\begin{aligned}
J\left(A(K_A)^+ &\dot{\otimes} A\left(K_{B_1}\right)^{*+} \dot{\otimes} A\left(K_{B_2}\right)^{*+} \right) \\
&= A(K_A)^+ \dot{\otimes} A\left(K_1 \times K_{B_2}\right)^{*+}.
\end{aligned}$$

Proof. Let $(\xi_1, \xi_2) \in A(K_A)^+ \dot{\otimes} A(K_{B_1} \times K_{B_2})^{*+}$, then as in the proof of Proposition 8 we have

$$\xi_1 = \sum_{i=1}^{n} f_i \otimes \psi_i, \quad \xi_2 = \sum_{i=1}^{n} f_i \otimes \varphi_i,$$

but now we have $f_i \geqslant 0$, $\psi_i \geqslant 0$, and $\varphi_i \geqslant 0$ for $i \in \{1, \ldots, n\}$. It follows by the same construction as in the proof of Proposition 8 that we can construct $\Xi = \sum_{i=1}^{n} k_i^{-1} f_i \otimes \psi_i \otimes \varphi_i$ and we get $\Xi \in A(K_A)^+ \dot{\otimes} A(K_{B_1})^{*+} \dot{\otimes} A(K_{B_2})^{*+}$.

Let $\Xi \in A(K_A)^+ \dot{\otimes} A(K_{B_1})^{*+} \dot{\otimes} A(K_{B_2})^{*+}$, then we have $\Xi = \sum_{i=1}^{n} f_i \otimes \psi_i \otimes \varphi_i$ such that $f_i \geqslant 0$, $\psi_i \geqslant 0$, $\varphi_i \geqslant 0$ for all $i \in \{1, \ldots, n\}$, moreover, without lack of generality we can assume $\langle \psi_i, 1_{B_1} \rangle = \langle \varphi_i, 1_{B_2} \rangle = 1$. We have

$$J(\Xi) = \sum_{i=1}^{n} f_i \otimes (\psi_i, \varphi_i) \in A(K_A)^+ \dot{\otimes} A\left(K_{B_1} \times K_{B_2}\right)^{*+},$$

which concludes the proof. $\blacksquare$

It would be very useful to know what is the image of the cone $A(K_A)^+ \hat{\otimes} A(K_{B_1} \tilde{\otimes} K_{B_2})^{*+}$ when mapped by J. We will denote the resulting cone $Q = J(A(K_A)^+ \hat{\otimes} A(K_{B_1} \tilde{\otimes} K_{B_2})^{*+})$. The cone is important due to the following:

Corollary 1. The channels Φ_1, Φ_2 are compatible if and only if

$$(\Phi_1, \Phi_2) \in Q = J\left(A(K_A)^+ \hat{\otimes} A\left(K_{B_1} \tilde{\otimes} K_{B_2}\right)^{*+} \right).$$

Proof. Follows from Proposition 7. ∎

Proposition 10. $A(K_A)^+ \dot{\otimes} A(K_{B_1} \times K_{B_2})^{*+} \subset Q$.

Proof. Since

$$A(K_A)^+ \dot{\otimes} A\left(K_{B_1} \dot{\otimes} K_{B_2}\right)^{*+} \subset A(K_A)^+ \hat{\otimes} A\left(K_{B_1} \tilde{\otimes} K_{B_2}\right)^{*+},$$

we must have

$$J\left(A(K_A)^+ \dot{\otimes} A\left(K_{B_1} \dot{\otimes} K_{B_2}\right)^{*+}\right) \subset Q.$$

The result follows from Proposition 9. ∎

Proposition 11. $Q \subset A(K_A)^+ \hat{\otimes} A(K_{B_1} \times K_{B_2})^{*+}$.

Proof. Since we have

$$A(K_A)^+ \hat{\otimes} A\left(K_{B_1} \tilde{\otimes} K_{B_2}\right)^{*+} \subset A(K_A)^+ \hat{\otimes} A\left(K_{B_1} \hat{\otimes} K_{B_2}\right)^{*+},$$

we must have

$$Q \subset J\left(A(K_A)^+ \hat{\otimes} A\left(K_{B_1} \hat{\otimes} K_{B_2}\right)^{*+}\right).$$

Let $\Xi \in A(K_A)^+ \hat{\otimes} A(K_{B_1} \hat{\otimes} K_{B_2})^{*+}$, then for $\psi \in A(K_A)^{*+}$ and $(f_1, f_2) \in A(K_{B_1} \times K_{B_2})^+$ we get

$$\langle J(\Xi), x \otimes f \rangle = \langle ((\mathrm{id} \otimes 1) \circ \Xi, (1 \otimes \mathrm{id}) \circ \Xi, x \otimes f \rangle$$
$$= \langle \Xi(x), f_1 \otimes 1 \rangle + \langle \Xi(x), 1 \otimes f_2 \rangle \geqslant 0,$$

that shows we have $J(A(K_A)^+ \hat{\otimes} A(K_{B_1} \hat{\otimes} K_{B_2})^{*+}) \subset A(K_A)^+ \hat{\otimes} A(K_{B_1} \times K_{B_2})^{*+}$, which concludes the proof. ∎

We can also construct Q as the cone we get when we factorize the cone $A(K_A)^+ \hat{\otimes} A(K_{B_1} \tilde{\otimes} K_{B_2})^{*+}$ with respect to the relation of equivalence given as follows: $\Xi_1 \approx \Xi_2$ if and only if $J(\Xi_1) = J(\Xi_2)$ or, equivalently, if and only if $\Xi_1 = \Xi_2 + \Xi$, such that $J(\Xi) = 0$.

Note that since J is a linear map, as we showed in Proposition 6, it is clear that Q is a convex cone. For two given channels $\Phi_1 : K_A \to K_{B_1}$, $\Phi_2 : K_A \to K_{B_1}$ one may write a primal linear program that would check the condition for compatibility given by Corollary 1. We will write such linear program for quantum channels later.

V. COMPATIBILITY OF MEASUREMENTS

We will apply the results of Sec. IV to the problem of compatibility of measurements. We will obtain the same results that were recently presented in Ref. [26], that are a generalization of [44].

Let K_A be a state space and let S_1, S_2 be simplexes and let $m_1 : K_A \to S_1$, $m_2 : K_A \to S_2$ be measurements. According to Proposition 7, the measurements m_1, m_2 are compatible if and only if

$$(m_1, m_2) \in J(A(K_A)^+ \hat{\otimes} A(S_1 \tilde{\otimes} S_2)^{*+}).$$

Since both S_1 and S_2 are simplexes, then we have $S_1 \tilde{\otimes} S_2 = S_1 \hat{\otimes} S_2$ and the condition for compatibility reduces according to Proposition 9 to

$$(m_1, m_2) \in A(K_A)^+ \dot{\otimes} A(S_1 \times S_2)^{*+}.$$

Due to the simpler structure of simplexes, one may get even more specific results about measurements (see [26]).

For demonstration of the derived conditions we will reconstruct the result of [44] about compatibility of two-outcome measurements. According to our definition, a measurement is two-outcome if the simplex it has as a target space has two vertexes, i.e., it is a line segment. Let K be a state space,

$f, g \in E(K)$ and $m_f : K \to S$, $m_g : K \to S$ be two-outcome measurements given as

$$m_f = f \otimes \delta_{\omega_1} + (1 - f) \otimes \delta_{\omega_2},$$
$$m_g = g \otimes \delta_{\omega_1} + (1 - g) \otimes \delta_{\omega_2}.$$

The state space given by $A(S \times S)^{*+}$ is a square given as $\mathrm{conv}((\delta_{\omega_1}, \delta_{\omega_1}), (\delta_{\omega_1}, \delta_{\omega_2}), (\delta_{\omega_2}, \delta_{\omega_1}), (\delta_{\omega_2}, \delta_{\omega_2}))$, that is just affinely isomorphic to $S \times S$. We have

$$(m_1, m_2) = f \otimes \left(\delta_{\omega_1}, 0\right) + (1 - f) \otimes \left(\delta_{\omega_2}, 0\right)$$
$$+ g \otimes \left(0, \delta_{\omega_1}\right) + (1 - g) \otimes \left(0, \delta_{\omega_2}\right)$$
$$= f \otimes \left(\delta_{\omega_1}, \delta_{\omega_2}\right) + (1 - f) \otimes \left(\delta_{\omega_2}, \delta_{\omega_2}\right)$$
$$+ g \otimes \left(0, \delta_{\omega_1} - \delta_{\omega_2}\right),$$

where in the second step we have used the basis $(\delta_{\omega_1}, \delta_{\omega_2})$, $(\delta_{\omega_2}, \delta_{\omega_2})$, $(0, \delta_{\omega_1} - \delta_{\omega_2})$ of $A(S \times S)^*$ to express (m_1, m_2) in a more reasonable form. To have $(m_1, m_2) \in A(K)^+ \dot{\otimes} A(S \times S)^{*+}$ we must have

$$(m_1, m_2) = h_{11} \otimes \left(\delta_{\omega_1}, \delta_{\omega_1}\right) + h_{12} \otimes \left(\delta_{\omega_1}, \delta_{\omega_2}\right)$$
$$+ h_{21} \otimes \left(\delta_{\omega_2}, \delta_{\omega_1}\right) + h_{22} \otimes \left(\delta_{\omega_2}, \delta_{\omega_2}\right)$$
$$= (h_{11} + h_{12}) \otimes \left(\delta_{\omega_1}, \delta_{\omega_2}\right)$$
$$+ (h_{21} + h_{22}) \otimes \left(\delta_{\omega_2}, \delta_{\omega_2}\right)$$
$$+ (h_{11} + h_{21}) \otimes \left(0, \delta_{\omega_1} - \delta_{\omega_2}\right)$$

for some $h_{11}, h_{12}, h_{21}, h_{22} \in E(K)$. This implies the standard conditions for the compatibility of two-outcome measurements m_f, m_g:

$$f = h_{11} + h_{12},$$
$$1 - f = h_{21} + h_{22},$$
$$g = h_{11} + h_{21}$$

(see, e.g., [45]).

