Incompatible measurements in a class of general probabilistic theories

Anna Jenčová[*]

Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia

(Received 24 January 2018; revised manuscript received 10 April 2018; published 25 July 2018)

We study incompatibility of measurements and its relation to steering and nonlocality in a class of finite-dimensional general probabilistic theories (GPTs). The basic idea is to represent finite collections of measurements as affine maps of a state space into a polysimplex and show that incompatibility is characterized by properties of these maps. We introduce the notion of an incompatibility witness and show its relation to incompatibility degree. We find the largest incompatibility degree attainable by pairs of two-outcome quantum measurements and characterize state spaces for which incompatibility degree attains maximal values possible in GPTs. As examples, we study the spaces of classical and quantum channels and show their close relation to polysimplices. This relation explains the superquantum nonclassical effects that were observed on these spaces.

DOI: 10.1103/PhysRevA.98.012133

I. INTRODUCTION

Incompatibility of measurements is one of the fundamental features of quantum mechanics. As a key ingredient in quantum information protocols, incompatibility and related nonclassical effects such as Bell nonlocality and steering are viewed as important resources in quantum information theory.

General probabilistic theories (GPTs) form a framework for description of physical theories admitting probabilistic processes. Within this framework, quantum theory is specified by several axioms [1,2]. Some of these axioms (causality, perfect distinguishability, and local tomography) define a class of theories that can be studied within the setting of ordered vector spaces and their tensor products.

Many of the basic features that distinguish quantum mechanics from any classical theory are shared by a large class of GPTs. It is a natural question what the properties that characterize quantum mechanics are. This question has been studied for many years, see, e.g., [3–5], but recently the advance of quantum information theory led to a renewed interest in this topic [6–9]. To answer this question, it is important to understand the nature of the nonclassical features and relations between them.

Such relations were already observed: it is well known that steering and nonlocality require both entanglement and incompatible observables. The relations between nonlocality, steering, and incompatibility were studied in [7,10–13], both for quantum theory and in GPTs. On the other hand, there exist unsteerable entangled quantum states [14] and incompatible sets of quantum measurements that cannot lead to violation of Bell inequalities [12]. To understand these relations, the more general setting is useful because it allows one to recognize which nonclassical manifestations are consequences of convexity and the tensor product structure, and which are inherently quantum.

In this contribution, we study incompatibility of measurements and its relation to steering and nonlocality in a class of finite-dimensional GPTs, using the tools of convex geometry. The basic idea is to represent finite collections of measurements as affine maps of a state space into a polysimplex (that is, a Cartesian product of simplices) and show that incompatibility is characterized by properties of these maps. A generalization of this idea was already used to describe incompatibility of channels [15].

We introduce the notion of an incompatibility witness and show its relation to incompatibility degree, defined in [16,17]. The concept of a witness as a functional separating convex sets or cones is not new; indeed it is the same as in the case of entanglement witnesses, but it seems that it was not explicitly used before in the context of incompatibility. Besides, our setting provides a geometric representation of incompatibility witnesses as certain maps from a polysimplex into the positive cone generated by the state space, which suggests a close relation of incompatibility to geometry of polysimplices. Using this representation, we find the largest incompatibility degree attainable by pairs of two-outcome quantum measurements, generalizing the results of [11], and characterize state spaces for which incompatibility degree attains maximal values possible in GPTs. This completes the results of [18] and [19], where maximal incompatibility of pairs of two-outcome measurements is considered.

The representation of collections of measurements as maps enables us to tie incompatibility directly to steering and nonlocality of states of composite systems. The concept of incompatibility witnesses helps to explain the relation of incompatibility degree and maximal violation of Bell inequalities, as well as the observed limitations of these relations. Here the geometry of polysimplices again serves as the main mathematical tool for treating these nonclassical features in a unified manner.

Besides the classical and quantum state spaces, we study the spaces of classical and quantum channels. It was observed in [19,20] that these spaces admit maximally incompatible measurements, which is known to be impossible in finite-dimensional quantum theory [16]. Moreover, it was shown

[*]jenca@mat.savba.sk

 ©2018 American Physical Society

that causal bipartite quantum channels can be used to obtain maximal violation of the CHSH inequality [21,22], in fact, all kinds of nonsignaling correlations [23]. We prove these results as easy consequences of the fact that the set of classical channels is a retract of the set of quantum channels and is affinely isomorphic to a polysimplex.

The paper is organized as follows. In the next section we briefly describe the main components of GPTs in our setting and present the main examples of spaces of classical and quantum states and channels. The structure of polysimplices is detailed in Sec. III. The relations to spaces of channels are also proved. Section IV is devoted to incompatibility of measurements, incompatibility witnesses, and incompatibility degree. The last section deals with steering and Bell nonlocality.

II. GENERAL PROBABILISTIC THEORIES

We present a brief overview of GPTs in the finite-dimensional setting, explain our overall assumptions, and introduce the mathematical tools needed in the sequel. Let us remark that the GPT framework is much broader; see, e.g., [6,24–26] for more details.

We first recall a few definitions and facts about convex sets and ordered vector spaces that will be needed below. For a subset $X \subseteq V$ of a finite-dimensional vector space V, we denote by $co(X)$ the convex hull and $aff(X)$ the affine span of X in V. For a convex subset $C \subseteq V$, $ri(C)$ denotes the relative interior of C in $aff(C)$ and $\dim(C) := \dim(aff(V))$. The set of all affine maps (that is, preserving the convex structure) between convex sets C_1 and C_2 will be denoted by $\mathcal{A}(C_1,C_2)$.

For the purposes of this paper, an ordered vector space is a pair (V,V^+), where V is a real finite-dimensional vector space and V^+ is a closed convex cone satisfying $V^+ \cap -V^+ = \{0\}$ and $V = V^+ - V^+$. This induces a partial order in V as $v \leqslant w$ if $w - v \in V^+$. Let V^* be the vector space dual with duality $\langle \cdot,\cdot \rangle$. The order dual of (V,V^+) is the ordered vector space $(V^*,(V^+)^*)$, with the closed convex cone of positive functionals

$$(V^+)^* := \{v^* \in V^*,\ \langle v^*,v\rangle \geqslant 0,\ \forall v \in V^+\}.$$

Note that we have $(V^+)^{**} = V^+$. A linear map between ordered vector spaces is called positive if it preserves the positive cone. We say that the cone V^+ is weakly self-dual if it is affinely isomorphic to $(V^+)^*$.

Let (V,V^+) and (W,W^+) be ordered vector spaces. There are two distinguished ways to define a positive cone in the tensor product $V \otimes W$:

$$V^+ \otimes_{\min} W^+ := \left\{ \sum_i v_i \otimes w_i,\ v_i \in V^+, w_i \in W^+ \right\},$$

$$V^+ \otimes_{\max} W^+ := ((V^+)^* \otimes_{\min} (W^+)^*)^*.$$

We have $V^+ \otimes_{\min} W^+ \subseteq V^+ \otimes_{\max} W^+$. The elements of $V^+ \otimes_{\min} W^+$ are called separable.

A. States, effects, and measurements

The framework of GPTs is built on basic notions of states, representing preparation procedures of a given system, and effects, assigning to each state the corresponding probabilities of outcomes in yes/no experiments. The state spaces have a natural convex structure, expressing the possibility of forming probabilistic mixtures of states. The effects must respect this structure and are therefore represented by affine functions from the state space into the unit interval. Throughout this paper, we will assume that any state space is a compact convex subset K of a finite-dimensional real vector space. Moreover, we adopt the no-restriction hypothesis, requiring that all affine functions $K \to [0,1]$ correspond to physical effects. For a discussion of these assumptions in GPTs see [2,24,27].

A state space K determines a pair of dual-ordered vector spaces as follows. Let $A(K) := \mathcal{A}(K,\mathbb{R})$ and $A(K)^+ := \mathcal{A}(K,\mathbb{R}^+)$. Let also 1_K be the constant map $1_K(x) \equiv 1$. Then $(A(K),A(K)^+)$ is an ordered vector space. The function 1_K is an interior element in $A(K)^+$ and hence is an order unit: for any $f \in A(K)$ we have $-t1_K \leqslant f \leqslant t1_K$ for some $t > 0$. The set of effects is thus given by

$$E(K) := \{f \in A(K),\ 0 \leqslant f \leqslant 1_K\}.$$

For $x \in K$, the evaluation map $f \mapsto f(x)$ defines a linear functional on $A(K)$ that is clearly positive and unital: $1_K(x) = 1$. The converse is also true [28], so that K can be identified with the set of positive unital functionals, or states, on $A(K)$. With this identification, K is a base of the dual cone $(A(K)^+)^*$, in the sense that each $0 \neq \varphi \in (A(K)^+)^*$ can be expressed in a unique way as a multiple of some element in K. We therefore have $(A(K)^+)^* \equiv V(K)^+ := \cup_{\lambda \geqslant 0}\lambda K$ and $A(K)^* \equiv V(K) := V(K)^+ - V(K)^+$. These identifications will be used throughout. For $\psi \in V(K)$, let

$$\|\psi\|_K = \inf\{a + b,\ \psi = ax - by,\ a,b \geqslant 0, x,y \in K\}.$$

Then $\|\cdot\|_K$ is a norm in $V(K)$, called the base norm. It is the dual of the supremum norm $\|\cdot\|_{\max}$ in $A(K)$.

Similarly to two-outcome measurements, any measurement on a system with state space K is fully described by its outcome statistics in each state. A measurement with $n+1$ outcomes is therefore identified with a map $f \in \mathcal{A}(K,\Delta_n)$, where Δ_n is the n-dimensional simplex of probability measures over $\{0,\dots,n\}$. The measurement f is determined by $n+1$ effects $f_0,\dots,f_n \in E(K)$, satisfying $\sum_i f_i = 1_K$. Here $f_i(x) = f(x)(i)$ is the ith component of the probability vector $f(x)$ and is interpreted as the probability of obtaining the outcome i if the system is in the state x. As before, we assume that each element of $\mathcal{A}(K,\Delta_n)$ describes a valid measurement.

Example 1 (classical state spaces). The state space of a classical system is an m-dimensional simplex Δ_m. We have $A(\Delta_m) \simeq V(\Delta_m) = \mathbb{R}^{m+1}$, with $V(\Delta_m)^+ \simeq A(\Delta_m)^+$ the positive cone of vectors with non-negative entries and $E(\Delta_m)$ is the set of vectors with entries in [0,1]. The base norm in this case is the l_1-norm in $\mathbb{R}^{m+1}$. Measurements $f \in \mathcal{A}(\Delta_m,\Delta_n)$ are classical channels and can be identified with $(n+1)\times(m+1)$ stochastic matrices $\{T(i|j)\}_{i,j}$, where $T(\cdot|j) \in \Delta_n$, $j = 0,\dots,m$, are determined by the values of f on the vertices of Δ_m.

Example 2 (quantum state spaces). A quantum state space is the set of density operators $\mathfrak{S}(\mathcal{H})$ on a finite-dimensional Hilbert space $\mathcal{H}$. We will sometimes use labels $\mathcal{H}_A$, $\mathcal{H}_B$, etc., for the Hilbert spaces; then we use the notations $d_A :=$

$\dim(\mathcal{H}_A)$, $\mathfrak{S}_A := \mathfrak{S}(\mathcal{H}_A)$. Any $f \in A(\mathfrak{S}(\mathcal{H}))$ has the form

$$f(\rho) = \mathrm{Tr}\, M\rho, \quad \rho \in \mathfrak{S}(\mathcal{H})$$

for some $M \in B_h(\mathcal{H})$, the set of Hermitian operators on $\mathcal{H}$. In this way, we have $A(\mathfrak{S}(\mathcal{H})) \simeq V(\mathfrak{S}(\mathcal{H})) = B_h(\mathcal{H})$, $A(\mathfrak{S}(\mathcal{H}))^+ \simeq V(\mathfrak{S}(\mathcal{H}))^+ = B(\mathcal{H})^+$, the cone of positive operators on $\mathcal{H}$, $1_{\mathfrak{S}(\mathcal{H})} = I$, the identity operator, and $E(\mathfrak{S}(\mathcal{H})) \simeq E(\mathcal{H})$, the set of quantum effects. The base norm $\|\cdot\|_{\mathfrak{S}(\mathcal{H})}$ is the trace norm $\|X\|_1 = \mathrm{Tr}\,|X|$. The measurements are given by positive operator valued measures (POVMs) on $\mathcal{H}$, that is, tuples of effects $M_0,\dots,M_n \in B(\mathcal{H})^+$, $\sum_i M_i = I$.

Example 3 (spaces of quantum channels). Let $\mathcal{H}_A,\mathcal{H}_{A'}$ be finite-dimensional Hilbert spaces. We will denote by $\mathcal{C}_{A,A'}$ the set of all quantum channels $\mathcal{H}_A \to \mathcal{H}_{A'}$, that is, all completely positive and trace-preserving linear maps $B(\mathcal{H}_A) \to B(\mathcal{H}_{A'})$. We now describe the corresponding cones and measurements for $\mathcal{C}_{A,A'}$; see [29] for details.

By the Choi representation, $\mathcal{C}_{A,A'}$ is isomorphic to a compact convex subset of the quantum state space $\mathfrak{S}_{A'A}$. Using this isomorphism, $V(\mathcal{C}_{A,A'})$ can be identified with the subspace

$$V(\mathcal{C}_{A,A'}) \equiv \{X \in B_h(\mathcal{H}_{A'A}),\ \mathrm{Tr}_{A'}X \in \mathbb{R}I\},$$

where $\mathrm{Tr}_{A'}$ is the partial trace over $\mathcal{H}_{A'}$. We then have

$$\mathcal{C}_{A,A'} = V(\mathcal{C}_{A,A'}) \cap \mathfrak{S}_{A'A}.$$

Consequently, $A(\mathcal{C}_{A,A'})$ is a quotient of $B_h(\mathcal{H}_{A'A})$ and $A(\mathcal{C}_{A,A'})^+$ is the set of equivalence classes containing some positive element. The base norm $\|\cdot\|_{\mathcal{C}_{A,A'}}$ is identified with the diamond norm $\|\cdot\|_\diamond$ [30]. One can also show that any measurement $f \in \mathcal{A}(\mathcal{C}_{A,A'},\Delta_n)$ has the form

$$f_i(\Phi) = \mathrm{Tr}\, M_i(\Phi \otimes id_R)(\rho_{AR})$$

for some POVM $M_0,\dots,M_n$ on $\mathcal{H}_{A'R}$ and some state $\rho \in \mathfrak{S}_{AR}$ where $\mathcal{H}_R$ is an ancilla, $d_R \leqslant d_A$, but the representation in this form is not unique; see also [31]. In particular, the unit effect $1_{\mathcal{C}_{A,A'}}$ is obtained from any state ρ_{AR} and the trivial measurement $M_0 = I_{A'R}$.

Example 4 (spaces of classical channels). The set of classical channels $\mathcal{A}(\Delta_m,\Delta_n)$ is isomorphic to a subset of $\mathcal{C}_{A,A'}$, with $d_A = m+1$, $d_{A'} = n+1$. Such an isomorphism is obtained by fixing orthonormal bases $|i\rangle_A$ of $\mathcal{H}_A$ and $|j\rangle_{A'}$ of $\mathcal{H}_{A'}$ and putting for any stochastic matrix $T \in \mathcal{A}(\Delta_m,\Delta_n)$,

$$\Phi_T(\sigma) = \sum_{i,j} \langle i|\sigma|i\rangle_A T(j|i)|j\rangle\langle j|_{A'}, \quad \sigma \in \mathfrak{S}(\mathcal{H}). \tag{1}$$

Quantum channels of this form are called classical-to-classical, or c-c, channels. The cones and measurements for this state space will be identified later (cf. Proposition 3).

