

No-free-information principle in general probabilistic theories

Teiko Heinosaari¹, Leevi Leppäjärvi¹, and Martin Plávala²

¹QTF Centre of Excellence, Department of Physics and Astronomy, University of Turku, Turku 20014, Finland

²Mathematical Institute, Slovak Academy of Sciences, Štefánikova 49, Bratislava, Slovakia

June 17, 2019

In quantum theory, the no-information-without-disturbance and no-free-information theorems express that those observables that do not disturb the measurement of another observable and those that can be measured jointly with any other observable must be trivial, i.e., coin tossing observables. We show that in the framework of general probabilistic theories these statements do not hold in general and continue to completely specify these two classes of observables. In this way, we obtain characterizations of the probabilistic theories where these statements hold. As a particular class of state spaces we consider the polygon state spaces, in which we demonstrate our results and show that while the no-information-without-disturbance principle always holds, the validity of the no-free-information principle depends on the parity of the number of vertices of the polygons.

1 Introduction

Quantum theory implies three simple, yet significant and powerful theorems: the no-broadcasting theorem [1], the no-information-without-disturbance theorem [2], and the no-free-information theorem (which can be extracted e.g. from [3, Prop. 3.25]). The no-broadcasting theorem says that quantum states cannot be copied; the no-information-without-disturbance theorem states that a quantum observable that can be measured without any disturbance must be trivial, meaning that it does not give any information on the input state; and the no-free-information theorem states that a quantum observable that can be measured jointly with any other observable must be a trivial observable. In other words, there is no free information, in the sense that a measurement of any non-trivial observable precludes the measurement of some other observable.

Each of the previous three statements can be formulated in the framework of general probabilistic theories (GPTs for short). GPTs constitute a wide class of theories that are based on operational notions such as states, measurements and transformations, where many of the key features of quantum theory, such as non-locality and incompatibility, can be formulated more generally. Including both quantum and classical theory as well as countless toy theories, GPTs then allow us to compare these theories to each other based on their features and quantify their properties.

Teiko Heinosaari: teiko.heinosaari@utu.fi

Leevi Leppäjärvi: leille@utu.fi

Martin Plávala: martin.plavala@mat.savba.sk

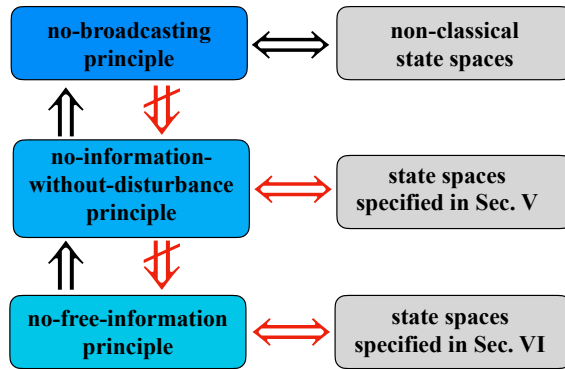


Figure 1: The three principles form a hierarchy, where the no-free-information is the most stringent principle. The main results of this paper (red color) are, firstly, to prove that the three principles are not equivalent and, secondly, to provide full characterizations of the state spaces where the no-free-information and no-information-without-disturbance principles are valid.

In the context of GPTs we find it better to call the previous statements as *principles* instead of theorems as they are not valid in all probabilistic theories. In particular, the no-broadcasting principle is known to be valid in any non-classical general probabilistic theory [4, 5]. In this work, we concentrate on the latter two principles and investigate their validity in the realm of GPTs. The no-information-without-disturbance principle has been shown to hold within GPTs with some additional assumptions, such as purification [6]; however, the validity of this principle has only been mentioned in [7] but never fully investigated in all probabilistic theories. The reverse of the principle was studied in [8]. The no-free-information principle seems to have not been investigated at all in any other theory than quantum theory.

Amongst these principles, no-free-information principle is conceptually the strongest, with no-broadcasting the weakest: If the no-free-information principle is valid in some GPT—that is, for every non-trivial observable there exists another incompatible with it—then the no-information-without-disturbance principle must also be valid, as a non-disturbing observable would be compatible with every other observable. Furthermore, if the no-information-without-disturbance principle is valid and hence no non-trivial observable is non-disturbing, then the no-broadcasting principle has to hold, otherwise we would be capable of using the broadcasting map to create non-trivial non-disturbing observables.

We will define three classes of observables, the first one consisting of those observables that always yield a constant outcome independent of the measured state, the second one consisting of those observables that can be measured without any disturbance and the third one consisting of those observables that are compatible with any other observable. We will then characterize these classes, enabling us to show that the properties are different in some GPTs. We will also derive a necessary and sufficient criterion for a GPT to have both the no-information-without-disturbance principle and no-free-information principle be valid. Finally, we demonstrate the difference between the three principles by analyzing them in polygon state spaces. The main results of our investigation are summarized in Fig. 1.

2 Motivating example

In this section we will present a simple example to motivate our current investigation. A proper mathematical formulation of the general framework will follow in later sections; in the following example we are going to work with the set $B_h(\mathcal{H})$ of square self-adjoint matrices over a finite dimensional Hilbert space \mathcal{H} . We denote by $\mathbb{1}$ the identity matrix and 0 the zero matrix. For $A \in B_h(\mathcal{H})$, we write $A \geq 0$ if A is positive-semidefinite. Let $A, B \in B_h(\mathcal{H})$, then if $A \geq 0$ and $\text{Tr}(A) = 1$, then A is a state and if $0 \leq B \leq \mathbb{1}$, then B is an effect. We refer the reader to [3] for a more throughout treatment of states and effects and their operational meanings in quantum theory.

Imagine that we have an imperfect state preparation device that is meant to prepare qubits in a state ρ , but may malfunction and prepare a qutrit in a state σ . Moreover we assume that the machine malfunctions with a probability p_e , thereby the final state should be a mixture of ρ and σ with probabilities $1 - p_e$ and p_e , respectively. This means that the machine is going to output a state Ψ that should formally be given as $\Psi = (1 - p_e)\rho + p_e\sigma$. But how does one understand the mixture of the 2×2 matrix ρ and the 3×3 matrix σ ? And how does one describe the output state-space of such a machine? We are going to present one possible way to handle this situation; in some cases one should consider the qubit Hilbert space as a subspace of the qutrit Hilbert space, but in other cases (as e.g. when dealing with bosons and fermions) one cannot.

Qubits are effectively a spin- $\frac{1}{2}$ systems and qutrits a spin-1 systems, hence the joint Hilbert space \mathcal{H} containing both representations of the group $\text{SU}(2)$ is going to be 5 dimensional and divided into two superselection sectors [9] of dimensions 2 and 3, corresponding to the qubit and qutrit respectively. The output state Ψ is going to be a block-diagonal 5×5 matrix given as

$$\Psi = \begin{pmatrix} (1 - p_e)\rho & 0 \\ 0 & p_e\sigma \end{pmatrix}.$$

Let M be an effect on \mathcal{H} , then M is of the form

$$M = \begin{pmatrix} M_1 & M_3 \\ M_3^* & M_2 \end{pmatrix},$$

where M_1, M_2, M_3 are matrices of corresponding sizes. We have

$$\text{Tr}(\Psi M) = \text{Tr} \begin{pmatrix} (1 - p_e)\rho M_1 & (1 - p_e)\rho M_3 \\ p_e\sigma M_3^* & p_e\sigma M_2 \end{pmatrix} = (1 - p_e) \text{Tr}(\rho M_1) + p_e \text{Tr}(\sigma M_2),$$

hence from the operational viewpoint we may set $M_3 = 0$ without loss of generality.

Let N be an effect given as

$$N = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$$

then N and $\mathbb{1} - N$ form a projective POVM. Moreover, both N and $\mathbb{1} - N$ commute with all other block-diagonal effects, hence we conclude that the observable corresponding to the POVM $N, \mathbb{1} - N$ is compatible with every other measurement.

This is hardly a surprise, rather a known property of the superselection sectors. Yet this opens the questions of whether this is the only case when an observable is compatible with every other observable; whether no-information-without-disturbance still holds; and whether an observable does not disturb any other observables if it is compatible with them all.

As we saw in this example, we need to at least describe the set of states containing only block-diagonal matrices. For this reason we will work in the GPT formalism as it will provide a unified, cleaner and better suited apparatus for our calculations.

3 Preliminaries

The formulation of GPTs that we adopt here is also called the convex operational framework [7]. In this framework we assume that a state space is convex as we want to interpret convex combinations as mixtures of states. To describe observables, we will introduce effects as functions that assign probabilities to states.

3.1 Structure of general probabilistic theories

A state space \mathcal{S} is a compact convex subset of an ordered real finite-dimensional vector space \mathcal{V} such that \mathcal{S} is a compact base for a generating positive cone $\mathcal{V}_+ = \{x \in \mathcal{V} \mid x \geq 0\}$. Let \mathcal{V}^* denote the dual vector space to \mathcal{V} , then the effect algebra $\mathcal{E}(\mathcal{S}) \subset \mathcal{V}^*$ is the set of linear functionals $e : \mathcal{V} \rightarrow \mathbb{R}$ such that $0 \leq e(x) \leq 1$ for every $x \in \mathcal{S}$. The zero and the unit effects $o \in \mathcal{E}(\mathcal{S})$ and $u \in \mathcal{E}(\mathcal{S})$ are the unique effects satisfying $o(x) = 0$ and $u(x) = 1$ for all $x \in \mathcal{S}$. We note that here we assume the No-Restriction Hypothesis so that every mathematically valid functional is assumed to be a physical effect in the theory [6].

The state space can be expressed as

$$\mathcal{S} = \{x \in \mathcal{V} \mid x \geq 0, u(x) = 1\},$$

i.e. as an intersection of the positive cone \mathcal{V}_+ and an affine hyperplane determined by the unit effect u on \mathcal{V} . Similarly we can define subnormalised states as

$$\mathcal{S}^{\leq 1} = \{x \in \mathcal{V} \mid x \geq 0, u(x) \leq 1\}.$$

If $\dim(\text{aff}(\mathcal{S})) = d$, we say that the state space \mathcal{S} is d -dimensional, and then we can choose \mathcal{V} such that $\dim(\mathcal{V}) = \dim(\mathcal{V}^*) = d + 1$. It follows that the effects can be expressed as linear functionals on \mathcal{V} such that

$$\mathcal{E}(\mathcal{S}) = \{e \in \mathcal{V}^* \mid o \leq e \leq u\},$$

where the partial order in the dual space is the dual order defined by the positive dual cone $\mathcal{V}_+^* = \{f \in \mathcal{V}^* \mid f(x) \geq 0 \text{ for all } x \in \mathcal{V}_+\}$ of \mathcal{V}_+ . In fact $\mathcal{E}(\mathcal{S})$ is then just the intersection of the positive dual cone \mathcal{V}_+^* and the set $u - \mathcal{V}_+^*$.

We say that a non-zero effect $e \in \mathcal{E}(\mathcal{S})$ is indecomposable if a decomposition $e = e_1 + e_2$ for some effects $e_1, e_2 \in \mathcal{E}(\mathcal{S})$ is possibly only if e_1 and e_2 are positive scalar multiples of e [10]. The indecomposable effects are exactly the ones that lie on the extreme rays of the positive dual cone \mathcal{V}_+^* . Indecomposable effects are also known as fine-grained effects (see e.g. [11]).

When dealing with systems composed of several systems we have to prescribe a procedure for how to construct a joint state space of the composed system. Mathematically, this amounts of specifying a tensor product. We are going to use a tensor product only in cases where the other state space is classical. Therefore, there is a unique choice known as the minimal tensor product [12].

Definition 1. Let $\mathcal{S}_1, \mathcal{S}_2$ be state spaces, then their minimal tensor product, denoted as $\mathcal{S}_1 \dot{\otimes} \mathcal{S}_2$, is given as

$$\mathcal{S}_1 \dot{\otimes} \mathcal{S}_2 = \text{conv}(\{x_1 \otimes x_2 \mid x_1 \in \mathcal{S}_1, x_2 \in \mathcal{S}_2\}).$$

3.2 Observables and channels

In this section we will introduce the main objects of interest to us - observables, channels and compatibility. We will begin with observables and their compatibility, and build our way towards channels.

Definition 2. An observable A with a finite outcome set Ω_A on a state space \mathcal{S} is a mapping $A : x \mapsto A_x$ from the outcome set Ω_A to the set of effects $\mathcal{E}(\mathcal{S})$ such that $\sum_{x \in \Omega_A} A_x = u$. The set of observables on \mathcal{S} is denoted by $\mathcal{O}(\mathcal{S})$. For each $A \in \mathcal{O}(\mathcal{S})$ we refer to Ω_A as the outcome set of A .

Let $A, B \in \mathcal{O}(\mathcal{S})$ with respective outcome sets Ω_A, Ω_B . We say that B is a post-processing of A , denoted by $A \rightarrow B$, if there is a right-stochastic matrix ν with elements $\nu_{xy}, x \in \Omega_A, y \in \Omega_B, 0 \leq \nu_{xy} \leq 1, \sum_{y \in \Omega_B} \nu_{xy} = 1$ such that

$$B_y = \sum_{x \in \Omega_A} \nu_{xy} A_x,$$

in which case we also write $B = \nu \circ A$. The operational interpretation is straightforward: we have $A \rightarrow B$ only if we can obtain the probabilities given by B from the probabilities given by A . The condition $\sum_{y \in \Omega_B} \nu_{xy} = 1$ follows from $\sum_{y \in \Omega_B} B_y = u$.

Definition 3. A collection of m observables $A^{(1)}, \dots, A^{(m)} \in \mathcal{O}(\mathcal{S})$ is compatible if there exists an observable $J_{A^{(1)}, \dots, A^{(m)}} \in \mathcal{O}(\mathcal{S})$ such that $J_{A^{(1)}, \dots, A^{(m)}} \rightarrow A^{(i)}$ for all $i = 1, \dots, m$. If two observables A and B are compatible we denote it $A \bowtie B$.

Compatibility of observables and of observables and channels will play a central role in our calculations.

Definition 4. Let $\mathcal{S}_1, \mathcal{S}_2$ be a state spaces. An operation is an affine map $\Psi : \mathcal{S}_1 \rightarrow \mathcal{S}_2^{\leq 1}$. A channel is an affine map $\Phi : \mathcal{S}_1 \rightarrow \mathcal{S}_2$. The set of channels from \mathcal{S}_1 to \mathcal{S}_2 is denoted by $\mathfrak{C}(\mathcal{S}_1, \mathcal{S}_2)$ and in the special case where $\mathcal{S}_1 = \mathcal{S}_2 \equiv \mathcal{S}$ we denote it by $\mathfrak{C}(\mathcal{S})$.

Definition 4 is, in a way, no-restriction hypothesis for channels. In quantum theory we also require channels to be completely positive, but we omit this within GPTs as in general it is problematic to specify what complete positivity means since it is not clear which ancillary state space is to be used in order to define it. Even if in quantum theory the set of channels is smaller than in Definition 4, our results are still valid. This is clarified in Remark 1 in Sec. 5.

Let \mathcal{S} be a state space and let $A \in \mathcal{O}(\mathcal{S})$ with an outcome set Ω_A of n elements. We can identify the points of Ω_A with the extreme points of a simplex, which allows us to form convex combinations of the points of Ω_A . Moreover we will denote this simplex $\mathcal{P}(\Omega_A)$ and its extreme points $\delta_1, \dots, \delta_n$ as they correspond to classical measures on Ω_A supported on a single point. Now we can see the observable A as a channel $A : \mathcal{S} \rightarrow \mathcal{P}(\Omega_A)$, where a state $s \in \mathcal{S}$ is mapped into a probability distribution $\sum_{i=1}^n A_i(s) \delta_i$. Furthermore, a post-processing ν can be seen as a channel mapping the classical state spaces corresponding to outcome sets of observables.

As mentioned above, similarly to compatibility of measurements, we can introduce the compatibility of a measurement and a channel. The central role is going to be played by a generalization of partial trace, which is as follows: let $\mathcal{S}_1, \mathcal{S}_2$ be state spaces and let $x \in \mathcal{S}_1 \otimes \mathcal{S}_2$, then by definition we have $x = \sum_{i=1}^n \lambda_i x_i^1 \otimes x_i^2$ for some $x_i^1 \in \mathcal{S}_1, x_i^2 \in \mathcal{S}_2, \lambda_i \geq 0$

for $i \in \{1, \dots, n\}$ and $\sum_{i=1}^n \lambda_i = 1$. We then define the linear maps $u_1 : \mathcal{S}_1 \dot{\otimes} \mathcal{S}_2 \rightarrow \mathcal{S}_2$ and $u_2 : \mathcal{S}_1 \dot{\otimes} \mathcal{S}_2 \rightarrow \mathcal{S}_1$ as

$$u_1(x) = \sum_{i=1}^n \lambda_i u(x_i^1) x_i^2 = \sum_{i=1}^n \lambda_i x_i^2,$$

$$u_2(x) = \sum_{i=1}^n \lambda_i u(x_i^2) x_i^1 = \sum_{i=1}^n \lambda_i x_i^1.$$

The maps u_1, u_2 are direct generalizations of partial traces. u_1 and u_2 are well-defined and independent of the convex decomposition of the state $x \in \mathcal{S}_1 \dot{\otimes} \mathcal{S}_2$: Let $\mathcal{S}_1 \subset \mathcal{V}_1$ and $\mathcal{S}_2 \subset \mathcal{V}_2$ where \mathcal{V}_1 and \mathcal{V}_2 are the real finite-dimensional vector spaces. Then $u_1 : \mathcal{V}_1 \otimes \mathcal{V}_2 \rightarrow \mathcal{V}_2$ is the unique linear map such that for $v^1 \in \mathcal{V}_1$ and $v^2 \in \mathcal{V}_2$ we have

$$u_1(v^1 \otimes v^2) = u(v^1)v^2$$

and by extending it to $\mathcal{V}_1 \otimes \mathcal{V}_2$ as a linear map it is well-defined and independent of the decomposition. Similar result holds also for $u_2 : \mathcal{V}_1 \otimes \mathcal{V}_2 \rightarrow \mathcal{V}_1$.

Definition 5. A channel $\Phi : \mathcal{S} \rightarrow \mathcal{S}$ is compatible with an observable $A \in \mathcal{O}(\mathcal{S})$ with outcome set Ω_A if and only if there is a channel $\tilde{\Phi} : \mathcal{S} \rightarrow \mathcal{S} \dot{\otimes} \mathcal{P}(\Omega_A)$ such that for all $x \in \mathcal{S}$ we have

$$\Phi(x) = (u_2 \circ \tilde{\Phi})(x),$$

$$A(x) = (u_1 \circ \tilde{\Phi})(x),$$

where \circ denotes the composition of maps.

If the channel Φ were an observable, we would obtain a definition of compatibility of observables which can be shown to be equivalent to Def. 3; see [13]. In a similar fashion one may also formulate the definition of compatibility of channels [14].

We will start with a simple lemma for the compatibility of an observable and a channel.

Lemma 1. A channel $\Phi \in \mathfrak{C}(\mathcal{S})$ and an n -outcome observable $A \in \mathcal{O}(\mathcal{S})$ are compatible if and only if for $i \in \{1, \dots, n\}$ there are operations $\Phi_i : \mathcal{S} \rightarrow \mathcal{V}_+$ such that

$$\Phi = \sum_{i=1}^n \Phi_i, \tag{1}$$

$$A_i = u \circ \Phi_i. \tag{2}$$

Proof. Let Ω_A denote the outcome space with n points and let $\mathcal{P}(\Omega_A) = \text{conv}(\delta_1, \dots, \delta_n)$ be the set of probability distributions on Ω_A , where δ_i for $i \in \{1, \dots, n\}$ is the Dirac measure supported on i -th point of Ω_A . Moreover let b_1, \dots, b_n denote the dual base of affine functions $\mathcal{P}(\Omega_A) \rightarrow \mathbb{R}$, such that $b_i(\delta_j) = 1$ if and only if $i = j$. It is easy to see that all of the functions b_1, \dots, b_n are positive on $\mathcal{P}(\Omega_A)$. If Φ and A are compatible, then there exists a channel $\tilde{\Phi} : \mathcal{S} \rightarrow \mathcal{S} \dot{\otimes} \mathcal{P}(\Omega_A)$ such that $\Phi = u_2 \circ \tilde{\Phi}$ and $A = u_1 \circ \tilde{\Phi}$.