VI. COMPATIBILITY OF QUANTUM CHANNELS

In this section, we will derive results more specific to the compatibility of quantum channels. Let $\Phi_1 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, $\Phi_2 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$ be quantum channels, then, according to Proposition 7 they are compatible if and only if there is a channel $\Phi : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ such that for all $\rho \in \mathfrak{D}_\mathcal{H}$ we have

$$(\Phi_1(\rho), \Phi_2(\rho)) = (\mathrm{Tr}_2(\Phi(\rho)), \mathrm{Tr}_1(\Phi(\rho))). \tag{4}$$

This is equivalent to the definition of compatibility of quantum channels already stated in Ref. [21]. It is straightforward that Eq. (4) implies that

$$(C(\Phi_1), C(\Phi_2)) = (\mathrm{Tr}_2(C(\Phi)), \mathrm{Tr}_1(C(\Phi))),$$

we will show that they are equivalent. This will help us to get rid of the state ρ in Eq. (4).

Proposition 12. The channels $\Phi_1 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, $\Phi_2 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$ are compatible if and only if there exists a channel $\Phi : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ such that

$$(C(\Phi_1), C(\Phi_2)) = (\mathrm{Tr}_2(C(\Phi)), \mathrm{Tr}_1(C(\Phi))).$$

Proof. Let $\rho \in \mathfrak{D}_\mathcal{H}$, then we have

$$\text{Tr}_2(\Phi(\rho)) = \text{Tr}_{2,E}(C(\Phi)\mathbb{1} \otimes \mathbb{1} \otimes \rho^T)$$

$$= \text{Tr}_E(\text{Tr}_2(C(\Phi))\mathbb{1} \otimes \rho^T)$$

$$= \text{Tr}_E(C(\Phi_1)\mathbb{1} \otimes \rho^T) = \Phi_1(\rho).$$

The same follows for Φ_2. ∎

As we already showed in Sec. IV, the cone $Q = J(A(\mathfrak{D}_\mathcal{H})^+ \hat{\otimes} A(\mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}})^{*+})$ is of interest for the compatibility of channels. In the case of quantum channels we will use Proposition 12 to formulate similar cone in terms of Choi matrices of the channels and we will write a semidefinite program for the compatibility of quantum channels based on this approach.

Denote $P = \{(\text{Tr}_2(C), \text{Tr}_1(C)) : C \in \mathfrak{C}_{\mathcal{H} \to \mathcal{H} \otimes \mathcal{H}}\}$, then according to Proposition 12 the channels $\Phi_1 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, $\Phi_2 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$ are compatible if and only if

$$(C(\Phi_1), C(\Phi_2)) \in P.$$

Note that, by our definition, P is not a cone, but it generates some cone just by adding all of the operators of the form λC, where $C \in P$ and $\lambda \in \mathbb{R}, \lambda \geqslant 0$.

It would be very interesting to obtain more specific results on the structure of P, but the task is not trivial. To make it simpler, we will investigate the structure of the dual cone P^* given as

$$P^* = \{(A, B) \in B_h(\mathcal{H}) \times B_h(\mathcal{H}) :$$

$$\langle C, (A, B) \rangle \geqslant 0, \forall C \in P\}.$$

Notice that $(A, B) \in B_h(\mathcal{H}) \times B_h(\mathcal{H})$ is simply a block-diagonal matrix having blocks A and B. Also, every $C \in P$ is a block-diagonal matrix, let $C = (C_1, C_2)$, then

$$\langle (C_1, C_2), (A, B) \rangle = \text{Tr}(C_1 A) + \text{Tr}(C_2 B).$$

Let $C \in P$, then by definition there exists a channel $\Phi : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ such that

$$C = (\text{Tr}_2(C(\Phi)), \text{Tr}_1(C(\Phi))).$$

Let $(A, B) \in P^*$, then we have

$$\langle C, (A, B) \rangle = \text{Tr}(\text{Tr}_2(C(\Phi))A + \text{Tr}_1(C(\Phi))B)$$

$$= \text{Tr}(C(\Phi)(\tilde{A} + \mathbb{1} \otimes B)) \geqslant 0,$$

where $\tilde{A}$ is the operator such that $\text{Tr}(\text{Tr}_2(C(\Phi))A) = \text{Tr}(C(\Phi)\tilde{A})$. If $A = A_1 \otimes A_2$, then $\tilde{A} = A_1 \otimes \mathbb{1} \otimes A_2$. In general, one can write A as a sum of factorized operators and express $\tilde{A}$ in such way because the map $A \mapsto \tilde{A}$ is linear.

The result is that $\tilde{A} + \mathbb{1} \otimes B$ must correspond to a positive function on quantum channels, hence, we must have $\tilde{A} + \mathbb{1} \otimes B \geqslant 0$ (see [46,47]). We have proved the following:

Proposition 13. The channels $\Phi_1 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, $\Phi_2 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$ are compatible if and only if

$$\text{Tr}(C(\Phi_1)A) + \text{Tr}(C(\Phi_2)B) \geqslant 0$$

for all $A, B \in B_h(\mathcal{H} \otimes \mathcal{H})$ such that

$$\tilde{A} + \mathbb{1} \otimes B \geqslant 0.$$

This allows us to formulate the semidefinite program [38] for the compatibility of quantum channels as follows:

Proposition 14. Given channels $\Phi_1 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, $\Phi_2 : \mathfrak{D}_\mathcal{H} \to \mathfrak{D}_\mathcal{H}$, the semidefinite program for the compatibility of quantum channels is

$$\inf_{A,B} \text{Tr}(C(\Phi_1)A) + \text{Tr}(C(\Phi_2)B), \quad \tilde{A} + \mathbb{1} \otimes B \geqslant 0,$$

where $\tilde{A}$ is given as above.

If the reached infimum is negative, then the channels are incompatible; if the reached infimum is 0, then the channels are compatible.

Proof. The result follows from Proposition 13. One may see that the infimum is at most 0 because one may always chose $A = B = 0$. ∎

VII. PRELUDE TO STEERING AND BELL NONLOCALITY

We will propose a possible test for the compatibility of channels that will not work, but it will motivate our definitions of steering and Bell nonlocality.

Let K_A, K_{B_1}, K_{B_2} be state spaces and let $\Phi_1 : K_A \to K_{B_1}$, $\Phi_2 : K_A \to K_{B_2}$ be channels. The channels Φ_1, Φ_2 are compatible if Eq. (3) is satisfied for some channel $\Phi : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$. This is the same as saying the channels Φ_1, Φ_2 are compatible if for all $x \in K_A$ we have

$$(\Phi_1(x), \Phi_2(x)) = (((\text{id} \otimes 1) \circ \Phi)(x), ((1 \otimes \text{id}) \circ \Phi)(x)). \quad (5)$$

If the channels Φ_1 and Φ_2 are compatible, then for every $x \in K_A$ there must exist a state $y \in K_{B_1} \tilde{\otimes} K_{B_2}$ such that

$$\Phi_1(x) = (\text{id} \otimes 1)(y), \quad (6)$$

$$\Phi_2(x) = (1 \otimes \text{id})(y). \quad (7)$$

Would it be a reasonable test for the compatibility of the channels Φ_1 and Φ_2 if we considered the state $x \in K_A$ fixed and we would test whether, for the fixed state x, there exists $y \in K_{B_1} \tilde{\otimes} K_{B_2}$ such that Eqs. (6) and (7) are satisfied? It would not, because for a fixed $x \in K_A$ one always has $\Phi_1(x) \otimes \Phi_2(x) \in K_{B_1} \tilde{\otimes} K_{B_2}$ that satisfies Eqs. (6) and (7).

Still, throwing away this line of thinking would not be a good choice because going further, one may ask the following: If there would be another system K_C, such that $K_C \tilde{\otimes} K_A$ is defined, then what if we would use the entanglement between the systems K_A and K_C to obtain a better condition for the compatibility of the channels Φ_1, Φ_2 using the very same line of thinking? As we will see, this approach leads to the notions of steering and Bell nonlocality.

VIII. STEERING

Steering is one of the puzzling phenomena we find in quantum theory but not in classical theory. It is usually described as a two-party protocol, that allows one side to alter the state of the other in a way that would not be possible in classical theory by performing a measurement and announcing the outcome. Although originally discovered by Schrödinger [1], steering was formalized in Ref. [27]. Recently, there was introduced a new formalism for steering in Ref. [26].

So far, it was always only considered that during steering one party performs a measurement. Since a measurement is a special case of a channel, one may ask whether it is possible

to define steering by channels. We will use our formalism for compatibility of channels to introduce steering by channels by continuing the line of thoughts presented in Sec. VII. We will have to formulate steering in a little different way than it usually is formulated for measurements, but we will show that for measurements we will obtain the known results.

Let $K_A, K_{B_1}, K_{B_2}, K_C$ be finite-dimensional state spaces, such that $K_C \tilde{\otimes} K_A$ is defined and let

$$\Phi_1 : K_A \to K_{B_1},$$
$$\Phi_2 : K_A \to K_{B_2}$$

be channels. We can construct channels

$$\mathrm{id} \otimes \Phi_1 : A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+} \to A(K_C)^{*+} \hat{\otimes} A(K_{B_1})^{*+},$$

$$\mathrm{id} \otimes \Phi_2 : A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+} \to A(K_C)^{*+} \hat{\otimes} A(K_{B_2})^{*+}.$$

Moreover, we can construct the conditional channel

$$\mathrm{id} \otimes (\Phi_1, \Phi_2) : A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+}$$
$$\to A(K_C)^{*+} \hat{\otimes} A(K_{B_1} \times K_{B_2})^{*+}.$$

These channels play a central role in steering and we will keep this notation throughout this section. First, we will introduce a handy name for the output state of $\mathrm{id} \otimes (\Phi_1, \Phi_2)$.

