

# Complete extension: the non-signalling analog of quantum purification

Marek Winczewski<sup>1,2</sup>, Tamoghna Das<sup>1,3</sup>, John H. Selby<sup>1</sup>, Karol Horodecki<sup>4,1</sup>, Paweł Horodecki<sup>1,5,6</sup>, Łukasz Pankowski<sup>7</sup>, Marco Piani<sup>8,9</sup>, and Ravishankar Ramanathan<sup>10</sup>

<sup>1</sup>International Centre for Theory of Quantum Technologies, University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland

<sup>2</sup>Institute of Theoretical Physics and Astrophysics and National Quantum Information Centre in Gdańsk, University of Gdańsk, 80-952 Gdańsk, Poland

<sup>3</sup>Department of Physics, Indian Institute of Technology Kharagpur, Kharagpur-721302, India

<sup>4</sup>Institute of Informatics and National Quantum Information Centre in Gdańsk, Faculty of Mathematics, Physics and Informatics, University of Gdańsk, 80-952 Gdańsk, Poland

<sup>5</sup>Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, 80-233 Gdańsk, Poland

<sup>6</sup>National Quantum Information Centre, University of Gdańsk, ul. Jana Bażyńskiego 8, 80-309 Gdańsk, Poland

<sup>7</sup>VOICELAB.AI, Al. Grunwaldzka 135A; 80-264 Gdańsk, Poland,

<sup>8</sup>evolutionQ Inc., Waterloo, Ontario, N2L 3L3, Canada

<sup>9</sup>SUPA and Department of Physics, University of Strathclyde, Glasgow, G4 0NG, UK

<sup>10</sup>Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong

Deriving quantum mechanics from information-theoretic postulates is a recent research direction taken, in part, with the view of finding a beyond-quantum theory; once the postulates are clear, we can consider modifications to them. A key postulate is the purification postulate, which we propose to replace by a more generally applicable postulate that we call the complete extension postulate (CEP), i.e., the existence of an extension of a physical system from which one can generate any other extension. This new concept leads to a plethora of open questions and research directions in the study of general theories satisfying the CEP (which may include a theory that hyper-decoheres to quantum theory). For example, we show that the CEP implies the impossibility of bit-commitment. This is exemplified by a case study of the theory of non-signalling behaviors which we show satisfies the CEP. We moreover show that in certain cases the complete extension will not be pure, highlighting the key divergence from the purification postulate.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Generalised probabilistic theories . . . . .	4
<b>2</b>	<b>The purification postulate</b>	<b>6</b>
2.1	No-go theorem for purification in discrete theories . . . . .	6
<b>3</b>	<b>The complete extension postulate</b>	<b>8</b>
3.1	Impossibility of bit-commitment . . . . .	10
3.2	No no-go for hyperdecoherence . . . . .	12
<b>4</b>	<b>Case study - theory of non-signalling behaviours</b>	<b>13</b>
4.1	The complete extension in the theory of non-signalling behaviours . . . . .	16
4.2	On the dimensionality of the complete extension . . . . .	21
<b>5</b>	<b>Conclusions</b>	<b>22</b>

<b>A</b>	<b>Proofs</b>	<b>30</b>
A.1	Proof of Proposition 13	30
A.2	Proof of Theorem 16	31
A.3	Proof of Corollary 20	33
A.4	Proof of Theorem 23	33
<b>B</b>	<b>Explicit examples in the theory of no-signaling behaviour</b>	<b>35</b>
B.1	Complete extensions of binary input-output behaviours	35
B.1.1	Example: the non-signaling complete extension that is not a vertex	38
B.1.2	Example: NSCE of $P_A$ gives ACCESS to any PME of $P_A$	40
B.1.3	Example: NSCE of $P_A$ gives ACCESS to any mixed ensemble of $P_A$	40
B.1.4	NSCE can generate any extension	42
B.1.5	Quantifying non-locality introduced in NSCE	42
B.1.6	Complete extension of the conjugate box	43
B.2	Dimensionality of the No-signaling complete extension	44
B.3	Complete extensions of three-cycle contextual behaviour	46
B.4	Complete extension of the behaviours lying on the isotropic line	50
B.4.1	Minimal ensembles for isotropic behaviours	50
B.4.2	Dimension of the extending party for the behaviours lying on the isotropic line	51
B.4.3	Minimal ensembles for non-local isotropic behaviours	53

## 1 Introduction

Quantum mechanics appears to be one of the best validated physical theories and, at the same time, lacking a single agreed-on interpretation [1]. It can be expressed via several mathematical axioms [2–4] and leads to numerous phenomena, which have been confirmed experimentally, and now form the basis for many of our everyday technologies. However, it cannot explain some of the visible (or detectable) parts of the universe – the key example being the realm of the theory of gravity. For this reason, an important scientific effort is underway to find a beyond-quantum theory which would explain phenomena beyond quantum theory [5–10]. However, it is not clear which axiom of quantum mechanics should be replaced or modified in order to keep the theory as expressive as quantum mechanics whilst explaining phenomena, such as gravity, which seem to be beyond quantum theory.

It is known that the standard axioms of quantum mechanics are not independent from one another [11–13], i.e., they can be used to derive one another. It is therefore difficult to modify the standard postulates of quantum theory as they cannot be independently modified. In order to examine modifications of quantum theory, it is therefore useful to first re-express quantum mechanics in the form of a new set of postulates which can be independently modified. Recently, there has been a great deal of interest in such *reformulations*, particularly, via information theoretic postulates within the framework of so-called Generalised Probabilistic Theories (GPTs). This idea led to quite a number of reformulations of quantum theory [8, 14–27] in which sets of axioms are proposed which single out quantum mechanics from the space of all GPTs.

One of the key insights in the reconstruction programme came from [28] which introduced the purification postulate. This was an essential postulate, in that it was the sole postulate that distinguished quantum from classical theory. Since its introduction the purification postulate has been used extensively in the literature to prove many results, for example, pertaining to computation [29–31], cryptography [32, 33], thermodynamics [34, 35], and interference [36]. Essentially the purification postulate is a generalisation of the notion of purification within quantum theory to arbitrary GPTs:

**Quantum Purifications** *In quantum theory, for any state  $\rho_A$  there exists a system  $B$  and a bipartite pure state  $|\psi\rangle_{AB}$ , which satisfies  $\text{tr}_B(|\psi\rangle\langle\psi|) = \rho_A$ . The bipartite state  $|\psi\rangle_{AB}$  is known as a purification of  $\rho_A$ , and these are essentially unique, in that there is always an isometry mapping between any pair of purifications.*

We formally introduce purifications in Sec. 2 and the formalism of GPTs necessary to talk about it in Sec. 1.1. At this point, however, let us observe that there are various issues with the purification postulate that may lead one to believe that it will have to be modified in order to find a more fundamental theory of nature.

1. There is a no-go theorem proven in Ref. [37] which shows that if one believes that a more fundamental theory should be causal, and should reduce to quantum theory by a decoherence-like mechanism, then one must give up on the purification postulate.
2. We prove a theorem in Sec. 2.1 which shows that the purification postulate must fail in any discrete theory, whilst in Refs. [38–40] ideas coming from quantum gravity are used to argue that on a fundamental level the quantum state space will become discrete.
3. The purification postulate also fails to hold in super-selected quantum systems [24, 27, 41], and so if we believe in any fundamental superselection rules then the postulate must be modified in some way.
4. There are also questions about whether the purification postulate is suitable in a theory which permits indefinite causal structure [42–45], as it is shown in Ref. [46] that a purification-like postulate fails to hold for all quantum process matrices. That a theory should permit indefinite causal structure can again be motivated by ideas coming from quantum gravity.

For all of these reasons one can then ask if the purification postulate can be weakened or replaced by another postulate, resulting in a physical theory with more explanatory power. In this manuscript, we propose such a replacement for the purification postulate called the Complete Extension Postulate (CEP) and study its consequences in various contexts.

In order to understand the complete extension postulate we begin by observing that, within quantum mechanics, purifications are examples of complete extensions [47, Exercise 2.8.2].