In general, we have $\tilde{\Phi} \in \mathcal{V}^* \otimes \mathcal{V} \otimes \mathcal{P}(\Omega_A)$, i.e.

$$\tilde{\Phi} = \sum_{i=1}^n \sum_{j \in J} f_{ij} \otimes \psi_j \otimes \delta_i$$

for some $f_{ij} \in \mathcal{V}^*$ and $\psi_j \in \mathcal{V}$ and for some index j from a finite index set J . Denote $\Phi_i = \sum_{j \in J} f_{ij} \otimes \psi_j$ and notice that Φ_i are linear maps $\mathcal{V} \rightarrow \mathcal{V}$.

Since $\tilde{\Phi}$ must be a channel then $b_i \circ \tilde{\Phi} : \mathcal{S} \rightarrow \mathcal{S}$ must also be a positive map and since $b_i \circ \tilde{\Phi} = \Phi_i$, we see that Φ_i are positive maps. Since $\tilde{\Phi}$ is a joint channel of Φ and A we must have

$$\begin{aligned}\Phi &= u_2 \circ \tilde{\Phi} = \sum_{i=1}^n \Phi_i, \\ A &= u_1 \circ \tilde{\Phi} = \sum_{i=1}^n (u \circ \Phi_i) \otimes \delta_i.\end{aligned}$$

$\sum_{i=1}^n (u \circ \Phi_i)(x) = 1$ for all $x \in \mathcal{S}$ implies that Φ_i are operations.

If there exist operations Φ_i satisfying (1) and (2), then define $\tilde{\Phi} = \sum_{i=1}^n \Phi_i \otimes \delta_i$. Positivity and normalisation of $\tilde{\Phi}$ follows from the positivity of Φ_i and (2). The fact that $\tilde{\Phi}$ is a joint channel of Φ and A follows from (1) and (2). \square

4 Formulation of the two principles

The purpose of measuring an observable is to learn something about the input state via the obtained measurement outcome probability distribution. An observable is called *trivial* if it cannot provide any information on input states. More precisely, this means that a trivial observable T assigns the same measurement outcome probability distribution to all states, i.e., $T = pu$ for some probability distribution p on Ω_T . Physically speaking, a measurement of a trivial observable can be implemented simply by rolling a dice and producing a probability distribution independently of the input state. We denote by \mathcal{T}_1 the set of all trivial observables, i.e.,

$$\begin{aligned}\mathcal{T}_1 &= \{T \in \mathcal{O}(\mathcal{S}) \mid T_x(s) = T_x(s') \ \forall s, s' \in \mathcal{S}\} \\ &= \{T \in \mathcal{O}(\mathcal{S}) \mid \exists p \in \mathcal{P}(\Omega_T) : T_x = p(x)u \ \forall x \in \Omega_T\}.\end{aligned}$$

From the banal structure of trivial observables it follows that any such observable is compatible with every other observable. Formally, if $T = pu$ is a trivial observable and A is some other observable, then we can define an observable $J_{T,A}$ with effects $J_{T,A}(x, y) = p(x)A_y$, and we have $\sum_x J_{T,A}(x, y) = A_y$ and $\sum_y J_{T,A}(x, y) = T_x$.

Furthermore, a trivial observable is compatible with every channel. Namely, if $T = pu$ is a trivial observable and Φ is a channel, then we can define operations $\Phi_i : \mathcal{S} \rightarrow \mathcal{V}_+$ as $\Phi_i = p(i)\Phi$ for all $i \in \Omega_T$. Clearly, then $\sum_{i \in \Omega_T} \Phi_i = \Phi$ and $(u \circ \Phi_i)(x) = p(i) = T_x(x)$ for all $i \in \Omega_T$ so that by Lemma 1 we conclude that T and Φ are compatible.

These two features of trivial observables raise natural questions: are there observables other than trivial ones that have these features? If so, what is the structure of such observables? As we have seen in Sec. 2, the answer to the first question is affirmative, hence the second question urges an investigation.

To properly analyze the two mentioned features, we consider them as independent properties that determine a subclass of observables. Hence, for a state space \mathcal{S} , we define the following subsets of observables:

$$\begin{aligned}\mathcal{T}_2 &= \{T \in \mathcal{O}(\mathcal{S}) \mid T \infty \Phi \ \forall \Phi \in \mathcal{C}(\mathcal{S})\}, \\ \mathcal{T}_3 &= \{T \in \mathcal{O}(\mathcal{S}) \mid T \infty A \ \forall A \in \mathcal{O}(\mathcal{S})\}.\end{aligned}$$

If an observable T is compatible with the identity channel id , then T is compatible with any channel $\Phi \in \mathcal{C}(\mathcal{S})$. Namely, suppose that T is compatible with id , so there exist operations $\Psi_i : \mathcal{S} \rightarrow \mathcal{V}_+$ such that $\sum_{i \in \Omega_T} \Psi_i = id$ and $u \circ \Psi_i = T_i$. Then we can define a new set of operations as $\Phi \circ \Psi_i$, and these operations give $\sum_{i \in \Omega_T} \Phi \circ \Psi_i = \Phi \circ id = \Phi$ and $u \circ (\Phi \circ \Psi_i) = (u \circ \Phi) \circ \Phi_i = u \circ \Phi_i = T_i$. Therefore, we can concisely write

$$\mathcal{T}_2 = \{T \in \mathcal{O}(\mathcal{S}) \mid T \infty id\}$$

so that there exist measurement set-ups for observables in \mathcal{T}_2 such that the measured states remain unchanged but nevertheless we get the outcome probability distribution of the observable. We conclude that \mathcal{T}_2 is the set of observables that can be measured without causing any disturbance.

Now, suppose that $T \in \mathcal{T}_2$, so there exist operations $\Phi_i : \mathcal{S} \rightarrow \mathcal{V}_+$ such that $\sum_{i \in \Omega_T} \Phi_i = id$ and $u \circ \Phi_i = T_i$ for all $i \in \Omega_T$. If $A \in \mathcal{O}(\mathcal{S})$, we define a joint observable G of A and T by $G_{ij} = A_j \circ \Phi_i$ for all $i \in \Omega_T$ and $j \in \Omega_A$. We then see that

$$\begin{aligned} \sum_j G_{ij} &= \sum_j (A_j \circ \Phi_i) = \left(\sum_j A_j \right) \circ \Phi_i = u \circ \Phi_i = T_i, \\ \sum_i G_{ij} &= \sum_i (A_j \circ \Phi_i) = A_j \circ \left(\sum_i \Phi_i \right) = A_j \circ id = A_j \end{aligned}$$

for all $i \in \Omega_T$ and $j \in \Omega_A$. Thus, $A = \nu^A \circ G$ and $T = \nu^T \circ G$, where $\nu^A : \Omega_T \times \Omega_A \rightarrow \Omega_A$ and $\nu^T : \Omega_T \times \Omega_A \rightarrow \Omega_T$ are defined as $\nu_{(i,j)k}^A = \delta_{jk}$ and $\nu_{(i,j)l}^T = \delta_{il}$ for all $j, k \in \Omega_A$ and $i, l \in \Omega_T$, so that A and T are compatible, and since A was an arbitrary observable, it follows that $T \in \mathcal{T}_3$. We conclude that

$$\mathcal{T}_1 \subseteq \mathcal{T}_2 \subseteq \mathcal{T}_3.$$

These three sets and the previous chain of inclusions allows us to give a simple and concise formulation of the two principles: *The no-information-without-disturbance principle means that $\mathcal{T}_2 = \mathcal{T}_1$, while the no-free-information principle means that $\mathcal{T}_3 = \mathcal{T}_1$.*

Indeed, these formulations capture the ideas behind the principles so that observables that can be measured without any disturbance, i.e. observables in \mathcal{T}_2 , should be trivial and similarly observables that can be measured jointly with any other observable, i.e. observables in \mathcal{T}_3 , should be a trivial as well so that only the measurement of a trivial observable allows for the joint measurement of any other observable.

5 Characterization of \mathcal{T}_2

The aim of this section is to characterize non-disturbing observables and the structure of the state spaces they may exist on. We will have to introduce additional mathematical results to provide the full description of such state spaces.

5.1 Direct sum of state spaces

We will introduce a direct sum of state spaces as a generalized description of using only block-diagonal quantum states. Our aim is to mathematically formalize the operational idea of having an ordered pair of weighted states from two different state spaces.

Definition 6. Let $\mathcal{V}_1, \mathcal{V}_2$ be real finite-dimensional vector spaces and let $\mathcal{S}_1 \subset \mathcal{V}_1$ and $\mathcal{S}_2 \subset \mathcal{V}_2$ be state spaces. We define a state space $\mathcal{S}_1 \oplus \mathcal{S}_2 \subset \mathcal{V}_1 \times \mathcal{V}_2$ as the set of ordered and weighted pairs of states from \mathcal{S}_1 and \mathcal{S}_2 , i.e.,

$$\mathcal{S}_1 \oplus \mathcal{S}_2 = \{(\lambda x_1, (1 - \lambda)x_2) \mid x_1 \in \mathcal{S}_1, x_2 \in \mathcal{S}_2, \lambda \in [0, 1]\}.$$

Given state spaces $\mathcal{S}_1, \dots, \mathcal{S}_n$ one can define $\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n$ in a similar fashion as a subset of $\mathcal{V}_1 \times \dots \times \mathcal{V}_n$, i.e., one would have

$$\mathcal{S}_1 \oplus \dots \oplus \mathcal{S}_n = \left\{ (\lambda_1 x_1, \dots, \lambda_n x_n) \mid x_i \in \mathcal{S}_i, \lambda_i \geq 0, \forall i \in \{1, \dots, n\}, \sum_{i=1}^n \lambda_i = 1 \right\}.$$

In what follows we will present a few basic results about $\mathcal{S}_1 \oplus \mathcal{S}_2$. We will limit only to direct sum of two state spaces for the sake of not drowning in a sea of symbols, but it will be straightforward to see that all of the results hold for any finite direct sum as well.

Proposition 1. $\mathcal{E}(\mathcal{S}_1 \oplus \mathcal{S}_2) = \mathcal{E}(\mathcal{S}_1) \times \mathcal{E}(\mathcal{S}_2)$, where $\mathcal{E}(\mathcal{S}_1) \times \mathcal{E}(\mathcal{S}_2) = \{(e_1, e_2) \mid e_1 \in \mathcal{E}(\mathcal{S}_1), e_2 \in \mathcal{E}(\mathcal{S}_2)\}$.

Proof. $\mathcal{S}_1 \oplus \mathcal{S}_2 \subset \mathcal{V}_1 \times \mathcal{V}_2$ so we must have $\mathcal{E}(\mathcal{S}_1 \oplus \mathcal{S}_2) \subset \mathcal{V}_1^* \times \mathcal{V}_2^*$. Let $(e_1, e_2) \in \mathcal{V}_1^* \times \mathcal{V}_2^*$ and let $(\lambda x_1, (1 - \lambda)x_2) \in \mathcal{S}_1 \oplus \mathcal{S}_2$, then from

$$(e_1, e_2)((\lambda x_1, (1 - \lambda)x_2)) = \lambda e_1(x_1) + (1 - \lambda)e_2(x_2) \quad (3)$$

it follows that $\mathcal{E}(\mathcal{S}_1) \times \mathcal{E}(\mathcal{S}_2) \subset \mathcal{E}(\mathcal{S}_1 \oplus \mathcal{S}_2)$. Assuming $(e_1, e_2) \in \mathcal{E}(\mathcal{S}_1 \oplus \mathcal{S}_2)$ and setting $\lambda = 0$ and $\lambda = 1$ in (3) we get $e_1 \in \mathcal{E}(\mathcal{S}_1)$ and $e_2 \in \mathcal{E}(\mathcal{S}_2)$. \square

It follows that if $A \in \mathcal{O}(\mathcal{S}_1 \oplus \mathcal{S}_2)$, then we have $A_i = (A_i^1, A_i^2)$ for some $A^1 \in \mathcal{O}(\mathcal{S}_1)$, $A^2 \in \mathcal{O}(\mathcal{S}_2)$.

Proposition 2. Let $A, B \in \mathcal{O}(\mathcal{S}_1 \oplus \mathcal{S}_2)$, such that $A_i = (A_i^1, A_i^2)$, $B_j = (B_j^1, B_j^2)$, then $A \otimes B$ if and only if $A^1 \otimes B^1$ and $A^2 \otimes B^2$.

Proof. If $A^1 \otimes B^1$ and $A^2 \otimes B^2$ then $A \otimes B$ as we can form the joint observable as $(J_{A,B})_k = ((J_{A^1, B^1})_k, (J_{A^2, B^2})_k)$ and apply the respective post-processings to the respective observables, hence $A \otimes B$. Note that to make the observables have the same number of outcomes, we can always pad out one with zero effects corresponding to some extra outcomes that never happen.

If $A \otimes B$, then by restricting the state space only to states of the form $(x_1, 0) \in \mathcal{S}_1 \oplus \mathcal{S}_2$, where $x_1 \in \mathcal{S}_1$ it follows that $A^1 \otimes B^1$ are compatible as we can obtain J_{A^1, B^1} from $J_{A, B}$. $A^2 \otimes B^2$ follows in the same manner. \square

This explains our motivational example in Sec. 2. One can also prove a similar result for the compatibility of an observable and a channel, but we will leave that for the next section, where we will investigate the conditions for the compatibility of an observable and the identity channel $id : \mathcal{S} \rightarrow \mathcal{S}$, where direct sums of state spaces will play a role.

This last result will help us identify the direct sum structure of a state space.

Proposition 3. Let \mathcal{S} be a state space and let $\mathcal{S}_1, \mathcal{S}_2 \subset \mathcal{S}$ be convex, closed sets, such that $\text{conv}(\mathcal{S}_1 \cup \mathcal{S}_2) = \mathcal{S}$ and for every $x \in \mathcal{S}$ there are unique $x_1 \in \mathcal{S}_1$, $x_2 \in \mathcal{S}_2$ and $\lambda \in [0, 1]$ such that $x = \lambda x_1 + (1 - \lambda)x_2$. It follows that $\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2$.

Proof. Let \mathcal{V}_1 and \mathcal{V}_2 denote the subspaces of \mathcal{V} generated by \mathcal{S}_1 and \mathcal{S}_2 respectively. Define map $P : \mathcal{S} \rightarrow \mathcal{V}_1 \times \mathcal{V}_2$ given for $x \in \mathcal{S}$, $x = \lambda x_1 + (1 - \lambda)x_2$, $x_1 \in \mathcal{S}_1$, $x_2 \in \mathcal{S}_2$ as $P(x) = (\lambda x_1, (1 - \lambda)x_2)$. It follows that we have $P : \mathcal{S} \rightarrow \mathcal{S}_1 \oplus \mathcal{S}_2$, moreover one can easily see that P is an affine isomorphism. It follows that \mathcal{S} is affinely isomorphic to $\mathcal{S}_1 \oplus \mathcal{S}_2$, the result follows by simply omitting the isomorphism. \square

5.2 Compatibility of an observable and the identity channel

We are going to derive conditions for an observable to be compatible with the identity channel $id : \mathcal{S} \rightarrow \mathcal{S}$. Our results will be similar to the results mentioned in [7, 15], but we will approach the problem from a different angle and with a different objective in mind.

Lemma 2. *An observable A with an n -outcome space Ω_A is compatible with the identity channel $id : \mathcal{S} \rightarrow \mathcal{S}$ if and only if there is a channel $\Phi : \mathcal{S} \rightarrow \mathcal{S} \dot{\otimes} \mathcal{P}(\Omega_A)$ such that for every extreme point $y \in \mathcal{S}$ we have*

$$\Phi(y) = \sum_{i=1}^n A_i(y)y \otimes \delta_i. \quad (4)$$

Proof. Assume that an observable A is compatible with id , then due to Lemma 1 we must have operations Φ_1, \dots, Φ_n such that $id = \sum_{i=1}^n \Phi_i$ and $A_i = u \circ \Phi_i$. To prove our claim we will use the defining property of extreme points. We have

$$y = id(y) = \sum_{i=1}^n \Phi_i(y)$$

that implies $\Phi_i(y) = \lambda_i(y)y$, where $\lambda_i(y) \in [0, 1]$ may in general depend on i and y . From $A_i = u \circ \Phi_i$ we obtain $\lambda_i(y) = A_i(y)$. For the joint channel Φ of id and A we have

$$\Phi(y) = \sum_{i=1}^n \Phi_i(y) \otimes \delta_i = \sum_{i=1}^n A_i(y)y \otimes \delta_i.$$

Now assume that for a channel $\Phi : \mathcal{S} \rightarrow \mathcal{S} \dot{\otimes} \mathcal{P}(\Omega_A)$ the equation (4) holds. For every extreme point $y \in \mathcal{S}$ we have

$$\begin{aligned} (u_2 \circ \Phi)(y) &= \sum_{i=1}^n A_i(y)y = y, \\ (u_1 \circ \Phi)(y) &= \sum_{i=1}^n A_i(y) \otimes \delta_i = A(y). \end{aligned}$$

Since this holds for every extreme point of \mathcal{S} it follows that Φ is a joint channel of A and id . \square

Proposition 4. *Observable A is compatible with id if and only if there is a set of affinely independent extreme points of \mathcal{S} , denoted by x_j , where $j \in \{1, \dots, d\}$, such that $\mathcal{S} \subset \text{aff}(\{x_1, \dots, x_d\})$ and for every extreme point $y \in \mathcal{S}$, $y = \sum_{j=1}^d \alpha_j x_j$ it holds that*

$$\alpha_j(A_i(x_j) - A_i(y)) = 0. \quad (5)$$

Proof. Assume that an observable A is compatible with id and let $x_1, \dots, x_d \in \mathcal{S}$ be a set of affinely independent extreme points, such that $\mathcal{S} \subset \text{aff}(\{x_1, \dots, x_d\})$. Let $y \in \mathcal{S}$ be an extreme point, then $y = \sum_{j=1}^d \alpha_j x_j$, where $\sum_{j=1}^d \alpha_j = 1$. According to Lemma 2 there is a channel Φ such that (4) holds. Plugging in the expression $y = \sum_{j=1}^d \alpha_j x_j$ we obtain

$$\Phi(y) = \sum_{j=1}^d \alpha_j \Phi(x_j) = \sum_{j=1}^d \alpha_j \sum_{i=1}^n A_i(x_j)x_j \otimes \delta_i$$

which implies

$$\sum_{i=1}^n \mathbf{A}_i(y)y \otimes \delta_i = \sum_{j=1}^d \sum_{i=1}^n \alpha_j \mathbf{A}_i(x_j)x_j \otimes \delta_i.$$

Since $\delta_1, \dots, \delta_n$ are linearly independent we must have $\mathbf{A}_i(y)y = \sum_{j=1}^d \alpha_j \mathbf{A}_i(x_j)x_j$ which yields

$$\sum_{j=1}^d \alpha_j (\mathbf{A}_i(x_j) - \mathbf{A}_i(y))x_j = 0.$$

Eq. (5) follows by affine independence of x_1, \dots, x_d .