Definition 10. Let $\psi \in K_C \tilde{\otimes} K_A$ be a bipartite state, then we call $(\mathrm{id} \otimes (\Phi_1, \Phi_2))(\psi)$ a bipartite conditional state.

Steering may be seen as a three-party protocol that tests the compatibility of channels. The parties in question will be named Alice, Bob, and Charlie. Alice and Charlie share a bipartite state $\psi \in K_C \tilde{\otimes} K_A$ and Alice has the channels Φ_1 and Φ_2 at her disposal, that would send her part of the state ψ to Bob. Since Alice can choose between the channels Φ_1 and Φ_2, she will be, in our formalism, applying the conditional channel (Φ_1, Φ_2) and the resulting state will be a bipartite state from $A(K_C)^{*+} \tilde{\otimes} A(K_{B_1} \times K_{B_2})^{*+}$. The structure of the resulting bipartite conditional state $(\mathrm{id} \otimes (\Phi_1, \Phi_2))(\psi)$ will not only depend on the input state ψ, but also on the compatibility of the channels Φ_1 and Φ_2. Let us assume that the channels Φ_1 and Φ_2 are compatible, then there is a channel $\Phi : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$ such that $(\Phi_1, \Phi_2) = J(\Phi)$ and we have

$$(id \otimes (\Phi_1, \Phi_2))(\psi) = (\mathrm{id} \otimes J(\Phi))(\psi)$$
$$= (\mathrm{id} \otimes J')((\mathrm{id} \otimes \Phi)(\psi)),$$

where $J' : A(K_{B_1} \tilde{\otimes} K_{B_2})^* \to A(K_{B_1} \times K_{B_2})^*$, $J'(\psi) = ((id \otimes 1)(\psi), (1 \otimes id)(\psi))$. The calculation shows that if the channels Φ_1, Φ_2 are compatible, then we must have

$$(\mathrm{id} \otimes (\Phi_1, \Phi_2))(\psi) \in (\mathrm{id} \otimes J')(K_C \tilde{\otimes} K_{B_1} \tilde{\otimes} K_{B_2})$$

which does not have to hold in general if the channels are not compatible. This shows that we can define steering of a state by channels as an entanglement-assisted incompatibility test.

Definition 11. The bipartite state $\psi \in A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+}$ is steerable by channels $\Phi_1 : A(K_A)^{*+} \to A(K_{B_1})^{*+}$, $\Phi_2 : A(K_A)^{*+} \to A(K_{B_2})^{*+}$ if

$$(\mathrm{id} \otimes (\Phi_1, \Phi_2))(\psi) \notin (\mathrm{id} \otimes J')(K_C \tilde{\otimes} K_{B_1} \tilde{\otimes} K_{B_2}).$$

Now, we present the standard result about the connection between compatibility of the channels and steering. The result follows from our definition immediately.

Corollary 2. The bipartite state $\psi \in A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+}$ is not steerable by channels $\Phi_1 : A(K_A)^{*+} \to A(K_B)^{*+}$, $\Phi_2 : A(K_A)^{*+} \to A(K_B)^{*+}$ if the channels Φ_1 and Φ_2 are compatible.

Proof. If the channels Φ_1, Φ_2 are compatible, then we have $(\Phi_1, \Phi_2) = J(\Phi)$ for some $\Phi : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$ and for every $\psi \in K_C \tilde{\otimes} K_A$ we have

$$(\mathrm{id} \otimes (\Phi_1, \Phi_2))(\psi) \in (\mathrm{id} \otimes J')\big(K_C \tilde{\otimes} (K_{B_1} \tilde{\otimes} K_{B_2})\big).$$

∎

Proposition 15. The bipartite state $\psi \in A(K_C)^{*+} \tilde{\otimes} A(K_A)^{*+}$ is not steerable by channels $\Phi_1 : A(K_A)^{*+} \to A(K_B)^{*+}$, $\Phi_2 : A(K_A)^{*+} \to A(K_B)^{*+}$ if $\psi \in A(K_C)^{*+} \dot{\otimes} A(K_A)^{*+}$, i.e., if ψ is separable.

Proof. Every separable state is by definition a convex combination of product states, i.e., of states of the form $x_C \otimes x_A$, where $x_A \in K_A$, $x_C \in K_C$. Since the maps $\mathrm{id} \otimes (\Phi_1, \Phi_2)$ and $\mathrm{id} \otimes J'$ are linear, it is sufficient to prove that for every product state $x_C \otimes x_A \in K_C \dot{\otimes} K_A$ we have $(\mathrm{id} \otimes (\Phi_1, \Phi_2))(x_C \otimes x_A) \in (\mathrm{id} \otimes J')(K_C \tilde{\otimes} K_{B_1} \tilde{\otimes} K_{B_2})$. It follows by our construction in Sec. VII that product states are not steerable by any channels as one can always take $x_C \otimes \Phi_1(x_A) \otimes \Phi_2(x_A)$. Remember that during steering, we fix not only the channels, but also the bipartite state, so the presented construction is valid. ∎

IX. STEERING BY MEASUREMENTS

We will show that the definition of steering given by Definition 11 follows the standard definition of steering [27] in the formalism introduced in Ref. [26], when we replace measurements by channels.

Proposition 16. Let S_1, S_2 be simplexes and let $m_1 : K_A \to S_1, m_2 : K_A \to S_2$ be measurements, then, a state $\psi \in K_C \tilde{\otimes} K_A$ is steerable by m_1, m_2 if and only if

$$(\mathrm{id} \otimes (m_1, m_2))(\psi) \notin K_C \dot{\otimes} (S_1 \times S_2).$$

Proof. The result follows from the fact that $K_C \tilde{\otimes} S_1 \tilde{\otimes} S_2 = K_C \dot{\otimes} S_1 \dot{\otimes} S_2$. ∎

To obtain the standard definition of steering, one only needs to note that if $\xi \in K_C \dot{\otimes} (S_1 \times S_2)$, then there are $x_i \in K_C$, $s_i \in S_1 \times S_2$ and $0 \leqslant \lambda_i \leqslant 1$ for $i \in \{1, \ldots, n\}$ such that

$$\xi = \sum_{i=1}^{n} \lambda_i x_i \otimes s_i, \tag{8}$$

where the interpretation of s_i is that it is a conditional probability, conditioned by the choice of the measurement. At this point, it is straightforward to see that Eq. (8) corresponds to [27, Eq. (5)].

X. STEERING BY QUANTUM CHANNELS

Steering plays an important role in quantum theory. It has found so far applications in quantum cryptography [48] as an intermediate step between quantum key distribution and device-independent quantum key distribution.

We will prove several results and present a simple example of steering by quantum channels. Given the standard, operational, interpretation of steering by measurements the example may seem strange, but rather expected.

Let $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ be channels and let $|\psi^+\rangle = [\dim(\mathcal{H})]^{-\frac{1}{2}} \sum_{i=1}^{\dim(\mathcal{H})} |ii\rangle$ be the maximally entangled vector. We will show that the maximally entangled state $|\psi^+\rangle\langle\psi^+|$ is steerable by the channels Φ_1, Φ_2 whenever they are incompatible.

The proof is rather simple as the bipartite conditional state we obtain is $(id \otimes (\Phi_1, \Phi_2))(|\psi^+\rangle\langle\psi^+|)$. If the channels Φ_1, Φ_2 are compatible, then the state $|\psi^+\rangle\langle\psi^+|$ is not steerable by compatible channels. Now, let us assume that there is a state in $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that we have

$$(id \otimes (\Phi_1, \Phi_2))(|\psi^+\rangle\langle\psi^+|) = (id \otimes J')(\rho), \quad (9)$$

i.e., that the state $|\psi^+\rangle\langle\psi^+|$ is not steerable by the channels Φ_1, Φ_2. Equation (9) implies that we must have

$$(id \otimes \Phi_1)(|\psi^+\rangle\langle\psi^+|) = \mathrm{Tr}_3(\rho),$$

that, after taking trace over the second Hilbert space, gives

$$\frac{1}{\dim(\mathcal{H})}\mathbb{1} = \mathrm{Tr}_{23}(\rho). \quad (10)$$

Now, the picture becomes clear: $(id \otimes \Phi_1)(|\psi^+\rangle\langle\psi^+|)$ is isomorphic to the Choi matrix $C(\Phi_1)$ and Eq. (10) implies that the state ρ must be isomorphic to a Choi matrix of some channel Φ. This together with Proposition 12 means that Eq. (9) holds if and only if the channels are compatible. Thus, we have proved the following:

Proposition 17. The maximally entangled state $|\psi^+\rangle\langle\psi^+|$ is steerable by channels $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ if and only if they are incompatible.

We will investigate steering by unitary channels. We will see a phenomenon that is impossible to happen for steering by measurements: it is possible to steer a state when the two channels we are testing for incompatibility are two copies of the same channel. Let U, V be unitary matrices, i.e., $UU^* = VV^* = \mathbb{1}$, where U^* denotes the conjugate transpose matrix to U and let Φ_U, Φ_V be the corresponding unitary channels, that is, for $\rho \in \mathfrak{D}_{\mathcal{H}}$ we have

$$\Phi_U(\rho) = U\rho U^*, \quad \Phi_V(\rho) = V\rho V^*.$$

Note that we have $\Phi_{\mathbb{1}} = id$, i.e., the unitary channel given by an identity matrix is the identity channel.

Proposition 18. The bipartite state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ is steerable by the unitary channels Φ_U, Φ_V if and only if it is steerable by two copies of the identity channel id.