**Quantum Complete Extensions** *A complete extension for a state  $\rho_A$  in quantum theory, is an extension  $\sigma_{AB}^*$  of  $\rho_A$ , that is it satisfies  $\text{tr}_B(\sigma_{AB}^*) = \rho_A$ , where, moreover, any other extension  $\sigma$  can be reached via local operations on  $B$ , and any ensemble for  $\rho_A$  can be reached via a measurement on  $B$ .*

It can be straightforwardly seen that, within quantum theory, all quantum purifications are quantum complete extensions but not vice versa. It is precisely these complete extensions which we generalise to GPTs in order to define the complete extension postulate in Sec. 3. As this is a weaker notion within quantum theory it is therefore plausible that the associated postulate will apply more broadly and avoid some of the aforementioned issues with the purification postulate. Indeed, we show that this is the case for the first three issues raised above:

1. In Sec. 3.2 we show the proof of the no-go theorem for hyperdecoherence provided in [37] no longer holds if we use the complete extension postulate.
2. We show that the complete extension postulate can hold in discrete theories, in particular, in Sec. 4 we demonstrate this for the widely studied theory of non-signalling behaviours [5, 48, 49] colloquially known as *boxworld*.
3. In Sec. 3 we show that the complete extension postulate holds in super-selected quantum systems and within classical systems.

This may therefore lead one to wonder whether perhaps the complete extension postulate is too weak and does not have any meaningful consequences. We show that this is not the case. In particular, we show that there are theories which have been studied in the literature, such as those found in [50], which fail to satisfy the CEP. This means that it is not a trivial postulate and does offer explanatory power. Indeed, there are two key reasons that a given GPT may fail to satisfy the CEP. On the one hand, it may not have sufficient bipartite states, in which case it would not admit of an extension that enables access to all ensembles. On the other hand, it may not have sufficient dynamics, as would be

the case for highly asymmetric state spaces. Namely, in the latter case, even if an extension enables access to all ensembles, it may not give a possibility for the generation of any other extension from it by this dynamics. Moreover, in Sec. 3.1 We show that any theory which satisfies the CEP cannot allow for the cryptographic primitive of bit-commitment, and, moreover, following the techniques of [33], we find a non-trivial lower bound on the success probability for any protocol. As an added benefit, this unifies the proof of impossibility of bit-commitment for the quantum and classical cases which previously were treated separately.

We also study the complete extensions within the theory of non-signalling behaviours in some depth in Sec. 4 and in App. B. In particular, we show how a complete extension can be constructed within this theory, and that the dimensions of these complete extensions are always finite. In this way, we show that CEP is not an empty postulate in general theories, as the complete extensions can actually be constructed. Furthermore, we characterise the minimal ensembles of the non-local isotropic behaviours, which was already proved useful in the context of the non-signalling device independent secure key agreement [51].

The introduction of a new postulate opens up many doors for future exploration. We touch on just a few of them in this paper and point out many others as future research directions as we go. In particular, we collect together many of these in Sec. 5.

## 1.1 Generalised probabilistic theories

The purification postulate and complete extension postulate that we introduce are expressed in the language of generalised probabilistic theories (GPTs), hence, in order to understand these we first provide a brief introduction to this formalism. For a more in-depth introduction see, for example, Refs. [52–54]

The formalism of GPTs [8, 48] is a formalism for describing the operational predictions of essentially arbitrary conceivable theories of physics. In the case of quantum theory, this operational description is isomorphic to the standard quantum informatic approach to quantum theory (that is, of completely positive trace preserving maps, density matrices, and positive operator value measures).

The primitive building blocks of any GPT,  $\mathcal{G}$  are the *systems*, denoted  $A, B, \dots \in \text{Syst}[\mathcal{G}]$ , these come equipped with an associative binary composition rule  $\otimes : \text{Syst}[\mathcal{G}] \times \text{Syst}[\mathcal{G}] \rightarrow \text{Syst}[\mathcal{G}]$  which allow us to build composite systems out of simpler system. In quantum theory these systems can be labeled by natural numbers, i.e., by their dimension, and then these labels compose simply by multiplication, as the dimension of a tensor product of two Hilbert spaces is the product of the individual dimensions.

Each system  $A$  corresponds to a finite dimensional real vector space  $V_A$ . The fact that it is finite can either be viewed as simply a choice for technical convenience, or can be motivated operationally as the idea that in practice we can only ever perform a finite number of measurements to do tomography on a system (even if in principle there are an infinite number of degrees of freedom). In quantum theory these real vector spaces are the spaces of Hermitian operators on some finite dimensional Hilbert space, whilst in classical theory these real vector spaces are the function spaces  $\mathbb{R}^\Lambda$  for some finite sample space  $\Lambda$ . It is important to note at this point that  $V_{A \otimes B}$  is not necessarily equal to  $V_A \otimes V_B$ . These are equal only under the assumption of tomographic locality [8] – that every multipartite state can be characterised by means of local measurements on its parts. Note that both quantum and classical theory are tomographically local theories.

Within these vector spaces  $V_A$  there is a specified convex set  $\Omega_A \subset V_A$  which describes the *state space* for the system. This convex set must be compact and closed, and, moreover, must have an affine dimension one less than the linear dimension of the vector space such that the affine span of  $\Omega_A$  does not intersect the origin. We can also define the convex state cone  $K_A$  which includes also subnormalised and supernormalised states as  $K_A := \{rs | r \in \mathbb{R}^+, s \in \Omega_A\}$ . In quantum theory this state space corresponds to the space of density matrices and the state cone to positive semidefinite operators.

Within the dual vector space  $V_A^*$  – that is, the vector space of linear functionals on  $V_A$  – there is also a specified convex set  $\mathcal{E}_A \subset V_A^*$  which describes the *effect space* for the system. This convex set must also be compact and closed, however, it will be full dimensional and contain the origin. The effects  $e \in \mathcal{E}_A$  assign probabilities to measurement outcomes when the measurement is performed on

an arbitrary state in  $\Omega_A$ , they therefore must satisfy  $e(\Omega_A) \subseteq [0, 1]$ . The effect space  $\mathcal{E}_A$  must also contain the unique unit effect  $u_A \in \mathcal{E}_A$  which is defined as  $u_A(\Omega_A) = \{1\}$ . The uniqueness of this effect is built into the setup of our framework, however, there are modifications to the framework in which this is instead an additional principle which captures the notion of causality [15]. In quantum theory this effect space corresponds to  $\text{tr}(\sigma \cdot)$  where  $\sigma$  is a POVM element and the unit corresponds to  $\sigma = \mathbb{1}$ .

Like the systems themselves, these state and effect spaces have an associative composition rule,  $\Omega_A \otimes \Omega_B := \Omega_{A \otimes B}$ ,  $\mathcal{E}_A \otimes \mathcal{E}_B := \mathcal{E}_{A \otimes B}$  and  $u_A \otimes u_B := u_{A \otimes B}$ . Note that in general  $\otimes$  is not related to the tensor product of vector spaces, the symbol is used in analogy to the role of the tensor product in composing quantum systems. This composition must be bilinear, and satisfy  $e \otimes f(s \otimes t) = e(s)f(t)$  for all  $e \in \mathcal{E}_A$ ,  $f \in \mathcal{E}_B$ ,  $s \in \Omega_A$  and  $t \in \Omega_B$ . These conditions, in particular, ensure that the unit effect  $u_B$  gives a way to uniquely define a kind of “partial trace” and hence marginal states – this in turn ensures that the theory does not permit signalling without sending a physical system [15, 55, 56].

Next we consider transformations between systems. For a pair of systems  $A$  and  $B$  there is a space of transformations  $\mathcal{T}_A^B$  from system  $A$  to system  $B$ . These again form a closed compact and convex set. In this case, there are two relevant associative composition rules. The first, parallel composition,  $\mathcal{T}_A^B \otimes \mathcal{T}_C^D := \mathcal{T}_{A \otimes C}^{B \otimes D}$  and the second, sequential composition,  $\mathcal{T}_B^C \circ \mathcal{T}_A^B := \mathcal{T}_A^C$ . These are both bilinear and must satisfy the condition that  $(T_1 \otimes T_2) \circ (T_3 \otimes T_4) = (T_1 \circ T_3) \otimes (T_2 \circ T_4)$ . Note that states (and effects) can be viewed as particular kinds of transformations, namely those that have a trivial input (resp. output) system. Denoting this trivial system by  $\star$  and noting that  $\star$  is a unit of system composition ( $A \otimes \star = A = \star \otimes A$ ) we can therefore write that  $\Omega_A = \mathcal{T}_\star^A$  and  $\mathcal{E}_A = \mathcal{T}_A^\star$ . Note that for every system,  $A$ , there is an identity transformation  $\mathbb{1}_A$  which for every  $T : A \rightarrow B$  must satisfy  $\mathbb{1}_B \circ T = T = T \circ \mathbb{1}_A$ . One can then observe that what we are defining here is nothing but a monoidal category where the objects are the GPT systems and the morphisms are the GPT transformations.