B. Composition of state spaces: Tensor products

Let K_A and K_B be state spaces, corresponding to two systems in a GPT. To describe the state space of the joint system, we need the notion of a tensor product of state spaces. Let the composite state space be denoted by $K_A\tilde{\otimes}K_B$. Assuming the local tomography axiom [1,2,24], $K_A\tilde{\otimes}K_B$ is a subset of the tensor product $V(K_A)\otimes V(K_B)$ such that

(a) $x_A \otimes x_B \in K_A\tilde{\otimes}K_B$ for all $x_A \in K_A$, $x_B \in K_B$,
(b) $f_A\otimes f_B \in E(K_A\tilde{\otimes}K_B)$ for all $f_A \in E(K_A)$, $f_B \in E(K_B)$,
(c) $1_{K_A\tilde{\otimes}K_B} = 1_A \otimes 1_B$, here $1_A := 1_{K_A}$, $1_B := 1_{K_B}$.

This is based on the requirement that for the composite system, all product states and all product effects are valid. These conditions determine the minimal and the maximal tensor product of state spaces. Let

$$K_A \otimes_{\min} K_B := co\{x_i \otimes y_i,\ x_i \in K_A,\ y_i \in K_B\},$$
$$K_A \otimes_{\max} K_B := \{y \in V(K_A)\otimes V(K_B), \langle f_A \otimes f_B, y\rangle \geqslant 0,$$
$$\forall f_A \in E(K_A), f_B \in E(K_B),\ \langle 1_A \otimes 1_B, y\rangle = 1\}.$$

Note that both sets satisfy the conditions for a composite state space and we always have

$$K_A \otimes_{\min} K_B \subseteq K_A\tilde{\otimes}K_B \subseteq K_A \otimes_{\max} K_B.$$

The states in $K_A \otimes_{\min} K_B$ are called separable; all other states in $K_A\tilde{\otimes}K_B$ are called entangled. The particular form of the composite state space is specified by the theory in question; see the examples below. In terms of the related spaces and cones, we have

$$\begin{aligned}
V(K_A \otimes_{\min} K_B) &\simeq V(K_A \otimes_{\max} K_B)\\
&\simeq V(K_A)\otimes V(K_B),\\
A(K_A \otimes_{\min} K_B) &\simeq A(K_A \otimes_{\max} K_B)\\
&\simeq A(K_A)\otimes A(K_B),\\
V(K_A \otimes_{\min} K_B)^+ &\simeq V(K_A)^+ \otimes_{\min} V(K_B)^+,\\
V(K_A \otimes_{\max} K_B)^+ &\simeq V(K_A)^+ \otimes_{\max} V(K_B)^+,\\
A(K_A \otimes_{\min} K_B)^+ &\simeq A(K_A)^+ \otimes_{\max} A(K_B)^+,\\
A(K_A \otimes_{\max} K_B)^+ &\simeq A(K_A)^+ \otimes_{\min} A(K_B)^+;
\end{aligned}$$

this follows easily from the definitions and duality relations.

Example 5. For classical state spaces, we have

$$\Delta_{n_A} \otimes_{\min} \Delta_{n_B} = \Delta_{n_A} \otimes_{\max} \Delta_{n_B} = \Delta_{n_{AB}} =: \Delta_{n_A} \otimes \Delta_{n_B};$$

$n_{AB} := n_A n_B + n_A + n_B$ is the set of probability measures on $\{0,\dots,n_A\}\times\{0,\dots,n_B\}$. In fact, we have $K \otimes_{\min} \Delta_n = K \otimes_{\max} \Delta_n =: \Delta_n \otimes K$ for any state space K and this property characterizes the simplices in a general infinite-dimensional setting; see [32].

Example 6. For quantum state spaces, we have

$$\mathfrak{S}_A\tilde{\otimes}\mathfrak{S}_B = \mathfrak{S}_{AB},$$

with the usual tensor product of Hilbert spaces $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$. Here the minimal tensor product $\mathfrak{S}_A \otimes_{\min} \mathfrak{S}_B$ is the subset of separable states and the maximal tensor product $\mathfrak{S}_A \otimes_{\max} \mathfrak{S}_B$ is the set of entanglement witnesses (with unit trace).

Example 7. For spaces of quantum channels, let $\Phi \in \mathcal{C}_{A,A'}\tilde{\otimes}\mathcal{C}_{B,B'}$. It is natural to require that $\Phi \in \mathcal{C}_{AB,A'B'}$, so that Φ is a bipartite quantum channel. By the condition (b), each product of effects is a valid effect; in particular, $1_{\mathcal{C}_{A,A'}} \otimes f_B$ is a valid effect for any $f_B \in E(\mathcal{C}_{B,B'})$. By Example 3, $1_{\mathcal{C}_{A,A'}}$ is obtained from any state ρ_{RA} and the identity $I_{RA'}$. Let f_B be given by σ_B and an effect $M_{B'}$; then

$$\langle 1_{\mathcal{C}_{A,A'}} \otimes f_B, \Phi\rangle = \mathrm{Tr}\,(I_{RA} \otimes M_{B'})(id_R \otimes \Phi)(\rho_{RA} \otimes \sigma_B),$$

and this expression does not depend on ρ_{RA}. It follows that

$$\sigma_B \mapsto \mathrm{Tr}_{RA}(id_R \otimes \Phi)(\rho_{RA} \otimes \sigma_B)$$

defines a channel in $\mathcal{C}_{B,B'}$ that does not depend on ρ; similarly, we obtain a channel in $\mathcal{C}_{A,A'}$ by applying the unit effect on the second part. Channels with this property are called causal or no-signaling bipartite channels, see [21]; the set of all such channels is denoted by $\mathcal{C}^{\text{caus}}_{AB,AB'}$. We define the composite state space as

$$\mathcal{C}_{A,A'}\widetilde{\otimes}\mathcal{C}_{B,B'} := \mathcal{C}^{\text{caus}}_{AB,A'B'}.$$

The minimal tensor product is the set of local bipartite channels, which are convex combinations of channels prepared by each party separately. The maximal tensor product is strictly larger than $\mathcal{C}^{\text{caus}}_{AB,A'B'}$, since its elements are not necessarily completely positive.

Example 8. It is clear that

$$\mathcal{A}(\Delta_{m_A},\Delta_{n_A})\otimes_{\max}\mathcal{A}(\Delta_{m_B},\Delta_{n_B})\subset\mathcal{A}(\Delta_{m_{AB}},\Delta_{n_{AB}}),$$

where $\Delta_{m_{AB}} = \Delta_{m_Am_B+m_A+m_B} = \Delta_{m_A}\otimes\Delta_{m_B}$ and similarly for $\Delta_{n_{AB}}$. We will see later (Sec. V E) that the maximal tensor product is the set of all classical bipartite causal channels, characterized by the no-signaling conditions (16) and (17).

C. Channels and positive maps

Channels in a GPT describe transformations of the systems allowed in the theory and are represented by affine maps between state spaces. Although all affine maps between simplices are classical channels, we do not assume in general that all elements in $\mathcal{A}(K,K')$ for state spaces K and K' correspond to valid channels. For the spaces of quantum states and channels, it is required that the maps have completely positive extensions. Completely positive maps $B(\mathcal{H}_{A'A})\to B(\mathcal{H}_{B'B})$ that map $\mathcal{C}_{AA'}$ into $\mathcal{C}_{BB'}$ are called quantum supermaps [33] or quantum combs [34] and belong to a hierarchy describing quantum networks.

Any $T\in\mathcal{A}(K,V(K'))$ extends uniquely to a linear map $T: V(K)\to V(K')$ and $\mathcal{A}(K,V(K'))$ has the structure of a real vector space. The subset $\mathcal{A}(K,V(K')^+)\subseteq\mathcal{A}(K,V(K'))$ is a closed convex cone of elements whose extensions are positive maps. With this cone, $\mathcal{A}(K,V(K'))$ is an ordered vector space.

Let $T_A\in\mathcal{A}(K_A,V(K'_A)^+)$ and $T_B\in\mathcal{A}(K_B,V(K'_B)^+)$. It is easy to see that $T_A\otimes T_B$ is positive with respect to both the maximal and minimal tensor product cones, that is,

$$T_A\otimes T_B\in\mathcal{A}(K_A\otimes_{\min}K_B,V(K'_A\otimes_{\min}K'_B)^+)$$

and

$$T_A\otimes T_B\in\mathcal{A}(K_A\otimes_{\max}K_B,V(K'_A\otimes_{\max}K'_B)^+).$$

Indeed, the first inclusion follows from the definition of the minimal tensor product and the second one from

$$\langle (T_A\otimes T_B)(y), f'_A\otimes f'_B\rangle = \langle y, T_A^*(f'_A)\otimes(T_B)^*(f'_B)\rangle,$$

for all $y\in K_A\otimes_{\max}K_B$, $f'_A\in A(K'_A)^+$, and $f'_B\in A(K'_B)^+$; here T^* denotes the adjoint of the linear extension of T. We say that T_A is entanglement breaking (ETB) if for any state space K_B, we have $T_A\otimes id_{K_B}\in\mathcal{A}(K_A\otimes_{\max}K_B,V(K'_A\otimes_{\min}K_B)^+)$. The set of all ETB maps will be denoted by $\mathcal{A}_{\text{sep}}(K_A,V(K'_A)^+)$; it is a closed convex subcone in $\mathcal{A}(K_A,V(K'_A)^+)$.

There is a well known relation between linear maps and tensor products of vector spaces, with respect to which the positive maps correspond to elements of the maximal tensor product and ETB maps to elements of the minimal one. Details on these relations, as well as the proofs of the following results, are given in Appendix A.

Proposition 1. Let $T\in\mathcal{A}(K,V(K')^+)$. Then T is ETB if and only if T factorizes through a simplex: there are a simplex Δ_n and maps $T_0\in\mathcal{A}(K,V(\Delta_n)^+)$ and $T_1\in\mathcal{A}(\Delta_n,V(K')^+)$ such that $T=T_1T_0$. If T is a channel, T_0 and T_1 may be chosen to be channels as well.

It is clear that any constant map factorizes through the 1-dimensional simplex Δ_0 and hence must be ETB.

We now look at the dual spaces and cones. For $T\in\mathcal{A}(K,V(K))$, let $\operatorname{Tr}T$ denote the usual trace of its linear extension. It is not difficult to see that the dual space of $\mathcal{A}(K,V(K'))$ can be identified with $\mathcal{A}(K',V(K))$, with duality $\langle S,T\rangle=\operatorname{Tr}ST$.

Proposition 2. The dual cone to $\mathcal{A}(K,V(K')^+)$ is $\mathcal{A}_{\text{sep}}(K',V(K)^+)$.

III. POLYSIMPLICES AND THEIR STRUCTURE

Let $k,l_0,\dots,l_k\in\mathbb{N}$. A polysimplex is a Cartesian product of simplices

$$\mathsf{S}_{l_0,\dots,l_k} := \Delta_{l_0}\times\cdots\times\Delta_{l_k},$$

with pointwise defined convex structure. This is a compact convex set, more precisely a convex polytope. Elements of $\mathsf{S}_{l_0,\dots,l_k}$ represent states of a device determined by a set of inputs indexed by $0,\dots,k$, each of which has an allowed set of outputs $0,\dots,l_i$. Such devices were introduced by Popescu and Rohrlich [35] as toy theories exhibiting superquantum correlations. In the framework of GPTs, theories with state spaces of this form were studied in [6,27].

If $l_0=\cdots=l_k=n$, the polysimplex will be denoted by Δ_n^{k+1}. The $(k+1)$-hypercube Δ_1^{k+1} will be denoted by $\Box_{k+1}$. If $l_0,\dots,l_k\in\mathbb{N}$ are assumed fixed, we often drop the multi-index $l_0,\dots,l_k$ and denote the polysimplex by S.

Let $f^0,\dots,f^k$ be a collection of measurements on K, such that $f^i\in\mathcal{A}(K,\Delta_{l_i})$. By definition of the Cartesian product, such collections correspond precisely to elements of $\mathcal{A}(K,\mathsf{S}_{l_0,\dots,l_k})$. Explicitly, the relations between $f^i\in\mathcal{A}(K,\Delta_{l_i})$, $i=0,\dots,k$, and $F=(f^0,\dots,f^k)\in\mathcal{A}(K,\mathsf{S}_{l_0,\dots,l_k})$ are given by

$$F(x)=(f^0(x),\dots,f^k(x)),\quad f^i=\mathsf{m}^iF,\ \forall i,\tag{2}$$

where $\mathsf{m}^i:\mathsf{S}_{l_0,\dots,l_k}\to\Delta_{l_i}$ is the projection onto the ith component. Since our main results are based on the relation of properties of such maps to incompatibility, it will be necessary to describe the structure of the polysimplices and related spaces and cones.

The vertices of $\mathsf{S}=\mathsf{S}_{l_0,\dots,l_k}$ are the $(k+1)$-tuples

$$\mathsf{s}_{n_0,\dots,n_k} := \left(\delta^0_{n_0},\dots,\delta^k_{n_k}\right),\quad n_i=0,\dots,l_i,\quad i=0,\dots,k,$$

where $\delta^i_{n_i}$ denotes the n_ith vertex of the ith simplex. If $u_i\in\Delta_{l_i}$ is the uniform probability distribution for all i, then

$$\bar{\mathsf{s}} := (u_0,\dots,u_k) = \frac{1}{\Pi_i(l_i+1)}\sum_{n_0,\dots,n_k}\mathsf{s}_{n_0,\dots,n_k}$$

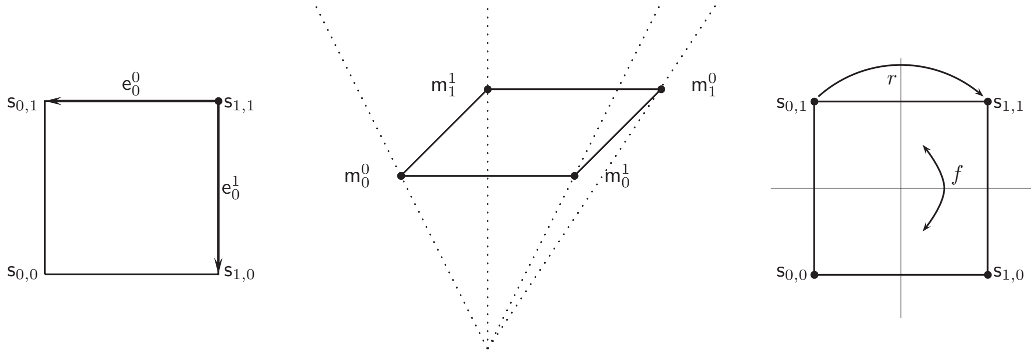

FIG. 1. The square $\square_2$ with basis elements $\{s_{1,1}, e_0^0, e_0^1\}$ and the dual cone $A(\square_2)^+$ generated by the effects m_j^i. The last figure shows the two generators of the group of automorphisms of the square: the clockwise rotation r and the horizontal flip f. These generate the dihedral group $D4$ of order 8.

is the barycenter of S. Further, note that each projection m^i is a measurement on S, with effects determined by

$$m_j^i(s_{n_0,\dots,n_k}) = \begin{cases} 1, & \text{if } n_i = j, \\ 0, & \text{otherwise.} \end{cases}$$

Since all faces of S have the form $F_0 \times \cdots \times F_k$, where F_i is a face of Δ_{l_i}, it is clear that the maximal faces are precisely the null spaces of m_j^i.