Assume that (5) holds for an observable \mathbf{A} and define a map $\Phi : \mathcal{S} \rightarrow \mathcal{S} \dot{\otimes} \mathcal{P}(\Omega_{\mathbf{A}})$ given for $j \in \{1, \dots, d\}$ as

$$\Phi(x_j) = \sum_{i=1}^n \mathbf{A}_i(x_j)x_j \otimes \delta_i$$

and extended by affinity to all of \mathcal{S} . One can show that the map Φ is well defined and does not depend on the choice of the points x_1, \dots, x_d and the proof relies on Eq. (5). Let $y \in \mathcal{S}$ be an extreme point, then we have $y = \sum_{j=1}^d \alpha_j x_j$, $\sum_{j=1}^d \alpha_j = 1$ and

$$\Phi(y) = \sum_{j=1}^d \alpha_j \Phi(x_j) = \sum_{j=1}^d \alpha_j \sum_{i=1}^n \mathbf{A}_i(x_j)x_j \otimes \delta_i = \sum_{j=1}^d \sum_{i=1}^n \alpha_j \mathbf{A}_i(y)x_j \otimes \delta_i = \sum_{i=1}^n \mathbf{A}_i(y)y \otimes \delta_i$$

where we have used (5) in the third step. By lemma 2 it follows that \mathbf{A} is compatible with id . \square

Note that if \mathcal{S} is a simplex, then the set $\{x_1, \dots, x_d\}$ is unique and contains all extreme points of \mathcal{S} , hence the requirement of Prop. 4 is trivially satisfied.

It is important to note that Prop. 4 provides a condition on the effects \mathbf{A}_i , not on \mathbf{A} as a whole. Therefore it will be interesting to investigate the set of effects that satisfy the condition (5).

Definition 7. We denote \mathcal{ET}_2 set of effects on a state space \mathcal{S} that satisfy the condition (5), i.e. $f \in \mathcal{ET}_2$ if there is some set $\{x_1, \dots, x_d\}$ of affinely independent extreme points of \mathcal{S} such that $\mathcal{S} \subset \text{aff}(\{x_1, \dots, x_d\})$ and for every extreme point $y \in \mathcal{S}$, $y = \sum_{j=1}^d \alpha_j x_j$ it holds that

$$\alpha_j (f(x_j) - f(y)) = 0. \quad (6)$$

The following is straightforward.

Lemma 3. \mathcal{ET}_2 is a convex subeffect algebra of $\mathcal{E}(\mathcal{S})$, i.e., if $f, g \in \mathcal{ET}_2$ and $0 \leq \lambda \leq 1$, then

1. $o, u \in \mathcal{ET}_2$ and if $f + g \in \mathcal{E}(\mathcal{S})$, then we must have $f + g \in \mathcal{ET}_2$,
2. $\lambda f + (1 - \lambda)g \in \mathcal{ET}_2$.

Proof. The results follow immediately from linearity of (6). \square

Proposition 5. Let $0 < \lambda \leq 1$ and $0 \leq \mu \leq 1$, then $f \in \mathcal{ET}_2$ if and only if $\lambda f + (1 - \lambda)\mu u \in \mathcal{ET}_2$.

Proof. If $f \in \mathcal{ET}_2$ then $\lambda f + (1-\lambda)\mu u \in \mathcal{ET}_2$ follows by Lemma 3. If $\lambda f + (1-\lambda)\mu u \in \mathcal{ET}_2$, then

$$\alpha_j((\lambda f + (1-\lambda)\mu)(x_j) - (\lambda f + (1-\lambda)\mu)(y)) = 0$$

which is the same as

$$\lambda\alpha_j(f(x_j) - f(y)) + \alpha_j(1-\lambda)\mu(u(x_j) - u(y)) = 0$$

and $\alpha_j(f(x_j) - f(y)) = 0$ follows because $\lambda \neq 0$ and $u(x_j) = u(y) = 0$. \square

The result of Prop. 5 is non-trivial. As we will see, there are observables that are compatible with all other observables because they are “noisy enough”. But according to Prop. 5 this is not the case for compatibility with the identity channel id . Loosely speaking Prop. 5 together with the next result show that the structure of \mathcal{T}_2 is more like \mathcal{T}_1 , than \mathcal{T}_3 in the sense that observables in \mathcal{T}_2 are in some sense classical; such as was the case in Sec. 2.

Corollary 1. *Observable $A \in \mathcal{T}_2$ if and only if $A_i \in \mathcal{ET}_2$ for all i .*

Proof. Follows from Prop. 4. \square

Theorem 1. *$f \in \mathcal{ET}_2$ if and only if $\mathcal{S} = \bigoplus_{k=1}^N \mathcal{S}_k$ and f is constant on each \mathcal{S}_k .*

Proof. If \mathcal{S} is a simplex, then there is only one set $\{x_1, \dots, x_d\}$ of affinely independent points and we have $\mathcal{S} = \bigoplus_{j=1}^d x_j$. The claim follows.

Let x_1, \dots, x_d be a set of affinely independent extreme points of \mathcal{S} and let $y \in \mathcal{S}$ be an extreme point, then we have $y = \sum_{j=1}^d \alpha_j x_j$, $\sum_{j=1}^d \alpha_j = 1$. Since \mathcal{S} is not a simplex we can find a pure state $y \in \mathcal{S}$ such that $\alpha_{j'} \neq 0$ and $\alpha_{j''} \neq 0$ for some $j', j'' \in \{1, \dots, d\}$. Eq. (6) implies $f(x_{j'}) = f(y)$ and $f(x_{j''}) = f(y)$, which gives $f(x_{j'}) = f(x_{j''})$.

Denote $\mathcal{S}_c = \text{conv}(\{z \in \mathcal{S} : f(z) = c, z \text{ is extreme}\})$. We have just proved that there is only finite number of the sets \mathcal{S}_c , $\mathcal{S}_c \subset \text{aff}(\{x_j : f(x_j) = c\})$.

Let $z \in \mathcal{S}$, then we have already proved that we have

$$z = \sum_{c \in [0,1]} \lambda_c y_c, \tag{7}$$

where $0 \leq \lambda_c \leq 1$, $\sum_{c \in [0,1]} \lambda_c = 1$ and $y_c \in \mathcal{S}_c$. Note that y_c is not necessarily an extreme point of \mathcal{S} . We will show that the decomposition (7) is unique. Assume there is another decomposition $z = \sum_{c \in [0,1]} \lambda'_c y'_c$, where again $0 \leq \lambda'_c \leq 1$, $\sum_{c \in [0,1]} \lambda'_c = 1$ and $y'_c \in \mathcal{S}_c$. Moreover assume that $\lambda_{c'} \neq 0$, then from $\sum_{c \in [0,1]} \lambda_c y_c = \sum_{c \in [0,1]} \lambda'_c y'_c$ we have

$$y_{c'} = \frac{1}{\lambda_{c'}} \left(\sum_{c \in [0,1]} \lambda'_c y'_c - \sum_{c \in [0,1] \setminus \{c'\}} \lambda_c y_c \right).$$

We can decompose $y_{c'} = \sum_{k=1}^n \mu_k y_{c',k}$, where $0 \leq \mu_k \leq 1$, $\sum_{k=1}^n \mu_k = 1$ and $y_{c',k}$ are extreme points of $\mathcal{S}_{c'}$. Moreover assume that $\mu_{k'} \neq 0$, then we have

$$y_{c',k'} = \frac{1}{\mu_{k'}} \left(\frac{1}{\lambda_{c'}} \left(\sum_{c \in [0,1]} \lambda'_c y'_c - \sum_{c \in [0,1] \setminus \{c'\}} \lambda_c y_c \right) - \sum_{k=1, k \neq k'}^n \mu_k y_{c',k} \right).$$

It follows that the right-hand side must be an affine combination of x_j , $j \in \{1, \dots, n\}$ such that $f(x_j) = c'$. This implies that for $c \neq c'$ we must have $\lambda_c y_c = \lambda'_{c'} y'_{c'}$ as otherwise the aforementioned result would be violated. We get

$$y_{c',k'} = \frac{1}{\mu_{k'}} \left(\frac{\lambda'_{c'}}{\lambda_{c'}} y'_{c'} - \sum_{k=1, k \neq k'}^n \mu_k y_{c',k} \right).$$

It follows that

$$y_{c'} = \frac{\lambda'_{c'}}{\lambda_{c'}} y'_{c'},$$

hence the two decompositions of z are the same. The result follows from Prop. 3. \square

By combining Cor. 1 and Thm. 1 we get our main result regarding \mathcal{T}_2 :

Corollary 2. *Observable $A \in \mathcal{T}_2$ if and only if one can represent the state space \mathcal{S} as a direct sum $\mathcal{S} = \bigoplus_{k=1}^N \mathcal{S}_k$ such that each effect A_x is constant on each \mathcal{S}_k .*

Remark 1. In quantum theory, channels and operations are required to be completely positive. As we have earlier taken all affine maps to be operations, the additional requirement of complete positivity potentially changes our previous results on compatibility. However, the crucial point is that also in quantum theory, the compatibility of an observable with all channels is equivalent to the compatibility of the observable with the identity channel. It is easy to see that the statement of Lemma 2 is valid in quantum theory even if operations are required to be completely positive. The results, including Thm. 1, were based on Lemma 2 and hence they are true in quantum theory.

In general, assume that we have $\mathcal{S} = \bigoplus_{k=1}^N \mathcal{S}_k$ and let $x \in \mathcal{S}$ be given as $x = \sum_{k=1}^N \lambda_k x_k$ where $x_k \in \mathcal{S}_k$, $\lambda_k \geq 0$ for all $k \in \{1, \dots, N\}$ and $\sum_{k=1}^N \lambda_k = 1$. We can easily see that the decomposition of x is unique, which is a result similar to Prop. 3. Define the projections $P_k : \mathcal{S} \rightarrow \mathcal{S}_k$ as $P_k(x) = \lambda_k x_k$, one can see that this is a well-defined and positive map and that $u \circ P_k = A_k \in \mathcal{ET}_2$ for all $k \in \{1, \dots, N\}$. One can show that if the projections P_k are completely positive, then the joint channel of the observable $A = \sum_{k=1}^N A_k \otimes \delta_k$ and the identity channel id is completely positive, because the projections P_k are the needed decomposition of id , i.e. $id = \sum_{k=1}^N P_k$.

It is natural to assume that the projections P_k are completely positive due to our interpretation of direct sum of state spaces; we see $\bigoplus_{k=1}^N \mathcal{S}_k$ as a randomization of some underlying state spaces \mathcal{S}_k that form the superselection sectors. Especially in superselected quantum theory as in Sec. 2 it is easy to see that this is the case and that the projections are completely positive.

Using Thm. 1 we can easily characterize all two-dimensional state spaces that have observables compatible with the identity channel, i.e. that have information without disturbance. Remember that if a state space is two-dimensional, then $\dim(\mathcal{V}) = 3$ where \mathcal{V} is the vector space containing the cone \mathcal{V}^+ which has the base \mathcal{S} .

Corollary 3. *Let $\dim(\mathcal{V}) = 3$, then $\mathcal{S} = \mathcal{S}^1 \oplus \mathcal{S}^2$ if and only if \mathcal{S} is the triangle state space.*

Proof. Assume that $\mathcal{S} = \mathcal{S}^1 \oplus \mathcal{S}^2$, then $\mathcal{V} = \mathcal{V}^1 \times \mathcal{V}^2$, where $\mathcal{V}^1, \mathcal{V}^2$ are the vector spaces that contain \mathcal{S}^1 and \mathcal{S}^2 respectively. This implies $\dim(\mathcal{V}^1) + \dim(\mathcal{V}^2) = \dim(\mathcal{V}) = 3$ and we can assume that $\dim(\mathcal{V}^1) = 1$, $\dim(\mathcal{V}^2) = 2$. This implies that \mathcal{S}^1 contains only one point and \mathcal{S}^2 is a line segment, i.e. it has two extreme points. It then follows that \mathcal{S} must have three extreme points, hence it is a triangle state space, which is a simplex. \square

In a similar fashion one can show that every three-dimensional state space that has information without disturbance is pyramid shaped, where the base of the pyramid can be any two-dimensional state space.

6 Characterization of \mathcal{T}_3

6.1 Simulability of observables

Simulation of observables is a method to produce a new observable from a given collection of observables by a classical procedure, that is, by mixing measurement settings and post-processing the outcome data [16–19]. For a subset $\mathcal{B} \subseteq \mathcal{O}(\mathcal{S})$, we denote by $\mathbf{sim}(\mathcal{B})$ the set of observables that can be simulated by using the observables from \mathcal{B} , i.e., $A \in \mathbf{sim}(\mathcal{B})$ if there exists a probability distribution p , a finite collection of post-processing matrices $\nu^{(i)}$ and observables $B^{(i)} \in \mathcal{B}$ such that

$$A = \sum_i p_i \left(\nu^{(i)} \circ B^{(i)} \right).$$

We will also denote $\mathbf{sim}(B) \equiv \mathbf{sim}(\{B\})$. Clearly,

$$\mathbf{sim}(B) = \{A \in \mathcal{O}(\mathcal{S}) : B \rightarrow A\}.$$

We recall from [18] that an observable A is called *simulation irreducible* if for any subset $\mathcal{B} \subset \mathcal{O}$, we have $A \in \mathbf{sim}(\mathcal{B})$ only if there is $B \in \mathcal{B}$ such that $A \in \mathbf{sim}(B)$ and $B \in \mathbf{sim}(A)$. Thus, a simulation irreducible observable can only be simulated by (essentially) itself. Equivalently, an observable is simulation irreducible if and only if it has indecomposable effects and is post-processing equivalent with an extreme observable. We denote by $\mathcal{O}_{irr}(\mathcal{S})$ the set of simulation irreducible observables. It was shown in [18] that for every observable there exists a finite collection of simulation irreducible observables from which it can be simulated.

It is worth mentioning that simulation irreducible observables are always incompatible, and in fact, a state space is non-classical if and only if there exists at least two inequivalent simulation irreducible observables [18].

6.2 Intersections of simulation sets

A trivial observable can be simulated by any other observable, and therefore

$$\mathcal{T}_1 = \bigcap_{B \in \mathcal{O}(\mathcal{S})} \mathbf{sim}(B). \quad (8)$$

The following stronger statement is less obvious, although not too surprising.

Proposition 6.

$$\mathcal{T}_1 = \bigcap_{B \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1} \mathbf{sim}(B). \quad (9)$$

Proof. Since $\mathcal{T}_1 \subseteq \bigcap_{B \in \mathcal{O}(\mathcal{S})} \mathbf{sim}(B)$, it is clear that $\mathcal{T}_1 \subseteq \bigcap_{B \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1} \mathbf{sim}(B)$. On the other hand, suppose that the inclusion is strict so that (w.l.o.g.) there exist a dichotomic observable $T \in \bigcap_{B \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1} \mathbf{sim}(B)$ such that $T \notin \mathcal{T}_1$. This means that the effects T_+ and T_- are not proportional to the unit effect u so that especially T_+ and u are linearly independent.

We take $\lambda, q \in (0, 1)$ and define another dichotomic observable A by $A = \lambda T + (1 - \lambda)Q$, where $Q \in \mathcal{T}_1$ is defined as $Q_+ = qu$ and $Q_- = (1 - q)u$. Since $\lambda \neq 0$ and $T \notin \mathcal{T}_1$, we have that $A \notin \mathcal{T}_1$ so that in fact $A \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$. Hence, because $T \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$ we have that in particular $T \in \mathbf{sim}(A)$, i.e. there exists two real numbers $\nu_1, \nu_2 \in [0, 1]$ such that $T_+ = \nu_1 A_+ + \nu_2 A_-$. When we expand A_+ and A_- , we find that

$$\begin{aligned} T_+ &= \nu_1(\lambda T_+ + (1 - \lambda)qu) + \nu_2(\lambda T_- + (1 - \lambda)(1 - q)u) \\ &= \lambda(\nu_1 - \nu_2)T_+ + (1 - \lambda)(\nu_1 - \nu_2)qu + \nu_2u, \end{aligned}$$

where on the second line we have used the fact that $T_- = u - T_+$. From the linear independence of u and T_+ it follows that we must have $\lambda(\nu_1 - \nu_2) = 1$, which is a contradiction since $0 < \lambda < 1$ and $\nu_1 - \nu_2 \leq 1$. \square

The equations (8) and (9) make one to wonder if the set $\mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$ can still be shrunk without altering the resulting set of the intersection of their simulation sets. Remarkably, taking $\mathcal{O}_{irr}(\mathcal{S})$ instead of $\mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$ changes the intersection, and leads to the following characterization for the set \mathcal{T}_3 .

Proposition 7.

$$\mathcal{T}_3 = \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \mathbf{sim}(B). \quad (10)$$

Proof. Let first $T \in \mathcal{T}_3$. Since T is compatible with every other observable, it is in particular compatible with every simulation irreducible observable. Thus, for every $B \in \mathcal{O}_{irr}(\mathcal{S})$ there exists $G^B \in \mathcal{O}(\mathcal{S})$ such that $\{B, T\} \subseteq \mathbf{sim}(G^B)$. Since B is simulation irreducible it follows from the definition that $B \leftrightarrow G^B$ so that $\mathbf{sim}(B) = \mathbf{sim}(G^B)$. Thus, $T \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$.

Now let $A \in \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \mathbf{sim}(B)$ so that $A \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$. We must show that A is compatible with every other observable. Thus, let $C \in \mathcal{O}(\mathcal{S})$. For C there exists a finite set of simulation irreducible observables $\mathcal{B} = \{B^{(i)}\}_{i=1}^n$ such that $C \in \mathbf{sim}(\mathcal{B})$. Thus, there exists a probability distribution $(p_i)_{i=1}^n$ and a post-processing $\nu : \{1, \dots, n\} \times \Omega_{\mathcal{B}} \rightarrow \Omega_C$ such that

$$C_y = \sum_{i,x} p_i \nu_{(i,x)y} B_x^{(i)} \quad (11)$$

for all $y \in \Omega_C$. If we denote by \tilde{B} the (generalized) mixture observable with outcomes set $\{1, \dots, n\} \times \Omega_{\mathcal{B}}$ defined by $\tilde{B}_{(i,x)} = p_i B_x^{(i)}$ for all $i \in \{1, \dots, n\}$ and $x \in \Omega_{\mathcal{B}}$, we see that actually $C_y = (\nu \circ \tilde{B})_y$ for all $y \in \Omega_C$ so that $C \in \mathbf{sim}(\tilde{B})$.

Since $A \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$, we have that $A \in \mathbf{sim}(B^{(i)})$ for all $i = 1, \dots, n$. Thus, there exists post-processings $\mu^{(i)} : \Omega_{\mathcal{B}} \rightarrow \Omega_A$ such that $A = \mu^{(i)} \circ B^{(i)}$ for all $i = 1, \dots, n$. If we use the same probability distribution $(p_i)_i$ as before, we have that for all $z \in \Omega_A$

$$A_z = \sum_i p_i A_z = \sum_i p_i \sum_x \mu_{xz}^{(i)} B_x^{(i)} = \sum_{i,x} \mu_{(i,x)z} p_i B_x^{(i)} = (\mu \circ \tilde{B})_z,$$

where we have defined a new post-processing $\mu : \{1, \dots, n\} \times \Omega_{\mathcal{B}} \rightarrow \Omega_A$ by setting $\mu_{(i,x)z} = \mu_{xz}^{(i)}$ for all $i \in \{1, \dots, n\}$, $x \in \Omega_{\mathcal{B}}$ and $z \in \Omega_A$. Hence, also $A \in \mathbf{sim}(\tilde{B})$ so that A and C are compatible. \square

As was shown in Prop. 7, the observables that are compatible with every other observable are exactly those that can be post-processed from every simulation irreducible observable. However, we note that it is enough to consider only post-processing inequivalent simulation irreducible observables since two observables B and B' are post-processing equivalent, $B \leftrightarrow B'$, if and only if $\text{sim}(B) = \text{sim}(B')$. Thus, when we consider the intersection of the simulation sets of simulation irreducible observables, we only need to select some representative for each post-processing equivalence class.