It is convenient to work with the convention whereby every GPT contains classical systems, denoted  $\Delta_I$ , in which measurement outcomes can be encoded, and which can act as control variables in descriptions of experiments. See, for example, [24, 57]. In particular, a demolition measurement of system  $A$  with outcome set  $I$  can then be viewed as a transformation in the set  $\mathcal{T}_A^{\Delta_I}$ . In order to see the connection between this and the description of measurements in terms of effects, we first set up some notation for describing classical systems.

Consider a classical system  $\Delta_I$  which is a system corresponding to, for example, some outcome degree of freedom on some measurement device where  $I$  labels the set of possible outcomes. The vector space  $V_{\Delta_I}$  will correspond to the real vector space of real functions from  $I \rightarrow \mathbb{R}$ , denoted  $\mathbb{R}^I$ . The state space  $\Omega_{\Delta_I}$  is the space of probability distributions over  $I$ , that is, real functions  $p : I \rightarrow \mathbb{R}$  such that  $p(i) \in [0, 1]$  and  $\sum_i p(i) = 1$  for all  $i \in I$ . Note that geometrically this is a simplex with vertices labelled by the elements of  $I$ . These vertices correspond to delta function distributions which we denote as  $\delta_i$ . The effect space  $\mathcal{E}_{\Delta_I}$  lives in the dual vector space, however, here we make use of the Riesz representation theorem via the inner product  $\sum_{i \in I} f(i)g(i)$  in order to view the effects as living in  $\mathbb{R}^I$ . Using this representation then the effects correspond to functions  $e : I \rightarrow \mathbb{R}$  such that  $e(i) \in [0, 1]$ . Geometrically these form a hypercube which contains the simplex of states. The vertices of the simplex, when interpreted as effects via the Riesz representation will be denoted as  $\epsilon_i$  such that  $\epsilon_i(\delta_j) = \sum_{k \in I} \delta_i(k)\delta_j(k) = \delta_{ij}$ . These systems compose via  $\Delta_I \otimes \Delta_J := \Delta_{I \times J}$ . Transformations between classical systems correspond to stochastic linear maps.

Now, note an important property of classical theory, that identity transformations  $\mathbb{1}_{\Delta_I}$  can be decomposed as  $\sum_{i \in I} \delta_i \circ \epsilon_i$ . Hence, any measurement  $M : A \rightarrow \Delta_I$  satisfies  $M = \mathbb{1}_{\Delta_I} \circ M = \sum_{i \in I} \delta_i \circ \epsilon_i \circ M$  then, noting that  $e_i := \epsilon_i \circ M \in \mathcal{T}_A^\star = \mathcal{E}_A$  we can rewrite this as  $\sum_{i \in I} \delta_i \circ e_i$ . One can then, recalling that  $\epsilon_i(\delta_j) = \delta_{ij}$ , see that it is possible to construct an isomorphism between measurements as transformations to a classical system and measurements as a collection of effects.

---

<sup>1</sup>Nondemolition measurements would live in the space  $\mathcal{T}_A^{A \otimes \Delta_I}$

## 2 The purification postulate

The purification postulate [15] has been widely studied within the literature on generalised probabilistic theories, from the study of thermodynamics [34, 35], cryptography [32, 33], and computation [29–31] through to higher-order interference [36] and reconstructions of quantum theory [28]. One of the reasons for the wide use of this principle is the fact that it is mathematically very powerful. For instance, one immediate corollary is that the state space is transitive [15] and there is a continuum of pure states (see Theorem 7 below).

Despite the utility of the purification postulate, as we discussed in Sec. 1, there are several criticisms which can be made of it. This motivates the search for alternative postulates such as the complete extension postulate that we propose here. Nonetheless, the purification postulate served as the inspiration behind the complete extension so it is useful to introduce it in some detail here.

The purification postulate takes the key ideas behind quantum purification and recasts them in the language of GPTs such that it can be taken as a postulate. We will slowly build up the relevant GPT concepts here before presenting the postulate itself.

**Definition 1** (Pure states). *A state  $s$  of a system  $A$  is said to be pure if and only if it is extremal in  $\Omega_A$ .*

In the case of quantum theory this coincides with the usual notion of purity, namely that of rank-1 density matrices. More generally, pure states are thought of as states of maximal knowledge whilst mixed states can be thought of as describing classical uncertainty of knowledge about which pure state has been prepared.

We can then define the notion of a purification of a given state:

**Definition 2** (Purifications of a state). *A purification of a state  $\omega_A$  of system  $A$  is a non-signalling extension to system  $B$ . That is, a state  $\varepsilon_{AB}$  of the composite system  $A \otimes B$ , satisfying*

$$(a) [\mathbb{1}_A \otimes u_B](\varepsilon_{AB}) = \omega_A.$$

$$(b) \varepsilon_{AB} \text{ is a pure state.}$$

*If  $s$  is pure by itself, we assume that it is one of its own purifications, by taking  $B$  system to be trivial. We denote the set of purifications of a given state  $\omega_A$  as  $\mathbf{Purif}[\omega_A]$ .*

According to Definition 2, a state may possess more than one state that purifies it. That is,  $\mathbf{Purif}[s]$  typically is not a singleton set. In quantum mechanics, for example  $|\psi_{AB}\rangle$  and  $|\psi_{AB}\rangle \otimes |\phi_{B'}\rangle$  purify the same state, i.e.,  $\text{tr}_B(|\psi_{AB}\rangle\langle\psi_{AB}|)$ . In quantum mechanics, however, purifications are unique up to isometries on the purifying system. This motivates the following:

**Definition 3** (Essential uniqueness of purifications). *The purifications of a given state  $s$  are said to be essentially unique, if and only if the elements in  $\mathbf{Purif}[\omega_A]$  with the same purifying system can be related by reversible transformations on the purifying systems.*

With these ideas in place we are then in the position that we can succinctly define the purification postulate:

**Definition 4** (Purification Postulate). *A GPT  $\mathcal{G}$  satisfies the Purification Postulate if and only if for all systems  $A$  and all states  $\omega_A \in \Omega_A$ , there exists purifications (Def. 2) which are essentially unique (Def. 3).*

### 2.1 No-go theorem for purification in discrete theories

In this subsection, we will prove that there cannot exist purifications of arbitrary states in any discrete theory. Later on, we will see that this implies that in such theories the complete extensions that we have introduced will generically not be pure. This fact has been proven in case of classical theory and for the theory of non-signalling behaviours in [15, 28]. We show below a different, direct argument, which holds in general for all non-signalling, convex discrete theories. In fact, the result developed

here is valid in any convex, non-signalling theory, provided that in each dimension number of different systems' types (state spaces with distinct shapes) is countable. We show that there is no single finite dimensional discrete theory [58–60], which vertices can purify all the states from a theory of a smaller dimension. We recall below the formal definition of a discrete theory:

**Definition 5** (Discrete theory). *A GPT  $\mathcal{G}$  is said to be discrete, if and only if, for each system  $T \in \text{Syst}[\mathcal{G}]$  there is a discrete number of pure states, that is, where each state space,  $\Omega_T$ , is a polytope.*

The demand for the theory to be convex presumes that the state space lives inside a vector space, the dimension of which is also the dimension of a theory. These assumptions together with the non-signalling condition allow for the well-definiteness of the partial trace as a linear map  $\text{Tr}_B(\cdot) : \Omega_A \otimes \Omega_B \equiv \Omega_{A \otimes B} \mapsto \Omega_A$ <sup>2</sup>, which for normalized states (in a sense of probabilistic outcomes)  $\rho_A, \sigma_B$  satisfies  $\text{Tr}_B(\rho_A \otimes \sigma_B) = (\mathbb{1}_A \otimes u_B)(\rho_A \otimes \sigma_B) = \rho_A$ , where  $u_B$  is the unit effect on the system B.

**Lemma 6.** *In any convex (nontrivial) theory, the cardinality of the set of states is at least of power of the continuum  $\mathfrak{c}$ .*

*Proof.* Let us take  $\omega_A \neq \omega'_A$  in  $\Omega_A$ . Since the theory is convex, any state of the form  $p\omega_A + (1-p)\omega'_A$ , for  $p \in [0, 1]$  is still a state in  $\Omega_A$ . The set of states  $\{p\omega_A + (1-p)\omega'_A\}_{p=0}^{p=1}$ , forms an interval in  $\Omega_A$ . Since there is a bijection between any interval and the set of real numbers that has cardinality of continuum, the state space has at least cardinality of continuum.  $\square$

We are ready to state the main theorem of this section now.