It will be convenient to fix a pair of dual bases of the spaces $A(S)$ and $V(S)$, such that the basis of $A(S)$ consists of the effects 1_S and m_j^i. Since $\sum_j m_j^i = 1_S$ for all i, we will fix a linearly independent subset. For the dual basis, we need to describe the edges of S. Since the edges are 1-dimensional faces, they have the form

$$\{\delta_{n_0}^0\} \times \cdots \times \{\delta_{n_{i-1}}^{i-1}\} \times E_i \times \{\delta_{n_{i+1}}^{i+1}\} \times \cdots \times \{\delta_{n_k}^k\},$$

where E_i is an edge of Δ_{l_i}. We see that the vertices adjacent to a vertex $s_{n_0,\dots,n_k}$ are those that differ from $s_{n_0,\dots,n_k}$ in exactly one index, that is, the vertices

$$s_{n_0,\dots,n_{i-1},j,n_{i+1},\dots,n_k}, \quad j \neq n_i, i = 0,\dots,k.$$

Pick the vertex $s_{l_0,\dots,l_k}$ and let

$$e_j^i := s_{l_0,\dots,l_{i-1},j,l_{i+1},\dots,l_k} - s_{l_0,\dots,l_k}$$

denote the vectors given by the adjacent edges.

Lemma 1.

(i) The extreme rays of the cone $A(S)^+$ are generated by the effects m_j^i, $i = 0,\dots,k$, $j = 0,\dots,l_i$.

(ii) The effects

$$1_S, m_0^0, \dots, m_{l_0-1}^0, m_0^1, \dots, m_{l_1-1}^1, \dots, m_0^k, \dots, m_{l_k-1}^k \qquad (3)$$

form a basis of the vector space $A(S)$.

(iii) The elements

$$s_{l_0,\dots,l_k}, e_0^0, \dots, e_{l_0-1}^0, e_0^1, \dots, e_{l_1-1}^1, \dots, e_0^k, \dots, e_{l_k-1}^k \qquad (4)$$

form a basis of the vector space $V(S)$, dual to (3).

Proof. Since the null spaces $(m_j^i)^{-1}(0)$ are exactly the maximal faces of S, these effects generate the extreme rays of $A(S)^+$. Further, let $f_1, f_2, \dots$ denote the elements in (3) and $x_1, x_2, \dots$ the elements of (4). It is easy to see that $\langle f_i, x_j\rangle = \delta_{ij}$, so that both sets are linearly independent. The statements (ii) and (iii) now follow from the fact that $\dim(A(S)) = \sum_{i=0}^k l_i + 1$. ■

The basis elements (4) are visualized in Figs. 1 and 2. With respect to this basis, the vertices are expressed as

$$s_{n_0,\dots,n_k} = s_{l_0,\dots,l_k} + \sum_{i=0}^k e_{n_i}^i. \qquad (5)$$

Remark 1. We can get another pair of dual bases using any vertex $s_{n_0,\dots,n_k}$ and its adjacent edges for a basis of $V(S)$ and

$$\{1_S\} \cup \{m_j^i,\ j \neq n_i, i = 0,\dots,k\}$$

for the dual basis of $A(S)$.

Example 9 (the square). The simplest example is the square $\square_2 = \Delta_1 \times \Delta_1$ (Fig. 1). The vertices $s_{0,0}, s_{0,1}, s_{1,0}, s_{1,1}$ satisfy the relation

$$s_{0,0} + s_{1,1} = s_{0,1} + s_{1,0}.$$

This state space is also called the gbit or square-bit; see [6,18,35]. The dual cone $A(\square_2)^+$ is generated by four effects $m_0^0, m_1^0, m_0^1, m_1^1$. Since we have

$$m_0^0 + m_1^0 = 1_{\square_2} = m_0^1 + m_1^1,$$

these effects again form a square, so that $V(\square_2)^+$ is weakly self-dual. Note that the square is the only polysimplex with this property. [This follows from the fact that the extreme rays of $V(S)^+$ are generated by vertices of S, whereas the extreme rays of $A(S)^+$ correspond to maximal faces of S. Therefore, $V(S)^+ \simeq A(S)^+$ implies $\Pi_{i=0}^k(l_i+1) = \sum_{i=0}^k(l_i+1)$; this holds only for the square.]

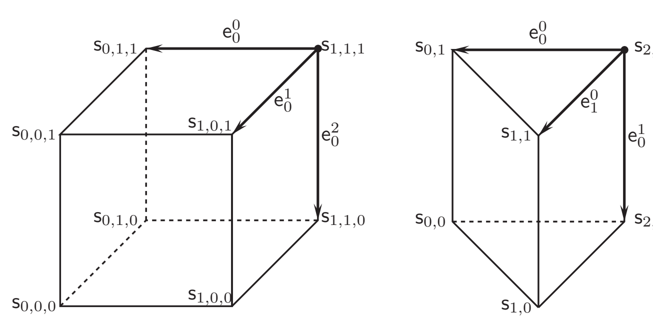

FIG. 2. Basis elements $\{s_{1,1,1}, e_0^0, e_0^1, e_0^2\}$ for the cube $\square_3$, and $\{s_{2,1}, e_0^0, e_1^0, e_0^1\}$ for the polysimplex $S_{2,1}$.

The following relation of polysimplices and spaces of classical channels is immediate (see also Examples 1 and 4).

Proposition 3. We have $\Delta_n^{k+1} \simeq \mathcal{A}(\Delta_k,\Delta_n)$. More generally, any polysimplex $\mathsf{S}_{l_0,\dots,l_k}$ is isomorphic to a face of $\mathcal{A}(\Delta_k,\Delta_n)$, with $n \geqslant \max_i l_i$. This isomorphism is given by

$$s \mapsto T_s \in \mathcal{A}(\Delta_k,\Delta_n), \quad T_s(j|i) = \begin{cases} \mathsf{m}^i_j(s), & \text{if } j \leqslant l_i, \\ 0, & \text{otherwise.} \end{cases}$$

There is also a relation of polysimplices and spaces of quantum channels. To describe this relation, we will need the following notion.

Let K and K' be state spaces. A map $R \in \mathcal{A}(K,K')$ such that there is a map $S \in \mathcal{A}(K',K)$ with $RS = id_{K'}$ is called a retraction. The map $S : K' \to K$ is then called a section. For any retraction-section pair (R,S), the map $P = SR$ is a projection on K onto the range of S, that is, an affine idempotent map $K \to K$ such that $P(K) = S(K')$. Moreover, any map in $\mathcal{A}(K',C)$ for a convex set C has an extension to a map in $\mathcal{A}(K,C)$.

Proposition 4. There exists a retraction-section pair $R \in \mathcal{A}(C_{A,A'},\Delta^{d_A}_{d_{A'}-1})$ and $S \in \mathcal{A}(\Delta^{d_A}_{d_{A'}-1},C_{A,A'})$, determined by

$$\mathsf{m}^i_j R(\Phi) = \langle j|\Phi(|i\rangle\langle i|_A)|j\rangle_{A'}, \quad \forall i,j;\Phi \in \mathcal{C}_{A,A'}$$

and

$$S(s) = \sum_{i,j} \langle i| \cdot |i\rangle_A \mathsf{m}^i_j(s)|j\rangle\langle j|_{A'}, \quad s \in \Delta^{d_A}_{d_{A'}-1}.$$

Proof. Since $\mathsf{m}^i_j R(\Phi) \geqslant 0$ for all i,j and $\sum_j \mathsf{m}^i_j R(\Phi) = \operatorname{Tr}\Phi(|i\rangle\langle i|) = 1$ for all i, R is a well defined element in $\mathcal{A}(C_{A,A'},\Delta^{d_A}_{d_{A'}-1})$. For each s, $S(s)$ is a c-c channel and is therefore completely positive. It is quite clear that

$$\mathsf{m}^i_j RS(s) = \mathsf{m}^i_j(s), \quad \forall i,j,$$

so that $RS = id$. ■

Remark 2. The above proposition implies that there is a projection of $\mathcal{C}_{A,A'}$ onto a set of c-c channels and that any map in $\mathcal{A}(\Delta^{d_A}_{d_{A'}-1},C)$ can be extended to a map in $\mathcal{A}(\mathcal{C}_{A,A'},C)$, for any convex set C. The consequences of this fact will become clear later on.

IV. INCOMPATIBILITY OF MEASUREMENTS

Let K be a state space and let $f^i \in \mathcal{A}(K,\Delta_{l_i})$ be a measurement with values in $\{0,\dots,l_i\}$, $i = 0,\dots,k$. We say that $f^0,\dots,f^k$ are compatible if they are the marginals of a single joint measurement with values in $\{0,\dots,l_0\} \times \cdots \times \{0,\dots,l_k\}$. For an exposition of incompatibility in our setting, see [17].

The joint measurement is described by a map $g \in \mathcal{A}(K,\Delta_L)$ with $L := \Pi_i(l_i+1) - 1$ (note that $\Delta_L \simeq \bigotimes_i \Delta_{l_i}$) and can be parametrized as

$$g(x) = \sum_{n_i \in \{0,\dots,l_i\}} g_{n_0,\dots,n_k}(x)\delta_{n_0,\dots,n_k},$$

where $g_{n_0,\dots,n_k} \in E(K)$ and $\delta_{n_0,\dots,n_k}$ is the probability measure concentrated at $(n_0,\dots,n_k)$. We then have

$$f^i_j = \sum_{n_0,\dots,n_{i-1},n_{i+1},\dots,n_k} g_{n_0,\dots,n_{i-1},j,n_{i+1},\dots,n_k} \tag{6}$$

for the effects of f^i. Let $J \in \mathcal{A}(\Delta_L,\mathsf{S})$ be determined by $J(\delta_{n_0,\dots,n_k}) = \mathsf{s}_{n_0,\dots,n_k}$. Then it is easy to see that (6) can be written as $f^i_j = \mathsf{m}^i_j Jg$. In other words, $f^0,\dots,f^k$ is compatible if and only if the corresponding $F \in \mathcal{A}(K,\mathsf{S})$ satisfies

$$F = Jg, \quad \text{for some} \quad g \in \mathcal{A}(K,\Delta_L). \tag{7}$$

The following observation is simple but important.

Theorem 1. The measurements $f^0,\dots,f^k$ are compatible if and only if the corresponding channel F is entanglement breaking.

Proof. If $f^0,\dots,f^k$ are compatible, then F is ETB by (7) and Proposition 1. Conversely, let F be ETB. Proposition 1 implies that there is some simplex Δ_n and channels $g' \in \mathcal{A}(K,\Delta_n)$, $T \in \mathcal{A}(\Delta_n,\mathsf{S})$ such that $F = Tg'$. The channel T corresponds to a collection of measurements $t^i := \mathsf{m}^i T \in \mathcal{A}(\Delta_n,\Delta_{l_i})$. Since all measurements on a simplex are compatible, there is some $h \in \mathcal{A}(\Delta_n,\Delta_L)$ such that $T = Jh$. Putting $g = hg'$ finishes the proof. ■

Remark 3. The above characterization of incompatible measurements as non-ETB channels suggests that these channels should be admissible in the GPT in question, which also means that we need to include the polysimplices into the theory. For quantum theory this might seem strange, since the polysimplices are certainly not quantum state spaces. On the other hand, the retraction-section pairs of Proposition 4 allow us to include maps in $\mathcal{A}(\mathfrak{S}(\mathcal{H}),\mathsf{S})$ into the larger setting of quantum networks. If $F \in \mathcal{A}(\mathfrak{S}(\mathcal{H}),\Delta^{k+1}_n)$ is a collection of quantum measurements and S is the section of Proposition 4, then SF is a map from quantum states into quantum channels. Using the Choi representation, one can see that this map is also completely positive, hence a quantum comb [34]. Moreover, since $RS = id$, SF is ETB iff F is. One should be aware that "entanglement breaking" has a different meaning here than for the usual completely positive maps: F is compatible iff $(SF \otimes id)(\rho)$ is a local bipartite channel for any bipartite state ρ.

A. Incompatibility witnesses

Let $F = (f^0,\dots,f^k) \in \mathcal{A}(K,\mathsf{S})$ be a collection of measurements. By Proposition 2, F is non-ETB if and only if there is some $W \in \mathcal{A}(\mathsf{S},V(K)^+)$ such that $\operatorname{Tr} FW < 0$. Such a W will be called an incompatibility witness. As can be seen from Proposition A1 (ii), this notion has a close relation to entanglement witnesses.

Any $W \in \mathcal{A}(\mathsf{S},V(K))$ is determined by the images of the vertices of S. The elements

$$w_{n_0,\dots,n_k} := W\big(\mathsf{s}_{n_0,\dots,n_k}\big)$$

will be called the vertices of W [although not all of these points must be vertices of the image $W(\mathsf{S})$]. The map W is positive if and only if all its vertices are in $V(K)^+$. The image of the barycenter of S, $\bar{w} := W(\bar{\mathsf{s}})$ will be called the barycenter of W. We say that W is degenerate if $\dim(W(\mathsf{S})) < \dim(\mathsf{S})$. A description of the cones $\mathcal{A}(\mathsf{S},V(K)^+)$ and $\mathcal{A}_{\text{sep}}(\mathsf{S},V(K)^+)$ can be found in Appendix B.

It is clear that an ETB map cannot be an incompatibility witness. As we shall see (Fig. 3 below), not all non-ETB maps are witnesses. For a characterization of witnesses, we will need the following notion. Let $W,\tilde{W} \in \mathcal{A}(\mathsf{S},V(K)^+)$. We say that

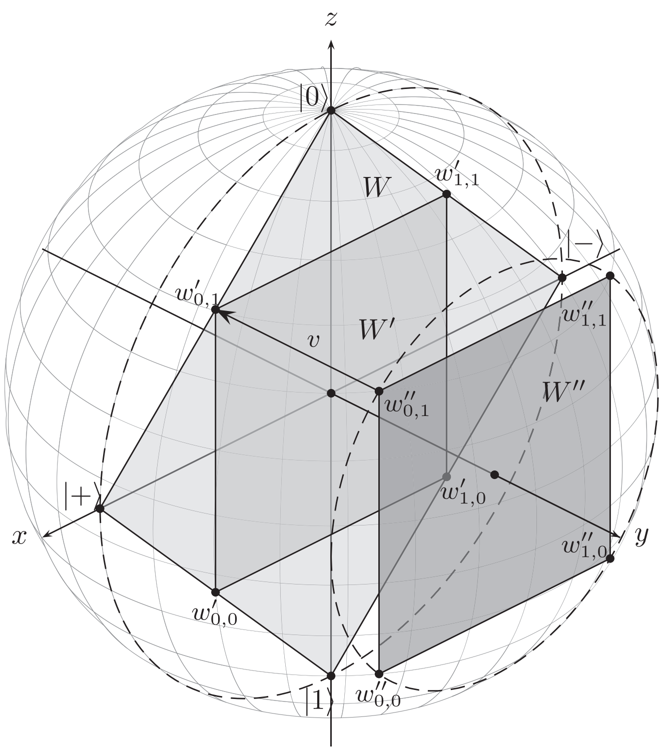

FIG. 3. Incompatibility witnesses for qubit states. Three examples of maps W, W', W'' from the square into the Bloch ball. The vertices of W are the pure states $|0\rangle$, $|1\rangle$, $|+\rangle$, $|-\rangle$ so that W is extremal and not ETB by Corollary 1 (i). It is easy to see that there is no nontrivial translation of W along K; hence W is a witness. The map W' is ETB, since the vertices $w'_{i,j}$ have a decomposition as in Proposition B1, Appendix B (where the elements ψ^i_j are the vertices of W multiplied by $\frac{1}{2}$). The map W'' has extremal vertices $w''_{i,j}$, so that it is again extremal and not ETB by Corollary 1 (i). But W'' is not a witness by Theorem 2, since the ETB map W' is a translation of W'' along K.