The natural choice for the representative was presented in [18]: we take it to be the extreme simulation irreducible observable as it has linearly independent indecomposable effects with the minimal number of outcomes in the respective post-processing equivalence class. Furthermore, it was shown that such extreme observable exists in every equivalence class for simulation irreducible observables. We denote the set of extreme simulation irreducible observables by $\mathcal{O}_{irr}^{ext}(\mathcal{S})$ so that

$$\mathcal{T}_3 = \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{sim}(B) = \bigcap_{B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})} \text{sim}(B).$$

Corollary 4. *An observable $A \in \mathcal{O}(\mathcal{S})$ on a state space \mathcal{S} is included in \mathcal{T}_3 if and only if*

$$A_y \in \bigcap_{B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})} \text{cone}(\{B_x\}_{x \in \Omega_B}) \quad \forall y \in \Omega_A. \quad (12)$$

Proof. Let first $A \in \mathcal{T}_3$. By Prop. 7 for all $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$ there exists a post-processing ν^B such that $A = \nu^B \circ B$, i.e.,

$$A_y = \sum_{x \in \Omega_B} \nu_{xy}^B B_x \quad (13)$$

for all $y \in \Omega_A$. Since $\nu_{xy}^B \geq 0$ for all $x \in \Omega_B$, $y \in \Omega_A$ for all $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$, we have that

$$A_y \in \text{cone}(\{B_x\}_{x \in \Omega_B}) \quad (14)$$

for all $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$ for all $y \in \Omega_A$, which proves the necessity part of the claim.

Then let Eq. (12) hold. Thus, for each $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$ there exists positive numbers $\mu_{xy}^B \geq 0$ such that

$$A_y = \sum_{x \in \Omega_B} \mu_{xy}^B B_x$$

for all $y \in \Omega_A$. From the normalization of observables A and B it follows that

$$\sum_{x \in \Omega_B} B_x = u = \sum_{y \in \Omega_A} A_y = \sum_{x \in \Omega_B} \left(\sum_{y \in \Omega_A} \mu_{xy}^B \right) B_x. \quad (15)$$

Since each $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$, we have that each B consists of linearly independent effects B_x [18], so that $\sum_{y \in \Omega_A} \mu_{xy}^B = 1$ for all $x \in \Omega_B$. Thus, we can define post-processings μ^B for each $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$ with elements μ_{xy}^B so that $A \in \text{sim}(B)$ for all $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$. \square

6.3 Example showing that $\mathcal{T}_2 \neq \mathcal{T}_3$

We will present an example of a two-dimensional state space \mathcal{S} , such that there is an observable $A \in \mathcal{O}(\mathcal{S})$ with $A \in \mathcal{T}_3$ but $A \notin \mathcal{T}_2$.

Let

$$\mathcal{S} = \text{conv} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.5 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right),$$

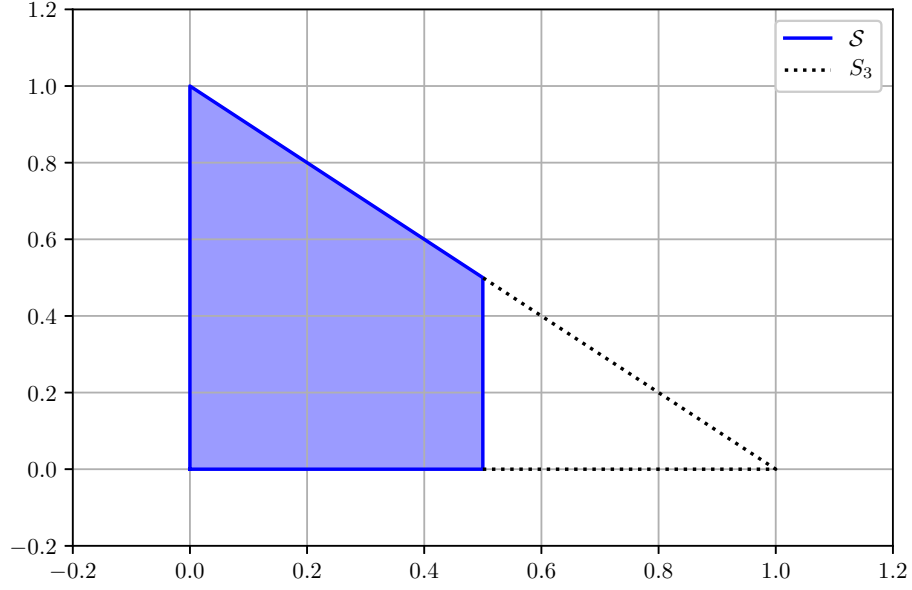


Figure 2: The blue solid line is the boundary of the state space \mathcal{S} used in the example. The black dotted line shows that \mathcal{S} can be considered as a subset of the simplex S_3 .

where the z -coordinate is used to identify \mathcal{S} with a base of a cone. Let

$$S_3 = \text{conv} \left(\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right)$$

be a simplex, then we have $\mathcal{S} \subset S_3$ as shown in Fig. 2.

Let us define functionals x, y, u given as

$$x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad u = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

The points are shown in Fig. 3.

According to Prop. 16 from appendix A there are 4 indecomposable effects corresponding to the 4 maximal faces of \mathcal{S} . They are ξ_1, ξ_2, ξ_3 and $u - \xi_3$, where

$$\xi_1 = y, \quad \xi_2 = u - x - y, \quad \xi_3 = u - 2x.$$

It was shown in [18, Corollary 1] that simulation irreducible observables must consist of indecomposable effects. We are going to find all simulation irreducible observables on \mathcal{S} as we know that $A \in \mathcal{T}_3$ if and only if A is simulable by every simulation irreducible observable; see Prop. 7.

Assume that there would be a simulation irreducible observable with the effects $\alpha_1 \xi_1, \alpha_2 \xi_2, \alpha_3 \xi_3$ and $\alpha'_3 (u - \xi_3)$, where $\alpha_1, \alpha_2, \alpha_3, \alpha'_3 \in \mathbb{R}$, then we must have

$$\alpha_1 \xi_1 + \alpha_2 \xi_2 + \alpha_3 \xi_3 + \alpha'_3 (u - \xi_3) = u$$

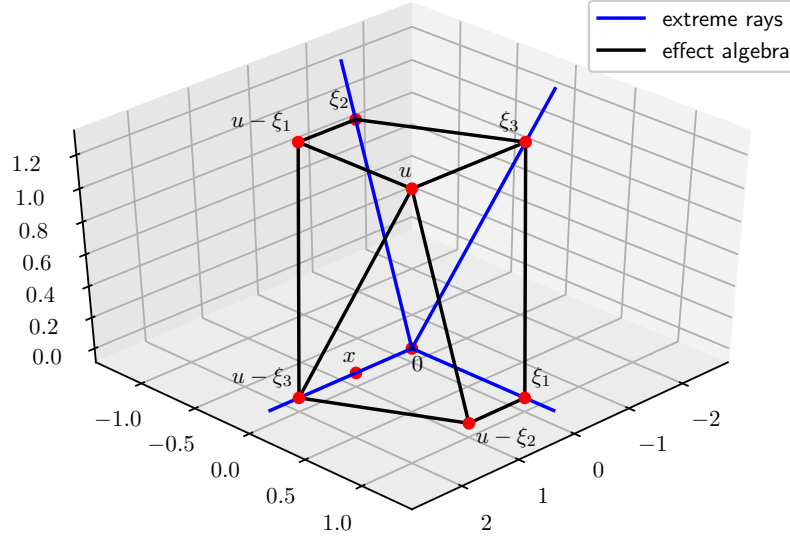


Figure 3: The effect algebra $\mathcal{E}(\mathcal{S})$ used in the example. The black lines represent the wireframe model of $\mathcal{E}(\mathcal{S})$, the blue lines are the extreme rays of the cone of positive functions and the red dots denote the effects that we are using in the example (with $\xi_1 = y$).

which yields

$$\begin{aligned} -\alpha_2 - 2\alpha_3 + 2\alpha'_3 &= 0, \\ \alpha_1 - \alpha_2 &= 0, \\ \alpha_2 + \alpha_3 &= 1. \end{aligned}$$

Since the effects of simulation irreducible observables must be linearly independent, we know that at least one of the coefficients must be equal to zero.

Assuming $\alpha_1 = 0$, we get $\alpha_2 = 0$ and $\alpha_3 = \alpha'_3 = 1$ and we obtain a dichotomic observable \mathbf{B} with effects

$$\begin{aligned} \mathbf{B}_1 &= \xi_3, \\ \mathbf{B}_2 &= u - \xi_3. \end{aligned}$$

Assuming $\alpha_2 = 0$ yields $\alpha_1 = 0$ and $\alpha_3 = \alpha'_3 = 1$, i.e. the same observable \mathbf{B} . Assuming $\alpha_3 = 0$ gives $\alpha_1 = \alpha_2 = 1$ and $\alpha'_3 = \frac{1}{2}$ and gives us a three-outcome observable \mathbf{C} with effects

$$\begin{aligned} \mathbf{C}_1 &= \xi_1, \\ \mathbf{C}_2 &= \xi_2, \\ \mathbf{C}_3 &= x. \end{aligned}$$

Finally assuming $\alpha'_3 = 0$ leads to a contradiction.

Let \mathbf{A} be a dichotomic observable given as

$$\begin{aligned} \mathbf{A}_1 &= x, \\ \mathbf{A}_2 &= u - x. \end{aligned}$$

Note that we have $x = \frac{1}{2}(u - \xi_3)$, which shows that A is simulable by B and we have $x = u - \xi_1 - \xi_2$, which shows that A is simulable by C . This shows that $A \in \mathcal{T}_3$.

We are going to use Prop. 5 to see that $A \notin \mathcal{T}_2$. Assume that $A \in \mathcal{T}_2$, then $A_1 \in \mathcal{ET}_2$, which by Prop. 5 implies also $u - \xi_3 \in \mathcal{ET}_2$ as $u - \xi_3 = 2x$. This would imply that B would be compatible with every other observable, but it is straightforward to see that B is incompatible with C as they are the only two simulation irreducible observables and if they would be compatible, then all of the observables on \mathcal{S} would be compatible. This would in turn yield that \mathcal{S} would have to be simplex [20] which it clearly is not.

An insight into how we obtained this example is provided by the simplex S_3 : ξ_1, ξ_2 and x are effects on the simplex S_3 so that the compatibility of A and C follows. Moreover, the fact that $u - \xi_3 = 2x \geq x$ gives the compatibility of A and B .

7 State spaces satisfying $\mathcal{T}_1 = \mathcal{T}_2 = \mathcal{T}_3$

Next we will consider conditions under which the no-information-without-disturbance principle ($\mathcal{T}_2 = \mathcal{T}_1$) and the no-free-information principle ($\mathcal{T}_3 = \mathcal{T}_1$) hold and when they do not. First we note that, as was mentioned earlier, in general we have that $\mathcal{T}_1 \subseteq \mathcal{T}_2 \subseteq \mathcal{T}_3$ so that if the no-free-information principle holds, and therefore we have that $\mathcal{T}_3 = \mathcal{T}_1$, it follows that also $\mathcal{T}_2 = \mathcal{T}_1$ so that the no-information-without-disturbance principle must hold as well.

7.1 Conditions for $\mathcal{T}_1 = \mathcal{T}_3$

With the help of Prop. 7 we can show the following.

Proposition 8. *The following conditions are equivalent:*

- i) $\mathcal{T}_1 = \mathcal{T}_3$
- ii) $\bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{cone}(\{\mathbf{B}_x\}_{x \in \Omega_B}) = \text{cone}(u)$
- iii) $\bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{conv}(\{\{\mathbf{B}_x\}_{x \in \Omega_B}, o, u\}) = \text{conv}(\{o, u\})$.

Proof. i) \Rightarrow iii): It is clear that $\text{conv}(\{o, u\}) \subseteq \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{conv}(\{\{\mathbf{B}_x\}_x, o, u\})$. Now take

$$e \in \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{conv}(\{\{\mathbf{B}_x\}_x, o, u\})$$

and define a dichotomic observable E with effects $E_+ = e$ and $E_- = u - e$. Since $E_+ \in \text{conv}(\{\{\mathbf{B}_x\}_x, o, u\})$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$, it follows from Prop. 8 in [18] that $E \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$. From Prop. 7 it follows that $E \in \mathcal{T}_3 = \mathcal{T}_1$ so that actually $e \in \text{conv}(\{o, u\})$.

iii) \Rightarrow ii): It is clear that $\text{cone}(u) \subseteq \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{cone}(\{\mathbf{B}_x\}_x)$. Now let us take $g \in \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{cone}(\{\mathbf{B}_x\}_x)$ so that for all $B \in \mathcal{O}_{irr}(\mathcal{S})$ there exists positive real numbers $(\alpha_x^B)_x \subset \mathbb{R}_+$ such that $g = \sum_x \alpha_x^B \mathbf{B}_x$. We denote $\alpha = \sup_{B \in \mathcal{O}_{irr}(\mathcal{S})} \sum_x \alpha_x^B$. If $\alpha = 0$, then $g = o \in \text{cone}(u)$; otherwise we define an effect $f \in \mathcal{E}(\mathcal{S})$ by $f = \frac{1}{\alpha}g$. Now

$$f \in \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{conv}(\{\{\mathbf{B}_x\}_x, o\}) \subseteq \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{conv}(\{\{\mathbf{B}_x\}_x, o, u\}) = \text{conv}(\{o, u\}) \quad (16)$$

so that $f = pu$ for some $p \in (0, 1]$. Thus, $g = \alpha pu \in \text{cone}(u)$.

ii) \Rightarrow i): As noted before, we always have $\mathcal{T}_1 \subseteq \mathcal{T}_3$ so that it suffices to show that $\mathcal{T}_3 \subseteq \mathcal{T}_1$. Thus, take $A \in \mathcal{T}_3$. By Prop. 7, $A \in \mathbf{sim}(B)$ for all $B \in \mathcal{O}_{irr}(\mathcal{S})$ so that for each $B \in \mathcal{O}_{irr}(\mathcal{S})$ there exists a post-processing $\nu^B : \Omega_B \rightarrow \Omega_A$ such that $A_y = \sum_{x \in \Omega_B} \nu_{xy}^B B_x$ for all $y \in \Omega_A$. Since all the post-processing elements are positive for each $B \in \mathcal{O}_{irr}(\mathcal{S})$, we have that $A_y \in \text{cone}(\{B_x\}_{x \in \Omega_B})$ for all $y \in \Omega_A$ and $B \in \mathcal{O}_{irr}(\mathcal{S})$. Thus,

$$A_y \in \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{cone}(\{B_x\}_{x \in \Omega_B}) = \text{cone}(u) \quad (17)$$

for all $y \in \Omega_A$ from which it follows that $A \in \mathcal{T}_1$. □

Proposition 9. *Let \mathcal{S} be a d -dimensional state space. If $|\mathcal{O}_{irr}^{ext}(\mathcal{S})| < \infty$ and all the extreme simulation irreducible observables have $d + 1$ outcomes, then $\mathcal{T}_1 \neq \mathcal{T}_3$.*

Proof. Since \mathcal{S} is d -dimensional (i.e. $\dim(\text{aff}(\mathcal{S})) = d$), the effect space is contained in a $d + 1$ -dimensional vector space. Suppose that, on the contrary $\mathcal{T}_1 = \mathcal{T}_3$. From Prop. 8 it follows then that

$$\bigcap_{B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})} \text{cone}(\{B_x\}_x) = \bigcap_{B \in \mathcal{O}_{irr}(\mathcal{S})} \text{cone}(\{B_x\}_{x \in \Omega_B}) = \text{cone}(u).$$

Since $\dim(\mathcal{V}^*) = d + 1$ and each extreme simulation irreducible observable consists of $d + 1$ linearly independent effects, it follows that $\text{cone}(\{B_x\}_x)$ has a non-empty interior, denoted by $\text{int}(\text{cone}(\{B_x\}_x))$, in \mathcal{V}^* for all $B \in \mathcal{O}_{irr}(\mathcal{S})$. In particular, $u \in \text{int}(\text{cone}(\{B_x\}_x))$ for all $B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$, so that

$$\emptyset = \text{int}(\text{cone}(u)) = \text{int}\left(\bigcap_{B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})} \text{cone}(\{B_x\}_x)\right) = \bigcap_{B \in \mathcal{O}_{irr}^{ext}(\mathcal{S})} \text{int}(\text{cone}(\{B_x\}_x)) \neq \emptyset \quad (18)$$

which is a contradiction. □

Proposition 10. *If there exist at least two post-processing inequivalent dichotomic simulation irreducible observables on \mathcal{S} , then $\mathcal{T}_1 = \mathcal{T}_2 = \mathcal{T}_3$.*

Proof. By the assumption there exist two dichotomic observables $E, F \in \mathcal{O}_{irr}(\mathcal{S})$ such that $E \leftrightarrow F$. Take $A \in \mathcal{T}_3$ so that by Prop. 7 we have that $A \in \mathbf{sim}(E)$ and $A \in \mathbf{sim}(F)$. From Prop. 11 in [18] it follows that $A_x \in \text{conv}(\{E_+, E_-, o, u\})$ and $A_x \in \text{conv}(\{F_+, F_-, o, u\})$ for all $x \in \Omega_A$. Since E and F are inequivalent, it follows that the set $\{u, E_+, F_+\}$ is linearly independent, so that $A_x \in \text{conv}(\{E_+, E_-, o, u\}) \cap \text{conv}(\{F_+, F_-, o, u\}) = \text{conv}(\{o, u\})$ for all $x \in \Omega_A$. Thus, $A \in \mathcal{T}_1$ so that $\mathcal{T}_1 = \mathcal{T}_3$. □

With the previous proposition we can show that the no-free-information principle holds in any point-symmetric state space, i.e., in a state space \mathcal{S} where there exists a state s_0 such that for all $s \in \mathcal{S}$ we have that

$$s' := 2s_0 - s \in \mathcal{S}. \quad (19)$$

This means that for each state s there exists another state s' such that s_0 is an equal mixture of s and s' , i.e., $s_0 = \frac{1}{2}(s + s')$. Point-symmetric state spaces include the classical bit, the qubit and polygon state spaces with even number of vertices.

One can show that the effect space structure is also symmetric for symmetric state spaces. Firstly, all the non-trivial extreme effects are seen to lie on a single affine hyperplane. Namely, if $e \in \mathcal{E}(\mathcal{S})$ is an extreme effect, $e \neq o, u$, there exists a (pure) state $s \in \mathcal{S}$

such that $e(s) = 0$ [10]. For s , there exists another state s' such that $s_0 = \frac{1}{2}(s + s')$ so that $e(s_0) = \frac{1}{2}e(s')$. Similarly there exists a (pure) state $t \in \mathcal{S}$ such that $e(t) = 1$ [10]. For t , we can find t' such that $e(s_0) = \frac{1}{2}(e(t) + e(t')) = \frac{1}{2}(1 + e(t'))$. Combining these two expressions for $e(s_0)$ we find that $e(s') = 1 + e(t')$ from which it follows that $e(t') = 0$ and $e(s') = 1$ so that $e(s_0) = \frac{1}{2}$ for all extreme effects e . Thus, all the non-trivial extreme effects lie on an affine hyperplane determined by the state s_0 .

Secondly, we see that all the non-trivial extreme effects must actually be indecomposable. If $e \in \mathcal{E}(\mathcal{S})$ is an extreme effect, $e \neq o, u$, then we can find some decomposition into indecomposable extreme effects $\{e_i\}_{i=1}^r$ for some $r \in \mathbb{N}$ so that $e = \sum_{i=1}^r \alpha_i e_i$ for some numbers $\{\alpha_i\}_{i=1}^r \subset [0, 1]$ [10]. Since all extreme effects give probability $\frac{1}{2}$ on the state s_0 , we have that $1 = 2e(s_0) = \sum_{i=1}^r \alpha_i$. Since e is extreme, it follows that $r = 1$ so that e is indecomposable.