**Theorem 7.** *In any discrete theory (Def. 5) there are no (non-trivial) systems that have purifications for all states. Hence, theories with purifications (Def. 4) cannot be discrete.*

*Proof.* As the theory is a discrete theory, any system,  $T$  within the theory contains only finite number of  $v$ , vertices, and so the vertices can only purify a finite number of states  $f$  (as each pure state purifies a single mixed state). On the other hand according to Lemma 6, for all (nontrivial) systems,  $T'$ , in the theory, the cardinality of the set of states is at least  $\mathfrak{c}$ . So according to set-theoretic fact [63]  $f < \aleph_0 < \mathfrak{c}$ , in the system  $T$  there are not enough vertices to purify all states from the system  $T'$ . As it is true for all systems  $T$  in the theory, there exist no discrete theory, that can purify all states.  $\square$

Note that the proof of theorem 7 does not strictly rely on convexity, but just that the cardinality of the state space is greater than  $\aleph_0$ . One could therefore replace the assumption of convexity simply with the assumption that the number of mixed states in a considered discrete theory is greater than  $\aleph_0$ .

As a consequence of Theorem 7, we have the following Proposition. Note that this result does not rely on the essential uniqueness of purifications, only their existence.

**Proposition 8.** *For any theory (with a countable number of system types) where purifications exist for all mixed states, there must be at least one system with a continuum of pure states.*

*Proof.* In any non-trivial theory there are state spaces which correspond to convex sets other than the singleton set. Such convex sets necessarily contain a continuum of mixed states, i.e., a continuum of non-extremal states. If purifications exist then each of these mixed states must purify to a distinct pure state in the theory (as marginalisation is described by a mapping from the bipartite states to the local states).

Now suppose, for the sake of contradiction, that there is a countable number of system types, that the state spaces of these systems live in finite dimensional vector spaces (as per the definition of a GPT), and, that these state spaces all have a discrete number of vertices (i.e., pure states). Then, the theory as a whole has a countable number of pure states, and hence there are not enough pure states to purify the continuum of mixed states – in contradiction to our initial assumption.

If we want to hold on to the countable number of system types, the finite dimensional vector spaces, and the existence of purifications, it must therefore be the case that there exists a system with a continuum of pure states.  $\square$

<sup>2</sup>The theory of composition in GPTs was developed in various places, for example, Refs. [15, 61, 62].

### 3 The complete extension postulate

The idea of the complete extension is that it should be taken as a competitor to the purification postulate and is, *prima facie*, a mathematically weaker postulate. Indeed, one cannot derive either transitivity of the state space or the necessity for a continuum of pure states from this postulate whereas these are consequences of the purification postulate. Consequently, the complete extension postulate has the potential to be much more broadly applicable in scope and, hence, any results which can be derived using this postulate will hold in a broader class of theories than those derived using the purification postulate.

In this section we will formalise the definition of the complete extension postulate in the language of GPTs introduced in Sec. 2. We will work towards this by first defining ensembles, extensions and the principles of ACCESS and GENERATION within this language.

To begin, suppose we have some state,  $s$ , a vector in the convex set of states  $\Omega_A$  of a GPT system  $A$ . An *ensemble* for  $s$ , is some set of pairs  $\{(p_i, s_i)\}_{i \in I}$  such that  $s_i \in \Omega_A$  and  $\{p_i\}$  define a probability distribution ( $p_i \in \mathbb{R}$ ,  $p_i \geq 0$ , and  $\sum_i p_i = 1$ ), satisfying:

$$s = \sum_{i \in I} p_i s_i \quad (1)$$

in other words, it is some convex decomposition of  $s$  within the state space. We will denote the set of all possible ensembles for a state  $s$  as  $\mathbf{Ens}[s]$ .

The set of states which appear in some ensemble in  $\mathbf{Ens}[s]$  is, will be, a face of the state space  $\Omega_A$ , we denote this as  $\mathbf{Face}[s]$ . A state is *pure* if and only if  $\mathbf{Face}[s] = \{s\}$  and is said to be an *interior* state if and only if  $\mathbf{Face}[s] = \Omega_A$ .

We can then define a *pure ensemble* for  $s$  as an ensemble  $\{(p_i, s_i)\}_{i \in I} \in \mathbf{Ens}[s]$  in which all of the  $s_i$  are pure, that is,  $\mathbf{Face}[s_i] = \{s_i\} \forall i \in I$ . Let us denote the set of such pure ensembles as  $\mathbf{Ens}_P[s] \subseteq \mathbf{Ens}[s]$ , this is necessarily non-empty as we can always decompose a point inside a compact and closed convex set in terms of the vertices of the convex set.

An *extension* of a state  $s \in \Omega_A$  of system  $A$  is a state  $\sigma \in \Omega_{A \otimes E}$  of a bipartite system  $A \otimes E$  for which we have:

$$s = [\mathbf{1}_A \otimes u_E](\sigma) \quad (2)$$

we will denote the set of extensions,  $\sigma$ , of a state  $s$  as  $\mathbf{Ext}[s]$ , note that, in general, there will be extensions with many different extending systems  $E$ . A *purification*, should it exist, is simply some  $\sigma \in \mathbf{Ext}[s]$  which is pure, that is,  $\mathbf{Purif}[s] \subseteq \mathbf{Ext}[s]$ .

Now we can define what we mean for an extension to satisfy GENERATION. An extension  $\sigma^* \in \mathbf{Ext}[s]$  with extending system  $E^*$  is said to be *generating* if and only if for every  $\sigma \in \mathbf{Ext}[s]$  with arbitrary extending system  $E$  there is some transformation  $T_\sigma : E^* \rightarrow E$  such that:

$$\mathbf{1}_A \otimes T_\sigma(\sigma^*) = \sigma \quad (3)$$

Next we show an equivalence between the notion of an ensemble and a particular class of extensions. This particular class of extensions are extensions in which the extending system,  $E$ , is taken to be a *classical* system,  $\Delta_d$  for some  $d \in \mathbb{N}$ . We will denote this class of extensions for a state  $s$  as  $\mathbf{Ext}_{class}[s]$  and show that there is an isomorphism between  $\mathbf{Ext}_{class}[s]$  and  $\mathbf{Ens}[s]$ .

**Proposition 9.** *The space of extensions with classical extending system is isomorphic to the space of ensembles, i.e.,  $\mathbf{Ext}_{class}[s] \cong \mathbf{Ens}[s]$*

*Proof.* Starting with an ensemble:

$$\{(p_i, s_i)\}_{i \in I} \quad (4)$$

we can construct an extension:

$$\sum_{i \in I} p_i s_i \otimes \delta_i \in \Omega_A \otimes \Delta_I \quad (5)$$

where  $\delta_i \in \text{Vert}[\Delta_I]$ .



Whilst, on the other hand, if we start with an extension

$$\sigma \in \Omega_A \otimes \Delta_J \quad (6)$$

we can construct an ensemble:

$$\left\{ \left( u_A \otimes \epsilon_j(\sigma) , \frac{1}{u_A \otimes \epsilon_j(\sigma)} \mathbb{1}_A \otimes \epsilon_j(\sigma) \right) \right\}_{j \in J} \quad (7)$$

It is then simple to verify that these two constructions are the inverse of one another and hence establishes an isomorphism between the two sets.  $\square$

Using this isomorphism we can easily express the property of ACCESS in a way which makes it manifestly a special case of GENERATION. That is, an extension  $\sigma^* \in \mathbf{Ext}[s]$  with extending system  $E^*$  is an *extension with access* if and only if, for any ensemble  $\sigma \in \mathbf{Ext}_{class}[s]$  with extending system  $\Delta_I$  we can find a measurement  $M_\sigma : E^* \rightarrow \Delta_I$  such that:

$$\mathbb{1}_A \otimes M_\sigma(\sigma^*) = \sigma \quad (8)$$

Then it is clear that:

**Proposition 10.** *If an extension  $\sigma$  has the property of GENERATION then it necessarily has the property of ACCESS.*

*Proof.* This immediately follows from inspection of the two definitions, and noting that  $\mathbf{Ens}[s] \cong \mathbf{Ext}_{class}[s] \subseteq \mathbf{Ext}[s]$ .  $\square$

The converse however will not always be true, that is, from ACCESS we cannot generally derive GENERATION.

**Proposition 11.** *If an extension  $\sigma$  has the property of ACCESS then it does not necessarily have the property of GENERATION.*

*Proof.* We prove this by providing an example of a GPT which has extensions which allow for ACCESS but do not allow for generation.

Specifically, consider the GPT which has the same states and measurements as quantum theory. Then, the static part of the purification postulate, namely the existence of purifications, is satisfied. Hence, we have that any purification of a state allows us to access the ensembles of that state – that is they have the property ACCESS.