$\tilde{W}$ is a translation of W in the direction $v \in V(K)$ if $\tilde{W} = W + L_v$ where L_v is the constant map $L_v(s) \equiv v$. Equivalently, the vertices of $\tilde{W}$ satisfy $\tilde{w}_{n_0,\dots,n_k} = w_{n_0,\dots,n_k} + v$ for all $n_0, \dots, n_k$. If v is such that $\langle 1_K, v\rangle = 0$, we say that $\tilde{W}$ is a translation of W along K.

Theorem 2. A map $W \in \mathcal{A}(\mathsf{S}, V(K)^+)$ is an incompatibility witness if and only if no translation of W along K is ETB.

Proof. The set of channels $\mathcal{A}(K,\mathsf{S})$ is a compact convex subset in the cone $\mathcal{A}(K, V(\mathsf{S})^+)$ of positive maps, generating a proper subcone $\mathcal{V}^+ \subset \mathcal{A}(K, V(\mathsf{S})^+)$. Let us also denote $\mathcal{V} := \mathcal{V}^+ - \mathcal{V}^+$ the vector subspace in $\mathcal{A}(K, V(\mathsf{S}))$ spanned by $\mathcal{A}(K,\mathsf{S})$. In the notation of Sec. II A, we have $\mathcal{V} = V(\mathcal{A}(K,\mathsf{S}))$ and $\mathcal{V}^+ = V(\mathcal{A}(K,\mathsf{S}))^+$. Then

$$\mathcal{V} = \{T \in \mathcal{A}(K, V(\mathsf{S})), 1_{\mathsf{S}} T \in \mathbb{R} 1_K\}$$

and $\mathcal{V}^+ = \mathcal{A}(K, V(\mathsf{S})^+) \cap \mathcal{V}$. Let $\mathcal{V}^\perp$ be the annihilator of $\mathcal{V}$ in the dual space $\mathcal{A}(\mathsf{S}, V(K))$; then it is not difficult to see that

$$\mathcal{V}^\perp = \{L_v,\ v \in V(K), \langle 1_K, v\rangle = 0\}.$$

Since $int(\mathcal{V}^+) \neq \emptyset$ [for example, any constant map of K onto $s \in ri(\mathsf{S})$ is in $int(\mathcal{V}^+)$], Krein's theorem [36] implies that any positive functional on $(\mathcal{V}, \mathcal{V}^+)$ extends to an element in the dual cone $\mathcal{A}(K, V(\mathsf{S})^+)^* = \mathcal{A}_{\text{sep}}(\mathsf{S}, V(K)^+)$. If $W \in \mathcal{A}(\mathsf{S}, V(K)^+)$ is not a witness, then $F \mapsto \operatorname{Tr} FW$ extends to a positive functional on $(\mathcal{V}, \mathcal{V}^+)$, so that there is some $\tilde{W} \in \mathcal{A}_{\text{sep}}(\mathsf{S}, V(K)^+)$ such that

$$\operatorname{Tr} FW = \operatorname{Tr} F\tilde{W}, \quad F \in \mathcal{A}(K,\mathsf{S}).$$

Hence $W - \tilde{W} \in \mathcal{V}^\perp$, so that $\tilde{W}$ is an ETB translation of W along K. Conversely, assume that W is a witness. Let $F \in \mathcal{A}(K,\mathsf{S})$ be such that $\operatorname{Tr} FW < 0$; then for any translation $\tilde{W}$ of W along K, we have $\operatorname{Tr} F\tilde{W} = \operatorname{Tr} FW < 0$. It follows that $\tilde{W}$ is a witness as well and cannot be ETB. ■

We will find another characterization of incompatibility witnesses for two-outcome measurements later (Corollary 2).

For detection of incompatibility, it suffices to use witnesses that are extremal in the cone $\mathcal{A}(\mathsf{S}, V(K)^+)$. More precisely, for an ordered vector space (V, V^+), we say that and element $v \in V^+$ is extremal if it is nonzero and lies on an extreme ray of V^+. Alternatively, v is extremal if $v \neq 0$ and $v' \leqslant v$ for any $v' \in V^+$ implies that $v' = tv$ for some $t \geqslant 0$. A description of extremal witnesses will be also useful in the next section.

So far, we can characterize extremal elements in $\mathcal{A}(\mathsf{S}, V(K)^+)$ in the simplest case when $\mathsf{S} = \square_2$. Such witnesses detect incompatibility of pairs of two-outcome measurements. Any $W \in \mathcal{A}(\square_2, V(K)^+)$ is given by four vertices $w_{i,j} \in V(K)^+$, $i,j = 0,1$, satisfying

$$w_{0,0} + w_{1,1} = w_{0,1} + w_{1,0} = 2\bar{w}. \tag{8}$$

The proof of the following results can be found in Appendix C.

Proposition 5. Let $W \in \mathcal{A}(\square_2, V(K)^+)$ have vertices $w_{i,j}$, $i,j = 0,1$, and barycenter $\bar{w} \neq 0$. Let $F_{i,j}$ denote the face of $V(K)^+$ generated by $w_{i,j}$ and let $L_{i,j} = F_{i,j} - F_{i,j}$ be the generated subspace. Then W is extremal if and only if

$$L_{0,0} \cap L_{1,1} = L_{0,1} \cap L_{1,0} = \{0\}, \tag{9}$$

$$(L_{0,0} \oplus L_{1,1}) \cap (L_{0,1} \oplus L_{1,0}) = \mathbb{R}\bar{w}. \tag{10}$$

Corollary 1. Let $W \in \mathcal{A}(\square_2, V(K)^+)$.

(i) Assume that W is nondegenerate and each vertex is extremal in $V(K)^+$. Then W is non-ETB and extremal in $\mathcal{A}(\square_2, V(K)^+)$.

(ii) Let $\dim(K) = 2$ and assume that W is non-ETB. Then W is extremal if and only if all its vertices are extremal in $V(K)^+$.

Example 10 (the square). Since $\dim(\square_2) = 2$, all non-ETB extremal maps must have extremal vertices, by Corollary 1 (ii). Therefore, $w_{0,0}$ and $w_{1,1}$ must be some multiples of opposite vertices; similarly $w_{0,1}$ and $w_{1,0}$ must be multiples of the other pair of opposite vertices. Applying effects m^i_j to the equality (8), we see that all coefficients must be the same. It follows that W is (a multiple of) an automorphism of $\square_2$. Hence there are 8 extremal rays in $\mathcal{A}(\square_2, V(\square_2)^+)$ that are non-ETB, corresponding to the elements of the dihedral group $D4$ (Fig. 1). It is easy to see that elements in these rays are witnesses, since they have no nontrivial translations along $\square_2$.

The next example shows that all extremal witnesses for quantum state spaces have vertices of rank one. Note that this implies that all vertices are supported by a two-dimensional subspace, and hence to detect incompatibility of pairs of quantum two-outcome measurements, it is enough to look at all their restrictions to such subspaces.

Example 11 (quantum state spaces). Let $W \in \mathcal{A}(\square_2, B(\mathcal{H})^+)$ be extremal and let $\rho = \bar{w}$ be the barycenter. Let $P = \mathrm{supp}(\rho)$ be the support projection of ρ. Let $0 \leqslant E_{i,j} \leqslant P$ be effects such that $\frac{1}{2}w_{i,j} = \rho^{1/2}E_{i,j}\rho^{1/2}$. We will show that all $E_{i,j}$ are projections. Indeed, let M be an effect majorized by both $E_{0,0}$ and $I - E_{0,0}$. Then $\sigma := \rho^{1/2}M\rho^{1/2} \leqslant w_{0,0}, w_{1,1}$, so that $\sigma \in L_{0,0} \cap L_{1,1} = \{0\}$, where we used the notations of Proposition 5. Since $M \leqslant P$, it follows that $M = 0$, so that $E_{0,0} \wedge (I - E_{0,0}) = 0$ and this implies that $E_{0,0}$ is a projection. Then $E_{1,1} = P - E_{0,0}$ is a projection as well, orthogonal to $E_{0,0}$. Similarly for $E_{0,1}$ and $E_{1,0}$.

Let $A \in B_h(P\mathcal{H})$ be such that A commutes with both $E_{0,0}$ and $E_{0,1}$. Then $\rho^{1/2}AE_{i,j}\rho^{1/2} \in L_{i,j}$ and $\rho^{1/2}A\rho^{1/2} \in (L_{0,0} + L_{1,1}) \cap (L_{0,1} + L_{1,0})$. By (10), this implies $\rho^{1/2}A\rho^{1/2} = t\rho$ for some $t \in \mathbb{R}$, so that $A = tP$. Hence any element commuting with both $E_{0,0}$ and $E_{1,1}$ must be a multiple of P. But the commutant of two projections is always nontrivial, unless one of the following two cases occurs:

(a) P is rank one and all $E_{i,j}$ are either P or 0. Then ρ is rank one and $W = \mathsf{m}^i_j(\cdot)\rho$ for some $i, j \in \{0,1\}$; these are precisely the ETB extremal maps.

(b) P is rank 2 and $E_{0,0}$ and $E_{0,1}$ are rank one noncommuting projections. In that case, ρ is rank 2 and all $w_{i,j}$ are rank 1 operators.

It follows that a non-ETB map $W \in \mathcal{A}(\square_2, B(\mathcal{H})^+)$ is extremal if and only if all its vertices are extremal in $B(\mathcal{H})^+$.

B. Incompatibility degree

The incompatibility degree of a collection of measurements can be defined as the least amount of noise that has to be added to obtain a compatible collection. Following [17], the noise will have the form of coin-toss measurements; see [10,18,37] for related definitions.

For $p \in \Delta_l$, a coin-toss measurement is defined as a constant map $f_p(x) \equiv p$. It is immediate that f_p is compatible with any $g \in \mathcal{A}(K, \Delta_l)$ [since the map (g, f_p) factorizes through $\Delta_l \times \Delta_0 \simeq \Delta_l$]. Let us again fix $l_0, \dots, l_k \in \mathbb{N}$ and let $p^i \in \Delta_{l_i}$, $0 = 1, \dots, k$. Then the channel given by the collection of coin tosses $(f_{p^0}, \dots, f_{p^k})$ is the constant map $F_s(x) \equiv s := (p^1, \dots, p^k) \in \mathsf{S}$.

For $F \in \mathcal{A}(K, \mathsf{S})$ and $s \in \mathsf{S}$, we define the incompatibility degree as

$$I_s(F) := \min\{\lambda \in [0,1],\ (1-\lambda)F + \lambda F_s \text{ is ETB}\}.$$

We also define

$$I(F) := \inf_{s \in \mathsf{S}} I_s(F).$$

It is proved in Appendix D (Lemma D1) that $I(F)$ is attained at an interior point of S.

The next result shows that the incompatibility degree can be obtained using incompatibility witnesses. Note that the minimum $q_s(F)$ below is attained at an extremal element in $\mathcal{A}(\mathsf{S}, V(K)^+)$.

For pairs of two-outcome measurements, the following expression for incompatibility degree is related to the dual linear program of [19,38]. In the quantum case, similar results using semidefinite programming were obtained in [10].

Proposition 6. Let $s \in ri(\mathsf{S})$. Let us denote

$$\mathcal{W}_s := \{W \in \mathcal{A}(\mathsf{S}, V(K)^+),\ W(s) \in K\},$$

and for $F \in \mathcal{A}(K, \mathsf{S})$,

$$q_s(F) := \min_{W \in \mathcal{W}_s} \mathrm{Tr}\, FW.$$

Then

$$I_s(F) = \begin{cases} 0, & \text{if } q_s(F) > 0, \\ \frac{-q_s(F)}{1-q_s(F)}, & \text{otherwise.} \end{cases}$$

Proof. Note that $F_s = 1_K(\cdot)s$ is an interior element in the cone $\mathcal{A}_{\mathrm{sep}}(K, V(\mathsf{S})^+)$, hence an order unit, and we have $\mathrm{Tr}\, F_s W = \langle 1_K, W(s)\rangle$. Clearly, if $q_s(F) > 0$ then F is compatible. Otherwise, by Lemma D2, $-q_s(F)$ is the smallest $t \geqslant 0$ such that $F + tF_s$ is ETB, so that $I_s(F) = \frac{-q_s(F)}{1-q_s(F)}$. ■

It is an important question what is the largest value of the incompatibility degree that can be attained by collections of measurements on a given state space. By Proposition 6, we see that this value can be obtained by minimizing $\mathrm{Tr}\, FW$ over all F and W or, in other words, under some normalization, by minimizing the trace of maps in $\mathcal{A}(K, V(K)^+)$ that factorize through a polysimplex.

While this is not an easy task in general, we at least find an expression for $\min_F \mathrm{Tr}\, FW$ for a fixed map W in the case when S is a hypercube.

We first use the pair of dual bases in (3) and (4) to find a suitable expression for $\mathrm{Tr}\, FW$. To shorten the notations, put $w^i_j := w_{l_0,\dots,l_{i-1},j,l_{i+1},\dots,l_k}$; note that $w^i_{l_i} = w_{l_0,\dots,l_k}$ for all i. Then

$$W(\mathsf{e}^i_j) = w^i_j - w^i_{l_i}, \quad j = 0, \dots, l_i,\ i = 0 \dots, k.$$

We have

$$\mathrm{Tr}\, FW = \langle 1_{\mathsf{S}}, FW(\mathsf{s}_{l_0,\dots,l_k})\rangle + \sum_{i=0}^{k}\sum_{j=0}^{l_i-1}\langle \mathsf{m}^i_j, FW(\mathsf{e}^i_j)\rangle \quad (11)$$

$$= \langle 1_K, w_{l_1,\dots,l_k}\rangle + \sum_{i=0}^{k}\sum_{j=0}^{l_i-1}\langle f^i_j, w^i_j - w^i_{l_i}\rangle$$

$$= \sum_{i=0}^{k}\sum_{j=0}^{l_i}\langle f^i_j, w^i_j\rangle - k\langle 1_K, w_{l_1,\dots,l_k}\rangle. \quad (12)$$

We now turn to the case $\mathsf{S} = \square_{k+1}$. The proof of the next result can be found in Appendix D. Recall that $\mathsf{e}^0_0, \dots, \mathsf{e}^k_0$ are edges adjacent to the vertex $\mathsf{s}_{1,\dots,1}$.

Proposition 7. Let $W \in \mathcal{A}(\square_{k+1}, V(K)^+)$. Then

$$\min_{F \in \mathcal{A}(K, \square_{k+1})} \mathrm{Tr}\, FW = \langle 1_K, \bar{w}\rangle - \frac{1}{2}\sum_{i=0}^{k}\|W(\mathsf{e}^i_0)\|_K.$$

Since W is a witness if and only if $\min_F \mathrm{Tr}\, FW < 0$, we immediately obtain the following characterization of witnesses for two-outcome measurements.

Corollary 2. Let $W \in \mathcal{A}(\square_{k+1}, V(K)^+)$. Then W is an incompatibility witness if and only if

$$\sum_{i=0}^{k}\|W(\mathsf{e}^i_0)\|_K > 2\langle 1_K, \bar{w}\rangle.$$

The largest value of $I_{\bar{s}}$ attainable by pairs of two-outcome quantum measurements was obtained in [11] and it was shown (cf. [10]) that it is related to the Tsirelson bound; see also Sec. V D below. We now prove this result by our method and show that I attains the same value.