Thirdly, the convex hull of all the extreme indecomposable effects (that lie on an affine hyperplane) is also point-symmetric: if $e \in \mathcal{E}(\mathcal{S})$ is a non-trivial extreme effect, then $e' := u - e$ is also a non-trivial extreme effect so that $e_0 := \frac{1}{2}u = \frac{1}{2}(e + e')$ acts as the inversion point of the set.

Corollary 5. *In every non-classical point-symmetric state space \mathcal{S} we have $\mathcal{T}_1 = \mathcal{T}_2 = \mathcal{T}_3$.*

Proof. Since \mathcal{S} is non-classical point-symmetric state space, there exists two non-trivial extreme effects e and f such that $e, f \neq o, u$, $e \neq f, u - f$. Namely, if this was not the case, there would be only two non-trivial extreme effects g and g' such that $g' = u - g$ so that $\mathcal{E}^{ext}(\mathcal{S}) = \{o, u, g, u - g\}$ which would mean that the state space would be a classical bit consisting of only two extreme points. We define two dichotomic observables \mathbf{E} and \mathbf{F} by setting $\mathbf{E}_+ = e$, $\mathbf{E}_- = u - e$, $\mathbf{F}_+ = f$ and $\mathbf{F}_- = u - f$. Since the state space is point-symmetric, the extreme effects $e, f, u - e$ and $u - f$ are indecomposable so that together with the fact that $\{e, u - e\}$ and $\{f, u - f\}$ are linearly independent sets it follows [18] that \mathbf{E} and \mathbf{F} are inequivalent dichotomic simulation irreducible observables. The claim follows from Prop. 10. \square

7.2 Alternative characterization of \mathcal{T}_1

Finally, we show that a seemingly different formulation of “free-information” does not lead to a new concept. Consider $\mathbf{T} \in \mathcal{T}_3$ and take an observable $\mathbf{A} \in \mathcal{O}(\mathcal{S})$ such that $\mathbf{A} \notin \mathcal{T}_1$. Since \mathbf{T} is compatible with \mathbf{A} there exists a joint observable $\mathbf{J}_{\mathbf{A}, \mathbf{T}}$ from which both \mathbf{A} and \mathbf{T} can be post-processed from. Since \mathbf{A} is non-trivial and \mathbf{T} is compatible with every other observable, we can ask whether measuring the joint observable $\mathbf{J}_{\mathbf{A}, \mathbf{T}}$ actually gives us any more information than just measuring \mathbf{A} . One way to consider this is to ask whether \mathbf{A} is actually post-processing equivalent to $\mathbf{J}_{\mathbf{A}, \mathbf{T}}$ so that both can be obtained from each other by classically manipulating their outcomes. If this is the case, there is no “free information” to be gained from measuring the joint observable. Thus, we consider one more set of observables:

$$\mathcal{T}_4 = \{\mathbf{T} \in \mathcal{T}_3 \mid \forall \mathbf{A} \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1 : \exists \mathbf{J}_{\mathbf{A}, \mathbf{T}} \in \mathcal{O}(\mathcal{S}) : \mathbf{J}_{\mathbf{A}, \mathbf{T}} \leftrightarrow \mathbf{A}\}.$$

We can show the following.

Proposition 11. $\mathcal{T}_1 = \mathcal{T}_4$.

Proof. Since $\mathcal{T}_1 \subseteq \mathcal{T}_4$ it suffices to show that $\mathcal{T}_4 \subseteq \mathcal{T}_1$. Thus, take $\mathbf{T} \in \mathcal{T}_4$ so that for all $\mathbf{A} \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$ we have that \mathbf{A} is post-processing equivalent with at least one of their joint observables $\mathbf{J}_{\mathbf{A}, \mathbf{T}}$. Thus, $\{\mathbf{A}, \mathbf{T}\} \subseteq \mathbf{sim}(\mathbf{J}_{\mathbf{A}, \mathbf{T}})$ and since $\mathbf{A} \leftrightarrow \mathbf{J}_{\mathbf{A}, \mathbf{T}}$ it follows that $\mathbf{T} \in \mathbf{sim}(\mathbf{A})$ for all $\mathbf{A} \in \mathcal{O}(\mathcal{S}) \setminus \mathcal{T}_1$. From Prop. 6 it follows that $\mathbf{T} \in \mathcal{T}_1$. \square

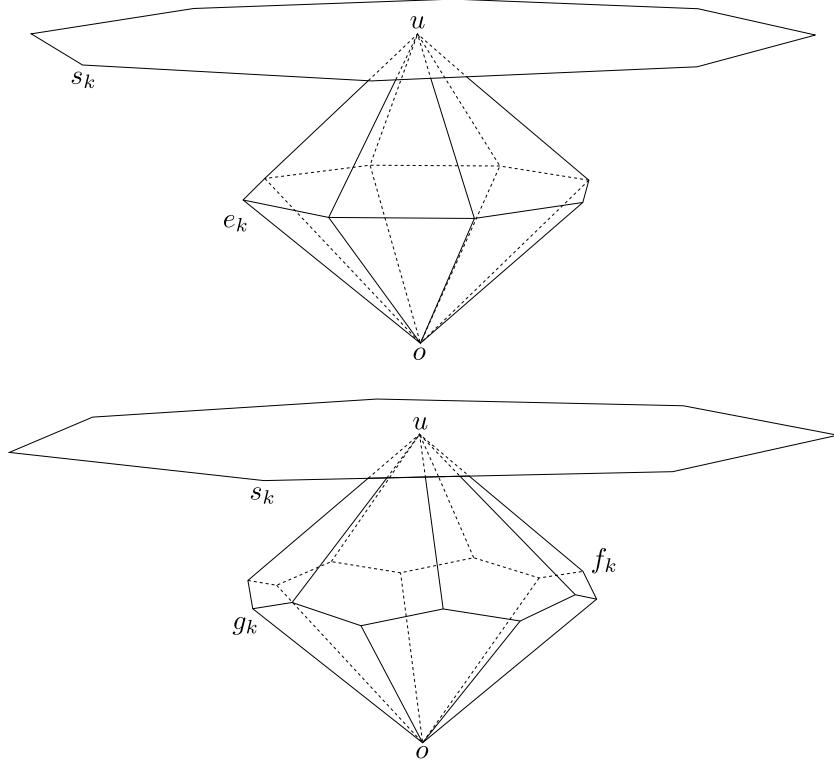


Figure 4: The even and odd polygon state spaces and their effects spaces.

8 Polygon state spaces

8.1 Characterization of polygons

A regular polygon with n vertices in \mathbb{R}^2 , or n -gon, is a convex hull of n points $\{\vec{x}_k\}_{k=1}^n$ such that $\|\vec{x}_k\| = \|\vec{x}_j\|$ and $\vec{x}_k \cdot \vec{x}_{k+1} = \|\vec{x}_k\|^2 \cos(\frac{\pi}{n})$ for all $j, k = 1, \dots, n$. As a state space \mathcal{S}_n , we consider the polygon to be embedded in \mathbb{R}^3 on the $z = 1$ - plane. Thus, we follow the notation of [21] and define the extreme points of \mathcal{S}_n as

$$s_k = \begin{pmatrix} r_n \cos\left(\frac{2k\pi}{n}\right) \\ r_n \sin\left(\frac{2k\pi}{n}\right) \\ 1 \end{pmatrix}, \quad k = 1, \dots, n,$$

where we have defined $r_n = \sec(\frac{\pi}{n})$.

As the polygons are two-dimensional, the effects can also be represented as elements in \mathbb{R}^3 . Hence, we can express each $e \in \mathcal{E}(\mathcal{S}_n)$ as a vector $e = (e_x, e_y, e_z)^T \in \mathbb{R}^3$. With this identification we have that $e(s) = e \cdot s$ for all $e \in \mathcal{E}(\mathcal{S}_n)$ and $s \in \mathcal{S}_n$ where \cdot is the Euclidean dot product. Clearly, we now have the zero effect $o = (0, 0, 0)^T$ and the unit effect $u = (0, 0, 1)^T$.

Depending on the parity of n , the state space may or may not have reflective point symmetry around the middle point $s_0 = (0, 0, 1)^T$. As a result of this, the effect space $\mathcal{E}(\mathcal{S}_n)$ has a different structure for odd and even n . For even n , we find (details for example

in [18]) that the effect space $\mathcal{E}(\mathcal{S}_n)$ has n non-trivial extreme points

$$e_k = \frac{1}{2} \begin{pmatrix} \cos\left(\frac{(2k-1)\pi}{n}\right) \\ \sin\left(\frac{(2k-1)\pi}{n}\right) \\ 1 \end{pmatrix}, \quad k = 1, \dots, n, \quad (20)$$

so that $\mathcal{E}(\mathcal{S}_n) = \text{conv}(\{o, u, e_1, \dots, e_n\})$. All the non-trivial extreme effects lie on a single (hyper)plane determined by those points e such that $e(s_0) = 1/2$.

In the case of odd n , the effect space has $2n$ non-trivial extreme effects

$$g_k = \frac{1}{1+r_n} \begin{pmatrix} \cos\left(\frac{2k\pi}{n}\right) \\ \sin\left(\frac{2k\pi}{n}\right) \\ 1 \end{pmatrix}, \quad f_k = u - g_k \quad (21)$$

for $k = 1, \dots, n$. Now $\mathcal{E}(\mathcal{S}_n) = \text{conv}(\{o, u, g_1, \dots, g_n, f_1, \dots, f_n\})$ and the non-trivial effects are scattered on two different planes determined by all those points g and f such that $g(s_0) = \sigma_n := \frac{1}{1+r_n}$ and $f(s_0) = 1 - \sigma_n = \frac{r_n}{1+r_n}$. The even and odd polygon state spaces and their respective effect spaces are depicted in Figure 4.

In order to give a simple characterization of polygons, let us define functions $\eta_e^n : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $\eta_o^n : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$\begin{aligned} \eta_e^n(\vec{x}) &= \max_{k \in \{1, \dots, n\}} r_n \left[\cos\left(\frac{2\pi k}{n}\right) x + \sin\left(\frac{2\pi k}{n}\right) y \right], \\ \eta_o^n(\vec{x}) &= \eta_e^n\left(R_{\frac{\pi}{n}} \vec{x}\right) = \max_{k \in \{1, \dots, n\}} r_n \left[\cos\left(\frac{(2k-1)\pi}{n}\right) x + \sin\left(\frac{(2k-1)\pi}{n}\right) y \right], \end{aligned}$$

for all $\vec{x} = (x, y)^T \in \mathbb{R}^2$, where

$$R_{\frac{\pi}{n}} = \begin{pmatrix} \cos\left(\frac{\pi}{n}\right) & -\sin\left(\frac{\pi}{n}\right) \\ \sin\left(\frac{\pi}{n}\right) & \cos\left(\frac{\pi}{n}\right) \end{pmatrix}$$

is the rotation matrix with a rotation angle π/n around the origin in \mathbb{R}^2 . We use the notation $\eta_{e/o}^n$ when we consider some properties that hold for both η_e^n and η_o^n .

We see that both $\eta_e^n(\vec{x})$ and $\eta_o^n(\vec{x})$ can be expressed as a maximization over an inner product of \vec{x} and a collection of unit vectors $\vec{b}_{e/o}^{(n,k)}$, i.e.

$$\eta_{e/o}^n(\vec{x}) = r_n \max_{k \in \{1, \dots, n\}} \vec{x} \cdot \vec{b}_{e/o}^{(n,k)}, \quad (22)$$

where we have defined

$$\vec{b}_e^{(n,k)} = \left(\cos\left(\frac{2\pi k}{n}\right), \sin\left(\frac{2\pi k}{n}\right) \right)^T, \quad (23)$$

$$\vec{b}_o^{(n,k)} = \left(\cos\left(\frac{(2k-1)\pi}{n}\right), \sin\left(\frac{(2k-1)\pi}{n}\right) \right)^T. \quad (24)$$

Thus, both η_e^n and η_o^n are polyhedral convex functions [22].

It is straightforward to see that $\eta_{e/o}^n$ satisfy the following properties for all $\vec{x}, \vec{y} \in \mathbb{R}^2$:

- i) $\eta_{e/o}^n(\vec{x}) \geq 0$,
- ii) $\eta_{e/o}^n(\vec{x}) = 0 \Leftrightarrow \vec{x} = \vec{0}$,
- iii) $\eta_{e/o}^n(\vec{x} + \vec{y}) \leq \eta_{e/o}^n(\vec{x}) + \eta_{e/o}^n(\vec{y})$.

Additionally we see that also the following is satisfied for all $x \in \mathbb{R}^2$:

- iv) $\eta_{e/o}^n(\alpha\vec{x}) = \alpha\eta_e^n(\vec{x})$ for all $\alpha \geq 0$.

Thus, both η_e^n and η_o^n almost satisfy the requirements of a norm; the only missing property is the requirement for a reflective point symmetry, i.e. $\eta_{e/o}^n(-\vec{x}) = \eta_{e/o}^n(\vec{x})$ for all $\vec{x} \in \mathbb{R}^2$. For even n , however, it is easy to confirm that both η_e^n and η_o^n are point symmetric so that they are norms on \mathbb{R}^2 . Similarly for odd n it is easy to see that the point symmetry does not hold.

Even though for general n the functions $\eta_{e/o}^n$ do not define a norm on \mathbb{R}^2 , we can still use them to define different sized polygons. As continuous polyhedral convex functions, η_e^n and η_o^n have closed polyhedral level sets

$$B_{e/o}^n(r) = \{\vec{x} \in \mathbb{R}^2 \mid \eta_{e/o}^n(\vec{x}) \leq r\}$$

which we will show to give rise to the polygons.

First of all, we see that the level sets $B_{e/o}^n(r)$ are bounded so that they actually describe polytopes: When we express $\vec{x} \in \mathbb{R}^2$ in its polar form $\vec{x} = (x, y)^T = \|\vec{x}\| (\cos(\theta), \sin(\theta))^T$, we have

$$\eta_e^n(\vec{x}) = r_n \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{2\pi k}{n} - \theta\right), \quad (25)$$

$$\eta_o^n(\vec{x}) = r_n \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{(2k-1)\pi}{n} - \theta\right). \quad (26)$$

Considering η_e^n first, we see that since the angles $\frac{2k\pi}{n}$ are an angle $\frac{2\pi}{n}$ apart from each other for consecutive k 's and since the maximization of cosine actually minimizes the angle $\frac{2\pi k'}{n} - \theta$, for the $k' \in \{1, \dots, n\}$ which minimizes the angle we have $\frac{2\pi k'}{n} - \theta \leq \frac{\pi}{n}$ so that $\cos\left(\frac{2\pi k'}{n} - \theta\right) \geq \cos\left(\frac{\pi}{n}\right)$. The same arguments hold for η_o^n as well so if $\vec{x} \in B_{e/o}^n(r)$ for some $r > 0$, then

$$\eta_{e/o}^n(\vec{x}) \leq r \Rightarrow \|\vec{x}\| \leq \frac{r}{r_n \cos\left(\frac{\pi}{n}\right)} = r. \quad (27)$$

Hence, the level sets $B_{e/o}^n(r)$ are bounded so together with being closed it means that they are compact (convex) polytopes for all $r > 0$. Furthermore, since $B_{e/o}^n(r)$ is polyhedral in \mathbb{R}^2 , it is a finite intersection of half-spaces in \mathbb{R}^2 so that it must have at most n extreme points [22].

The functions η_e^n and η_o^n have the following connection:

$$\eta_{e/o}^n(\vec{x}) \leq r_n \eta_{o/e}^n(\vec{x}) \quad (28)$$

for all $\vec{x} \in \mathbb{R}^2$ and $r \geq 0$. This can be seen using the expressions from (25) and (26); for

example

$$\begin{aligned}
\eta_o^n(\vec{x}) &= r_n \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{(2k-1)\pi}{n} - \theta\right) = r_n^2 \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{(2k-1)\pi}{n} - \theta\right) \cos\left(\frac{\pi}{n}\right) \\
&= \frac{r_n^2 \|\vec{x}\|}{2} \max_{k \in \{1, \dots, n\}} \left[\cos\left(\frac{2(k-1)\pi}{n} - \theta\right) + \cos\left(\frac{2k\pi}{n} - \theta\right) \right] \\
&\leq \frac{r_n}{2} \left[r_n \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{2(k-1)\pi}{n} - \theta\right) + r_n \|\vec{x}\| \max_{k \in \{1, \dots, n\}} \cos\left(\frac{2k\pi}{n} - \theta\right) \right] \\
&= r_n \eta_e^n(\vec{x}).
\end{aligned}$$

Let us consider the specific level set $B_o^n(r_n)$. For each $k \in \{1, \dots, n\}$, we define $\vec{s}_k = \left(r_n \cos\left(\frac{2k\pi}{n}\right), r_n \sin\left(\frac{2k\pi}{n}\right)\right)^T$ so that $s_k = (\vec{s}_k, 1)^T$. It is easy to see that $\eta_o^n(\vec{s}_k) = r_n$ so that $\vec{s}_k \in B_o^n(r_n)$ for all $k = 1, \dots, n$. Furthermore, we have that $\|\vec{s}_k\| = r_n$ for all k so that each \vec{s}_k lies on a circle of radius r_n centered at the origin. This shows that \vec{s}_k is extreme in $B_o^n(r_n)$ for all $k = 1, \dots, n$, since a non-trivial convex decomposition for \vec{s}_k would contradict the fact that $\|\vec{x}\| \leq r_n$ for all $\vec{x} \in B_o^n(r_n)$. This, combined with the fact that $B_o^n(r_n)$ has at most n extreme points, shows that the extreme points of $B_o^n(r_n)$ are exactly the vectors \vec{s}_k for all $k = 1, \dots, n$. Hence, $s = (\vec{s}, 1) \in \mathcal{S}_n$ if and only if $\vec{s} \in B_o^n(r_n)$.

By similar arguments, we see that also $B_e^n(r)$ is a regular polygon whose extreme points are rotated and scaled from \vec{s}_k . For example, in the case of even n , we see that the effects lying on the hyperplane that contains all the non-trivial extreme effects can be characterized in terms of $B_e^n(r)$; namely, $e = \left(\vec{e}, \frac{1}{2}\right)^T \in \text{conv}(\{e_1, \dots, e_n\})$ if and only if $\vec{e} \in B_e^n\left(\frac{1}{2}\right)$. Similarly for odd polygons we have that $g = (\vec{g}, \sigma_n)^T \in \text{conv}(\{g_1, \dots, g_n\})$ if and only if $\vec{g} \in B_o^n(\sigma_n)$.

Hence, we can characterize (both the odd and even) polygon state spaces with the polyhedral functions η_o^n :

$$\mathcal{S}_n = \{(\vec{s}, 1)^T \in \mathbb{R}^3 \mid \eta_o^n(\vec{s}) \leq r_n\}. \quad (29)$$

Furthermore, for even n we have that

$$\text{conv}(\{e_1, \dots, e_n\}) = \left\{ \left(\vec{e}, \frac{1}{2}\right)^T \in \mathbb{R}^3 \mid \eta_e^n(\vec{e}) \leq \frac{1}{2} \right\}, \quad (30)$$

and similarly for odd n

$$\text{conv}(\{g_1, \dots, g_n\}) = \left\{ (\vec{g}, \sigma_n)^T \in \mathbb{R}^3 \mid \eta_o^n(\vec{g}) \leq \sigma_n \right\}. \quad (31)$$

In both cases, the above sets serve as a compact bases for the positive dual cones in \mathbb{R}^3 .

8.2 Characterization of \mathcal{T}_2

The analysis of \mathcal{T}_2 on polygon state spaces is straight-forward. By Thm. 1, we can have $\mathcal{T}_2 \neq \mathcal{T}_1$ if and only if the state space can be represented as a (non-trivial) direct sum of state spaces such that some non-trivial observable takes constant values for each effect on each summand of the direct sum. Since polygons are 2-dimensional state spaces embedded in \mathbb{R}^3 , by Cor. 3 the state space can be represented as a non-trivial direct sum only in the case when $n = 3$. Thus, if $n = 3$ then the state space is a simplex and $\mathcal{T}_2 = \mathcal{O}(\mathcal{S}_3)$, and in all other cases we have $\mathcal{T}_1 = \mathcal{T}_2$.