On the other hand, let us assume that the set of possible dynamics is restricted relative to quantum theory, such that transformations between non-trivial systems are necessarily noisy, that is, mixed with some non-zero parameter  $\epsilon$  of the totally depolarising channel. Then we fail the dynamical, “essential uniqueness” part of the purification postulate and so these extensions will not have the property of GENERATION.  $\square$

We are now in a position such that we can formally define the complete extension postulate within the language of GPTs.

**Definition 12** (Complete Extension Postulate). *A GPT  $\mathcal{G}$  satisfies the Complete Extension Postulate (CEP) iff: for all systems  $A$  and all states  $s \in \Omega_A$ , there exists an extension  $\sigma^* \in \mathbf{Ext}[s]$  which is generating, that is, which has the GENERATION property.*

With this definition in place, we can now compare this to the Purification Postulate [15], Def. 4. The essential uniqueness property in the purification postulate is intuitively very closely related to the GENERATION property of the complete extension postulate. However, note an important subtlety – essential uniqueness is a property for extensions on the same extending system, whilst GENERATION is a property for extensions on arbitrary extending systems. In order to derive GENERATION from essential uniqueness we must therefore invoke one additional, relatively innocuous looking, assumption about the GPT. Specifically, that the parallel composition of pure states is necessarily pure.

**Proposition 13.** *For the set of GPTs in which the product of pure states is pure, the purification postulate implies the complete extension postulate.*

*Proof.* See App. A.1. □

An immediate corollary of this is that quantum theory satisfies the CEP, as it is a GPT satisfying the purification postulate and in which the product of pure states is pure. More specifically, in the case of quantum theory, a purification  $|\psi\rangle$  of a state  $\rho$  is a complete extension. In particular, it allows for GENERATION of arbitrary extensions and ACCESS of arbitrary ensembles.

It can also easily be shown that classical probability theory satisfies the CEP. That is, the complete extension of a probability distribution with finite support over some set  $X$  has a complete extension given by simply copying the variable  $X$ . See, for example, Ex. 3.4 in [64] for details. Specifically, a general extension of a probability distribution  $p(x)$  is simply a distribution  $q(x, y)$  which marginalises to  $p(x)$ , that is, such that  $\sum_y q(x, y) = p(x)$ . It is then straightforward to verify that the bipartite probability distribution  $p_{ce}(x, x') := p(x)\delta(x, x')$  (obtained by copying  $x$ ) is a complete extension, as any other extension can be generated by applying a stochastic map to  $x'$ . This is in stark contrast to the purification postulate, which does not hold in the case of classical theory. In this sense, CEP is a far more natural postulate to consider; indeed, it can be thought of as salvaging the most intuitive part of the purification postulate, which is common to both quantum and classical theory (even though the full purification postulate does not hold in classical theory). We also show in Sec. 4, that the complete extension postulate holds in the GPT of non-signalling boxes (colloquially known as Boxworld) which is also known to fail the purification postulate.

Another proposal to modify the purification postulate was put forward in Ref. [24] which provides a time-symmetric version of the purification postulate. It can be shown that this time-symmetric purification postulate, together with another postulate from Ref. [24], namely, the existence of cups & caps, suffices to derive the CEP. More specifically, if we take a time-symmetric purification of some state and then turn the input via a cup, this defines a complete extension of the state. Since both quantum and classical theory satisfy all of the postulates of Ref. [24], it therefore immediately follows that they both have complete extensions. Moreover, superselected quantum systems also satisfy the time-symmetric purification postulate and have cups & caps, hence they also satisfy CEP.

We now consider two results which have been proven using the purification postulate, and consider whether or not they can be reproven using the complete extension postulate instead.

### 3.1 Impossibility of bit-commitment

The task of *bit-commitment* is a two-party cryptographic task that can be used as a primitive in building up other important cryptographic protocols such as coin-flipping [65] and zero-knowledge proofs [66]. It is well known that bit-commitment and its generalisation to *integer-commitment* is impossible in both quantum and classical theory [67, 68]. However, this is not true in all GPTs, indeed in [50] they show that any GPT which is non-classical but does not have entanglement allows for bit-commitment. It is therefore interesting to study what are the features of quantum theory which make this task impossible.

There have been various works on this subject including [15, Corollary 45] and [33]. The former demonstrates the impossibility of bit-commitment under a set of physical postulates including the purification postulate. The latter goes beyond a strict impossibility proof and provides an analytic lower bound on the product of the two relevant cheating probabilities, again, however, relying on the purification postulate in the process. In this section we will show that we can derive exactly the same analytic lower bound but using the complete extension rather than the purification postulate.

Before getting to this result, however, let us first introduce the protocol. For honest parties this task consists of two phases: i) the commit phase, in which Alice chooses an integer  $j \in \{1, \dots, n\}$  via a uniform random distribution and “commits” it to Bob by passing him a “token”. After doing so it should be impossible for Alice to change the value  $j$  but it should also be impossible for Bob to learn the value  $j$ . ii) the reveal phase, in which Alice communicates the integer  $j$  to Bob and sends him a second token to verify the integer. Bob must be able to check that the integer that Alice revealed was indeed the integer that she committed to in the first phase. Succinctly, for a protocol to be secure,

it must be the case that Bob cannot learn anything about the integer  $j$  prior to the reveal phase and that Alice cannot change  $j$  after the commit phase. The extent to which they can deviate from this ideal is characterised by the cheating probabilities  $p_B^*$  and  $p_A^*$ , respectively.

More formally, the protocol can be described as follows: (i) In the commit phase, Alice chooses an integer  $j \in \{1, \dots, n\}$  via a uniform random distribution and creates a corresponding bipartite state  $s^j \in \Omega_{A \otimes B}$  of the bipartite system  $A \otimes B$ . This is her ‘‘commitment’’ to the integer  $j$ . She then sends the system  $B$  to Bob to complete this phase. The idea behind this being that is now having access to only one part of the state will prevent Alice from being able to change her commitment, whilst at the same time, the fact that Bob only has access to one part of the state will prevent him from being able to learn the integer value. (ii) In the reveal phase, Alice communicates the integer  $j$  to Bob and sends him the system  $A$ , Bob applies a two-outcome measurement  $(e_{\text{accept}}^j, e_{\text{reject}}^j)$  to the bipartite system  $AB$  to check that it is in fact in state  $s^j$  to verify that the integer that Alice reveals is indeed the integer that she originally committed to.

The cheating probabilities can then be described as follows.  $p_B^*$  is the maximum probability with which Bob can learn the value  $j$  prior to the reveal phase, that is, by performing some measurement on the reduced states on system  $B$  alone. In the case of quantum theory, this can be computed as a semi-definite program, whilst, as shown in [33], in the case of generic GPTs, they are computed via cone programs (a natural generalisation of SDPs).  $p_A^*$  is the maximum probability with which Alice can reveal an integer other than the value  $j$  that she committed to in the commit phase. There are many ways in which Alice can attempt to cheat but in order to find a lower bound on  $p_A^*$  we focus on the following particular cheating strategy: she applies some transformation to her part of the shared state between the two phases of the task thereby transforming  $s_j$  into some other state  $s'_{j'}$ . Alice’s success probability with this strategy is then the probability which Bob’s measurement  $(e_{\text{accept}}^{j'}, e_{\text{reject}}^{j'})$  will give the accept outcome on state  $s'_{j'}$ . A perfect protocol would be one such that  $p_B^* = p_A^* = 0$ , that is, in which there is no probability that Bob can learn anything about  $j$  prior to the reveal phase, and there is no probability that Alice can influence the value after the commit phase. In the case that the GPT satisfies the complete extension postulate and an additional assumption called the no-restriction hypothesis [15], then it can be shown that such a perfect protocol is impossible. The following theorem encapsulates the result:

**Theorem 14.** *In any GPT satisfying the complete extension postulate and the no-restriction hypothesis [15], and in any integer-commitment protocol within that GPT, Alice and Bob’s cheating probabilities satisfy*

$$p_A^* \cdot p_B^* \geq \frac{\alpha}{n} > \frac{1}{2n} \quad (9)$$

*Proof.* (Adapted from [33]). The proof is identical up to the paragraph containing eq. 14. At which point we must rewrite the proof as follows:

Since  $\rho^j$  is the marginal of  $s^j$ , we call  $s^j$  a *extension* let  $t^j$  be an extension of  $r^j$ . Then

$$\chi^j := \frac{1}{u_{\mathcal{B}}[x] \cdot n} (s^j + t^j) \in K_{A \otimes B} \quad (10)$$

is an extension of  $x'$  for each  $j$ .