Proof. The state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ is steerable by the channels Φ_U, Φ_V if and only if there is a state $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that

$$\mathrm{Tr}_3(\sigma) = (id \otimes \Phi_U)(\rho), \quad \mathrm{Tr}_2(\sigma) = (id \otimes \Phi_V)(\rho).$$

If such state σ exists, then for $\tilde{\sigma} = (id \otimes \Phi_{U^*} \otimes \Phi_{V^*})(\sigma)$ we have

$$\mathrm{Tr}_3(\tilde{\sigma}) = \rho, \quad \mathrm{Tr}_2(\tilde{\sigma}) = \rho,$$

i.e., the state ρ is not steerable by two copies of id. The same holds the other way around by almost the same construction; if the state ρ is not steerable by two copies of id, then it is not steerable by any unitary channels Φ_U, Φ_V. ∎

Note that a similar result would hold if only one of the channels would be unitary, but then only that one unitary channel would be replaced by the identity map id. Clearly, if

the state ρ would be separable, then it would not be steerable by any channel. The converse does not hold, even if the state ρ is entangled it still may not be steerable by any channels. We will provide a useful condition for the steerability of a given state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ that will help us to show that even if the state ρ is entangled, it does not have to be steerable by any pair of channels $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$.

Proposition 19. The state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ is steerable by the channels $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ only if it is steerable by two copies of the identity channel id $: \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$.

Proof. Assume that the state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ is not steerable by two copies of the identity channel id $: \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, then there exists a state $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that

$$\mathrm{Tr}_3(\sigma) = \rho, \quad \mathrm{Tr}_2(\sigma) = \rho.$$

Let $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ be any two channels and denote

$$\tilde{\sigma} = (id \otimes \Phi_1 \otimes \Phi_2)(\sigma),$$

then we have

$$\mathrm{Tr}_3(\tilde{\sigma}) = (id \otimes \Phi_1)(\rho), \quad \mathrm{Tr}_2(\tilde{\sigma}) = (id \otimes \Phi_2)(\rho),$$

so the state ρ is not steerable by the channels Φ_1, Φ_2. ∎

Note that one may get other conditions for steering by replacing only one of the channels by the identity map id. One may generalize this result to the general probabilistic theory, but it may be rather restrictive and not as general as one would wish. One may also use the idea of the proof of Proposition 19 together with the result of Proposition 17 to obtain the results on compatibility of channels that are concatenations of other channels, similar to the results obtained in Ref. [21].

We will present an example of an entangled state that is not steerable by any pair of channels.

Example 7. Let $\dim(\mathcal{H}) = 2$ with the standard basis $|0\rangle, |1\rangle$ and let $|W\rangle \in \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$ be given as

$$|W\rangle = \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle).$$

The projector $|W\rangle\langle W| \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ is known as W state. We have

$$\rho_W = \mathrm{Tr}_2(|W\rangle\langle W|) = \mathrm{Tr}_3(|W\rangle\langle W|) \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}},$$

that shows that the state ρ_W is not steerable by a pair of the identity channels id $: \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, which as a result of Proposition 19 means that it is not steerable by any channels $\Phi_1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$. Moreover it is known that the state ρ_W is entangled [43, Example 6.70].

Since it will be useful in later calculations, we will show that the state $|W\rangle\langle W|$ is the only state from $\mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that $\rho_W = \mathrm{Tr}_2(|W\rangle\langle W|) = \mathrm{Tr}_3(|W\rangle\langle W|)$. Let $|\varphi\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$, then we have

$$\rho_W = \tfrac{1}{3}|00\rangle\langle00| + \tfrac{2}{3}|\varphi\rangle\langle\varphi|.$$

Let $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ denote the state such that $\rho_W = \mathrm{Tr}_2(\sigma) = \mathrm{Tr}_3(\sigma)$. We have $\rho_W|11\rangle = 0$ that implies $\mathrm{Tr}(\sigma|11\rangle\langle11| \otimes \mathbb{1}) = \mathrm{Tr}(\sigma|1\rangle\langle1| \otimes \mathbb{1} \otimes |1\rangle\langle1|) = 0$ that implies $\langle111|\sigma|111\rangle = \langle110|\sigma|110\rangle = \langle101|\sigma|101\rangle = 0$ as $\sigma \geqslant 0$. We will show that this implies $\sigma|111\rangle = \sigma|110\rangle = \sigma|101\rangle = 0$.

Let $A \in B_h(\mathcal{H})$, $A \geqslant 0$, and let $|\psi\rangle \in \mathcal{H}$. Let $\|\psi\| = \sqrt{\langle\psi|\psi\rangle}$ denote the norm given by inner product. Assume that we have $\langle\psi|A|\psi\rangle = 0$, then

$$\|\sqrt{A}\psi\|^2 = \langle\sqrt{A}\psi|\sqrt{A}\psi\rangle = \langle\psi|A|\psi\rangle = 0$$

and in conclusion we have $\sqrt{A}|\psi\rangle = 0$ and

$$A|\psi\rangle = \sqrt{A}(\sqrt{A}|\psi\rangle) = 0.$$

Finally, let us denote $|\varphi^\perp\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$. We have $\rho_W|\varphi^\perp\rangle = 0$ that implies $\mathrm{Tr}(\sigma|\varphi^\perp\rangle\langle\varphi^\perp| \otimes \mathbb{1}) = 0$ which yields $\sigma|\varphi^\perp 0\rangle = \sigma|\varphi^\perp 1\rangle = 0$. We still use the shorthand $|\varphi^\perp 0\rangle = |\varphi^\perp\rangle \otimes |0\rangle$.

The eight vectors $|000\rangle$, $|001\rangle$, $|\varphi 0\rangle$, $|\varphi 1\rangle$, $|\varphi^\perp 0\rangle$, $|\varphi^\perp 1\rangle$, $|110\rangle$, $|111\rangle$ form an orthonormal basis of $\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}$. We have already showed that we must have

$$\sigma|\varphi^\perp 0\rangle = \sigma|\varphi^\perp 1\rangle = \sigma|110\rangle = \sigma|111\rangle = 0,$$

so in general we must have

$$\begin{aligned}
\sigma &= a_{00}|000\rangle\langle000| + a_{01}|001\rangle\langle001| + a_{\varphi0}|\varphi0\rangle\langle\varphi0| \\
&\quad + a_{\varphi1}|\varphi1\rangle\langle\varphi1| + b_1|000\rangle\langle001| + \bar{b}_1|001\rangle\langle000| \\
&\quad + b_2|000\rangle\langle\varphi0| + \bar{b}_2|\varphi0\rangle\langle000| + b_3|000\rangle\langle\varphi1| \\
&\quad + \bar{b}_3|\varphi1\rangle\langle000| + b_4|001\rangle\langle\varphi0| + \bar{b}_4|\varphi0\rangle\langle001| \\
&\quad + b_5|001\rangle\langle\varphi1| + \bar{b}_5|\varphi1\rangle\langle001| + b_6|\varphi0\rangle\langle\varphi1| \\
&\quad + \bar{b}_6|\varphi1\rangle\langle\varphi0|.
\end{aligned}$$

Using the above expression for σ we get

$$\begin{aligned}
\mathrm{Tr}_2(\sigma) &= a_{00}|00\rangle\langle00| + a_{01}|01\rangle\langle01| + \frac{a_{\varphi0}}{2}\mathbb{1} \otimes |0\rangle\langle0| \\
&\quad + \frac{a_{\varphi1}}{2}\mathbb{1} \otimes |1\rangle\langle1| + b_1|00\rangle\langle01| + \bar{b}_1|01\rangle\langle00| \\
&\quad + \frac{b_2}{\sqrt{2}}|00\rangle\langle10| + \frac{\bar{b}2}{\sqrt{2}}|10\rangle\langle00| + \frac{b_3}{\sqrt{2}}|00\rangle\langle11| \\
&\quad + \frac{\bar{b}_3}{\sqrt{2}}|11\rangle\langle00| + \frac{b_4}{\sqrt{2}}|01\rangle\langle10| + \frac{\bar{b}_4}{\sqrt{2}}|10\rangle\langle01| \\
&\quad + \frac{b_5}{\sqrt{2}}|01\rangle\langle11| + \frac{\bar{b}_5}{\sqrt{2}}|11\rangle\langle01| + \frac{b_6}{2}\mathbb{1} \otimes |0\rangle\langle1| \\
&\quad + \frac{\bar{b}_6}{2}\mathbb{1} \otimes |1\rangle\langle0|,
\end{aligned}$$

that implies $a_{00} = a_{\varphi1} = 0$, $a_{\varphi0} = \frac{2}{3}$, $a_{01} = \frac{1}{3}$, $b_1 = b_2 = b_3 = b_5 = b_6 = 0$, and $b_4 = \frac{\sqrt{2}}{3}$. In conclusion, we have

$$\begin{aligned}
\sigma &= \tfrac{1}{3}(|001\rangle\langle001| + 2|\varphi0\rangle\langle\varphi0| \\
&\quad + \sqrt{2}|001\rangle\langle\varphi0| + \sqrt{2}|\varphi0\rangle\langle001|) \\
&= |W\rangle\langle W|.
\end{aligned}$$

XI. BELL NONLOCALITY

Bell nonlocality is, similarly to steering, a phenomenon that we do not find in classical theory, but is often used in quantum theory. Bell nonlocality [2] was formulated as a response to the well-known EPR paradox [3]. Although in the original formulation the operational idea was different than

the one we will present, we will see that Bell nonlocality may be understood as an incompatibility test, in the same way as steering.

Let us assume that we have four parties: Alice, Bob, Charlie, and Dan. Alice has two channels $\Phi_1^A : K_A \to K_{B_1}$ and $\Phi_2^A : K_A \to K_{B_2}$ that she can use to send a state to Bob, and Charlie has two channels $\Phi_1^C : K_C \to K_{D_1}$ and $\Phi_2^C : K_C \to K_{D_2}$ that he can use to send a state to Dan. Assume that $K_C \tilde{\otimes} K_A$ is defined and let $\psi \in K_C \tilde{\otimes} K_A$ be a bipartite state shared by Alice and Charlie. The idea that we use to define Bell nonlocality is very simple: if we were able to use $(\mathrm{id} \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ and $((\Phi_1^C, \Phi_2^C) \otimes \mathrm{id})(\psi)$ as nontrivial incompatibility test, we may as well investigate whether $((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ provides an incompatibility test in the same manner.