8.3 Characterization of \mathcal{T}_3

The post-processing equivalence classes of simulation irreducible observables on polygon state spaces were characterized in [18] where it was found that for an n -gon state space there exists m dichotomic and $\frac{1}{3}m(m-1)(m-2)$ trichotomic extreme simulation irreducible observables when $n = 2m$ for some $m \in \mathbb{N}$ (even polygons) and $\frac{1}{6}m(m+1)(2m+1)$ trichotomic extreme simulation irreducible observables when $n = 2m+1$ for some $m \in \mathbb{N}$ (odd polygons).

For even polygons with $n = 2m$ where $m \geq 2$, there exists at least two inequivalent dichotomic simulation irreducible observables, so by Prop. 10 the set \mathcal{T}_3 coincides with the set of trivial observables.

For odd polygon state spaces we see that the extreme simulation irreducible observables have the same number of outcomes as the dimension of the effect space, so given that there are a finite number of them, it follows from Prop. 9 that $\mathcal{T}_3 \neq \mathcal{T}_1$. We continue to give a characterization of \mathcal{T}_3 for the odd polygon state spaces.

Let \mathcal{S}_n be an odd polygon state space so that $n = 2m+1$ for some $m \in \mathbb{N}$. There are $q_m := \frac{1}{6}m(m+1)(2m+1)$ extreme simulation irreducible observables that generate the cones generated by all the simulation irreducible observables. By using some enumeration $\mathbf{B}^{(1)}, \dots, \mathbf{B}^{(q_m)}$ for these observables, we have that $\mathcal{O}_{irr}^{ext}(\mathcal{S}_n) = \{\mathbf{B}^{(i)}\}_{i=1}^{q_m}$ so that for an observable $A \in \mathcal{O}(\mathcal{S}_n)$ we have

$$A \in \mathcal{T}_3 \quad \Leftrightarrow \quad A_x \in \bigcap_{j=1}^{q_m} \text{cone} \left(\{\mathbf{B}_x^{(j)}\}_{x \in \Omega_{\mathbf{B}^{(j)}}} \right) \quad \forall x \in \Omega_A.$$

We can show that there are certain extreme simulation irreducible observables that are enough to characterize the above intersection. Let $\mathbf{B} \in \mathcal{O}_{irr}^{ext}(\mathcal{S})$. Since for all $k \in \{1, 2, 3\}$ the effects \mathbf{B}_k are indecomposable, for each $k \in \{1, 2, 3\}$ there exists $0 < c_k \leq 1$ and effect $g_{i_k} \in \{g_1, \dots, g_{2m+1}\}$ such that $\mathbf{B}_k = c_k g_{i_k}$. We see that we only need to consider the case when $i_k \in \{j, j+m, j+m+1\}$ for all $k \in \{1, 2, 3\}$ for some $j \in \{1, \dots, 2m+1\}$, where the addition of the indices is taken modulo $2m+1$.

Proposition 12. *An observable $A \in \mathcal{O}(\mathcal{S}_{2m+1})$ on an odd polygon state space \mathcal{S}_{2m+1} is in \mathcal{T}_3 if and only if*

$$A_x \in \bigcap_{i=1}^{2m+1} \text{cone}(\{g_i, g_{i+m}, g_{i+m+1}\}) \quad \forall x \in \Omega_A.$$

The complete proof of the proposition can be found in the appendix but one can easily convince oneself by looking at Fig. 5 which shows the case of heptagon effect space. For each $\mathbf{B} \in \mathcal{O}_{irr}^{ext}(\mathcal{S}_n)$ we can consider the base of the cone $\text{cone}(\{\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3\})$ on the plane containing the indecomposable extreme effects $\{g_i\}_{i=1}^n$, where the base takes the form of a triangle that contains the middle point $\sigma_n u$. We can see that in order to characterize the intersection of such cones, it is enough to consider the intersection of their respective bases, or triangles containing $\sigma_n u$, equivalently. In the left of Fig. 5, the bases (coloured as blue and red) of two extreme simulation irreducible observables are shown with the whole effects space. On the right is depicted all the triangles (formed by dashed lines) of all the bases on the plane with the blue and red bases from the left figure also shown on the right. We see that the base of the intersection of the cones (darker blue area) is characterized by triangles with vertices g_i, g_{i+m} and g_{i+m+1} (like the blue triangle) so that their intersection is always contained in the intersection of other triangles (like the red triangle).

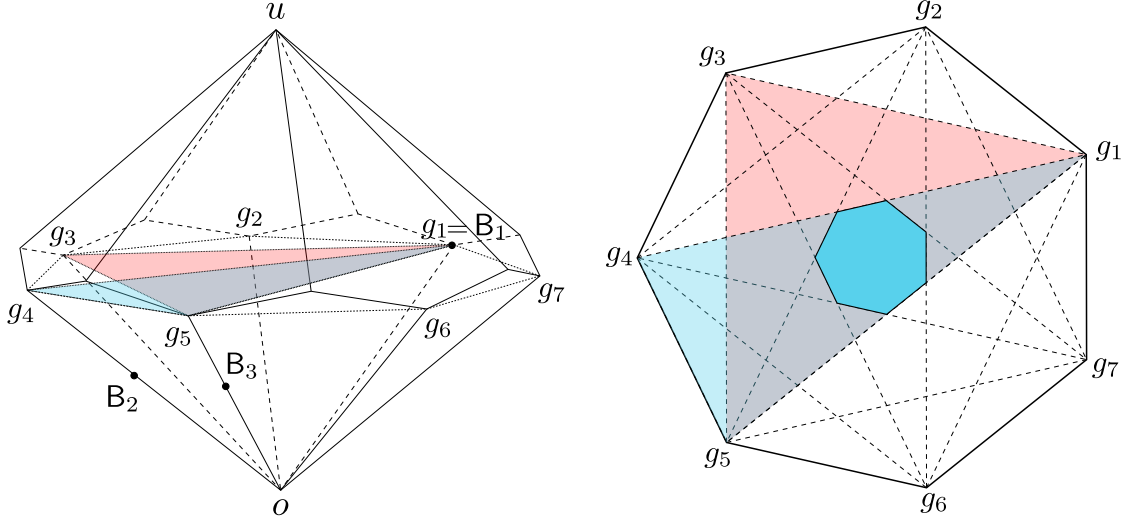


Figure 5: Simulation irreducible observable B on the heptagon state space with $B_1 = g_1$, $B_2 = 0.555g_4$ and $B_3 = 0.555g_5$. The base of the cone generated by the effects of B forms a (blue) triangle on the base of the positive cone (left). The intersection of the bases of all the simulation irreducible observables forms another polygon (right).

We are going to proceed with finding the base of $\bigcap_{i=1}^{2m+1} \text{cone}(\{g_i, g_{i+m}, g_{i+m+1}\})$ by identifying the extreme points of the base $\bigcap_{i=1}^{2m+1} \text{conv}(\{g_i, g_{i+m}, g_{i+m+1}\})$. Let us denote

$$L_i = \text{conv}(\{g_i, g_{i+m}\})$$

and

$$C_m = \bigcap_{i=1}^{2m+1} \text{conv}(\{g_i, g_{i+m}, g_{i+m+1}\}).$$

We will approach the problem as follows: at first, we will identify that C_m must be a polygon itself by looking at its relation with the line segments L_i . Then we will find the form of the extreme points of C_m and in the end we will identify them. During the calculations we will work only in the 2-dimensional vector space given by $\text{aff}(\{g_i\}_{i=1}^{2m+1})$.

It is very useful to realize that L_i generate hyperplanes in \mathbb{R}^2 and that C_m is an intersection of the halfspaces corresponding to the hyperplanes L_i that contain the point 0. It follows that we must have $L_i \cap C_m \neq \emptyset$, $\forall i \in \{1, \dots, 2m+1\}$, otherwise there would be hyperplanes separating L_i and C_m , which is a contradiction with C_m being given as an intersection of halfspace corresponding to L_i . Since there are only $2m+1$ different line segments L_i it follows that C_m must have exactly $2m+1$ edges and from the symmetry it also follows that C_m must be a polygon. Now the only thing we need to do is to identify the extreme points of C_m .

Since the line segments L_i must intersect C_m it follows that the extreme points of C_m must correspond to the intersections of these line segments. Let us denote

$$x_{i,j} = L_i \cap L_{i+j}$$

where $j \in \{1, \dots, m\}$, where if $i+j \geq 2m+1$, then we take $(i+j) \bmod (2m+1)$. Also note that considering $j \geq m+1$ would be redundant. The next key step is to characterize the relation of $x_{i,j}$ and C_m . We can show the following.

Lemma 4. $x_{i,1}$ are the extreme points of C .

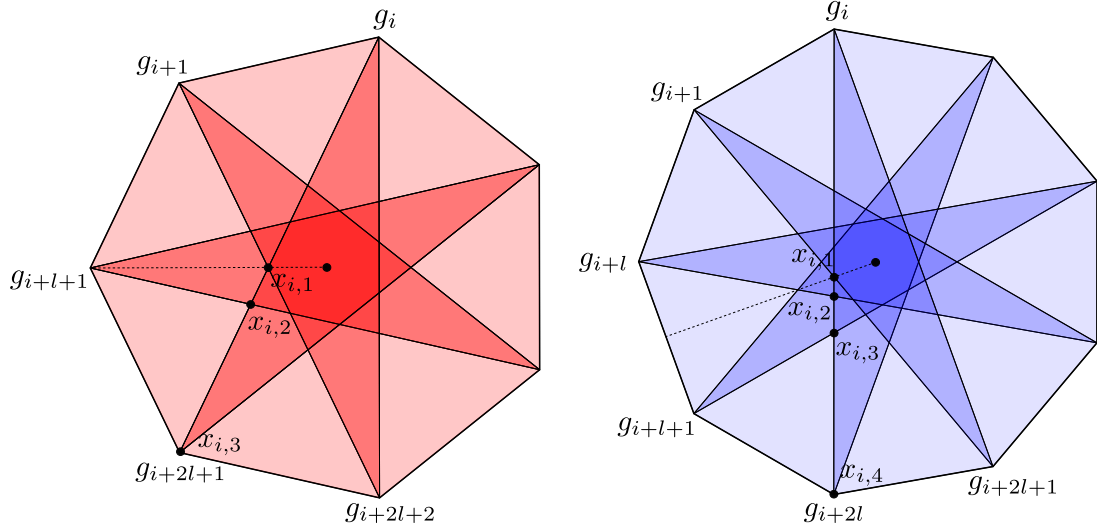


Figure 6: The points $\{x_{i,j}\}_{j=1}^m$ for a fixed i and the orientation of the inner polygon for odd (left, $m = 3$) and even (right, $m = 4$) m .

Again, the complete proof of the lemma can be found in the appendix, but one can easily convince oneself by looking at Fig. 6, where the points $\{x_{i,j}\}_{j=1}^m$ are depicted for a fixed i in the case of a heptagon (left) and nonagon (right) state space.

We are almost ready to move on to the complete characterization of \mathcal{T}_3 in odd polygon theories in terms of the previously defined $\eta_{e/o}^n$ functions. We will still make a few remarks on the inner polygons C_m .

Let $n = 2m + 1$. We will consider separately, although analogously, the cases of even and odd m . This is because of the orientation of the inner polygon C_m with respect to the outer polygon $\text{conv}(\{g_1, \dots, g_n\})$. To show the difference between even and odd m , let us consider the intersection point of the boundary of the outer polygon and the half-line through an extreme point $x_{i,1}$ of the inner polygon emanating from the centroid $(0, 0, \sigma_n)^T$. If this intersection point is also an extreme point of the outer polygon, then both the inner and outer polygons are similarly oriented; otherwise they are differently oriented.

As $x_{i,1} = L_i \cap L_{i+1} = \text{conv}(\{g_i, g_{i+m}\}) \cap \text{conv}(\{g_{i+1}, g_{i+m+1}\})$, it is clear that the half-line through $x_{i,1}$ that emanates from the centroid meets the boundary of the outer polygon at some of the line segments $\text{conv}(\{g_{i+1}, g_{i+2}\}), \dots, \text{conv}(\{g_{i+m-1}, g_{i+m}\})$.

For even m , i.e., for $m = 2l$ for some $l \in \mathbb{N}$, there exists an even number $2(l - 1)$ of vertices g_j between the vertices g_{i+1} and g_{i+m} so that there is an odd number of such edges. From the symmetry it follows that for even m , the intersection point must lie in the middle of the midmost edge $\text{conv}(\{g_{i+l}, g_{i+l+1}\})$. Thus, for even m , the inner polygon C_m is differently oriented with respect to the outer polygon $\text{conv}(\{g_1, \dots, g_n\})$.

By contrast, for odd m , i.e., for $m = 2l + 1$ for some $l \in \mathbb{N}$, there exists an even number of such edges, which together with the symmetry of the situation tells us that now the intersection point is exactly one of the vertices of the outer polygon, namely g_{i+l+1} . Thus, for odd m , the inner polygon is similarly oriented to the outer polygon. The orientations of the inner polygon for odd and even m are depicted in Fig. 6.

As we saw in the beginning of the section, the orientation of the polygon can also be characterized with the $\eta_{e/o}^n$ functions. Thus, in the characterization of \mathcal{T}_3 we must use either η_e^n or η_o^n depending on the parity of m .

Proposition 13. An observable $A \in \mathcal{O}(\mathcal{S}_{2m+1})$ with effects $A_x = \alpha_x(\vec{a}_x, \sigma_{2m+1})^T$ for all $x \in \Omega_A$ is compatible with every other observable if and only if for all $x \in \Omega_A$

$$\eta_e^n(\vec{a}_x) \leq \sigma_n r_n \sin\left(\frac{\pi}{2n}\right)$$

if $m = 2l$ for some $l \in \mathbb{N}$, or

$$\eta_o^n(\vec{a}_x) \leq \sigma_n r_n \sin\left(\frac{\pi}{2n}\right)$$

if $m = 2l + 1$ for some $l \in \mathbb{N}$.

Proof. By Prop. 6 it follows that $A \in \mathcal{T}_3$ if and only if $(\vec{a}_x, \sigma_n)^T \in C_n$ for all $x \in \Omega_A$, and from Lemma 4 we know that the $x_{i,1} = (\vec{x}_{i,1}, \sigma_n)^T$ are the extreme points of C_n . Thus, if we show that $\|\vec{x}_{i,1}\| = \eta_{e/o}^n(\vec{x}_{i,1}) = \sigma_n r_n \sin\left(\frac{\pi}{2n}\right)$, it follows that $C_n = \left\{(\vec{x}, \sigma_n)^T \in \mathbb{R}^3 \mid \eta_{e/o}^n(\vec{x}) \leq \sigma_n r_n \sin\left(\frac{\pi}{2n}\right)\right\}$ which will prove the claim.

From $x_{i,j} = L_i \cap L_{i+j}$ we have that $x_{i,1} = \lambda_1 g_i + (1 - \lambda_1)g_{i+m}$, where $\lambda_1 = 1 - \frac{1}{2}r_n = \frac{3\sigma_n - 1}{2\sigma_n}$. By using (rather a lot of) trigonometric identities we find that

$$x_{k,1} = \begin{pmatrix} -\frac{\sin\left(\frac{\pi}{2n}\right)}{1 + \cos\left(\frac{\pi}{n}\right)} \sin\left(\frac{(4k+1)\pi}{2n}\right) \\ \frac{\sin\left(\frac{\pi}{2n}\right)}{1 + \cos\left(\frac{\pi}{n}\right)} \cos\left(\frac{(4k+1)\pi}{2n}\right) \\ \sigma_n \end{pmatrix},$$

so that

$$\vec{x}_{k,1} = \sigma_n r_n \sin\left(\frac{\pi}{2n}\right) \begin{pmatrix} -\sin\left(\frac{(4k+1)\pi}{2n}\right) \\ \cos\left(\frac{(4k+1)\pi}{2n}\right) \end{pmatrix},$$

from which it is easy to see that $\|\vec{x}_{k,1}\| = \sigma_n r_n \sin\left(\frac{\pi}{2n}\right)$ for all $k \in \{1, \dots, n\}$.

We also see that (the simplified expressions of) $\eta_e^n(\vec{x}_{k,1})$ and $\eta_o^n(\vec{x}_{k,1})$ then read as

$$\eta_e^n(\vec{x}_{k,1}) = \sigma_n r_n^2 \sin\left(\frac{\pi}{2n}\right) \max_{j \in \{1, \dots, n\}} \sin\left(\frac{(4j - 4k - 1)\pi}{2n}\right) \quad (32)$$

and

$$\eta_o^n(\vec{x}_{k,1}) = \sigma_n r_n^2 \sin\left(\frac{\pi}{2n}\right) \max_{j \in \{1, \dots, n\}} \sin\left(\frac{(4j - 4k - 3)\pi}{2n}\right). \quad (33)$$

In both cases the maximum is attained when the expression inside the sine is closest to $\pi/2$. Now depending on the parity of m , this happens for different values of j resulting in different expressions. For $m = 2l$ for some $l \in \mathbb{N}$, we find that the maximum in Eq. (32) is attained for $j \in \{k+l, k+l+1\}$ and similarly the maximum in Eq. (33) is attained for $j = k+l+1$ so that for this case we have

$$\begin{aligned} \eta_e^{4l+1}(\vec{x}_{k,1}) &= \sigma_{4l+1} r_{4l+1}^2 \sin\left(\frac{\pi}{2(4l+1)}\right) \sin\left(\frac{(4l-1)\pi}{2(4l+1)}\right) \\ &= \sigma_{4l+1} r_{4l+1} \sin\left(\frac{\pi}{2(4l+1)}\right) = \|\vec{x}_{k,1}\|. \end{aligned}$$

However, for $m = 2l + 1$ for some $l \in \mathbb{N}$ we have that the maximum in Eq. (32) is attained for $j = k + l + 1$ and similarly the maximum in Eq. (33) is attained for $j \in \{k + l, k + l + 1\}$ so that for this case we have

$$\begin{aligned}\eta_o^{4l+3}(\vec{x}_{k,1}) &= \sigma_{4l+3} r_{4l+3}^2 \sin\left(\frac{\pi}{2(4l+3)}\right) \sin\left(\frac{(4l+1)\pi}{2(4l+3)}\right) \\ &= \sigma_{4l+3} r_{4l+3} \sin\left(\frac{\pi}{2(4l+3)}\right) = \|\vec{x}_{k,1}\|.\end{aligned}$$

□

8.4 Noise content

The noise content $w(\mathbf{A}; \mathcal{N})$ of an observable $\mathbf{A} \in \mathcal{O}(\mathcal{S})$ on a state space \mathcal{S} with respect to a noise set $\mathcal{N} \subset \mathcal{O}(\mathcal{S})$ is defined [13] as

$$w(\mathbf{A}; \mathcal{N}) = \sup\{\lambda \in [0, 1] \mid \mathbf{A} = \lambda \mathbf{N} + (1 - \lambda) \mathbf{B} \text{ for some } \mathbf{N} \in \mathcal{N} \text{ and } \mathbf{B} \in \mathcal{O}(\mathcal{S})\}.$$

When describing noisy observables, the noise is most commonly added externally to an observable, but the noise content describes the amount of noise that an observable already has intrinsically. Usually the noise set is taken to be the set of trivial observables \mathcal{T}_1 .

Examining Prop. 13 more closely, the set \mathcal{T}_3 seems to be quite noisy in the sense that the effects of observables in \mathcal{T}_3 are scattered quite closely to the trivial effects on the line segment $\text{conv}(\{o, u\})$. Our aim is to show this remark quantitatively by showing that an observable that is compatible with every other observable must have a quite high noise content with respect to the trivial observables. We also show that an observable with a high enough noise content is indeed compatible with every other observable on odd polygon state spaces.