Now, consider a complete extension of  $x'$ , denoted  $\tilde{\chi} \in K_{B \otimes E^*}$ . Then, the GENERATION property implies that, for each  $j$  there exists  $T_j : E^* \rightarrow A$  such that  $\mathbb{1}_A \otimes T_j(\tilde{\chi}) = \chi^j$ .

We now have the following cheating strategy for Alice: Alice prepares the state  $\tilde{\chi}$  and passes system  $B$  to Bob keeping hold of the system  $E^*$ . Then to reveal  $j$ , she ‘steers’ the state  $\tilde{\chi}$  to  $\chi^j$  using  $T_j$  before sending  $j$  and the system  $A$  to Bob.

The remainder of the proof is identical relying on the use of the tool from convex-optimisation known as *cone programming* [69]. This is the natural generalisation of the notion of semi-definite programs, which are ubiquitous in quantum information theory, to the setting of GPTs. To date there have only been a handful of papers which utilise this tool either in quantum theory [70–74] or in GPTs [33, 75–80].  $\square$

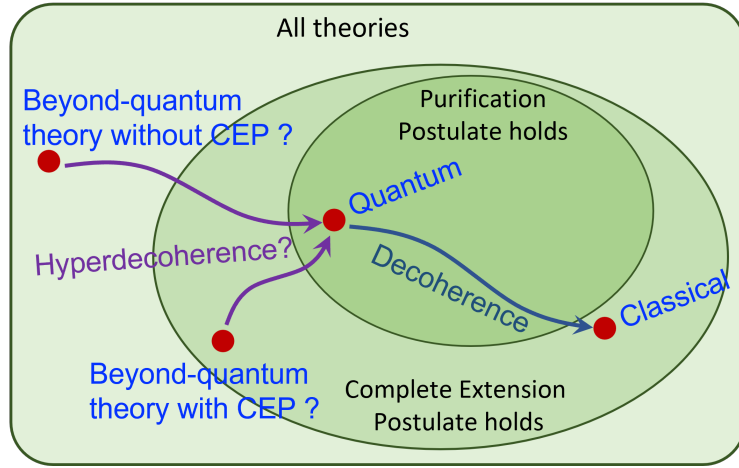


Figure 1: Schematic depiction of the relationship between the set of all theories, those are satisfying the complete extension postulate and those are satisfying the purification postulate. Also shown is the hyperdecoherence mechanism (purple arrows) relating hypothetical beyond quantum theories and quantum theory in analogy to the relationship between quantum and classical theory by a decoherence mechanism (blue arrow).

In the above, the no-restriction hypothesis [15] means that any logically possible measurement compatible with the state space of theory is physically realisable within the GPT. This is satisfied by both quantum and classical theory and is well motivated in adversarial scenarios when one does not want to make assumptions on the capabilities of the adversaries. Formally this means that, for every system  $A$  in the GPT, that the effect space  $\mathcal{E}_A$  is the dual of the state space  $\Omega$ . That is:

$$\mathcal{E}_A = \Omega_A^* := \{e \in V_A^* | e(s) \in [0, 1] \forall s \in \Omega_A \subset V_A\} \quad (11)$$

That is, any convex-linear functional which gives valid probabilities for all states corresponds to a physically realisable effect.

### 3.2 No no-go for hyperdecoherence

The question that any beyond-quantum theory must answer is the following: why do we not observe beyond-quantum phenomena in all of our current experimental tests of quantum theory? That is, within the beyond-quantum theory there must be some mechanism which explains the emergence of the quantum world. We can gain intuition for this by considering the beyond-classical theory that is quantum theory. Within quantum theory, the mechanism that best explains the emergence of the classical world is decoherence.

It has recently been understood [81, 82] how decoherence is not simply a way in which quantum states can be made effectively classical (as diagonal density matrices are isomorphic to classical probability distributions) but that this can be lifted to a mechanism which shows how the entirety of quantum theory (i.e., including states, measurements, transformations, and so on) can be made effectively classical (i.e., the theory of classical stochastic dynamics). In [37] this idea was generalised to define the notion of hyperdecoherence (a term coined in [9]) which similarly describes how the entirety of a beyond-quantum theory can be made effectively quantum.

The basic idea, is that for every beyond-quantum system,  $A$ , there exists some hyperdecoherence process,

$$\mathbf{H}_A : A \rightarrow A, \quad (12)$$

which will cause the system  $A$  to behave essentially as quantum systems. These hyperdecoherence processes must satisfy three basic properties:

1. they must be unit-effect preserving,

$$u_A \circ \mathbf{H}_A = u_A, \quad (13)$$

which is analogous to the trace-preservation condition for quantum decoherence processes; and

2. that they must be idempotent,

$$\mathbf{H}_A \circ \mathbf{H}_A = \mathbf{H}_A, \quad (14)$$

which means that once the beyond-quantum features have been lost, that hyperdecohering again does nothing.

3. Moreover, they must be chosen compositionally, that is, such that

$$\mathbf{H}_{A \otimes B} = \mathbf{H}_A \otimes \mathbf{H}_B, \quad (15)$$

such that if two systems hyperdecohere independently then the global system will behave as a quantum system.

We then can model the hyperdecoherence of a beyond-quantum theory by replacing the identity processes  $\mathbb{1}_A$  with hyperdecoherence processes  $\mathbf{H}_A$ , the intuition being that we are trying to describe a situation in which the hyperdecoherence happens on a time scale which is much shorter than any we can currently experimentally probe such that when we “do nothing” we are actually letting the system hyperdecohere. The conditions that we have imposed on the hyperdecoherence processes ensure that this can be described in consistent manner and that the result of this is a valid physical theory.

In order for such a procedure to describe the emergence of the quantum world from some beyond-quantum theory, it must be the case that the physical theory that we obtain by replacing the identities by hyperdecoherence processes in the beyond-quantum theory, is (isomorphic to) quantum theory. In [37], they demonstrate that any beyond-quantum theory which can hyperdecohere to quantum theory in this way, must violate either the causality principle – that information can only flow forwards in time, or the purification postulate. Letting go of purification seems like the more palatable option, and so this motivates the question of whether or not the complete extension postulate could be satisfied by a beyond-quantum theory which hyperdecoheres to quantum theory.

By examination of the proof of [37], however, it becomes clear that one cannot obtain the same result using the complete extension postulate – at least, not with the same proof technique. One may have suspected that the proof would still go through as it is clear that at no point does it crucially rely on the fact that a purification is actually pure. Instead, what is needed is the fact, which can be derived from the purification postulate, that pure extensions are generating. Or, in other words, that the purification postulate tells us that pure extensions are necessarily complete extensions. This fact, however, does not follow from the complete extension postulate alone, and so the proof does not go through. It therefore remains a possibility that there is a beyond-quantum theory satisfying the CEP, which moreover satisfies the causality principle and can hyperdecohere to quantum theory. An important direction for future research is therefore to try to formulate such a beyond-quantum theory.

## 4 Case study - theory of non-signalling behaviours

As a case study, in this section, we focus on the theory of non-signalling behaviours [5]. It exhibits stronger correlations between its subsystem than quantum theory, for example, it violates the Information Causality principle [83] and reaches the algebraic maximum for the CHSH inequality. It is also proved that, such behaviours can not be used for teleportation, and it does not allow for swapping of non-separable correlations [84,85] between two parties which were initially completely uncorrelated. This is not surprising as we already know that the theory of non-signalling behaviours violates the purification postulate (as it is a discrete theory), whilst in case of quantum theory, the ability to teleport an unknown state or entanglement swapping directly related to the existence of purifications [86,87].

In this section, we consider the lack of existence of purifications for the theory of non-signalling behaviours and try to bypass it with other entity fitted for the theory of non-signalling behaviours, namely, the complete extension that we introduced in Sec. 3. For a given behaviour  $P_A$ , we want to construct a non-signalling extension  $P_{AE}$ , which, whilst not being a purification, still satisfies the two key properties of purifications that we identified, namely:

ACCESS: The extension gives access to all ensembles of the extended system.

That is one can generate an arbitrary ensemble  $\{(p_i, P_A^i)\}$  of  $P_A$ , (such that  $\sum_i p_i P_A^i = P_A$  and  $\sum_i p_i = 1$ ) from the extension  $P_{AE}$ .

GENERATION: The extension can be transformed to any other extension of  $P_A$ .

That is, upon applying pre(post)-processing of inputs(outputs) (a classical channel)  $\mathcal{P}_E$  on system  $E$  of  $P_{AE}$ , it should be transformed to an arbitrary extension  $P'_{AE} = \mathcal{P}_E(P_{AE})$ , e.g.,  $\text{tr}_E(P'_{AE}) = P_A$ .