Definition 12. Let $\psi \in K_C \tilde{\otimes} K_A$ and let

$$\begin{aligned}
\Phi_1^A &: K_A \to K_{B_1}, \\
\Phi_2^A &: K_A \to K_{B_2}, \\
\Phi_1^C &: K_C \to K_{D_1}, \\
\Phi_2^C &: K_C \to K_{D_2}
\end{aligned}$$

be channels. We call the state $((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ bipartite biconditional state.

Assume that the channels Φ_1^A and Φ_2^A are compatible, so that we have $(\Phi_1^A, \Phi_2^A) = J(\Phi^A)$ for some channel $\Phi^A : K_A \to K_{B_1} \tilde{\otimes} K_{B_2}$ and also that the channels Φ_1^C and Φ_2^C are compatible, so there is a channel $\Phi^C : K_C \to K_{D_1} \tilde{\otimes} K_{D_2}$ such that $(\Phi_1^A, \Phi_2^A) = J(\Phi^A)$. Let $\psi \in K_C \tilde{\otimes} K_A$, then we have

$$\big((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A)\big)(\psi) = (J' \otimes J')((\Phi_C \otimes \Phi_A)(\psi)),$$

where the maps J' are defined as before, with the exception that we denote them the same even though they map different spaces.

We present a definition of Bell nonlocality using the same line of thinking as we used in Definition 11. For simplicity we will denote

$$Q_{DC} = (J' \otimes J')\big(K_{D_1} \tilde{\otimes} K_{D_2} \tilde{\otimes} K_{C_1} \tilde{\otimes} K_{C_2}\big).$$

Definition 13. Let $\psi \in K_C \tilde{\otimes} K_A$ be a bipartite state and let $\Phi_1^A : K_A \to K_{B_1}$, $\Phi_2^A : K_A \to K_{B_2}$, $\Phi_1^C : K_A \to K_{C_1}$, and $\Phi_2^C : K_A \to K_{D_2}$ be channels. We say that the bipartite biconditional state $((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ is Bell nonlocal if

$$\big((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A)\big)(\psi) \notin Q_{DC}.$$

Otherwise, we call the bipartite biconditional state Bell local.

The following result follows immediately from Definition 13.

Corollary 3. Let $\psi \in K_C \tilde{\otimes} K_A$ be a bipartite state and let $\Phi_1^A : K_A \to K_{B_1}$, $\Phi_2^A : K_A \to K_{B_2}$, $\Phi_1^C : K_A \to K_{C_1}$, and $\Phi_2^C : K_A \to K_{D_2}$ be channels. The bipartite biconditional state $((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ is Bell nonlocal only if the channels Φ_1^A, Φ_2^A and Φ_1^C, Φ_2^C are incompatible.

We will show that entanglement plays a key role in Bell nonlocality.

Proposition 20. Let $\psi \in K_C \dot{\otimes} K_A$ be a separable bipartite state and let $\Phi_1^A : K_A \to K_{B_1}$, $\Phi_2^A : K_A \to K_{B_2}$, $\Phi_1^C : K_A \to$

K_{C_1}, and $\Phi_2^C : K_A \to K_{D_2}$ be channels. The bipartite bicon-
ditional state $((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi)$ is Bell local.

Proof. It is again sufficient to consider $\psi = x_C \otimes x_A$ for
$x_A \in K_A$, $x_C \in K_C$ due to the linearity of the maps (Φ_1^A, Φ_2^A)
and (Φ_1^C, Φ_2^C). Consider the state $\varphi \in K_{D_1} \tilde{\otimes} K_{D_2} \tilde{\otimes} K_{C_1} \tilde{\otimes} K_{C_2}$
given as

$$\varphi = \Phi_1^C(x_C) \otimes \Phi_2^C(x_C) \otimes \Phi_1^A(x_A) \otimes \Phi_2^A(x_A),$$

then we have

$$((\Phi_1^C, \Phi_2^C) \otimes (\Phi_1^A, \Phi_2^A))(\psi) = (J' \otimes J')(\varphi).$$

∎

XII. BELL NONLOCALITY OF MEASUREMENTS

We will again show that Definition 13 follows the standard
definition of Bell nonlocality [27] in the formalism of [26].

Proposition 21. Let S_1^A, S_2^A, S_1^C, and S_2^C be sim-
plexes and let $m_1^A : K_A \to S_1^A$, $m_2^A : K_A \to S_2^A$, $m_1^C : K_C \to S_1^C$, $m_2^C : K_C \to S_2^C$ be measurements. Let $\psi \in K_C \tilde{\otimes} K_A$, then
the bipartite biconditional state $((m_1^C, m_2^C) \otimes (m_1^A, m_2^A))(\psi)$ is
Bell nonlocal if

$$((m_1^C, m_2^C) \otimes (m_1^A, m_2^A))(\psi) \notin (S_1^C \times S_2^C) \dot{\otimes} (S_1^A \times S_2^A).$$

Proof. By direct calculation we have

$$\begin{aligned} Q_{CD} &= (J' \otimes J')(S_1^C \dot{\otimes} S_2^C \dot{\otimes} S_1^A \dot{\otimes} S_2^A) \\ &= (S_1^C \times S_2^C) \dot{\otimes} (S_1^A \times S_2^A). \end{aligned}$$

∎

One may again use the interpretation that both $S_1^C \times S_2^C$ and
$S_1^A \times S_2^A$ are spaces of conditional measurement probabilities,
so if we have $\psi \in (S_1^C \times S_2^C) \dot{\otimes} (S_1^A \times S_2^A)$, then we must have
$0 \leqslant \lambda_i \leqslant 1$, for $i \in \{1, \ldots, n\}$, $\sum_{i=1}^n \lambda_i = 1$, such that

$$\psi = \sum_{i=1}^n \lambda_i s_i^C \otimes s_i^A,$$

where in standard formulations both $s_i^C \in S_1^C \times S_2^C$ and $s_i^A \in_1^A \times S_2^A$ are represented by probabilities, i.e., by numbers, so the
tensor product between them is omitted.

We will provide proof of the standard and well-known
result about connection of steering and Bell nonlocality of
measurements.

Proposition 22. Let S_1^A, S_2^A, S_1^C, and S_2^C be sim-
plexes and let $m_1^A : K_A \to S_1^A$, $m_2^A : K_A \to S_2^A$, $m_1^C : K_C \to S_1^C$, $m_2^C : K_C \to S_2^C$ be measurements. Let $\psi \in K_C \tilde{\otimes} K_A$. If

$$(\mathrm{id} \otimes (m_1^A, m_2^A))(\psi) \in K_C \dot{\otimes} (S_1^A \times S_2^A),$$

i.e., if the bipartite state is not steerable by measurements
m_1^A, m_2^A, then

$$((m_1^C, m_2^C) \otimes (m_1^A, m_2^A))(\psi) \in (S_1^C \times S_2^C) \dot{\otimes} (S_1^A \times S_2^A).$$

Proof. Let

$$(\mathrm{id} \otimes (m_1^A, m_2^A))(\psi) \in K_C \dot{\otimes} (S_1^A \times S_2^A),$$

then for $n \in \mathbb{N}$, $i \in \{1, \ldots, n\}$, there are $0 \leqslant \lambda_i \leqslant 1$, $x_i \in K_C$
and $s_i \in S_1^A \times S_2^A$, $\sum_{i=1}^n \lambda_i = 1$, such that we have

$$(\mathrm{id} \otimes (m_1^A, m_2^A))(\psi) = \sum_{i=1}^n \lambda_i x_i \otimes s_i.$$

We get

$$((m_1^C, m_2^C) \otimes (m_1^A, m_2^A))(\psi) = \sum_{i=1}^n \lambda_i (m_1^C, m_2^C)(x_i) \otimes s_i,$$

and since we have $(m_1^C, m_2^C)(x_i) = (m_1^C(x_i), m_2^C(x_i)) \in S_1^C \times S_2^C$ we have

$$((m_1^C, m_2^C) \otimes (m_1^A, m_2^A))(\psi) \in (S_1^C \times S_2^C) \dot{\otimes} (S_1^A \times S_2^A).$$

∎

Note that the same result would also hold for steering by
the measurements m_1^C, m_2^C.

One may think that steering is somehow half of Bell
nonlocality, or that it is some middle step towards Bell
nonlocality as even our constructions in Secs. VIII and XI
would point to such a result. We will show that this is not true
in general, as we will provide a counterexample using quantum
channels in Example 9.

XIII. BELL NONLOCALITY OF QUANTUM CHANNELS

Bell nonlocality of quantum measurements is a deeply
studied topic in quantum theory, with several applications
in various device-independent protocols [49–52], randomness
generation and randomness expansion [53,54], and others (for
a recent review on Bell nonlocality see [55]).

Bell nonlocality of quantum channels follows very similar
rules to steering by quantum channels. We will derive results
specific for quantum theory in the same manner as in Sec. X.

Proposition 23. Let $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ and let $\Phi_1^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_2^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_1^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_2^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ be
channels. The bipartite biconditional state $((\Phi_1^1, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell nonlocal only if the bipartite biconditional
state $((\mathrm{id}, \mathrm{id}) \otimes (\mathrm{id}, \mathrm{id}))(\rho)$ is Bell nonlocal.