For the noise set $\mathcal{N} = \mathcal{T}_1$, the noise content of an observable $\mathbf{A} \in \mathcal{O}(\mathcal{S})$ takes a rather simple form [13]:

$$w(\mathbf{A}; \mathcal{T}_1) = \sum_{x \in \Omega_{\mathbf{A}}} \min_{s \in \mathcal{S}} \mathbf{A}_x(s), \quad (34)$$

and furthermore if the state space is a polytope (as is in the case of polygons), we have that

$$w(\mathbf{A}; \mathcal{T}_1) = \sum_{x \in \Omega_{\mathbf{A}}} \min_{s \in \mathcal{S}^{ext}} \mathbf{A}_x(s), \quad (35)$$

where \mathcal{S}^{ext} denotes the set of extreme points of \mathcal{S} .

We start by making a connection between $\min_{s \in \mathcal{S}} \mathbf{A}_x(s)$ and $\eta_o^n(\vec{a}_x)$. As before, for each effect \mathbf{A}_x there exists $\alpha_x > 0$ such that $\mathbf{A}_x = \alpha_x a_x$ for some $a_x = (\vec{a}_x, \sigma_n)^T$, where $\vec{a}_x \in \mathbb{R}^2$. Since $a_x \in \text{conv}(\{g_1, \dots, g_n\})$ for all $x \in \Omega_{\mathbf{A}}$, we have that for all $x \in \Omega_{\mathbf{A}}$ there exists $\lambda_x \in [0, 1]$ such that $a_x = \lambda_x h_x + (1 - \lambda_x) \sigma_n u$ for some

$$h_x \in \partial \text{conv}(\{g_1, \dots, g_n\}) = \{(\vec{g}, \sigma_n)^T \in \text{conv}(\{g_1, \dots, g_n\}) \mid \eta_o^n(\vec{g}) = \sigma_n\}.$$

We note that since h_x lies on the boundary of the convex hull of the indecomposable effects, for all $x \in \Omega_{\mathbf{A}}$, there exists $i_x \in \{1, \dots, n\}$ such that $h_x \in \text{conv}(\{g_{i_x}, g_{i_x+1}\})$. Since g_{i_x} and g_{i_x+1} are indecomposable, by Prop. 16 they give zero for some maximal faces G_{i_x} and G_{i_x+1} of \mathcal{S}_n . Furthermore, it is easy to see that they must be adjacent maximal faces so that there exists an extreme state $s_{i_x} \in \mathcal{S}_n$ such that $h_x(s_{i_x}) = 0$. Thus,

$$\begin{aligned}\min_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s) &= \alpha_x \min_{s \in \mathcal{S}_n^{ext}} [\lambda_x h_x(s) + (1 - \lambda_x) \sigma_n u(s)] = \alpha_x \lambda_x \min_{s \in \mathcal{S}_n^{ext}} h_x(s) + \alpha_x (1 - \lambda_x) \sigma_n \\ &= \alpha_x \lambda_x h_x(s_{i_x}) + \alpha_x (1 - \lambda_x) \sigma_n = \alpha_x (1 - \lambda_x) \sigma_n\end{aligned}$$

for all $x \in \Omega_A$. If we denote $h_x = (\vec{h}_x, \sigma_n)^T$, we then see that $\vec{a}_x = \lambda_x \vec{h}_x$ and

$$\eta_o^n(\vec{a}_x) = \lambda_x \eta_o^n(\vec{h}_x) = \lambda_x \sigma_n = \sigma_n - \frac{1}{\alpha_x} \min_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s)$$

for all $x \in \Omega_A$. Thus, $\min_{s \in \mathcal{S}^{ext}} \mathbf{A}_x(s) = \alpha_x [\sigma_n - \eta_o^n(\vec{a}_x)]$ for all $x \in \Omega_A$.

We can now show the following.

Proposition 14. *Let $\mathbf{A} \in \mathcal{O}(\mathcal{S}_n)$ be an observable on an odd polygon state space \mathcal{S}_n with effects $\mathbf{A}_x = \alpha_x(\vec{a}_x, \sigma_n)$ for all $x \in \Omega_A$. If $\mathbf{A} \in \mathcal{T}_3$, then*

$$w(\mathbf{A}; \mathcal{T}_1) \geq 1 - r_n \sin\left(\frac{\pi}{2n}\right) \quad (36)$$

if $n = 4l + 3$ for some $l \in \mathbb{N}$, or

$$w(\mathbf{A}; \mathcal{T}_1) \geq 1 - r_n^2 \sin\left(\frac{\pi}{2n}\right) \quad (37)$$

if $n = 4l + 1$ for some $l \in \mathbb{N}$.

Proof. As was established above, we have that $\min_{s \in \mathcal{S}^{ext}} \mathbf{A}_x(s) = \alpha_x(\sigma_n - \eta_o^n(\vec{a}_x))$.

For $n = 4l + 3$, we have from Prop. 13 that $\eta_o^n(\vec{a}_x) \leq r_n \sigma_n \sin\left(\frac{\pi}{2n}\right)$ for all $x \in \Omega_A$ so that

$$\begin{aligned} w(\mathbf{A}; \mathcal{T}_1) &= \sum_{x \in \Omega_A} \min_{s \in \mathcal{S}^{ext}} \mathbf{A}_x(s) = \sum_{x \in \Omega_A} \alpha_x(\sigma_n - \eta_o^n(\vec{a}_x)) \\ &\geq \sum_{x \in \Omega_A} \alpha_x \sigma_n \left(1 - r_n \sin\left(\frac{\pi}{2n}\right)\right) \\ &= 1 - r_n \sin\left(\frac{\pi}{2n}\right), \end{aligned}$$

where on the last line we have used the fact that $\sum_{x \in \Omega_A} \alpha_x = 1/\sigma_n$ which follows from the normalization of \mathbf{A} .

For $n = 4l + 1$, we have from Prop. 13 that $\eta_e^n(\vec{a}_x) \leq r_n \sigma_n \sin\left(\frac{\pi}{2n}\right)$ for all $x \in \Omega_A$. From Eq. (28) we get that $\eta_o^n(\vec{a}_x) \leq r_n \eta_e^n(\vec{a}_x)$ for all $x \in \Omega_A$ so that from similar calculation as above we get that $w(\mathbf{A}; \mathcal{T}_1) \geq 1 - r_n^2 \sin\left(\frac{\pi}{2n}\right)$. \square

The lower bounds of the noise content from the previous proposition for the first few polygons are presented in Table 1. We see that for $n = 3$, i.e., when the state space is classical, Eq. (36) gives the trivial lower bound zero, but already for the pentagon ($n = 5$) Eq. (37) shows that the noise content of an observable in \mathcal{T}_3 must be more than 1/2. We see that as the number of vertices in the polygons increase, so does the noise content of observables in \mathcal{T}_3 for both Eq. (36) and (37). In the limit where $n \rightarrow \infty$ the right hand sides (R.H.S.) of both equations give the limit 1, so that the observables in \mathcal{T}_3 become trivial. Indeed, as the number of vertices approaches infinity, the state space becomes shaped like a disc, which is seen to be a point-symmetric state space so that by Cor. 5 we have $\mathcal{T}_1 = \mathcal{T}_3$.

From the other point of view, we can ask if sufficiently noisy observables are necessarily compatible with every other observable. For that, let us consider the binarizations of an observable $\mathbf{A} \in \mathcal{O}(\mathcal{S}_n)$, i.e., binary observables $\hat{\mathbf{A}}^{(x)}$ with effects $\hat{\mathbf{A}}_+^{(x)} = \mathbf{A}_x$ and $\hat{\mathbf{A}}_-^{(x)} = u - \mathbf{A}_x$ for all $x \in \Omega_A$. The noise content for these binarizations then read as

$$w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1) = \min_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s) + \min_{s \in \mathcal{S}_n^{ext}} (u - \mathbf{A}_x)(s) = 1 + \min_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s) - \max_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s)$$

n	3	5	7	9	11	13	15	...	$\rightarrow \infty$
R.H.S. of (36)	0	–	0.753	–	0.852	–	0.893	...	$\rightarrow 1$
R.H.S. of (37)	–	0.528	–	0.803	–	0.872	–	...	$\rightarrow 1$

Table 1: The lower bounds of Eq. (36) and (37) for the noise contents of observables in \mathcal{T}_3 for the first few odd polygons and the limit $n \rightarrow \infty$.

for all $x \in \Omega_x$.

Denoting the extreme points of the state space \mathcal{S}_{2m+1} by $s_k = (\vec{s}_k, 1)^T$, from the definition of η_e^n we see that

$$\eta_e^n(\vec{a}_x) = \max_{k \in \{1, \dots, n\}} \vec{a}_x \cdot \vec{s}_k = \frac{1}{\alpha_x} \max_{k \in \{1, \dots, n\}} \mathbf{A}_x(s_k) - \sigma_n = \frac{1}{\alpha_x} \max_{s \in \mathcal{S}_n^{ext}} \mathbf{A}_x(s) - \sigma_n$$

for all $x \in \Omega_A$. Hence, together with the previous expressions for $\min_{x \in \Omega_A} \mathbf{A}_x(s)$, we have shown the following for the binarizations $\hat{\mathbf{A}}^{(x)}$ of an observable \mathbf{A} :

$$w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1) = 1 - \alpha_x [\eta_e^n(\vec{a}_x) + \eta_o^n(\vec{a}_x)] \quad (38)$$

for all $x \in \Omega_A$. We can now show that observables that have a high enough noise content are indeed included in \mathcal{T}_3 .

Proposition 15. *Let $\mathbf{A} \in \mathcal{O}(\mathcal{S}_n)$ be an observable on an odd polygon state space \mathcal{S}_n with effects $\mathbf{A}_x = \alpha_x(\vec{a}_x, \sigma_n)$ for all $x \in \Omega_A$. If*

$$\frac{1 - w(\mathbf{A}; \mathcal{T}_1)}{\min_{x \in \Omega_A} \alpha_x} \leq \sin\left(\frac{\pi}{2n}\right), \quad (39)$$

then \mathbf{A} is compatible with every other observable on \mathcal{S}_n .

Proof. From the previous expression for the noise contents of the binarizations $\hat{\mathbf{A}}^{(x)}$ of \mathbf{A} , and by using Eq. (28), we have that

$$\eta_{e/o}^n(\vec{a}_x) = \frac{1 - w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1)}{\alpha_x} - \eta_{o/e}^n(\vec{a}_x) \leq \frac{1 - w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1)}{\alpha_x} - \frac{\eta_{e/o}^n(\vec{a}_x)}{r_n}.$$

Since \mathcal{T}_1 is closed under post-processing and since $\hat{\mathbf{A}}^{(x)}$ is clearly a post-processing of \mathbf{A} for each $x \in \Omega_x$, we have by the basic properties of the noise content [13] that $w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1) \geq w(\mathbf{A}; \mathcal{T}_1)$ for all $x \in \Omega_A$. Thus, by rearranging the previous expression we have that

$$\begin{aligned} \eta_{e/o}^n(\vec{a}_x) &\leq \left(1 + \frac{1}{r_n}\right)^{-1} \frac{1 - w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1)}{\alpha_x} = \sigma_n r_n \left(\frac{1 - w(\hat{\mathbf{A}}^{(x)}; \mathcal{T}_1)}{\alpha_x}\right) \\ &\leq \sigma_n r_n \left(\frac{1 - w(\mathbf{A}; \mathcal{T}_1)}{\alpha_x}\right) \leq \sigma_n r_n \left(\frac{1 - w(\mathbf{A}; \mathcal{T}_1)}{\min_{x \in \Omega_A} \alpha_x}\right) \end{aligned}$$

for all $x \in \Omega_A$, where we have noticed that $(1 + 1/r_n)^{-1} = \sigma_n r_n$. Now, if Eq. (39) holds, from Prop. 13 it then follows that $\mathbf{A} \in \mathcal{T}_3$. \square

9 Summary

We have considered the no-information-without-disturbance and no-free-information principles in general probabilistic theories. We defined three sets of observables that correspond

trivial measurements (\mathcal{T}_1), measurements that can be performed without disturbing the system (\mathcal{T}_2) and measurements that can be performed jointly with any other measurement (\mathcal{T}_3). Although in quantum theory these sets are seen to coincide, we show that in general only the inclusions $\mathcal{T}_1 \subseteq \mathcal{T}_2 \subseteq \mathcal{T}_3$ hold. This means that there are operationally valid theories – even other than classical theories – where one can get non-trivial information about the system without disturbing it and where one can always choose to measure a non-trivial observable when performing any other measurement. Some of these theories were illustrated by examples.

We continued to characterize the sets \mathcal{T}_2 and \mathcal{T}_3 . We showed that observable is non-disturbing, i.e., in \mathcal{T}_2 if and only if the state space can be represented as a direct sum of state spaces such that the observable is constant on the summands. The result can be interpreted that a non-disturbing observable is only able to give the somewhat classical information about which state space of the direct sum we are using in our system. However, as noted, it does not mean that the observable should be trivial or the state space classical. One example showing this is the superselected quantum theory where the nontrivial non-disturbing observable gives the classical information about which quantum system (or which superselection sector) we are using.

As for \mathcal{T}_3 , we showed that observable is compatible with every other observable if and only if it can be post-processed from every simulation irreducible observable. In the previous work [18], the simulation irreducible observables were seen to be the minimal set of observables from which any other observable can be obtained by the means of classical manipulations, i.e., by mixing the observables and/or by post-processing their classical measurement outcomes. Thus, the result shows that to see if an observable is compatible with every other observable, it suffices only to consider the compatibility with the simulation irreducible observables, which is a much simpler task. This was demonstrated with the help of regular polygon theories, where the set \mathcal{T}_3 was characterized. Furthermore, it was shown that even though there are non-trivial observables in \mathcal{T}_3 for polygons, also those observables must be noisy, i.e., they have a substantial amount of some trivial observables in them with respect to the convex noise robustness.

Acknowledgments

The authors are thankful to Anna Jenčová for drawing their attention to [7] and to Tom Bullock for useful comments on the manuscript. LL acknowledges financial support from University of Turku Graduate School (UTUGS). MP was supported by grant VEGA 2/0069/16 and by the grant of the Slovak Research and Development Agency under contract APVV-16-0073. MP acknowledges that this research was done during a PhD study at Faculty of Mathematics, Physics and Informatics of the Comenius University in Bratislava. This work was performed as part of the Academy of Finland Centre of Excellence program (Project No. 312058).

References

- [1] H. Barnum, C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher. Noncommuting mixed states cannot be broadcast. *Phys. Rev. Lett.*, 76:2818–2821, 1996. DOI: [10.1103/PhysRevLett.76.2818](https://doi.org/10.1103/PhysRevLett.76.2818).
- [2] P. Busch. “No Information Without Disturbance”: Quantum Limitations of Measurement. In J. Christian and W. Myrvold, editors, *Quantum Reality, Relativistic Causal-*

- ity, and Closing the Epistemic Circle*. Springer-Verlag, 2009. DOI: [10.1007/978-1-4020-9107-0_13](https://doi.org/10.1007/978-1-4020-9107-0_13).
- [3] T. Heinosaari and M. Ziman. *The Mathematical Language of Quantum Theory. From Uncertainty to Entanglement*. Cambridge University Press, 2012. DOI: [10.1017/CBO9781139031103](https://doi.org/10.1017/CBO9781139031103).
 - [4] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Cloning and Broadcasting in Generic Probabilistic Theories. 2006. URL <http://arxiv.org/abs/quant-ph/0611295>.
 - [5] H. Barnum, J. Barrett, M. Leifer, and A. Wilce. Generalized No-broadcasting theorem. *Phys. Rev. Lett.*, 99:240501, 2007. DOI: [10.1103/PhysRevLett.99.240501](https://doi.org/10.1103/PhysRevLett.99.240501).
 - [6] G. Chiribella, G. M. D’Ariano, and P. Perinotti. Probabilistic theories with purification. *Phys. Rev. A*, 81:062348, 2010. DOI: [10.1103/PhysRevA.81.062348](https://doi.org/10.1103/PhysRevA.81.062348).
 - [7] H. Barnum and A. Wilce. Information processing in convex operational theories. *Electron. Notes Theor. Comput. Sci.*, 270:3–15, 2011. DOI: [10.1016/j.entcs.2011.01.002](https://doi.org/10.1016/j.entcs.2011.01.002).
 - [8] C. Pfister and S. Wehner. An information-theoretic principle implies that any discrete physical theory is classical. *Nat. Commun.*, 4:1851, 2013. DOI: [10.1038/ncomms2821](https://doi.org/10.1038/ncomms2821).
 - [9] G. C. Wick, A. S. Wightman, and E. P. Wigner. The intrinsic parity of elementary particles. *Phys. Rev.*, 88:101–105, 1952. DOI: [10.1103/PhysRev.88.101](https://doi.org/10.1103/PhysRev.88.101).
 - [10] G. Kimura, K. Nuida, and H. Imai. Distinguishability measures and entropies for general probabilistic theories. *Rep. Math. Phys.*, 66:175 – 206, 2010. DOI: [10.1016/S0034-4877\(10\)00025-X](https://doi.org/10.1016/S0034-4877(10)00025-X).
 - [11] A. J. Short and S. Wehner. Entropy in general physical theories. *New J. Phys.*, 12:033023, 2010. DOI: [10.1088/1367-2630/12/3/033023](https://doi.org/10.1088/1367-2630/12/3/033023).
 - [12] I. Namioka and R. Phelps. Tensor products of compact convex sets. *Pacific J. Math.*, 31:469–480, 1969. DOI: [10.2140/pjm.1969.31.469](https://doi.org/10.2140/pjm.1969.31.469).
 - [13] S. N. Filippov, T. Heinosaari, and L. Leppäjärvi. Necessary condition for incompatibility of observables in general probabilistic theories. *Phys. Rev. A*, 95:032127, 2017. DOI: [10.1103/PhysRevA.95.032127](https://doi.org/10.1103/PhysRevA.95.032127).
 - [14] T. Heinosaari and T. Miyadera. Incompatibility of quantum channels. *J. Phys. A: Math. Gen.*, 50:135302, 2017. DOI: [10.1088/1751-8121/aa5f6b](https://doi.org/10.1088/1751-8121/aa5f6b).
 - [15] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. Nonlocal correlations as an information-theoretic resource. *Phys. Rev. A*, 71:022101, 2005. DOI: [10.1103/PhysRevA.71.022101](https://doi.org/10.1103/PhysRevA.71.022101).
 - [16] L. Guerini, J. Bavaresco, M. T. Cunha, and A. Acín. Operational framework for quantum measurement simulability. *J. Math. Phys.*, 58:092102, 2017. DOI: [10.1063/1.4994303](https://doi.org/10.1063/1.4994303).
 - [17] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín. Simulating Positive-Operator-Valued Measures with Projective Measurements. *Phys. Rev. Lett.*, 119:190501, 2017. DOI: [10.1103/PhysRevLett.119.190501](https://doi.org/10.1103/PhysRevLett.119.190501).
 - [18] S. N. Filippov, T. Heinosaari, and L. Leppäjärvi. Simulability of observables in general probabilistic theories. *Phys. Rev. A*, 97:062102, 2018. DOI: [10.1103/PhysRevA.97.062102](https://doi.org/10.1103/PhysRevA.97.062102).
 - [19] M. Oszmaniec, F. B. Maciejewski, and Z. Puchała. All quantum measurements can be simulated using projective measurements and postselection. 2018. URL <http://arxiv.org/abs/1807.08449>.
 - [20] M. Plávala. All measurements in a probabilistic theory are compatible if and only if the state space is a simplex. *Phys. Rev. A*, 94:042108, 2016. DOI: [10.1103/PhysRevA.94.042108](https://doi.org/10.1103/PhysRevA.94.042108).
 - [21] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner. Limits on nonlocal correlations

from the structure of the local state space. *New J. Phys.*, 13:063024, 2011. DOI: 10.1088/1367-2630/13/6/063024.