Corollary 3. For a quantum state space $\mathfrak{S} = \mathfrak{S}(\mathcal{H})$, we have

$$\max_{F\in\mathcal{A}(\mathfrak{S},\Box_2)} I(F) = 1 - \frac{1}{\sqrt{2}}.$$

Proof. Let $F \in \mathcal{A}(\mathfrak{S},\Box_2)$ be incompatible and let $W \in \mathcal{W}_{\bar{s}}$ be such that $\mathrm{Tr}\, FW = q_{\bar{s}}(F)$. We may assume that W is extremal in $\mathcal{A}(\Box_2, B(\mathcal{H})^+)$, so by Example 11, there are some unit vectors $x_{i,j} \in \mathcal{H}$ and a rank 2 density operator ρ with support projection P such that

$$|x_{0,0}\rangle\langle x_{0,0}| + |x_{1,1}\rangle\langle x_{1,1}| = |x_{0,1}\rangle\langle x_{0,1}| + |x_{1,0}\rangle\langle x_{1,0}| = P$$

and the vertices of W satisfy $\frac{1}{2}w_{i,j} = \rho^{1/2}|x_{i,j}\rangle\langle x_{i,j}|\rho^{1/2}$. By Hölder's inequality, we have

$$\begin{aligned}\tfrac{1}{2}\left\|W\left(\mathsf{e}_0^0\right)\right\|_1 &= \|\rho^{1/2}(|x_{0,1}\rangle\langle x_{0,1}| - |x_{1,1}\rangle\langle x_{1,1}|)\rho^{1/2}\|_1\\ &\leqslant \||x_{0,1}\rangle\langle x_{0,1}| - |x_{1,1}\rangle\langle x_{1,1}|\|\\ &= \sqrt{1-|\langle x_{0,1}|x_{1,1}\rangle|^2} =: c.\end{aligned}$$

Similarly,

$$\tfrac{1}{2}\left\|W\left(\mathsf{e}_0^1\right)\right\|_1 \leqslant \sqrt{1-|\langle x_{1,0}|x_{1,1}\rangle|^2} =: d.$$

Since $\langle x_{0,1},x_{1,0}\rangle = 0$, we have $c^2 + d^2 = 1$ and hence $c + d \leqslant \sqrt{2}$. By Proposition 7, it follows that

$$\begin{aligned}q_{\bar{s}}(F) = \mathrm{Tr}\, FW &\geqslant 1 - \tfrac{1}{2}\left[\left\|W\left(\mathsf{e}_0^0\right)\right\|_1 + \left\|W\left(\mathsf{e}_0^1\right)\right\|_1\right]\\ &\geqslant 1-(c+d) \geqslant 1-\sqrt{2}.\end{aligned}$$

On the other hand, let W_0 have vertices $\rho_{i,j} = |x_{i,j}\rangle\langle x_{i,j}|$, with unit vectors $x_{i,j}$ such that $|\langle x_{0,1}|x_{1,1}\rangle| = |\langle x_{1,0}|x_{1,1}\rangle| = \sqrt{2}/2$. In this case, $W(s) \in \mathfrak{S}$ for any $s \in \Box_2$. Let $F_0 = (f^0, f^1)$ be determined by the effects

$$f_0^1 := \mathrm{argmax}_{f\in E(\mathfrak{S})}\langle f, \rho_{1,1} - \rho_{0,1}\rangle,$$

$$f_0^2 := \mathrm{argmax}_{f\in E(\mathfrak{S})}\langle f, \rho_{1,1} - \rho_{1,0}\rangle.$$

Then we have

$$\begin{aligned}1-\sqrt{2} \leqslant q_{\bar{s}}(F_0) &\leqslant \mathrm{Tr}\, F_0 W_0\\ &= -\tfrac{1}{2}(\|\rho_{0,1}-\rho_{1,1}\|_1 + \|\rho_{1,0}-\rho_{1,1}\|_1) + 1\\ &= 1-\sqrt{2}.\end{aligned}$$

It follows that $\max_F I_{\bar{s}}(F) = 1 - \frac{1}{\sqrt{2}}$; this corresponds to the results obtained in [11]. Note further that the witness $W_0 \in \mathcal{W}_s$ for any $s \in ri(\Box_2)$, so that we have

$$q_s(F_0) = \min_{W\in\mathcal{W}_s} \mathrm{Tr}\, F_0 W \leqslant \mathrm{Tr}\, F_0 W_0 = q_{\bar{s}}(F_0).$$

It follows that

$$I_{\bar{s}}(F_0) = \frac{-q_{\bar{s}}(F_0)}{1-q_{\bar{s}}(F_0)} \leqslant \frac{-q_s(F_0)}{1-q_s(F_0)} = I_s(F_0)$$

so that $I_{\bar{s}}(F_0) = I(F_0)$ by Lemma D1. We then have

$$I(F) \leqslant I_{\bar{s}}(F) \leqslant I_{\bar{s}}(F_0) = I(F_0) = 1 - \frac{1}{\sqrt{2}}. \qquad \blacksquare$$

C. Maximally incompatible measurements

Let S be any polysimplex and let $F = (f^0,\dots,f^k) \in \mathcal{A}(K,\mathsf{S})$ be a collection of measurements. It is well known that

$$I_s(F) \leqslant \frac{k}{k+1}$$

for any $s \in \mathsf{S}$ (see, e.g., [17]): the joint measurement for $\frac{1}{k+1}F + \frac{k}{k+1}F_s$ can be defined by choosing one of the measurements uniformly at random and replacing all other measurements by the corresponding coin tosses. If $I(F) = \frac{k}{k+1}$, we say that F is maximally incompatible. We now give a general characterization of maximal incompatibility.

Theorem 3. Let $F \in \mathcal{A}(K,\mathsf{S})$. Then the following are equivalent:

(i) F is maximally incompatible.
(ii) $I_s(F) = \frac{k}{k+1}$ for all $s \in \mathsf{S}$.
(iii) $I_s(F) = \frac{k}{k+1}$ for some $s \in ri(\mathsf{S})$.
(iv) There is some $W \in \mathcal{A}(\mathsf{S},K)$ such that $\mathrm{Tr}\, FW = -k$.
(v) There is some $W \in \mathcal{A}(\mathsf{S},K)$ such that

$$\left\langle f_j^i, w_{n_1,\dots,n_{i-1},j,n_{i+1},\dots,n_k}\right\rangle = 0, \quad \forall i,j; \quad n_0,\dots,n_k.$$

Proof. (i) $\Rightarrow$ (ii) follows from the definition of $I(F)$ and the fact that $I_s(F) \leqslant \frac{k}{k+1}$ for all s; (ii) $\Rightarrow$ (iii) is trivial. Assume (iii); then by Proposition 6 there is some $W \in \mathcal{W}_s$ such that $\mathrm{Tr}\, FW = q_s(F) = -k$. For any $n_0,\dots,n_k$, choose the pair of dual bases of $V(\mathsf{S})$ and $A(\mathsf{S})$ by fixing the vertex $\mathsf{s}_{n_0,\dots,n_k}$ as in Remark 1; then exactly as in (12), we obtain

$$\begin{aligned}-k &= \sum_{i=0}^{k}\sum_{j=0}^{l_i}\left\langle f_j^i, w_{n_0,\dots,n_{i-1},j,n_{i+1},\dots,n_k}\right\rangle - k\left\langle 1_K, w_{n_0,\dots,n_k}\right\rangle\\ &\geqslant -k\left\langle 1_K, w_{n_0,\dots,n_k}\right\rangle.\end{aligned}$$

This implies that $\langle 1_K, w_{n_0,\dots,n_k}\rangle \geqslant 1$ for all $n_0,\dots,n_k$. On the other hand, since $W(s) \in K$ is a convex combination of all $w_{n_0,\dots,n_k}$ with nonzero coefficients, we must have $\langle 1_K, w_{n_0,\dots,n_k}\rangle = 1$; hence (iv) holds. Further, if W is as in (iv), the inequality in the above computation must be an equality, so that $\langle f_j^i, w_{n_0,\dots,n_{i-1},j,n_{i+1},\dots,n_k}\rangle = 0$ for all i and j; hence (v) holds.

Assume (v); then $W(s) \in K$ for any $s \in \mathsf{S}$ and $\mathrm{Tr}\, FW = -k$ by (12). It follows that $I_s(F)$ is maximal for all $s \in ri(\mathsf{S})$. By Lemma D1, this implies (i). $\blacksquare$

The map W in (iv) or, equivalently, (v) of the above theorem will be called a witness of maximal incompatibility for F.

Maximal incompatibility has a nice geometric interpretation for two-outcome measurements. For $k = 1$ the following results were essentially proved in [19]. Recall the definition of retraction-section pairs in Sec. III.

Corollary 4. Let $F \in \mathcal{A}(K,\Box_{k+1})$. Then F is maximally incompatible if and only if F is a retraction.

Proof. Assume that F is a retraction and let $S : \mathsf{S} \to K$ be the corresponding section. Let U be the automorphism of S given as $U(\mathsf{s}_{n_0,\dots,n_k}) = \mathsf{s}_{1-n_0,\dots,1-n_k}$ and put $W = SU$. Then $W \in \mathcal{A}(\mathsf{S},K)$ and we have

$$\mathrm{Tr}\, FW = \mathrm{Tr}\, FSU = \mathrm{Tr}\, U = 1 + \sum_{i=0}^{k}\left\langle \mathsf{m}_0^i, U\left(\mathsf{e}_0^i\right)\right\rangle = -k.$$

Hence F is maximally incompatible. Conversely, assume that F is maximally incompatible and let W be a witness of maximal incompatibility for F. Observe that then by Theorem 3 (v)

$$\mathsf{m}^i_j FW = f^i_j W = \mathsf{m}^i_{1-j};$$

it follows that $FW = U$. Putting $S = WU$ we obtain $FS = U^2 = id_{\mathsf{S}}$, so that F is a retraction. ■

Corollary 5. There exist $k+1$ maximally incompatible two-outcome measurements on K if and only if there exists a projection $K \to K$ whose range is affinely isomorphic to $\Box_{k+1}$.

Proof. If F is maximally incompatible, then by Corollary 4 there is a section $S \in \mathcal{A}(\Box_{k+1}, K)$ such that $FS = id$. It follows that $P := SF$ is a projection $K \to K$ such that $PS = S$ and $FP = F$; hence the restriction of F to the range of P is an isomorphism onto $\Box_{k+1}$ whose inverse is S.

Conversely, assume that $P : K \to K$ is such a projection and let $U : P(K) \to \Box_{k+1}$ be the isomorphism onto the cube. Then $F = UP$ is obviously a retraction, since then $FU^{-1} = UPU^{-1} = UU^{-1} = id_{\Box_{k+1}}$. ■

Example 12 (maximal incompatibility in S). We will show that any collection of effects $\mathsf{m}^0_{n_0}, \ldots, \mathsf{m}^k_{n_k}$ with $n_i \in \{0, \ldots, l_i\}$ is maximally incompatible. Indeed, let $F \in \mathcal{A}(\mathsf{S}, \Box_{k+1})$ be the corresponding channel and let $S \in \mathcal{A}(\Box_{k+1}, \mathsf{S})$ be determined by the collection of measurements $t^0, \ldots, t^k$, given by

$$t^i(s) = \mathsf{m}^i_0(s)\delta^i_{n_i} + [1 - \mathsf{m}^i_0(s)]\delta^i_{n'_i}, \quad s \in \Box_{k+1},$$

for some $n'_i \neq n_i$. Then $FS = id_{\Box_{k+1}}$, so that F is a retraction. Note that this also implies that the projections $\mathsf{m}^0, \ldots, \mathsf{m}^k$ of S are maximally incompatible as well, since they determine the identity map id_{S} and if W is a witness of maximal incompatibility for F, then $\mathrm{Tr}\, id_{\mathsf{S}} WF = \mathrm{Tr}\, FW = -k$, so that $WF \in \mathcal{A}(\mathsf{S}, \mathsf{S})$ is a witness of maximal incompatibility for id_{S}.

We next show that maximally incompatible measurements exist in the space of quantum channels; cf. [19,20]. This result is a simple consequence of Proposition 4 and Example 12; see also Remark 2.

Example 13 (maximal incompatibility in $\mathcal{C}_{A,A'}$). Let $R \in \mathcal{A}(\mathcal{C}_{A,A'}, \Delta^{d_A}_{d_{A'}-1})$ be the retraction as in Proposition 4 and let $F \in \mathcal{A}(\Delta^{d_A}_{d_{A'}-1}, \Box_{d_A})$ be as in Example 12. Then F is a retraction, so that $FR \in \mathcal{A}(\mathcal{C}_{A,A'}, \Box_{d_A})$ is a retraction as well. Hence FR is a maximally incompatible collection of d_A two-outcome measurements on $\mathcal{C}_{A,A'}$. Similarly to the previous example, R is itself maximally incompatible: if W' is a witness of maximal incompatibility for FR, then $W'F$ is such a witness for R.

Remark 4. By the same reasoning as in the above examples, any retraction $K \to \mathsf{S}$ is maximally incompatible. On the other hand, let $F' = (f^0, \ldots, f^k)$ be such that the collection G of two-outcome measurements determined by the effects $f^0_{n_0}, \ldots, f^k_{n_k}$ is maximally incompatible. Then F' is maximally incompatible. Indeed, with the notation of Example 12, $G = FF'$. Let W' be the witness of maximal incompatibility for G; then $W'F$ is such a witness for F'. Since the other effects are not involved, F' is not necessarily a retraction if S is not a hypercube.

V. STEERING AND NONLOCALITY

Quantum steering refers to the property of entangled quantum states which allows one to "steer" the state of one component by choosing suitable measurements on the other [39]. A rigorous operational definition was given in [14] and can be easily rephrased in the setting of GPTs.

A. Steering in GPTs

Let K_A, K_B be state spaces and let $y \in K_A \widetilde{\otimes} K_B$ be a joint state. If a measurement $f_A \in \mathcal{A}(K_A, \Delta_n)$ is applied on system A, then y is mapped onto some element $(f_A \otimes id_B)(y) \in \Delta_n \otimes K_B$. This means that there are some states $x_j \in K_B$ and a probability measure $p \in \Delta_n$ such that $(f_A \otimes id_B)(y) = \sum_{j=0}^n p(j)\delta_j \otimes x_j$. This has the interpretation that with probability $p(j)$, the outcome j is observed on A and the state of B turns into x_j. The collection $\{p(j), x_j\}$ of states and probabilities is called an ensemble. If B has no information about the outcome, the state of B is just the average state $\sum_j p(j)x_j = (1_A \otimes id_B)(y) =: y_B$.

Assume now that an observer on the system A can choose from a collection of measurements $f^i_A \in \mathcal{A}(K_A, \Delta_{l_i})$. Then we obtain a set of ensembles $\{p(j|i), x_{j|i}\}$ with a common average state y_B. Such a set is called an assemblage. According to [14], an assemblage does not demonstrate steering if there is a (finite) set Λ of "classical messages" distributed according to a probability measure q, corresponding elements $\{x_\lambda \in K_B, \lambda \in \Lambda\}$, and conditional probabilities $q(j|i,\lambda)$ such that

$$p(j|i)x_{j|i} = \sum_\lambda q(\lambda)q(j|i,\lambda)x_\lambda. \tag{13}$$

In this case, the assemblage can be explained by a local hidden state model; see [14] for more details. The next result shows that steering can be conveniently expressed in terms of the minimal and maximal tensor products of compact convex sets.

Theorem 4. Let K be a state space and S a polysimplex. Let $\beta \in \mathsf{S} \otimes_{\max} K$.

(i) There is an assemblage $\{p(j|i), x_{j|i}, j = 0, \ldots, l_i, i = 0, \ldots, k\}$ of elements in K with average state $\sum_j p(j|i)x_{j|i} = x \in K$, such that

$$\beta = \mathsf{s}_{l_0,\ldots,l_k} \otimes x + \sum_{i=0}^k \sum_{j=0}^{l_i-1} \mathsf{e}^i_j \otimes p(j|i)x_{j|i}. \tag{14}$$

(ii) The assemblage in (i) does not demonstrate steering if and only if β is separable.

Moreover, any element of the form (14) is in $\mathsf{S} \otimes_{\max} K$.