In the following, we show that this is indeed possible, and, hence, the theory of non-signalling behaviours satisfies the Complete Extension Postulate. In this way we show that CEP is not an empty postulate and complete extensions can be actually constructed in certain theories. Moreover, we explore the properties of such extensions. In particular, we show that there are infinitely many ensembles of each behaviour, therefore each behaviour can be extended to infinite number of different extensions. Each ensemble corresponds then to certain choice of input in the extending system. As a consequence, these extensions might possess an arbitrarily large number of inputs. But in the theory of non-signalling behaviours, one can have behaviours lying only in a finite-dimensional polytope. This shows that some of the ensembles (inputs) are redundant, in the sense that they can be obtained from others via adequate operations of pre(post)-processing of inputs(outputs), and others can be equivalent under reversible operations, e.g., relabelling of inputs and outputs (see Fig. 2). In this section, we will show that corresponding to each behaviour, there exist many finite-dimensional non-signalling extensions with the property of ACCESS and GENERATION. In fact, what we show, the above properties are equivalent, access to all possible ensembles is equivalent to generation of arbitrary extensions.

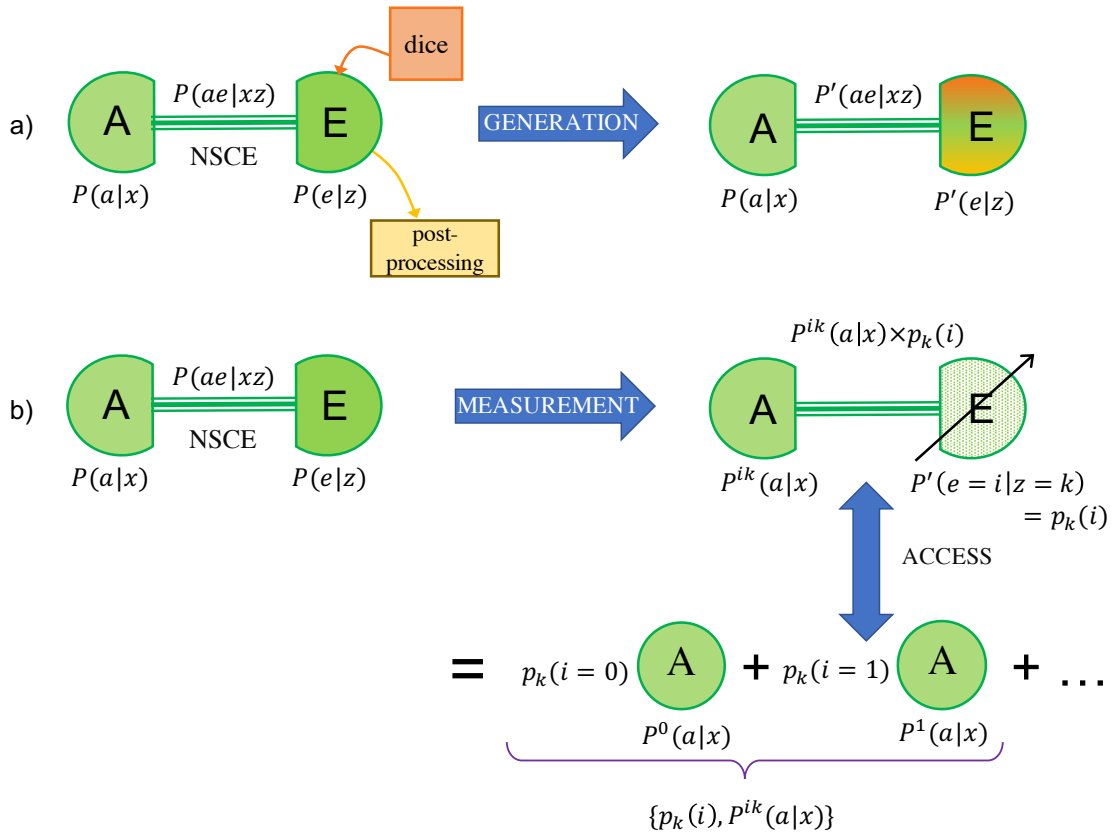


Figure 2: Schematic diagram. In any discrete convex theory, generic states do not have an extension which is a vertex (aka purification). This is different to the quantum theory that is not discrete. However, in non-signalling theory, the Complete Extension postulate is satisfied. a) From the Non-Signalling Complete Extension (NSCE) of a system,  $A$  one can generate any other extension by means of the input randomizer and output post-processing. b) The extending system  $E$  of the NSCE provides access to any ensemble  $\{p_k(i), P^{ik}(a|x)\}$  of the system  $A$ , which can be generated upon measurement  $z = k$  performed on the system  $E$ .

We also show that there is always a finite-dimensional complete extension, usually of a large dimension. We give an upper bound on this dimension as a function of the dimension of an extended

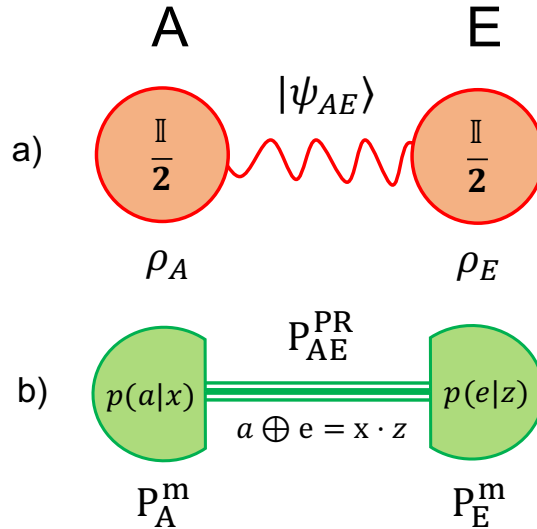


Figure 3: Schematic diagram of purifications of system  $A$  to an extended system  $E$ . a) Purification of any quantum state  $\rho_A$ , to  $|\psi_{AE}\rangle$  in the extended state space. If the quantum state is a single qubit maximally mixed state  $\frac{\mathbb{I}}{2}$ , then its purification is maximally entangled Bell state. b) Complete extension of a maximally mixed behaviour  $P_A^m$ , (given in Eq. (50)) is a pure behaviour - the well known PR box, which is however generically not the case.

system and the number of inputs and outputs. Among these possible complete extensions, we give a prescription of how to construct (up to relabeling) the minimal one, which we call the non-signalling extension with access (NSEA). This lies in the lowest dimensional polytope among all possible extensions with the above properties. The fact that access to all possible ensembles of a non-signalling system, has been considered operationally for the worst case extension someone might have, in context of device independent cryptography [88, 89] against non-signalling adversary by the Authors of Ref. [89–92], and more recently in private randomness [93–98]. The NSEA which we define is the structure responsible for this fact: access to this special extension, gives the non-signalling eavesdropper an ultimate operational power. That is, the extending system of NSEA naturally represents the minimal memory of a non-signalling adversary with maximal power in cryptography based on no-faster than light communication [51].

Interestingly, whilst we know that not all NSEAs can be pure, we show that some of them are. In particular, the aforementioned NSEA of a maximally mixed binary input binary output behaviour, i.e., the PR box [5] is a pure non-signalling behaviour. Observing that a PR box is an extension with access for a maximally mixed system can be viewed as a *derivation* of the PR box without referring to the notion of a Bell inequality – this is in stark contrast to the original approach of Popescu and Rohrlich [5]. The sense in which we derive the PR box is the following. We assume only a) the structure of single-partite systems, and b) that NSCE of any single-partite system is a valid state of the theory. In particular, we do *not* presuppose a) that all non-signalling behaviours belong to the given theory, or b) a particular composition rule, or c) the no-restriction hypothesis. This means that if the PR box was not the NSCE of any single-partite behaviour, then it would be “excluded” from the state space built with the above prescription – i.e., a priori, we do not know that the PR box is an allowed physical behaviour. What we show is that the PR box is the complete extension of the maximally mixed single-partite behaviour, and hence, by our assumptions, it is a valid bipartite state.

Since an NSEA  $P_{AE}$  is generically not extreme in the set of behaviours, it can be non-trivially extended to a new system  $E'$  giving a new NSEA  $P_{AE'E'}$ . Note that an interesting corollary of Theorem 7 is that for particular behaviours this process can be iterated indefinitely. Thus, identifying or constructing such behaviours and understanding their properties is an important direction for future research.