- [22] R. T. Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, 1997.

A Some results on the structure of GPTs

In this appendix we will prove several minor results about the structure of general probabilistic theories that were needed in the calculations. We will denote the interior of the set \mathcal{S} by $\text{intr}(\mathcal{S})$.

Definition 8. Let \mathcal{S} be a state space and let $E \subset \mathcal{S}$ be a convex subset. We say that E is a face of \mathcal{S} if $z \in E$ and $z = \lambda x + (1 - \lambda)y$ for some $x, y \in \mathcal{S}$ and $\lambda \in [0, 1]$ implies $x, y \in E$.

Definition 9. Let $E \subset \mathcal{S}$, we say that E is a maximal face if E is a face and for any $x \in \mathcal{S} \setminus E$ we have $\text{conv}(E \cup \{x\}) \cap \text{intr}(\mathcal{S}) \neq \emptyset$.

If \mathcal{S} is d -dimensional and has a finite number of extreme points, then maximal faces are the $(d - 1)$ -dimensional faces of \mathcal{S} . From a geometrical perspective their special properties all follow from the requirement that $\text{conv}(E \cup \{x\}) \cap \text{intr}(\mathcal{S}) \neq \emptyset$.

Lemma 5. Let \mathcal{S} be a state space, let $e \in \mathcal{E}(\mathcal{S})$ and let $E_0 = \{x \in \mathcal{S} : e(x) = 0\}$. If E_0 is a maximal face, then e is indecomposable.

Proof. Let $f \in \mathcal{E}(\mathcal{S})$, denote $F_0 = \{x \in \mathcal{S} : f(x) = 0\}$ and assume $e \geq f$. It follows that we must have $E_0 \subset F_0$ and since E_0 is maximal face it follows that either $F_0 = \mathcal{S}$ or $F_0 = E_0$.

If $F_0 = \mathcal{S}$ then $f = 0$. If $F_0 = E_0$, then pick $x \in \mathcal{S}$ such that $x \notin E_0$. Both e and f are uniquely defined by the values $e(x)$ and $f(x)$, because E_0 is a maximal face. This implies that we have $f = \frac{f(x)}{e(x)}e$, which shows that e is indecomposable. \square

Proposition 16. Assume that \mathcal{S} has only a finite number of extreme points. Let $e \in \mathcal{E}(\mathcal{S})$ and let $E_0 = \{x \in \mathcal{S} : e(x) = 0\}$, then e is indecomposable if and only if E_0 is a maximal face.

Proof. Assume that $E_0 = \{x \in \mathcal{S} : e(x) = 0\}$ is not a maximal face, then there is a maximal face F_0 such that $E_0 \subset F_0$ [20]. Moreover let $f \in \mathcal{E}(\mathcal{S})$ be such that $F_0 = \{x \in \mathcal{S} : f(x) = 0\}$ and denote $G = \{y \in \mathcal{S} : y \notin E_0, y \text{ is extreme}\}$ and

$$m = \min_{y \in G} e(y).$$

Clearly $m > 0$. We will show that $e \geq mf$; let $z \in \mathcal{S}$ be an extreme point, then either $z \in E_0$ or $z \in G$. If $z \in E_0 \subset F_0$, then $e(z) = 0 \geq 0 = mf(z)$. If $z \in G$, then $e(z) \geq m \geq mf(z)$. Since by construction $e \neq \alpha f$ for any $\alpha \in \mathbb{R}$ as that would imply $E_0 = F_0$ it follows that e can not be indecomposable. \square

B Proof of Prop. 12

We recall from [18] that the extreme simulation irreducible observables are characterized by triangles on the base $\text{conv}(\{g_1, \dots, g_n\})$ with vertices from the set $\{g_1, \dots, g_{2m+1}\}$ such that $\sigma_n u$ is included in the triangles. We show that such triangles are in one-to-one correspondence with the extreme simulation irreducible observables.

To see this, first let $\mathbf{B} \in \mathcal{O}_{irr}^{ext}(\mathcal{S}_{2m+1})$ so that $\Omega_{\mathbf{B}} = \{1, 2, 3\}$. Since for all $k \in \{1, 2, 3\}$ the effects \mathbf{B}_k are indecomposable, for each $k \in \{1, 2, 3\}$ there exists $0 < c_k \leq 1$ and effect $g_{i_k} \in \{g_1, \dots, g_{2m+1}\}$ such that $\mathbf{B}_k = c_k g_{i_k}$. From the normalization of \mathbf{B} it follows that

$$u = c_1 g_{i_1} + c_2 g_{i_2} + c_3 g_{i_3}$$

so that from the z -components of the vectors we get a requirement that $c_1 + c_2 + c_3 = \frac{1}{\sigma_n}$. Thus, if we denote the sum $c_1 + c_2 + c_3$ by c , we see that

$$\sigma_n u = \frac{c_1}{c} g_{i_1} + \frac{c_2}{c} g_{i_2} + \frac{c_3}{c} g_{i_3} \quad (40)$$

which shows that the vertices $\{g_{i_1}, g_{i_2}, g_{i_3}\}$ form a triangle $\text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\})$ on the base $\text{conv}(\{g_1, \dots, g_{2m+1}\})$ such that $\sigma_n u \in \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\})$.

To see the contrary, let j_1, j_2, j_3 be any three indices from the set $\{1, \dots, 2m+1\}$ such that $\sigma_n u \in \text{conv}(\{g_{j_1}, g_{j_2}, g_{j_3}\})$. Thus, there exists convex coefficients $\tilde{d}_1, \tilde{d}_2, \tilde{d}_3 \in [0, 1]$, $\tilde{d}_1 + \tilde{d}_2 + \tilde{d}_3 = 1$, such that $\sigma_n u = \tilde{d}_1 g_{j_1} + \tilde{d}_2 g_{j_2} + \tilde{d}_3 g_{j_3}$. If we denote $d_k = \tilde{d}_k / \sigma_n \in (0, 1]$ and $\mathbf{B}'_k = d_k g_{j_k}$ for all $k \in \{1, 2, 3\}$, we find that $\{\mathbf{B}'_1, \mathbf{B}'_2, \mathbf{B}'_3\}$ is a set of linearly independent indecomposable effects such that $\mathbf{B}'_1 + \mathbf{B}'_2 + \mathbf{B}'_3 = u$, which shows that an observable \mathbf{B}' defined with these effects is an extreme simulation irreducible observable.

Since the set $\text{conv}(\{g_1, \dots, g_{2m+1}\})$ is a base for the positive cone of the effects, for each effect \mathbf{A}_y of an observable $\mathbf{A} \in \mathcal{O}(\mathcal{S}_{2m+1})$ there exists $\alpha_y > 0$ and $a_y \in \text{conv}(\{g_1, \dots, g_{2m+1}\})$ such that $\mathbf{A}_y = \alpha_y a_y$. Similarly, for each $j \in \{1, \dots, q_m\}$ we have that $\mathbf{B}_k^{(j)} = c_k^{(j)} g_{i_k^{(j)}}$ for some $c_k^{(j)} \in (0, 1]$ and $i_k^{(j)} \in \{1, \dots, n\}$ for all $k \in \{1, 2, 3\}$. We then see that in order to characterize the intersection of the cones generated by the extreme simulation irreducible observables, i.e. essentially \mathcal{T}_3 , we need to only consider the intersection of the respective triangles on the base.

Lemma 6. *Observable $\mathbf{A} \in \mathcal{O}(\mathcal{S}_{2m+1})$ with effects $\mathbf{A}_y = \alpha_y a_y$, where we have now $a_y \in \text{conv}(\{g_1, \dots, g_{2m+1}\})$ for all $y \in \Omega_{\mathbf{A}}$, is in \mathcal{T}_3 if and only if*

$$a_y \in \bigcap_{j=1}^{q_m} \text{conv} \left(\left\{ g_{i_1^{(j)}}^{(j)}, g_{i_2^{(j)}}^{(j)}, g_{i_3^{(j)}}^{(j)} \right\} \right) \quad \forall y \in \Omega_{\mathbf{A}}.$$

Proof. By Cor. 4 we see that we need to show that

$$\mathbf{A}_y \in \bigcap_{j=1}^{q_m} \text{cone} \left(\left\{ g_{i_1^{(j)}}^{(j)}, g_{i_2^{(j)}}^{(j)}, g_{i_3^{(j)}}^{(j)} \right\} \right) \quad (41)$$

if and only if

$$a_y \in \bigcap_{j=1}^{q_m} \text{conv} \left(\left\{ g_{i_1^{(j)}}^{(j)}, g_{i_2^{(j)}}^{(j)}, g_{i_3^{(j)}}^{(j)} \right\} \right) \quad (42)$$

for all $y \in \Omega_{\mathbf{A}}$.

First let \mathbf{A}_y be in the intersection of cones, which itself is a cone, in (41) for some $y \in \Omega_{\mathbf{A}}$. Since $\mathbf{A}_y = \alpha_y a_y$ for some $\alpha_y > 0$, it follows that also a_y is included in the

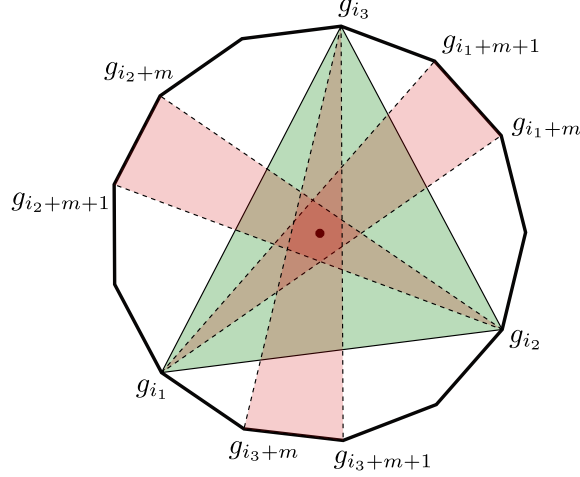


Figure 7: Illustration of Lemma 7 when $n = 13$. For a simulation irreducible observable B with $B_k = c_k g_{i_k}$ for $k = 1, 2, 3$, the sets $\text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\})$ are depicted in red and the set $\text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\})$ in green. One sees that the intersection of the red sets is contained in the green set just as the lemma states.

same intersection of cones. Thus, a_y can be expressed as a positive linear combination of $g_{i_1}^{(j)}, g_{i_2}^{(j)}, g_{i_3}^{(j)}$ for all $j \in \{1, \dots, q_m\}$. Since all the vectors $a_y, g_{i_1}^{(j)}, g_{i_2}^{(j)}, g_{i_3}^{(j)}$ lie on the same $z = \sigma_n$ -plane for all j , it follows that the positive linear combination must actually be a convex combination which shows (42).

Let then a_y be included in the intersection of the convex hulls in (42) for some $y \in \Omega_A$. Since a convex hull is just a special case of a conic hull, we see that a_y is also included in the intersection of cones in (41). By multiplying a_y by α_y we see that then (41) holds. \square

The smallest such triangles to contain the centroid $(0, 0, \sigma_n)^T$ have vertices g_i, g_{i+m} and g_{i+m+1} for $i = 1, \dots, 2m+1$, where the addition is modulo $2m+1$. We will show that the intersection of these smallest triangles gives us the whole intersection of all the triangles that represent the extreme simulation irreducible observables. We start with a small Lemma (see Fig. 7).

Lemma 7. *For an extreme simulation irreducible observable B such that $B_k = c_k g_{i_k}$ for $k \in \{1, 2, 3\}$ we have that*

$$\bigcap_{k=1}^3 \text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\}) \subseteq \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\}).$$

Proof. To see this, suppose that, contrary to this, there exists a point

$$x \in \bigcap_{k=1}^3 \text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\})$$

such that $x \notin \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\})$. Without loss of generality we assume that $i_1 < i_2 < i_3$.

If we consider a fixed vertex g_{i_k} for some $k \in \{1, 2, 3\}$, it is clear that the indices i_k+m and i_k+m+1 are contained in the set of indices $\{i_{k+1}, i_{k+1}+1, \dots, i_{k+2}-1, i_{k+2}\}$ (Fig. 7). This is because otherwise they would be contained in either $\{i_{k+2}, i_{k+2}+1, \dots, i_k-1, i_k\}$ or $\{i_k, i_k+1, \dots, i_{k+1}-1, i_{k+1}\}$ so that

$$\text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\}) \subset \text{conv}(\{g_{i_{k+2}}, \dots, g_{i_k}\})$$

or

$$\text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\}) \subset \text{conv}(\{g_{i_k}, \dots, g_{i_{k+1}}\})$$

both of which would contradict the fact that $\sigma_n u \in \text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\})$.

Since now $x \in \text{conv}(\{g_{i_k}, g_{i_k+m}, g_{i_k+m+1}\})$ for all $k \in \{1, 2, 3\}$ but $x \notin \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\})$, we must have for all $k' \in \{1, 2, 3\}$ that

$$x \notin \text{conv}(\{g_{i_{k'}}, g_{i_{k'}+m}, g_{i_{k'}+m+1}\}) \cap \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\}).$$

We have by the above statement about the indices that $\{i_k+m, i_k+m+1\} \subseteq \{i_{k+1}, i_{k+1}+1, \dots, i_{k+2}-1, i_{k+2}\}$ so that it then follows that

$$\text{conv}(g_{i_k}, g_{i_k+m}, g_{i_k+m+1}) \subseteq \text{conv}(\{g_{i_1}, g_{i_2}, g_{i_3}\}) \cup \text{conv}(\{g_{i_{k+1}}, g_{i_{k+1}+1}, \dots, g_{i_{k+2}}\})$$

which is a disjoint union for all $k \in \{1, 2, 3\}$. Thus, $x \in \text{conv}(\{g_{i_{k+1}}, g_{i_{k+1}+1}, \dots, g_{i_{k+2}}\})$ for all $k \in \{1, 2, 3\}$ which is a contradiction since the sets do not intersect. \square

Proof of Proposition 12. From Cor. 4 it is clear that in order to prove the statement we need to show that

$$\bigcap_{j=1}^{q_m} \text{cone}(\{B_x^{(j)}\}_{x \in \Omega_{B^{(j)}}}) = \bigcap_{i=1}^{2m+1} \text{cone}(\{g_i, g_{i+1}, g_{i+m+1}\}).$$

The above statement is about cones but by Lemma 6 we can equivalently consider it in terms of the triangles that represent the observables in $\mathcal{O}_{irr}^{ext}(\mathcal{S}_{2m+1})$. By using the previously introduced notation for the effects of the extreme simulation irreducible observables, the above statement about the triangles then reads as

$$\bigcap_{j=1}^{q_m} \text{conv}(\{g_{i_1}^{(j)}, g_{i_2}^{(j)}, g_{i_3}^{(j)}\}) = \bigcap_{i=1}^{2m+1} \text{conv}(\{g_i, g_{i+m}, g_{i+m+1}\}).$$

The inclusion “ \subseteq ” is clear since among the q_m triangles that represent the extreme simulation irreducible observables the triangles with vertices g_i, g_{i+m} and g_{i+m+1} for $i = 1, \dots, 2m+1$ are included.

For the inclusion “ \supseteq ”, we use Lemma 7 for observables $\{B^{(j)}\}_{j=1}^{q_m}$ which states that

$$\bigcap_{k=1}^3 \text{conv}(\{g_{i_k}^{(j)}, g_{i_k}^{(j)+m}, g_{i_k}^{(j)+m+1}\}) \subseteq \text{conv}(\{g_{i_1}^{(j)}, g_{i_2}^{(j)}, g_{i_3}^{(j)}\})$$

for all $j \in \{1, \dots, q_m\}$. By taking the intersection of all $j \in \{1, \dots, q_m\}$ we get

$$\begin{aligned} \bigcap_{i=1}^{2m+1} \text{cone}(\{g_i, g_{i+1}, g_{i+m+1}\}) &= \bigcap_{j=1}^{q_m} \bigcap_{k=1}^3 \text{conv}(\{g_{i_k}^{(j)}, g_{i_k}^{(j)+m}, g_{i_k}^{(j)+m+1}\}) \\ &\subseteq \bigcap_{j=1}^{q_m} \text{conv}(\{g_{i_1}^{(j)}, g_{i_2}^{(j)}, g_{i_3}^{(j)}\}) \end{aligned}$$

which proves the statement. \square

C Proof of Lemma 4

Proof of Lemma 4. We first see that either $x_{i,j}$ is an extreme point of C_m or $x_{i,j} \notin C_m$. Namely, assume that $x_{i,j} \in C_m$ but it is not an extreme point of C_m , then there exists some open line segment M , such that $x_{i,j} \in M$ and $M \subset C_m$. We must have $M \subset L_i$ since if M would intersect L_i , then we would get a contradiction with $M \subset C_m$. But then we must also have $M \subset L_{i+j}$ which is a contradiction with $L_i \neq L_{i+j}$.

Next fix $i \in \{1, \dots, 2m+1\}$. From $x_{i,j} = L_i \cap L_{i+j}$ we get $x_{i,j} = \lambda_j g_i + (1 - \lambda_j) g_{i+m}$, where

$$\lambda_j = \frac{\cos\left(\frac{(2j+1)\pi}{4m+2}\right)}{2 \cos\left(\frac{j\pi}{2m+1}\right) \cos\left(\frac{\pi}{4m+2}\right)} = \frac{1}{2} \left[1 - \tan\left(\frac{\pi}{4m+2}\right) \tan\left(\frac{j\pi}{2m+1}\right) \right].$$

Since $j\pi/(2m+1) \in [0, \pi/2]$ for all $j \in \{1, \dots, m\}$, we have that $\tan\left(\frac{j\pi}{2m+1}\right)$ is an increasing function of j so that $\lambda_j \leq \lambda_1$. We then see that

$$x_{i,j} = \lambda_j g_i + (1 - \lambda_j) g_{i+m} = \frac{\lambda_j}{\lambda_1} x_{i,1} + \left(1 - \frac{\lambda_j}{\lambda_1}\right) g_{i+m}$$

for all $j \in \{1, \dots, m\}$, where now $\frac{\lambda_j}{\lambda_1} \in [0, 1]$ so that $x_{i,j} \in \text{conv}(\{x_{i,1}, g_{i+m}\})$. Since

$$x_{i,1}, g_{i+m} \in \text{conv}(\{g_{i+1}, g_{i+2}, \dots, g_{i+m+1}\}),$$

it follows that also $x_{i,j} \in \text{conv}(\{g_{i+1}, g_{i+2}, \dots, g_{i+m+1}\})$ for all $j \in \{1, \dots, m\}$.

Clearly $x_{i,j} \in L_i$ for all $j \in \{1, \dots, m\}$ but $x_{i,j} \notin L_{i+1}$ for all $j \in \{2, \dots, m\}$, where L_{i+1} can be expressed as

$$\text{conv}(\{g_{i+1}, \dots, g_{i+m+1}\}) \cap \text{conv}(\{g_{i+1}, g_{i+m+1}, g_{i+m+2}\}),$$

so that $x_{i,j} \notin \text{conv}(\{g_{i+1}, g_{i+m+1}, g_{i+m+2}\})$ for $j \in \{2, \dots, m\}$. Thus, it follows that $x_{i,j} \notin C_m$ for $j \in \{2, \dots, m\}$.

The only candidates for the extreme points of C_m are then $x_{i,1}$ for all $i \in \{1, \dots, 2m+1\}$. From the symmetry it follows that all $x_{i,1}$ indeed must be extreme since if $x_{i',1}$ is not extreme for some $i' \in \{1, \dots, 2m+1\}$ it would follow that $x_{i,1}$ is not extreme for any $i \in \{1, \dots, 2m+1\}$. Hence, the claim follows. \square