Proof. If the bipartite biconditional state $((\mathrm{id}, \mathrm{id}) \otimes (\mathrm{id}, \mathrm{id}))(\rho)$ is Bell local, then there exist $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such
that

$$\begin{aligned} \mathrm{Tr}_{24}(\sigma) &= \rho, \\ \mathrm{Tr}_{23}(\sigma) &= \rho, \\ \mathrm{Tr}_{14}(\sigma) &= \rho, \\ \mathrm{Tr}_{13}(\sigma) &= \rho. \end{aligned}$$

Let

$$\tilde{\sigma} = (\Phi_1^1 \otimes \Phi_2^1 \otimes \Phi_1^2 \otimes \Phi_2^2)(\sigma),$$

then

$$\begin{aligned} \mathrm{Tr}_{24}(\tilde{\sigma}) &= (\Phi_1^1 \otimes \Phi_1^2)(\rho), \\ \mathrm{Tr}_{23}(\tilde{\sigma}) &= (\Phi_1^1 \otimes \Phi_2^2)(\rho), \\ \mathrm{Tr}_{14}(\tilde{\sigma}) &= (\Phi_2^1 \otimes \Phi_1^2)(\rho), \\ \mathrm{Tr}_{13}(\tilde{\sigma}) &= (\Phi_2^1 \otimes \Phi_2^2)(\rho). \end{aligned}$$

∎

Note that again we do not have to replace all of the channels by the identity channels id, but we may replace only some.

Proposition 24. Let $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ and let $\Phi_1^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_2^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_1^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$, $\Phi_2^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ be channels, moreover, let $\Phi_1^1 = \Phi_U$ be a unitary channel given by the unitary matrix U, then the bipartite biconditional state $((\Phi_U, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell nonlocal if and only if the bipartite biconditional state $((id, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell nonlocal.

Proof. Using the very same idea as before, if the bipartite biconditional state $((\Phi_U, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell local, then there is $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that

$$\mathrm{Tr}_{24}(\sigma) = (\Phi_U \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{23}(\sigma) = (\Phi_U \otimes \Phi_2^2)(\rho),$$
$$\mathrm{Tr}_{14}(\sigma) = (\Phi_2^1 \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{13}(\sigma) = (\Phi_2^1 \otimes \Phi_2^2)(\rho).$$

Let

$$\tilde{\sigma} = (\Phi_{U^*} \otimes \mathrm{id} \otimes \mathrm{id} \otimes \mathrm{id})(\sigma),$$

then we get

$$\mathrm{Tr}_{24}(\tilde{\sigma}) = (\mathrm{id} \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{23}(\tilde{\sigma}) = (\mathrm{id} \otimes \Phi_2^2)(\rho),$$
$$\mathrm{Tr}_{14}(\tilde{\sigma}) = (\Phi_2^1 \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{13}(\tilde{\sigma}) = (\Phi_2^1 \otimes \Phi_2^2)(\rho).$$

∎

One may obtain similar results if some other of the channels $\Phi_1^1, \Phi_2^1, \Phi_1^2, \Phi_2^2$ is unitary as well as if more or even all of them are unitary.

The most iconic and most studied aspects of Bell nonlocality are the Bell inequalities. We are going to present a version of CHSH inequality for quantum channels. Assume that $\dim(\mathcal{H}) = 2$ and let $|0\rangle, |1\rangle$ denote any orthonormal basis of $\mathcal{H}$. We will use the shorthand $|00\rangle = |0\rangle \otimes |0\rangle$. Let $i, j \in \{1, 2\}$ and let

$$\begin{aligned}
E(\Phi_i^1, \Phi_j^2) &= \langle 00|(\Phi_i^1 \otimes \Phi_j^2)(\rho)|00\rangle \\
&\quad - \langle 01|(\Phi_i^1 \otimes \Phi_j^2)(\rho)|01\rangle \\
&\quad - \langle 10|(\Phi_i^1 \otimes \Phi_j^2)(\rho)|10\rangle \\
&\quad + \langle 11|(\Phi_i^1 \otimes \Phi_j^2)(\rho)|11\rangle \\
&= \mathrm{Tr}((\Phi_i^1 \otimes \Phi_j^2)(\rho)A),
\end{aligned}$$

where

$$A = |00\rangle\langle 00| - |01\rangle\langle 01| - |10\rangle\langle 10| + |11\rangle\langle 11|.$$

The quantity $E(\Phi_i^1, \Phi_j^2)$ is to be interpreted as the correlation between the marginals $\mathrm{Tr}_1((\Phi_i^1 \otimes \Phi_j^2)(\rho))$ and $\mathrm{Tr}_2((\Phi_i^1 \otimes \Phi_j^2)(\rho))$. Since we have $-\mathbb{1} \leqslant A \leqslant \mathbb{1}$ it is straightforward that we have $-1 \leqslant E(\Phi_i^1, \Phi_j^2) \leqslant 1$. Define a quantity

$$X_\rho = E(\Phi_1^1, \Phi_1^2) + E(\Phi_1^1, \Phi_2^2) + E(\Phi_2^1, \Phi_1^2) - E(\Phi_2^1, \Phi_2^2),$$

we will show that X_ρ corresponds to the quantity used in CHSH inequality. It is straightforward to see that $-4 \leqslant X_\rho \leqslant 4$ is the algebraic bound on X_ρ.

Proposition 25. If the biconditional bipartite state $((\Phi_1^1, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell local, then we have $-2 \leqslant X_\rho \leqslant 2$.

Proof. If the biconditional bipartite state $((\Phi_1^1, \Phi_2^1) \otimes (\Phi_1^2, \Phi_2^2))(\rho)$ is Bell local then there is $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that

$$\mathrm{Tr}_{24}(\sigma) = (\Phi_1^1 \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{23}(\sigma) = (\Phi_1^1 \otimes \Phi_2^2)(\rho),$$
$$\mathrm{Tr}_{14}(\sigma) = (\Phi_2^1 \otimes \Phi_1^2)(\rho),$$
$$\mathrm{Tr}_{13}(\sigma) = (\Phi_2^1 \otimes \Phi_2^2)(\rho).$$

This yields

$$\begin{aligned}
E(\Phi_1^1, \Phi_1^2) &= \mathrm{Tr}((\Phi_1^1 \otimes \Phi_1^2)(\rho)A) = \mathrm{Tr}(\mathrm{Tr}_{24}(\sigma)A) \\
&= \mathrm{Tr}(\sigma(|0\rangle\langle 0| \otimes \mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} \\
&\quad - |0\rangle\langle 0| \otimes \mathbb{1} \otimes |1\rangle\langle 1| \otimes \mathbb{1} \\
&\quad - |1\rangle\langle 1| \otimes \mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} \\
&\quad + |1\rangle\langle 1| \otimes \mathbb{1} \otimes |1\rangle\langle 1| \otimes \mathbb{1})).
\end{aligned}$$

In the same manner, we get

$$\begin{aligned}
E(\Phi_1^1, \Phi_2^2) &= \mathrm{Tr}(\sigma(|0\rangle\langle 0| \otimes \mathbb{1} \otimes \mathbb{1} \otimes |0\rangle\langle 0| \\
&\quad - |0\rangle\langle 0| \otimes \mathbb{1} \otimes \mathbb{1} \otimes |1\rangle\langle 1| \\
&\quad - |1\rangle\langle 1| \otimes \mathbb{1} \otimes \mathbb{1} \otimes |0\rangle\langle 0| \\
&\quad + |1\rangle\langle 1| \otimes \mathbb{1} \otimes \mathbb{1} \otimes |1\rangle\langle 1|)), \\
E(\Phi_2^1, \Phi_1^2) &= \mathrm{Tr}(\sigma(\mathbb{1} \otimes |0\rangle\langle 0| \otimes |0\rangle\langle 0| \otimes \mathbb{1} \\
&\quad - \mathbb{1} \otimes |0\rangle\langle 0| \otimes |1\rangle\langle 1| \otimes \mathbb{1} \\
&\quad - \mathbb{1} \otimes |1\rangle\langle 1| \otimes |0\rangle\langle 0| \otimes \mathbb{1} \\
&\quad + \mathbb{1} \otimes |1\rangle\langle 1| \otimes |1\rangle\langle 1| \otimes \mathbb{1})),
\end{aligned}$$

and

$$\begin{aligned}
E(\Phi_2^1, \Phi_2^2) &= \mathrm{Tr}(\sigma(\mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} \otimes |0\rangle\langle 0| \\
&\quad - \mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} \otimes |1\rangle\langle 1| \\
&\quad - \mathbb{1} \otimes |1\rangle\langle 1| \otimes \mathbb{1} \otimes |0\rangle\langle 0| \\
&\quad + \mathbb{1} \otimes |1\rangle\langle 1| \otimes \mathbb{1} \otimes |1\rangle\langle 1|)).
\end{aligned}$$

Together we get

$$\begin{aligned}
X_\rho = 2\,\mathrm{Tr}(\sigma(&|0000\rangle\langle 0000| + |0001\rangle\langle 0001| - |0010\rangle\langle 0010| \\
&- |0011\rangle\langle 0011| + |0100\rangle\langle 0100| - |0101\rangle\langle 0101| \\
&+ |0110\rangle\langle 0110| - |0111\rangle\langle 0111| - |1000\rangle\langle 1000| \\
&+ |1001\rangle\langle 1001| - |1010\rangle\langle 1010| + |1011\rangle\langle 1011| \\
&- |1100\rangle\langle 1100| - |1101\rangle\langle 1101| + |1110\rangle\langle 1110| \\
&+ |1111\rangle\langle 1111|))
\end{aligned}$$

that implies $-2 \leqslant X_\rho \leqslant 2$. ∎

At this point, one may ask whether there exists an equivalent of Tsirelson bound [28] for the inequality given by Proposition 25, or what is the maximum violation of the

aforementioned inequality. We will show that the Tsirelson bound $2\sqrt{2}$ is both reachable and maximum violation by quantum channels.