Proof. Using the basis (4), we have

$$\beta = \mathsf{s}_{l_0,\ldots,l_k} \otimes \phi + \sum_{i=0}^k \sum_{j=0}^{l_i-1} \mathsf{e}^i_j \otimes \phi^i_j$$

for some $\phi, \phi^i_j \in V(K)$. By definition of $\mathsf{S} \otimes_{\max} K$, we must have $\langle \beta, \mathsf{m}^i_j \otimes f \rangle \geqslant 0$ for all i,j and $f \in E(K)$. For $j \neq l_i$ this is true if and only if $\phi^i_j \in V(K)^+$. We also have for all i and $f \in E(K)$

$$\sum_{j=0}^{l_i-1} \langle \phi^i_j, f \rangle \leqslant \sum_{j=0}^{l_i} \langle \beta, \mathsf{m}^i_j \otimes f \rangle = \langle \beta, 1_{\mathsf{S}} \otimes f \rangle = \phi;$$

hence $\sum_j \phi^i_j \leqslant \phi \in V(K)^+$ and $\langle \beta, 1_{\mathsf{S}} \otimes 1_K\rangle = \langle \phi, 1_K\rangle = 1$. Put $\phi^i_{l_i} := \phi - \sum_{j=0}^{l_i-1}\phi^i_j$ and $p(j|i) := \langle 1_K, \phi^i_j\rangle$, $x_{j|i} := p(j|i)^{-1}\phi^i_j$ [if $p(j|i) > 0$, otherwise $x(j|i)$ can be anything] for all i and j. Then $\{p(j|i), x_{j|i}\}$ is an assemblage with average state $x := \phi$. This proves (i).

For (ii), assume that (13) holds. Then $x = \sum_\lambda q(\lambda)x_\lambda$ and we have

$$\beta = \sum_\lambda q(\lambda)s_\lambda \otimes x_\lambda,$$

where $s_\lambda := \mathsf{s}_{l_0,\dots,l_k} + \sum_{i=0}^k \sum_{j=0}^{l_i-1} q(j|i,\lambda)\mathsf{e}^i_j \in \mathsf{S}$; this follows from $\langle \mathsf{m}^i_j, s_\lambda\rangle = q(j|i,\lambda) \geqslant 0$ and $\langle 1_{\mathsf{S}}, s_\lambda\rangle = \langle 1_{\mathsf{S}}, \mathsf{s}_{l_0,\dots,l_k}\rangle = 1$. Hence β is separable. Conversely, let

$$\beta = \sum_{n_0,\dots,n_k} \mathsf{s}_{n_0,\dots,n_k} \otimes \alpha_{n_0,\dots,n_k}$$

for some $\alpha_{n_0,\dots,n_k} \in V(K)^+$. Put $\Lambda := \{(n_0,\dots,n_k), n_i = 0,\dots,l_i\}$, $q(n_0,\dots,n_k) := \langle 1_K, \alpha_{n_0\dots n_k}\rangle$, $x_{n_0,\dots,n_k} := q(n_0,\dots,n_k)^{-1}\alpha_{n_0\dots n_k}$, and

$$q(j|i,(n_0,\dots,n_k)) := \begin{cases} 1, & \text{if } n_i = j,\\ 0, & \text{otherwise.}\end{cases}$$

Then for all i and j,

$$p(j|i)x_{j|i} = \langle \beta, \mathsf{m}^i_j \otimes \cdot\rangle = \sum_{n_0,\dots,n_k, n_i=j} \alpha_{n_0,\dots,n_k}$$
$$= \sum_{\lambda\in\Lambda} q(\lambda)q(j|i,\lambda)x_\lambda.$$

This proves (ii). The last statement follows from the proof of (i). ■

In view of the preceding theorem, any element in $\mathsf{S}\otimes_{\max} K$ will be called an assemblage. The common average state x will be called the barycenter of β.

Let us return to the steering scenario described in the paragraph before Theorem 4. It is already known that there is no steering if y is separable or the measurements f^i_A are compatible. This can be seen immediately from Theorem 4. Let $F_A = (f^0_A,\dots,f^k_A)$, then $\beta = (F_A \otimes id_B)(y) \in \mathsf{S}\otimes_{\max} K_B$ is the associated assemblage. If $y \in K_A \otimes_{\min} K_B$, then β must be separable and hence does not demonstrate steering. If the measurements are compatible, then β is separable by Theorem 1.

In general, not all assemblages in $\mathsf{S}\otimes_{\max} K_B$ are obtained from some collection F_A and a bipartite state $y \in K_A\widetilde{\otimes}K_B$. If this is the case, we say that the state space K_B admits steering. Note that quantum state spaces satisfy this condition. Somewhat stronger conditions were studied in [40] and their relations to homogeneity and weak self-duality of the state spaces were found.

B. Steering degree

Let $\beta \in \mathsf{S}\otimes_{\max} K$ be an assemblage with barycenter x. A steering degree can be defined similarly to incompatibility degree, as the smallest amount of noise that has to be added to β to obtain a separable element. For the noise, we use assemblages of the form $s\otimes x$ for $s \in \mathsf{S}$. This is a separable assemblage with $p(j|i) = \mathsf{m}^i_j(s)$ and $x_{j|i} = x$ for all i,j. We put

$$S_s(\beta) := \min\{\lambda \in [0,1], (1-\lambda)\beta + \lambda s \otimes x \in \mathsf{S}\otimes_{\min} K\}$$

and

$$S(\beta) := \inf_{s\in\mathsf{S}} S_s(\beta).$$

Observe that for any $F_A \in \mathcal{A}(K_A, \mathsf{S})$ and $y \in K_A\widetilde{\otimes}K_B$, we have

$$S_s((F_A\otimes id_B)(y)) \leqslant I_s(F_A). \tag{15}$$

To see this, note that the barycenter of $(F_A\otimes id_B)(y)$ is the marginal y_B and $s\otimes y_B = (F_s \otimes id_B)(y)$. We therefore have

$$(1-\lambda)(F_A\otimes id_B)(y) + \lambda s\otimes y_B$$
$$= \{[(1-\lambda)F_A + \lambda F_s]\otimes id_B\}(y),$$

and this is separable if $(1-\lambda)F_A + \lambda F_s$ is ETB. The possibility of attaining equality depends on the properties of K_A and the form of composite state spaces in the GPT. Assume that $V(K_A)^+$ is weakly self-dual, so that there is an affine isomorphism $\Psi : A(K_A)^+ \to V(K_A)^+$. With the notations of Appendix A, we have $(id\otimes\Psi)(\chi_{K_A}) \in K_A \otimes_{\max} K_A$; see Lemma A1. We are now prepared to state the following result.

Theorem 5. Assume that there is an isomorphism $\Psi : A(K_A)^+ \to V(K_A)^+$ such that $(id\otimes\Psi)(\chi_{K_A}) \in K_A\widetilde{\otimes}K_A$. Then for any polysimplex S, $F_A \in \mathcal{A}(K,\mathsf{S})$ and $s\in\mathsf{S}$, we have

$$\sup_{y\in K_A\widetilde{\otimes}K_A} S_s((F_A\otimes id_A)(y)) = I_s(F_A).$$

Proof. By (15), the supremum on the left is never larger than $I_s(F_A)$. Put $y = (id\otimes\Psi)(\chi_{K_A})$ and let $F_\lambda := (1-\lambda)F_A + \lambda F_s$. Then $(F_\lambda\otimes id)(y)$ is separable if and only if $(F_\lambda \otimes id)(\chi_{K_A})$ is separable, which by Proposition A1 (ii) means that F_λ is ETB. It follows that $S_s((F_A\otimes id_A)(y)) \geqslant I_s(F_A)$. ■

The conditions in the previous theorem are fulfilled in quantum state spaces, where y is a pure maximally entangled state. In this case, this result was proved in ([37], pp. 8-9).

Remark 5. Similarly to incompatibility, we may define steering witnesses and their relation to steering degree, maximal steering degree, etc. The witnesses will now be elements in $A(\mathsf{S}\otimes_{\min} K)^+$. We will not investigate this here, only remark that since the assemblages generate all of the positive cone $V(\mathsf{S}\otimes_{\max} K)^+$, any nonseparable element in $A(\mathsf{S}\otimes_{\min} K)^+$ is a witness.

C. Nonlocality and Bell's inequalities

Let $f^i_A \in \mathcal{A}(K_A, \Delta_{l^A_i})$, $i = 0,\dots,k_A$, and $f^i_B \in \mathcal{A}(K_B, \Delta_{l^B_i})$, $i = 0,\dots,k_B$, and let $y \in K_A\widetilde{\otimes}K_B$. If a measurement $f^{i_A}_A$ is chosen for A and $f^{i_B}_B$ for B, then the result is a pair (j_A, j_B) with probability

$$p(j_A, j_B|i_A, i_B) := \left\langle \left(f^{i_A}_A\right)_{j_A} \otimes \left(f^{i_B}_B\right)_{j_B}, y\right\rangle.$$

These conditional probabilities satisfy the no-signaling properties

$$\sum_{j_A} p(j_A, j_B|i_A, i_B) = p_B(j_B|i_B), \quad \forall i_A, \tag{16}$$

$$\sum_{j_B} p(j_A, j_B|i_A, i_B) = p_A(j_A|i_A), \quad \forall i_B, \tag{17}$$

where $p_A(j_A|i_A) := \langle (f_A^{i_A})_{j_A}, y_A\rangle$, $p_B(j_B|i_B) := \langle (f_B^{i_B})_{j_B}, y_B\rangle$. Following [14], we say that the state y is Bell local if for all measurements f_A^i and f_B^i, these probabilities admit a local hidden variable (LHV) model; that is, there is a probability distribution q on a set Λ and conditional probabilities $q_A(j_A|i_A,\lambda)$ and $q_A(j_B|i_B,\lambda)$ such that

$$p(j_A,j_B|i_A,i_B) = \sum_\lambda q(\lambda)q_A(j_A|i_A,\lambda)q_B(j_B|i_B,\lambda). \quad (18)$$

Let $F_A = (f_A^0,\dots,f_A^{k_A})$ and let S_A be the related polysimplex; similarly define F_B and S_B. Then $\gamma := (F_A \otimes F_B)(y) \in \mathsf{S}_A \otimes_{\max} \mathsf{S}_B$ and

$$(\mathsf{m}_{j_A}^{i_A} \otimes \mathsf{m}_{j_B}^{i_B})(\gamma) = \langle (f_A^{i_A})_{j_A} \otimes (f_B^{i_B})_{j_B}, y\rangle = p(j_A,j_B|i_A,i_B).$$

It can be seen by putting $K_B = \mathsf{S}_B$ in Theorem 4 that the elements $\gamma \in \mathsf{S}_A \otimes_{\max} \mathsf{S}_B$ are characterized by the property that $p(j_A,j_B|i_A,i_B) := (\mathsf{m}_{j_A}^{i_A} \otimes \mathsf{m}_{j_B}^{i_B})(\gamma)$ are no-signaling conditional probabilities and (18) describes precisely the separable elements. The tensor product $\mathsf{S}_A \otimes_{\max} \mathsf{S}_B$ is therefore called the no-signaling polytope and $\mathsf{S}_A \otimes_{\min} \mathsf{S}_B$ the local polytope.

The steering witnesses in this case (see Remark 5) will be called Bell witnesses. These are precisely the elements of $A(\mathsf{S}_A \otimes_{\min} \mathsf{S}_B)^+$ that are not separable. With some normalization, there is a finite number of extremal Bell witnesses $\mu_1,\dots,\mu_N$ that completely determine the local polytope: if $\gamma \in \mathsf{S}_A \otimes_{\max} \mathsf{S}_B$, then γ is local if and only if

$$\langle \mu_i,\gamma\rangle \geqslant 0, \quad i = 1,\dots,N. \quad (19)$$

These are the Bell inequalities. Lemma 5 (iv) shows that, similarly to the case of incompatibility, the Bell witnesses correspond to affine maps of the polysimplex S_A into a positive cone; this time it is the cone $A(\mathsf{S}_B)^+$. All Bell inequalities are given by extremal non-ETB elements in $\mathcal{A}(\mathsf{S}_A, A(\mathsf{S}_B)^+)$.

Example 14 (the CHSH inequality). Assume that there is a pair of two-outcome measurements on both sides, so that $\mathsf{S}_A = \mathsf{S}_B = \Box_2$. Since $V(\Box_2)^+ \simeq A(\Box_2)^+$, we see by Example 10 that all extremal witnesses are precisely (multiples of) the isomorphisms $\Psi_{i,j,k} \in \mathcal{A}(\Box_2, A(\Box_2)^+)$ that map the extreme points of $\Box_2$ to the four effects m_j^i:

$$\mathsf{s}_{0,0} \mapsto \mathsf{m}_j^i,\ \mathsf{s}_{1,1} \mapsto \mathsf{m}_{1-j}^i,\ \mathsf{s}_{0,1} \mapsto \mathsf{m}_k^{1-i},\ \mathsf{s}_{1,0} \mapsto \mathsf{m}_{1-k}^{1-i}.$$

Let $\mu_{i,j,k}$ be the witness corresponding to $\Psi_{i,j,k}$. Using the basis elements (3) and (4), we get

$$\mu_{i,j,k} = \mathsf{m}_{1-j}^i \otimes 1_{\Box_2} + \left(\mathsf{m}_k^{1-i} - \mathsf{m}_{1-j}^i\right) \otimes \mathsf{m}_0^0 + \left(\mathsf{m}_{1-k}^{1-i} - \mathsf{m}_{1-j}^i\right) \otimes \mathsf{m}_0^1.$$

Let $F_A = (f_A^0, f_A^1) \in \mathcal{A}(K_A,\Box_2)$, $F_B = (f_B^0, f_B^1) \in \mathcal{A}(K_B,\Box_2)$ and put

$$a_1 := 1 - 2\left(f_A^1\right)_0, \quad a_2 := 1 - 2\left(f_A^0\right)_0,$$
$$b_1 := 1 - 2\left(f_B^0\right)_0, \quad b2 := 1 - 2\left(f_B^1\right)_0;$$

then we can see that

$$\langle \mu_{0,1,0}, (F_A \otimes F_B)(y)\rangle = \langle (F_A \otimes F_B)^*(\mu_{0,1,0}), y\rangle = \tfrac{1}{2}\left(1 - \tfrac{1}{2}\mathbb{B}\right), \quad (20)$$

where $\mathbb{B} = \langle a_1 \otimes (b_1 + b_2) + a_2 \otimes (b_1 - b_2), y\rangle$, so that (19) becomes the CHSH inequality.

D. Bell inequalities and the incompatibility degree

The maximal value of $\mathbb{B}$ in Example 14 that can be attained by two-outcome measurements is called the CHSH bound. It is well known that the outcome probabilities satisfy the LHV model (18) if and only if $\mathbb{B} \leqslant 2$ and we always have $\mathbb{B} \leqslant 4$. In quantum state spaces, the Tsirelson bound holds: $\mathbb{B} \leqslant 2\sqrt{2}$. It was observed in [10] that the incompatibility degree for pairs of quantum effects is connected to this bound.

The relation of incompatibility degree and CHSH bound in GPTs was proved in [7]. We include a proof in our setting.

Theorem 6. Let $\mathsf{S}_A,\mathsf{S}_B$ be polysimplices. Let $F_A \in \mathcal{A}(K_A,\mathsf{S}_A)$, $F_B \in \mathcal{A}(K_B,\mathsf{S}_B)$, $y \in K_A\widetilde{\otimes}K_B$, and assume that F_A is incompatible. Then for any $\mu \in A(\mathsf{S}_A \otimes_{\min} \mathsf{S}_B)^+$ and $s \in ri(\mathsf{S}_A)$, we have

$$\langle \mu, F_A \otimes F_B(y)\rangle \geqslant \|\mu\|_{\max} q_s(F_A).$$

If K_A admits steering and $\mathsf{S}_A = \Box_2$, then there is some state space K_B, $F_B \in (K_B,\Box_2)$, and $y \in K_A\widetilde{\otimes}K_B$ such that

$$\langle \mu_{i,j,k}, F_A \otimes F_B(y)\rangle = \tfrac{1}{2} q_{\bar{\mathsf{s}}}(F_A)$$

for the witness $\mu_{i,j,k}$ as in Example 14.