Proposition 26. For any state $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ and any four channels $\Phi_1^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2^1 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_1^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}, \Phi_2^2 : \mathfrak{D}_{\mathcal{H}} \to \mathfrak{D}_{\mathcal{H}}$ we have

$$X_\rho \leqslant 2\sqrt{2}.$$

Proof. We define the adjoint channel Φ_1^{1*} to channel Φ_1^1 as the linear map $\Phi_1^{1*} : B_h(\mathcal{H}) \to B_h(\mathcal{H})$ such that for all $\sigma \in \mathfrak{D}_{\mathcal{H}}$ and $E \in B_h(\mathcal{H}), 0 \leqslant E \leqslant \mathbb{1}$, we have

$$\mathrm{Tr}\big(\Phi_1^1(\sigma)E\big) = \mathrm{Tr}\big(\sigma \Phi_1^{1*}(E)\big).$$

Since Φ_1^1 is a channel, we have $0 \leqslant \Phi_1^{1*}(E) \leqslant \mathbb{1}$ and $\Phi_1^{1*}(\mathbb{1}) = \mathbb{1}$. This approach of mapping effects instead of states is called the Heisenberg picture.

Let $i,j \in \{1,2\}$, then we have

$$\mathrm{Tr}\big((\Phi_i^1 \otimes \Phi_j^2)(\rho)|00\rangle\langle00|\big) = \mathrm{Tr}\big(\rho \Phi_i^{1*}(|0\rangle\langle0|) \otimes \Phi_j^{2*}(|0\rangle\langle0|)\big).$$

Denoting

$$M_i^1 = \Phi_i^{1*}(|0\rangle\langle0|),$$
$$M_j^2 = \Phi_j^{1*}(|0\rangle\langle0|),$$

we see that we have

$$
\begin{aligned}
E\big(\Phi_i^1, \Phi_j^2\big) &= \mathrm{Tr}\big(\rho M_i^1 \otimes M_j^2\big) - \mathrm{Tr}\big(\rho (\mathbb{1} - M_i^1) \otimes M_j^2\big) \\
&\quad - \mathrm{Tr}\big(\rho M_i^1 \otimes (\mathbb{1} - M_j^2)\big) \\
&\quad + \mathrm{Tr}\big(\rho (\mathbb{1} - M_i^1) \otimes (\mathbb{1} - M_j^2)\big) \\
&= E\big(M_i^1, M_j^2\big),
\end{aligned}
$$

where $E\big(M_i^1, M_j^2\big)$ is a correlation for the two-outcome measurements given by the effects M_i^1 and M_j^2. It is a well-known result [28] that we always have

$$E\big(M_1^1, M_1^2\big) + E\big(M_1^1, M_2^2\big) + E\big(M_2^1, M_1^2\big) - E\big(M_2^1, M_2^2\big)$$
$$\leqslant 2\sqrt{2}.$$

∎

It is very intuitive that the Tsirelson bound, reachable by measurements, will be also reachable by channels. To prove this, let $M,N \in B_h(\mathcal{H}), 0 \leqslant M \leqslant \mathbb{1}, 0 \leqslant N \leqslant \mathbb{1}$, and define channels $\Phi_M : B_h(\mathcal{H}) \to B_h(\mathcal{H}), \Phi_N : B_h(\mathcal{H}) \to B_h(\mathcal{H})$ such that for $\sigma \in \mathfrak{D}_{\mathcal{H}}$ we have

$$\Phi_M(\sigma) = \mathrm{Tr}(\sigma M)|0\rangle\langle0| + \mathrm{Tr}(\sigma(\mathbb{1} - M))|1\rangle\langle1|,$$
$$\Phi_N(\sigma) = \mathrm{Tr}(\sigma N)|0\rangle\langle0| + \mathrm{Tr}(\sigma(\mathbb{1} - N))|1\rangle\langle1|.$$

It is easy to verify that the maps Φ_M, Φ_N are quantum channels and that they are also measurements as they map the state space $\mathfrak{D}_{\mathcal{H}}$ to the simplex $\mathrm{conv}\{|0\rangle\langle0|, |1\rangle\langle1|\}$. Let $\rho \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$, then we have

$$
\begin{aligned}
\mathrm{Tr}((\Phi_M \otimes \Phi_N)(\rho)A) &= \mathrm{Tr}(\rho(M \otimes N - (\mathbb{1} - M) \otimes N) \\
&\quad - \mathrm{Tr}(\rho(M \otimes (\mathbb{1} - N))) \\
&\quad + \mathrm{Tr}(\rho((\mathbb{1} - M) \otimes (\mathbb{1} - N))) \\
&= E(M,N).
\end{aligned}
$$

This proves that any set of correlations and any violation of CHSH inequality reachable by measurements are also reachable by quantum channels as a violation of the bound given by Proposition 25.

To generalize the proposed inequality, one may replace the projectors $|0\rangle\langle0|$ and $|1\rangle\langle1|$ by any pair of effects $M,N \in B_h(\mathcal{H}), 0 \leqslant M \leqslant \mathbb{1}, 0 \leqslant N \leqslant \mathbb{1}$, and have

$$
\begin{aligned}
A &= M \otimes N - (\mathbb{1} - M) \otimes N - M \otimes (\mathbb{1} - N) + (\mathbb{1} - M) \\
&\quad \otimes (\mathbb{1} - N).
\end{aligned}
$$

From now on, we will consider a special case. Keep $\dim(\mathcal{H}) = 2$ and let

$$|\psi^+\rangle\langle\psi^+| = \tfrac{1}{2}(|00\rangle\langle00| + |11\rangle\langle00| + |00\rangle\langle11| + |11\rangle\langle11|)$$

be the maximally entangled state, let U_1, U_2, V_1, V_2 be unitary matrices and let $\Phi_1^1 = \Phi_{U_1}, \Phi_2^1 = \Phi_{U_2}, \Phi_1^2 = \Phi_{V_1}, \Phi_2^2 = \Phi_{V_2}$ be unitary channels given by the respective unitary matrices. We will consider the bipartite biconditional state $((\Phi_{U_1}, \Phi_{U_2}) \otimes (\Phi_{V_1}, \Phi_{V_2}))(|\psi^+\rangle\langle\psi^+|)$ and we will show that the correlations for the given bipartite biconditional state are of a particular nice form. We have

$$\big(\Phi_{U_i} \otimes \Phi_{V_j}\big)(|\psi^+\rangle\langle\psi^+|) = \big(\mathrm{id} \otimes \Phi_{V_j U_i^T}\big)(|\psi^+\rangle\langle\psi^+|),$$

where $i,j \in \{1,2\}$ and for U^T denotes the transpose of the matrix U. For the correlation we have

$$
\begin{aligned}
E\big(\Phi_{U_i}, \Phi_{V_j}\big) &= \mathrm{Tr}\big((\mathrm{id} \otimes \Phi_{V_j U_i^T})(|\psi^+\rangle\langle\psi^+|)A\big) \\
&= \tfrac{1}{2}\big(|\langle0|V_j U_i^T|0\rangle|^2 + |\langle1|V_j U_i^T|1\rangle|^2 \\
&\quad - |\langle0|V_j U_i^T|1\rangle|^2 - |\langle1|V_j U_i^T|0\rangle|^2\big). \quad (11)
\end{aligned}
$$

We will provide an example of a violation of the bound given by Proposition 25 by incompatible unitary channels.

Example 8. Let $\dim(\mathcal{H}) = 2$ and let $\vartheta \in \mathbb{R}$ be a parameter. Let U_1, U_2, V_1, V_2 be unitary matrices given as

$$U_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix},$$

$$U_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$V_1 = \frac{1}{\sqrt{1+\vartheta}}\begin{pmatrix} \sqrt{\vartheta} & 1 \\ 1 & -\sqrt{\vartheta} \end{pmatrix},$$

$$V_2 = \frac{1}{\sqrt{1+\vartheta}}\begin{pmatrix} 1 & \sqrt{\vartheta} \\ \sqrt{\vartheta} & -1 \end{pmatrix}.$$

Consider the bipartite biconditional state $(\Phi_{U_1}, \Phi_{U_2}) \otimes (\Phi_{V_1}, \Phi_{V_2})(|\psi^+\rangle\langle\psi^+|)$. Using Eq. (11), we can obtain $X_{|\psi^+\rangle\langle\psi^+|}$ as a function of ϑ. The function is plotted in Fig. 1, where it is shown that for certain values of ϑ the bipartite biconditional state violates the bound given by Proposition 25.

It is also easy to see that the bipartite biconditional state $((\mathrm{id},\mathrm{id}) \otimes (\mathrm{id},\mathrm{id}))(|\psi^+\rangle\langle\psi^+|)$ does not violate the bound given by Proposition 25 because all of the correlations are the same, yet according to Proposition 24 we know that it must be a Bell nonlocal bipartite biconditional state. This shows that not all Bell nonlocal bipartite biconditional states violate the inequality given by Proposition 25.
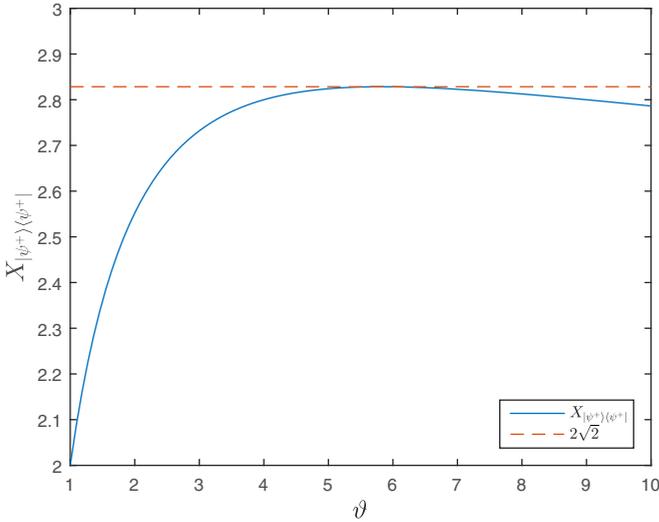
FIG. 1. The blue solid line is $X_{|\psi^+\rangle\langle\psi^+|}$ as a function of the parameter $\vartheta \in [1,10]$ when we consider the bipartite biconditional state $(\Phi_{U_1}, \Phi_{U_2}) \otimes (\Phi_{V_1}, \Phi_{V_2})(|\psi^+\rangle\langle\psi^+|)$ from Example 8. The red dashed line corresponds to the Tsirelson bound $2\sqrt{2}$.