Proof. By Lemma A1 (iv), there are some $T \in \mathcal{A}(K_B^\dagger, V(K_A)^+)$ and $M \in \mathcal{A}(\mathsf{S}_A, V(\mathsf{S}_B^\dagger)^+)$ such that $y = (T \otimes id)(\chi_{K_B^\dagger})$ and $\mu = (M^* \otimes id)(\chi_{\mathsf{S}_B})$. Then by using Lemma A1 (iii),

$$\langle \mu, F_A \otimes F_B(y)\rangle = \left\langle (M^* \otimes id)(\chi_{\mathsf{S}_B}), (F_A T \otimes F_B)\left(\chi_{K_B^\dagger}\right)\right\rangle$$
$$= \left\langle \chi_{\mathsf{S}_B}, (M F_A T F_B^* \otimes id)\left(\chi_{\mathsf{S}_B}^\dagger\right)\right\rangle$$
$$= \mathrm{Tr}\ F_A T F_B^* M.$$

Put $W := T F_B^* M$, then $W \in \mathcal{A}(\mathsf{S}_A, V(K_A^+))$. Moreover,

$$\mathrm{Tr}\ F_s W = \left\langle 1_{K_A}, W(s)\right\rangle = \left\langle F_B T^*\left(1_{K_A}\right), M(s)\right\rangle = \langle s', M(s)\rangle,$$

where $s' := F_B T^*(1_{K_A})$. It is easy to see that $T^*(1_{K_A}) \in K_B$, so that $s' \in \mathsf{S}_B$. It follows that

$$t := \mathrm{Tr}\ F_s W = \langle s', M(s)\rangle = \langle M^*(s'), s\rangle = \mu(s \otimes s') \leqslant \|\mu\|_{\max}.$$

We have $t^{-1}W \in \mathcal{W}_s$, so that

$$t^{-1}\langle \mu, F_A \otimes F_B(y)\rangle = t^{-1}\mathrm{Tr}\ F_A W \geqslant q_s(F_A).$$

The statement now follows by the assumption that F_A is incompatible, so that $q_s(F) < 0$.

Assume now that K_A admits steering. Let $W \in \mathcal{A}(\Box_2, V(K_A)^+)$ be a witness in $\mathcal{W}_{\bar{\mathsf{s}}}$ such that $\mathrm{Tr}\ F_A W = q_{\bar{\mathsf{s}}}(F_A)$. In view of the above proof, it is enough to show that $W = T F_B^* M$ for suitable T, F_B, and M. So let $M = \Psi_{i,j,k}$ be the isomorphism as in Example 14. Put $\beta := (W M^{-1} \otimes id)(\chi_{\Box_2^\dagger})$; then clearly $\beta \in V(K_A \otimes_{\max} \Box_2)^+$ and

$$\langle \beta, 1 \otimes 1\rangle = \langle 1_{K_A}, W M^{-1}(1_{\Box_2})\rangle = 2\langle 1_{K_A}, W(\bar{\mathsf{s}})\rangle = 2.$$

Hence $\frac{1}{2}\beta$ is an assemblage. Since K_A admits steering, there is some state space K_B, $F_B \in \mathcal{A}(K_B,\Box_2)$ and $y \in K_A\tilde{\otimes}K_B$ such that

$$\tfrac{1}{2}\beta = (id \otimes F_B)(y) = (T \otimes F_B)\big(\chi_{K_B^\dagger}\big) = (TF_B^* \otimes id)\big(\chi_{\Box_2^\dagger}\big).$$

It follows that $\frac{1}{2}W = TF_B^*M$; this finishes the proof. ∎

Note that the crucial part of the proof of the equality in the above theorem is that $V(\Box_2)^+$ is weakly self-dual, so we may chose M to be an isomorphism. This is not true for any other S_A and S_B. As we have seen in the proof, the Bell scenario provides incompatibility witnesses only of the form $W = TF_B^*M$, that is, factorizing through some $A(\mathsf{S}_B)^+$, which can be weaker for detection of some types of incompatibility. This seems to be the reason why already for three quantum effects, incompatibility in some cases cannot be detected by violation of Bell inequalities. This was observed in [12] in the case of qubit states, but the above arguments suggest that such effect exist in any nonclassical theory in our class of GPTs.

E. Nonlocality in spaces of quantum channels

It was proved in [21] that one can obtain probabilities maximally violating the CHSH inequality, that is, attaining the value $\mathbb{B} = 4$, by using causal bipartite quantum channels. In fact, it was shown recently in [23] that one can obtain all no-signaling probabilities in this way. In these works, the GPT setting was not used, but nevertheless it was shown that any element of the no-signaling polytope can be obtained by applying sets of channel measurements to both parts of an element of $\mathcal{C}^{\text{caus}}_{AB,A'B'}$. Maximal violation of the CHSH inequality in spaces of quantum channels was also proved using GPT in [22]. Note that the channel used in [21,22] was a bipartite c-c channel.

The aim of the present paragraph is to remark that this feature of quantum channels is immediate from Proposition 4. Indeed, the isomorphism $\Delta_n^{k+1} \simeq \mathcal{A}(\Delta_k,\Delta_n)$ also implies that

$$\Delta_{n_A}^{k_A+1} \otimes_{\max} \Delta_{n_B}^{k_B+1} \simeq \mathcal{A}\big(\Delta_{k_A},\Delta_{n_A}\big) \otimes_{\max} \mathcal{A}\big(\Delta_{k_B},\Delta_{n_B}\big).$$

In this way, any no-signaling polytope is isomorphic to a face in the space of causal classical bipartite channels and the local polytope corresponds to the local channels in this face (see also Example 7).

Let now (R_A,S_A), (R_B,S_B) be the retraction-section pairs as in Proposition 4 and let $\gamma \in \Delta_{n_A}^{k_A+1} \otimes_{\max} \Delta_{n_B}^{k_B+1}$ be a collection of no-signaling conditional probabilities. Then γ corresponds to a classical causal channel $T_\gamma : \Delta_{k_{AB}} \to \Delta_{n_{AB}}$ (cf. the notation in Example 5), given by $T_\gamma(j_A,j_B|i_A,i_B) = (\mathsf{m}_{j_A}^{i_A} \otimes \mathsf{m}_{j_B}^{i_B})(\gamma)$. Put $\Phi := \Phi_{T_\gamma}$ as in (1), choosing product bases in $\mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{H}_{A'} \otimes \mathcal{H}_{B'}$. Then $\Phi = (S_A \otimes S_B)(\gamma) \in \mathcal{C}_{A,A'} \otimes_{\max} \mathcal{C}_{B,B'}$, but since Φ is completely positive, we have $\Phi \in \mathcal{C}^{\text{caus}}_{AB,A'B'} = \mathcal{C}_{A,A'}\tilde{\otimes}\mathcal{C}_{B,B'}$. Clearly,

$$(R_A \otimes R_B)(\Phi) = (R_A \otimes R_B)(S_A \otimes S_B)(\gamma) = \gamma.$$

We have proved the following result (cf. [23]).

Theorem 7. For any collection of no-signaling conditional probabilities $\gamma \in \Delta_{n_A}^{k_A+1} \otimes_{\max} \Delta_{n_B}^{k_B+1}$, there is a causal bipartite quantum channel $\Phi \in \mathcal{C}^{\text{caus}}_{AB,A'B'}$, with $d_A = k_A + 1$, $d_{A'} = n_A + 1$ and $d_B = k_B + 1$, $d_{B'} = n_B + 1$, and collections of measurements $R_A \in \mathcal{A}(\mathcal{C}_{A,A'},\Delta_{n_A}^{k_A+1})$ and $R_B \in \mathcal{A}(\mathcal{C}_{B,B'},\Delta_{n_B}^{k_B+1})$ such that $(R_A \otimes R_B)(\Phi) = \gamma$.

Example 15. Let $\mathsf{S}_A = \mathsf{S}_B = \Box_2$ and choose

$$\gamma = \tfrac{1}{2}((\mathsf{s}_{0,0} - \mathsf{s}_{1,0}) \otimes \mathsf{s}_{0,0} + \mathsf{s}_{1,1} \otimes \mathsf{s}_{1,0} + \mathsf{s}_{1,0} \otimes \mathsf{s}_{01}) \in \Box_2 \otimes_{\max} \Box_2.$$

We obtain the same c-c bipartite channel and sets of measurements attaining maximal CHSH violation as in [21,22].

VI. CONCLUSION AND FURTHER QUESTIONS

We have studied incompatibility of measurements in a family of convex finite-dimensional GPTs by representing collections of measurements as affine maps into a polysimplex. We have shown how properties of these maps (like being ETB or a retraction) are tied to incompatibility. We introduced incompatibility witnesses and used them to characterize incompatibility degree. Our results suggest that incompatibility is closely related to the geometry of polysimplices. For example, the largest value of the incompatibility degree $I(F)$ that is attainable by collections of measurements on a given state space K can be obtained by considering positive maps on $(V(K),V(K)^+)$ that factorize through the polysimplex of the corresponding shape.

Our setting allows us to study the relations of incompatibility to steering and nonlocality through incompatibility witnesses. Here the geometry of polysimplices plays a key role as a unifying mathematical tool. We have shown that the Bell scenario provides incompatibility witnesses of a restricted type, more precisely, factorizing through the dual of a polysimplex. In general, the family of such witnesses is strictly smaller than the set of all witnesses and is therefore weaker for the detection of incompatibility. This explains the existence of incompatible collections of measurements that do not violate Bell inequalities and suggests that this feature is not specific for quantum theory, but is common in GPT.

There are a number of questions left for further research. For example, the incompatibility degree attainable by more general collections of quantum measurement can be investigated using witnesses as in Corollary 3. For this, a characterization of extremal maps of a polysimplex into $B(\mathcal{H})^+$ would be useful. For the study of relations between incompatibility and nonlocality, one could describe the witnesses that can be obtained from Bell inequalities as in Theorem 6. It is an interesting question to what extent are these witnesses weaker and how it depends on the theory in question. It might be also worthwhile to study collections of quantum measurements and their incompatibility, as well as steering and Bell nonlocality, within the framework of quantum networks as suggested in Remark 3.

Maps into a polysimplex can be used to describe collections of measurements only up to a fixed size and number of outcomes. The isomorphism with (faces of) classical channels (Proposition 3) suggests that it might be possible do include all collections of measurements as affine maps into the space of more general Markov kernels.

ACKNOWLEDGMENTS

I am grateful to the anonymous referees whose valuable comments helped me to improve the paper. This research was supported by the grants VEGA 2/0069/16 and APVV-16-0073.

APPENDIX A: POSITIVE AND ETB MAPS

We list some well known results on the cones of positive maps. Let K be a compact convex set. Let $x_0,\dots,x_n$ be a basis of $V(K)$ and let $e_0,\dots,e_n\in A(K)$ be the dual basis. Put

$$\chi_K := \sum_i x_i\otimes e_i \in V(K)\otimes A(K). \tag{A1}$$

Let $f\in A(K)$; then $f=\sum_i\langle f,x_i\rangle e_i$, and therefore we have

$$\langle \chi_K, f\otimes y\rangle = \sum_i \langle f,x_i\rangle e_i(y) = f(y), \quad \forall y\in K. \tag{A2}$$

Consequently, χ_K does not depend on the choice of the basis.

Lemma A1. Let $x'\in ri(K)$ and put $K^\dagger := \{f\in A(K)^+, f(x')=1\}$. Then,

(i) $K^\dagger$ is a compact convex set and we have $V(K^\dagger)^+ = A(K)^+$, $A(K^\dagger)^+ = V(K)^+$, and $1_{K^\dagger}=x'$.

(ii) $\chi_K\in K\otimes_{\max} K^\dagger$.

(iii) Let $T\in\mathcal{A}(K,V(K'))$; then $(T\otimes id)(\chi_K) = (id\otimes T^*)(\chi_{K'})$.

(iv) For any $\xi\in V(K'\otimes_{\max}K^\dagger)^+$, there is a unique $T\in\mathcal{A}(K,V(K')^+)$ such that $(T\otimes id)(\chi_K)=\xi$, determined by

$$\langle T(x), f'\rangle = \langle \xi, f'\otimes x\rangle, \quad \forall x\in K, f'\in A(K'). \tag{A3}$$

Proof. Since $x'\in ri(K)$, $f(x')=0$ for $f\in A(K)^+$ implies that $f=0$, so that $K^\dagger$ is a base of $A(K)^+$. This proves (i). The statement (ii) follows immediately from (A2) and (i). For (iii), let $f'\in A(K')$ and $y\in K$; then by (A2), we have

$$\begin{aligned}\langle (T\otimes id)(\chi_K), f'\otimes y\rangle &= \langle \chi_K, T^*(f')\otimes y\rangle = T^*(f')(y)\\ &= \langle f', T(y)\rangle = \langle \chi_{K'}, f'\otimes T(y)\rangle\\ &= \langle (id\otimes T^*)(\chi_{K'}), f'\otimes y\rangle.\end{aligned}$$

For (iv), it is clear that (A3) determines an element $T\in\mathcal{A}(K,V(K')^+)$ and $(T\otimes id)(\chi_K)=\xi$ holds by (A2). ■

We now have the following characterizations of ETB maps.

Proposition A1. Let $T\in\mathcal{A}(K,V(K')^+)$. The following are equivalent:

(i) T is ETB.

(ii) $(T\otimes id)(\chi_K)$ is separable.

(iii) T factorizes through a simplex: there are a simplex Δ_n and maps $T_0\in\mathcal{A}(K,V(\Delta_n)^+)$ and $T_1\in\mathcal{A}(\Delta_n,V(K')^+)$ such that $T=T_1T_0$.

If T is a channel, T_0 and T_1 in (iii) may be chosen to be channels as well.

Proof. (i) $\Rightarrow$ (ii) is clear. Assume (ii); then there are some $\phi_j\in V(K')^+$ and $f_j\in A(K)^+$ such that $(T\otimes id)(\chi_K)=\sum_{j=0}^n \phi_j\otimes f_j$. By (A3), we have for $y\in K$, $g\in A(K')$,

$$\langle T(y), g\rangle = \left\langle \sum_j \phi_j\otimes f_j, g\otimes y\right\rangle = \left\langle \sum_j f_j(y)\phi_j, g\right\rangle,$$

so that $T=\sum_j f_j(\cdot)\phi_j$. Let δ_j be the extreme points of Δ_n and put $T_0=\sum_j f_j(\cdot)\delta_j$, $T_1(\delta_j)=\phi_j$; then $T=T_1T_0$. (iii) $\Rightarrow$ (i) since any $T_0\in\mathcal{A}(K,V(\Delta_n)^+)$ is ETB; see Example 5.