One may wonder whether there is or is not a connection between steering and Bell nonlocality. As we have already showed in Proposition 22, for measurements Bell nonlocality implies steering. We will show that for channels the same does not hold.

Example 9. Let $\dim(\mathcal{H}) = 2$. Let $\rho_W \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H}}$ be given as in Example 7 as a partial trace over the state $|W\rangle\langle W|$. We already know that the state ρ_W is not steerable by any pair of channels. Consider the bipartite biconditional state $((\mathrm{id},\mathrm{id}) \otimes (\mathrm{id},\mathrm{id}))(\rho_W)$, if it is Bell local, then there must be a state $\sigma \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ such that

$$\mathrm{Tr}_{13}(\sigma) = \mathrm{Tr}_{14}(\sigma) = \mathrm{Tr}_{23}(\sigma) = \mathrm{Tr}_{24}(\sigma) = \rho_W.$$

Observe that $\mathrm{Tr}_1(\sigma) \in \mathfrak{D}_{\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H}}$ is such that $\mathrm{Tr}_3(\mathrm{Tr}_1(\sigma)) = \mathrm{Tr}_4(\mathrm{Tr}_1(\sigma)) = \rho_W$ which implies that, according to our calculations in Example 7, we must have

$$\mathrm{Tr}_1(\sigma) = |W\rangle\langle W|.$$

According to [56, Lemma 3] this implies that there is a state $\rho \in \mathfrak{D}_{\mathcal{H}}$ such that $\sigma = \rho \otimes |W\rangle\langle W|$. This implies that we have $\mathrm{Tr}_{23}(\sigma) = \rho \otimes \frac{1}{3}(2|0\rangle\langle 0| + |1\rangle\langle 1|)$ which is clearly a separable state. This is a contradiction as we should have had $\mathrm{Tr}_{23}(\sigma) = \rho_W$, which is an entangled state.

XIV. CONCLUSIONS

We have introduced the general definition of compatibility of channels in general probabilistic theory through the idea of conditional channels. We have also shown that a naive idea for a compatibility test leads to a simple and straightforward formulation of steering and Bell nonlocality. These formulations of steering and Bell nonlocality are different even when we consider only measurements instead of channels. Throughout the paper, we have shown that all of our definitions and results are in correspondence with the known result for measurements

and we have also provided several examples and results about the introduced concepts in quantum theory.

The paper has opened several questions and areas of research. For example, a possible area of research would be to look at the structure of conditional states and conditional channels and to try to connect them to Bayesian theory.

Concerning the compatibility of channels, one may formulate different notions of degree of (in)compatibility or of robustness of compatibility in general probabilistic theory and look at their properties, in a similar way as it was already done in quantum theory [57]. For quantum channels, one may wonder which types of channels are compatible. This would generalize the no broadcasting theorem [56,58] which states that two unitary channels can not be compatible.

One may also consider our formulations of steering and Bell nonlocality as a case of the problem of finding a multipartite state with given marginals. Such problems were studied in recent years [59,60], but not in the form that would be applicable to the problems of steering and Bell nonlocality as incompatibility tests. This opens questions as to whether one may characterize the structure of the cone Q_{CD} and of other cones of interest in quantum theory. From a geometrical viewpoint, this question is closely tied to the question of existence of other Bell inequalities for channels than the one we presented. Existence and exact form of the generalized Bell inequalities are also a very interesting possible area of research.

We may also consider the use of steering and Bell nonlocality of channels in the context of quantum information theory and quantum communication. Both steering and Bell nonlocality of measurement were used to formulate quantum protocols and it is of great interest whether exploiting the steering and Bell nonlocality of channels may lead to even better or more useful applications.

One may also try and clarify the lack of connection between steering and Bell nonlocality of channels. As we have showed in Example 9, even if two channels can not steer a state, when applied to both parts of the state the resulting biconditional bipartite state may be Bell nonlocal. This may even have interesting applications in quantum theory of information as so far steering has been considered to lead to one-side device-independent protocols that were seen as a middle step between the original protocol and device-independent protocol.

It may also be interesting to consider the resource theories of channel incompatibility, of steering by channels, and of Bell nonlocality of channels. Several similar resource theories were already constructed (see [61] for a review).

[1] E. Schrödinger, Math. Proc. Cambridge Philos. Soc. **32**, 446 (1936).

[2] J. S. Bell, Physics (College. Park. Md) **1**, 195 (1964).

[3] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[4] T. Heinosaari, D. Reitzner, and P. Stano, Found. Phys. **38**, 1133 (2008).

[5] L.-Y. Hsu, Quantum Inf. Proc. **15**, 5167 (2016).

[6] T. Heinosaari, J. Kiukas, and D. Reitzner, Phys. Rev. A **92**, 022115 (2015).

[7] T. Heinosaari, J. Schultz, A. Toigo, and M. Ziman, Phys. Lett. A **378**, 1695 (2014).

[8] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Phys. Rev. A **96**, 022110 (2017).

[9] D. Cavalcanti and P. Skrzypczyk, Rep. Prog. Phys. **80**, 024001 (2017).

[10] R. Uola, T. Moroder, and O. Gühne, Phys. Rev. Lett. **113**, 160403 (2014).

[11] R. Uola, C. Budroni, O. Gühne, and J. P. Pellonpää, Phys. Rev. Lett. **115**, 230402 (2015).

[12] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).

[13] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).

[14] M. Banik, Md. R. Gazi, S. Ghosh, and G. Kar, Phys. Rev. A **87**, 052125 (2013).

[15] F. Hirsch, M. T. Quintino, and N. Brunner, arXiv:1707.06960.

[16] B. Erika and T. Vértesi, arXiv:1705.10069.

[17] P. Girdhar and E. G. Cavalcanti, Phys. Rev. A **94**, 032317 (2016).

[18] D. Cavalcanti and P. Skrzypczyk, Phys. Rev. A **93**, 052112 (2016).

[19] T. Heinosaari, T. Miyadera, and M. Ziman, J. Phys. A: Math. Theor. **49**, 123001 (2016).

[20] M. Sedlák, D. Reitzner, G. Chiribella, and M. Ziman, Phys. Rev. A **93**, 052323 (2016).

[21] T. Heinosaari and T. Miyadera, J. Phys. A: Math. Theor. **50**, 135302 (2017).

[22] M. Banik, S. Das, and A. S. Majumdar, Phys. Rev. A **91**, 062124 (2015).

[23] S. N. Filippov, T. Heinosaari, and L. Leppäjärvi, Phys. Rev. A **95**, 032127 (2017).

[24] M. Banik, J. Math. Phys. **56**, 052101 (2015).

[25] N. Stevens and P. Busch, Phys. Rev. A **89**, 022123 (2014).

[26] A. Jenčová, arXiv:1705.08008.

[27] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).

[28] B. S. Cirel'son, Lett. Math. Phys. **4**, 93 (1980).

[29] J. Barrett, Phys. Rev. A **75**, 032304 (2007).

[30] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).

[31] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).

[32] P. Janotta and R. Lal, Phys. Rev. A **87**, 052131 (2013).

[33] W. Van Dam, Nat. Comput. **12**, 9 (2013).

[34] Y. Quek and P. W. Shor, Phys. Rev. A **95**, 052329 (2017).

[35] S. Aaronson, Phys. Rev. A **71**, 032325 (2005).

[36] N. Johansson and J.-Å. Larsson, Quantum Inf. Proc. **16**, 233 (2017).

[37] R. T. Rockafellar, *Convex Analysis*, Princeton Landmarks in Mathematics and Physics (Princeton University Press, Princeton, NJ, 1997).

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*, Berichte über verteilte messysteme (Cambridge University Press, Cambridge, 2004).

[39] D. J. Foulis and M. K. Bennett, Found. Phys. **24**, 1331 (1994).

[40] L. Asimow and A. J. Ellis, *Convexity Theory and its Applications in Functional Analysis*, L.M.S. Monographs (Academic Press, New York, 1980).

[41] A. Kissinger and S. Uijlen, arXiv:1701.04732.

[42] R. A. Ryan, *Introduction to Tensor Products of Banach Spaces*, Springer Monographs in Mathematics (Springer, London, 2013).

[43] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory. From Uncertainty to Entanglement* (Cambridge University Press, Cambridge, 2012).

[44] I. Namioka and R. Phelps, Pacific J. Math. **31**, 469 (1969).

[45] M. Plávala, Phys. Rev. A **94**, 042108 (2016).

[46] A. Jenčová, arXiv:1207.5420.

[47] M. Ziman, Phys. Rev. A **77**, 062112 (2008).

[48] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, Phys. Rev. A **85**, 010301(R) (2012).

[49] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[50] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, Phys. Rev. Lett. **98**, 230501 (2007).

[51] Č. Brukner, M. Zukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. **92**, 127901 (2004).

[52] W.-Y. Hwang, H.-Y. Su, and J. Bae, Phys. Rev. A **95**, 062313 (2017).

[53] R. Colbeck and A. Kent, J. Phys. A: Math. Theor. **44**, 095305 (2011).

[54] F. Bischof, H. Kampermann, and D. Bruß, Phys. Rev. A **95**, 062305 (2017).

[55] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).

[56] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, arXiv:quant-ph/0611295.

[57] E. Haapasalo, J. Phys. A: Math. Theor. **48**, 255303 (2015).

[58] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007).

[59] A. Klyachko, arXiv:quant-ph/0409113.

[60] N. Wyderka, F. Huber, and O. Gühne, Phys. Rev. A **96**, 010102 (2017).

[61] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2013).