Assume that T is an ETB channel, so that $T=\sum_j f_j(\cdot)\phi_j$ as above. Let $c_j := \langle \phi_j, 1_{K'}\rangle$. We may assume $c_j>0$; otherwise we may replace Δ_n by a smaller simplex. Put $\tilde f_j := c_j f_j$ and $\tilde\phi_j := c_j^{-1}\phi_j$; then the corresponding maps $\tilde T_0$ and $\tilde T_1$ are channels such that $T=\tilde T_1\tilde T_0$. ■

We next prove Proposition 2. Let $x_0,\dots,x_n\in K$ be a basis of $V(K)$ and $e_0,\dots,e_n$ the dual basis; then

$$\mathrm{Tr}\, T = \sum_i \langle e_i, T(x_i)\rangle = \langle (T\otimes id)\chi_K, \chi_{K^\dagger}\rangle.$$

Proof of Proposition 2. Let $S\in\mathcal{A}_{\rm sep}(K',V(K))$, so that $S=\sum_j f'_j(\cdot)\phi_j$ for some $f'_j\in A(K')^+$ and $\phi_j\in V(K)^+$. Then $\mathrm{Tr}\, TS = \sum_j\langle T(\phi_j), f'_j\rangle \geqslant 0$. Conversely, assume that for all $T\in\mathcal{A}(K,V(K')^+)$

$$\mathrm{Tr}\, ST = \langle (T\otimes id)\chi_K, (S^*\otimes id)\chi_{K^\dagger}\rangle \geqslant 0.$$

By Lemma A1 (iv), we have $(S^*\otimes id)\chi_{K^\dagger}\in A(K\otimes_{\max}K^\dagger)^+$, so it must be separable. By Proposition A1, S^*, and hence also S, is ETB. ■

APPENDIX B: THE CONES $\mathcal{A}(\mathsf{S},V(K)^+)$ AND $\mathcal{A}_{\rm sep}(\mathsf{S},V(K)^+)$

We describe the cone of positive maps $\mathcal{A}(\mathsf{S},V(K)^+)$ and characterize the ETB ones.

Proposition B1. The elements $w_{n_0,\dots,n_k}\in V(K)^+$, $n_i=0,\dots,l_i$, $i=0,\dots,k$ are vertices of some $W\in\mathcal{A}(\mathsf{S},V(K)^+)$ if and only if they satisfy

$$\begin{aligned}w_{n_0,\dots,n_k} + w_{n'_0,\dots,n'_k} &= w_{n_0,\dots,n_{i-1},n'_i,n_{i+1},\dots,n_k}\\ &\quad + w_{n'_0,\dots,n'_{i-1},n_i,n'_{i+1},\dots,n'_k}\end{aligned} \tag{B1}$$

for all $n_0,\dots,n_k$, $n'_0,\dots,n'_k$, and i. Moreover, W is ETB if and only if there are some $\psi^i_j\in V(K)^+$, $j=0,\dots,l_i$, $i=0,\dots,k$, such that

$$w_{n_0,\dots,n_k} = \sum_{i=0}^k \psi^i_{n_i}.$$

Proof. Let $w_{n_),\dots,n_k}$ be vertices of $W\in\mathcal{A}(\mathsf{S},V(K)^+)$. We have

$$\begin{aligned}&\tfrac12\big(\mathsf{s}_{n_0,\dots,n_k}+\mathsf{s}_{n'_0,\dots,n'_k}\big)\\ &= \big(\tfrac12\big(\delta_{n_0}+\delta_{n'_0}\big),\dots,\tfrac12\big(\delta_{n_k}+\delta_{n'_k}\big)\big)\\ &= \big(\tfrac12\big(\delta_{n_0}+\delta_{n'_0}\big),\dots,\tfrac12\big(\delta_{n'_i}+\delta_{n_i}\big),\dots,\tfrac12\big(\delta_{n_k}+\delta_{n'_k}\big)\big)\\ &= \tfrac12\big(\mathsf{s}_{n_0,\dots,n_{i-1},n'_i,n_{i+1},\dots,n_k}+\mathsf{s}_{n'_0,\dots,n'_{i-1},n_i,n'_{i+1},\dots,n'_k}\big);\end{aligned}$$

hence (B1) must hold. Conversely, assume $w_{n_0,\dots,n_k}$ satisfy (B1) and put

$$\begin{aligned}W\big(\mathsf{s}_{l_0,\dots,l_k}\big) &:= w_{l_0,\dots,l_k},\\ W\big(\mathsf{e}^i_j\big) &:= w_{l_0,\dots,l_{i-1},j,l_{i+1},\dots,l_k} - w_{l_0,\dots,l_k}.\end{aligned}$$

This determines a map $W\in\mathcal{A}(\mathsf{S},V(K))$. By (5), we have $W(\mathsf{s}_{n_0,\dots,n_k}) = w_{l_0,\dots,l_k} + \sum_{i=0}^k W(\mathsf{e}^i_{n_i})$. Using repeatedly the relations (B1), we get $W(\mathsf{s}_{n_0,\dots,n_k}) = w_{n_0,\dots,n_k}$. For the second statement, note that since the effects m^i_j generate $A(\mathsf{S})^+$, W is ETB if and only if there are $\psi^i_j \in V(K)^+$ such that

$$W = \sum_{i=1}^{k} \sum_{j=0}^{l_i} \mathsf{m}^i_j(\cdot)\psi^i_j. \tag{B2}$$

Applying this to the vertices of S, we obtain the statement. ■

APPENDIX C: EXTREMAL AND NON-ETB ELEMENTS IN $\mathcal{A}(\square_2, V(K)^+)$

We prove Proposition 5 and Corollary 1.

Proof of Proposition 5. Assume that the two conditions are fulfilled and let $W' \leqslant W$, with vertices $w'_{i,j}$ and barycenter $\bar{w}'$. Then clearly $w'_{i,j} \in F_{i,j}$. By (10) we must have $\bar{w}' = t\bar{w}$ for some $t \in [0,1]$, but then $w'_{i,j} = tw_{i,j}$ for all i,j by (9). It follows that W is extremal.

Conversely, let us denote the subspace on the left-hand side of (10) by L and assume that there is some $\psi \neq t\bar{w}$ in L. Then there are some $\psi_{i,j} \in L_{i,j}$ such that $\psi = \psi_{0,0} + \psi_{1,1} = \psi_{0,1} + \psi_{1,0}$. By definition of $L_{i,j}$, there is some $u > 0$ such that $w^{\pm}_{i,j} := \frac{1}{2}w_{i,j} \pm u\psi_{i,j} \in V(K)^+$. Obviously, $w^+_{i,j}$ and $w^-_{i,1}$ are vertices of some W^+ and W^-, which are not multiples of W, and we have $W = W^+ + W^-$. It follows that W is not extremal.

If (9) is not true, then there are some $\eta_{i,j} \in L_{i,j}$ such that not all of them are 0 and $\eta_{0,0} + \eta_{1,1} = \eta_{0,1} + \eta_{1,0} = 0$. We may then proceed as above to show that W is not extremal. ■

We are now interested in extremal elements that are non-ETB. We start with a simple observation.

Lemma C1. Let $W \in \mathcal{A}(\square_2, V(K)^+)$. If W is degenerate, then it is ETB.

Proof. W is degenerate iff $\dim(W(\square_2)) \leqslant 1$. If the dimension is 0, then W is constant, hence clearly ETB. Assume that the dimension is 1, then all vertices $w_{i,j}$ of W lie on a segment. It is easy to see that we may find a decomposition as in Proposition B1 using multiples of the end points of the segment, so that W is ETB. ■

We can now prove Corollary 1.

Proof of Corollary 1. We will use the notation of Proposition 5.

(i) Any ETB map which is extremal in $\mathcal{A}(\square_2, V(K)^+)$ has the form $\mathsf{m}^i_j(\cdot)\phi$ for some $i,j \in \{0,1\}$ and ϕ an extremal element in $V(K)^+$. Such a map is clearly degenerate. It is therefore enough to show that W is extremal. Since $w_{i,j}$ are extremal in $V(K)^+$, $\dim(L_{i,j}) = 1$ for all i,j. If (9) or (10) is not satisfied, then it is easy to see that $\dim(W(\square_2)) = 1$. Since W is nondegenerate, this is impossible.

(ii) Assume that W is extremal and that, say, $w_{1,0} = 0$. Then $w_{0,0} + w_{1,1} = w_{0,1} = 2\bar{w}$ and $F_{0,0}, F_{1,1} \subseteq F_{0,1}$. By (10), we must have $L_{0,1} = \mathbb{R}\bar{w}$, so that $\bar{w}$ must be extremal in $V(K)^+$. Consequently, both $w_{0,0}$ and $w_{1,1}$ are multiples of $\bar{w}$, but then W is degenerate and hence ETB. It follows that all vertices must be nonzero. Then it follows from (9), (10) by dimension counting that we must have $\dim(L_{i,j}) = 1$, so that all vertices are extremal in $V(K)^+$. The converse holds by statement (i). ■

APPENDIX D: SOME RESULTS AND PROOFS ON INCOMPATIBILITY DEGREE

Lemma D1. $I(F) = \inf_{s\in ri(\mathsf{S})} I_s(F)$.

Proof. Let $s_0 \in \partial\mathsf{S}$ and let $s_1 \in ri(\mathsf{S})$; then $s_t := ts_1 + (1-t)s_0 \in ri(\mathsf{S})$ for all $t \in (0,1]$. Put $\mu := \frac{\lambda(1-t)}{1-\lambda t}$. We have $F_{s_t} = tF_{s_1} + (1-t)F_{s_0}$ and

$$(1-\lambda)F + \lambda F_{s_t} = (1-\lambda t)\big[(1-\mu)F + \mu F_{s_0}\big] + \lambda t F_{s_1}.$$

Assume that $I_{s_0}(F) = \mu$; then $(1-\mu)F + \mu F_{s_0}$ is ETB, so that $(1-\lambda)F + \lambda F_{s_t}$ is ETB as well. It follows that

$$\inf_{s\in ri(\mathsf{S})} I_s(F) \leqslant I_{s_t}(F) \leqslant \lambda = \frac{I_{s_0}(F)}{1 - t\big[1 - I_{s_0}(F)\big]}.$$

Letting $t \to 0$ implies the result. ■

The following lemma is used in the proof of Proposition 6.

Lemma D2. Let (V,V^+) be an ordered vector space and $u \in V$ an order unit. Let $\mathfrak{S}(V,V^+,u) = \{\sigma \in (V^+)^*,\ \langle\sigma,u\rangle = 1\}$. Then for $v \in V$,

$$\inf\{t, v + tu \in V^+\} = \max_{\sigma\in\mathfrak{S}(V,V^+,u)} -\langle\sigma,v\rangle.$$

Proof. Let t_0 denote the infimum on the left-hand side and s_0 the supremum on the right-hand side. Let $t \in \mathbb{R}$ be such that $v + tu \in V^+$. Then for any $\sigma \in \mathfrak{S}(V,V^+,u)$, we have

$$0 \leqslant \langle\sigma, v + tu\rangle = \langle\sigma,v\rangle + t$$

so that $-\langle\sigma,w\rangle \leqslant t$. It follows that $s_0 \leqslant t_0$. Conversely, note that

$$\langle\sigma, v + s_0 u\rangle = \langle\sigma,v\rangle + s_0 \geqslant 0$$

for all $\sigma \in \mathfrak{S}(V,V^+,u)$; hence for all elements in $(V^+)^*$, this implies $t_0 \leqslant s_0$. ■

We next prove Proposition 7.

Proof of Proposition 7. We may assume that $\langle 1_K, \bar{w}\rangle = 1$ (so that $W \in \mathcal{W}_{\bar{\mathsf{S}}}$). Put $\mu^i_j := \langle 1_K, w^i_j\rangle$. Then for any $F \in \mathcal{A}(K, \square_{k+1})$, we have using (12)

$$\begin{aligned}\mathrm{Tr}\, FW &= \mathrm{Tr}\,(F - F_{\bar{\mathsf{S}}})W + \mathrm{Tr}\, F_{\bar{\mathsf{S}}}W\\ &= \sum_{i=0}^{k}\sum_{j=0}^{1}\left(\langle f^i_j, w^i_j\rangle - \frac{1}{2}\mu^i_j\right) + 1.\end{aligned}$$

Hence

$$\begin{aligned}&\min_{F\in\mathcal{A}(K,\square_{k+1})} \mathrm{Tr}\, FW\\ &= \sum_{i=0}^{k} \min_{f\in E(K)}\left(\langle f, w^i_0\rangle + \langle 1 - f, w^i_1\rangle - \frac{1}{2}(\mu^i_0 + \mu^i_1)\right) + 1\\ &= \sum_{i=0}^{k}\left(\frac{1}{2}(\mu^i_1 - \mu^i_0) - \max_{f\in E(K)}\langle f, w^i_1 - w^i_0\rangle\right) + 1\\ &= -\frac{1}{2}\sum_{i=0}^{k}\|W(\mathsf{e}^i_0)\|_K + 1.\end{aligned}$$

The last equality follows from the fact that $\max_{f\in E(K)}\langle f,\psi\rangle = \frac{1}{2}(\|\psi\|_K + \langle 1_K,\psi\rangle)$ for all $\psi \in V(K)$. ■

[1] L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012.
[2] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **84**, 012311 (2011).
[3] H. Araki, Commun. Math. Phys. **75**, 1 (1980).
[4] G. Birkhoff and J. von Neumann, Ann. Math. **37**, 823 (1936).
[5] G. W. Mackey, *The Mathematical Foundations of Quantum Theory* (W. A. Benjamin, Inc., New York, 1963).
[6] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[7] N. Stevens and P. Busch, Phys. Rev. A **89**, 022123 (2014).
[8] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Cloning and broadcasting in generic probabilistic theories, arXiv:quant-ph/0611295.
[9] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007).
[10] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).
[11] M. Banik, M. R. Gazi, S. Ghosh, and G. Kar, Phys. Rev. A **87**, 052125 (2013).
[12] M. T. Quintino, T. Vértesi, and N. Brunner, Phys. Rev. Lett. **113**, 160402 (2014).
[13] M. Banik, J. Math. Phys. **56**, 052101 (2015).
[14] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[15] M. Plávala, Phys. Rev. A **96**, 052127 (2017).
[16] T. Heinosaari, J. Schultz, A. Toigo, and M. Ziman, Phys. Lett. A **378**, 1695 (2014).
[17] T. Heinosaari, T. Miyadera, and M. Ziman, J. Phys. A: Math. Theor. **49**, 123001 (2016).
[18] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, EPL **103**, 10002 (2013).
[19] A. Jenčová and M. Plávala, Phys. Rev. A **96**, 022113 (2017).
[20] M. Sedlák, D. Reitzner, G. Chiribella, and M. Ziman, Phys. Rev. A **93**, 052323 (2016).
[21] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskil, Phys. Rev. A **64**, 052309 (2001).
[22] M. Plávala and M. Ziman, Popescu-Rohrlich box implementation in general probabilistic theory of processes, arXiv:1708.07425.
[23] M. J. Hoban and A. B. Sainz, New J. Phys. **20**, 053048 (2018).
[24] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
[25] P. Janotta and H. Hinrichsen, J. Phys. A: Math. Theor. **47**, 323001 (2014).
[26] H. Barnum and A. Wilce, Post-classical probability theory, arXiv:1205.3833.
[27] P. Janotta and R. Lal, Phys. Rev. A **87**, 052131 (2013).
[28] E. M. Alfsen, *Compact Convex Sets and Boundary Integrals* (Springer-Verlag, Berlin, Heidelberg, New York, 1971).
[29] A. Jenčová, J. Math. Phys. **53**, 012201 (2012).
[30] A. Jenčová, J. Math. Phys. **55**, 022201 (2014).
[31] M. Ziman, Phys. Rev. A **77**, 062112 (2008).
[32] I. Namioka and R. R. Phelps, Pacific J. Math. **31**, 469 (1969).
[33] G. Chiribella, G. M. D'Ariano, and P. Perinotti, EPL **83**, 30004 (2008).
[34] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **80**, 022339 (2009).
[35] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[36] M. A. Naimark, *Normed Rings* (P. Noordhoff N.V., Groningen, 1959).
[37] D. Cavalcanti and P. Skrzypczyk, Phys. Rev. A **93**, 052112 (2016).
[38] M. Plávala, Phys. Rev. A **94**, 042108 (2016).
[39] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936).
[40] H. Barnum, C. P. Gaebler, and A. Wilce, Found. Phys. **43**, 1411 (2013).