From our proof technique (heavily based on convex geometry) we expect that the complete extensions exemplified using the theory of non-signalling behaviours can also be defined in a similar

way in any discrete convex theory. Proving this, however, needs further work.

In Sec. 4.1, we provide the basic definition of the *Non-signalling extension with access* (NSEA) and prove that it is a complete extension as it satisfies ACCESS and GENERATION, and moreover, in Sec. 4.2 that it is a minimal complete extension. In App. B we construct many explicit examples and study their properties. In particular in App. B.1, we prove that the NSEA of the maximally mixed single party binary input and binary output behaviour is the Popescu-Rohrlich behaviour (See Fig. 3). We also give an explicit example showing that the non-signalling complete extension has the property of ACCESS and GENERATION. In App. B.4, we find the NSCE for two party binary input and binary output behaviour in the isotropic line, which enclose the Bell-Tsirelsen box and the NSCE for Specker's triangle has been found in Sec. B.3.

#### 4.1 The complete extension in the theory of non-signalling behaviours

In this section, we provide a definition of the *non-signalling extension with access* (NSEA) in the theory of non-signalling behaviours (NS), and show that NSEA satisfies the property ACCESS. We then show that NSEA also satisfies the GENERATION property, and, moreover, that in this theory these two properties are actually equivalent. This will allow us to refer to NSEA as a non-signalling complete extension (NSCE), and demonstrates that NS satisfies the complete extension postulate (CEP). What is interesting here is that it provides a constructive proof that NS satisfies the CEP, which proves useful for introducing explicitly the state of the system of the eavesdropper in the non-signalling device independent secure key agreement protocol [51].

Let us first fix the notation. We call a conditional probability distribution,  $P_A = p_{A|\mathcal{X}}(a|x)$  a box or behaviour, representing the state of system  $A$ , where  $\mathcal{X}$  stands for all possible measurement choices (inputs) and the index  $x \in \mathcal{X}$  is a particular choice.  $\mathcal{A}_x$  is the set of measurement outcomes (outputs), corresponding to input  $x$  and the index  $a \equiv a_x \in \mathcal{A}_x$  represents one particular instance. Note that if all of the outputs have the same outcome set then we will drop the subscript  $x$  and write  $\mathcal{A}$  instead. As  $P_A$  is a conditional probability distribution it must satisfy the positivity conditions,  $0 \leq p_{A|\mathcal{X}}(a|x) \leq 1$ ,  $\forall a \in \mathcal{A}_x$ ,  $x \in \mathcal{X}$ , and the normalization condition  $\sum_{a \in \mathcal{A}_x} p_{A|\mathcal{X}}(a|x) = 1$ ,  $\forall x \in \mathcal{X}$ <sup>3</sup>. Let  $P_{AE} = p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz)$  be a bipartite conditional probability distribution, where the system  $E$  has associated input and output sets  $\mathcal{Z}$  and  $\mathcal{E}$  respectively. We say that it is a non-signalling extension of the behaviour  $P_A$  if  $P_A$  is its marginal distribution

$$\sum_{e \in \mathcal{E}_z} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = p_{A|\mathcal{X}}(a|x), \quad \forall a \in \mathcal{A}_x, x \in \mathcal{X}, z \in \mathcal{Z}, \quad (16)$$

and it satisfies non-signalling conditions

$$\sum_{e \in \mathcal{E}_z} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = \sum_{e \in \mathcal{E}_{z'}} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz'), \quad \forall a \in \mathcal{A}_x, x \in \mathcal{X}, z, z' \in \mathcal{Z}, \quad (17)$$

$$\sum_{a \in \mathcal{A}_x} p_{AE|\mathcal{X}\mathcal{Z}}(ae|xz) = \sum_{a \in \mathcal{A}_{x'}} p_{AE|\mathcal{X}\mathcal{Z}}(ae|x'z), \quad \forall e \in \mathcal{E}_z, z \in \mathcal{Z}, x, x' \in \mathcal{X}, \quad (18)$$

For the sake of simplicity we omit the subscript, and we assign  $p_{A|\mathcal{X}}(a|x) \equiv P_A(a|x)$ . The set of all non-signalling behaviours of system  $A$  is denoted by  $\Omega_A$ , mutatis mutandis for multipartite systems. The  $P_A$  can have inner non-signalling structure [5], i.e., system  $A$  can be compound (e.g.,  $A = A_1 A_2$ ) itself, where the system  $A_1$  can not signal to system  $A_2$  and vice versa. We will encounter this later when we consider the extension of the binary two input two output behaviour.

All  $n$ -partite non-signalling behaviours, with the same cardinalities of inputs and outputs satisfy a set of linear equations and inequalities (constraints). These constraints define a convex, bounded polyhedron (a polytope) that is a subset of  $\mathbb{R}^N$  for some  $N \in \mathbb{N}$ , with a finite number  $D$  of vertices. Each point within this polytope represents a different non-signalling behaviour. In such a polytope of behaviours, we will distinguish those which are *extremal* (vertices), i.e., the behaviours which can not be expressed as a convex combination of other non-signalling behaviours, distinguishing them by subscript  $\mathbf{E}$ :  $P_{\mathbf{E}}$ .

<sup>3</sup>The summation over the output here is an analogue of partial trace.



An ensemble  $\{(p_i, P^i)\}_i$ , of a behaviour  $P$ , where  $P = \sum_i p_i P^i$ , we will denote by  $\mathcal{E}(P)$ . In this paper, we will consider only those ensembles which consists of finite number of elements. We will also need the notion of the set of *members* of an ensemble  $\mathcal{E}(P) = \{(p_i, P^i)\}_i$ , this is defined by  $V(\mathcal{E}(P)) = V(\{(p_i, P^i)\}_i) := \{P^i : p_i > 0\}$ , and its *distribution*, which is  $\{p_i\}$ . If *all* the members,  $P^i \in V(\mathcal{E})$ , are extremal (pure),  $P_E^i$ , then we call the ensemble a *pure members ensemble* (PME) and denoted as  $\mathcal{E}_{\text{pure}}(P)$ . We say that an ensemble  $\mathcal{E}_{\text{pure}}(P)$  is generated on system  $A$  by measurement  $\mathcal{M}$  on system  $E$ , if upon this measurement on the extending system, the outcome  $e = i$  is obtained on  $E$  with probability  $p_i$  and conditionally upon it, the state of the system  $A$  is described by behaviour  $P_E^i$ . By  $S_P$  we will denote the set of all PME of a behaviour  $P$ .

Our aim is to obtain the minimal extension with properties ACCESS and GENERATION. To begin with, however, as an intermediate step used to simplify further proofs, we will define an extension, satisfying properties ACCESS and GENERATION yet of a larger dimension hence called *overcomplete non-signalling extension with access* (ONSEA), defined as:

**Definition 15** (Overcomplete non-signalling extension with access - ONSEA). *Given a behaviour  $P_A : P_A(a|x)$ , we say that a behaviour  $P_{AE} : P_{AE}(ae|xz)$  is its overcomplete non-signalling extension with access (extension to system  $E$  with access to system  $A$ ), if for any input choice  $z = k$  and outcome  $e = i$  obtained in the extending part, there holds*

$$P_{AE}(a, e = i|x, z = k) = P_E^{i,k}(a|x)p(e = i|z = k), \quad (19)$$

such, that, for each  $k$ , the ensemble  $\left\{ \left( p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$  is a pure members ensemble of the behaviour  $P_A$ , and corresponding to each pure members ensemble of  $P_A$ , there is exactly one input  $z = k$ , in the extending system which generates it.

It is simple to see that the overcomplete extension ONSEA of an arbitrary behaviour  $P_A$  exists. Indeed,  $P_A$  belongs to a polytope. That is it belongs to a set with a finite number of pure behaviours. Any pure member ensemble of  $P_A$  is a subset of the set of pure behaviours. Hence, there is finite number of the latter ensembles and we can construct extension where for each of such ensemble we have  $k$  such that  $z = k$  generates it. This extension is precisely the ONSEA. We will denote it as  $\tilde{\mathcal{E}}(P)_{AE}$ . Notice that when a particular input  $z = k$  is chosen in the extending part, an outcome  $e = i$  occurs with the probability  $p(e = i|z = k)$ , resulting in a pure behaviour  $P_E^{i,k}(a|x)$  in part of  $A$ . Eq. (19), expresses the partially measured behaviour with the probability times the conditional pure behaviour. Moreover,  $\left\{ \left( p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$  is a pure members ensemble of  $P_A$ , for each  $k$ , as

$$\sum_i p(e = i|z = k) P_E^{i,k}(a|x) = \sum_i \tilde{\mathcal{E}}(P)_{AE}(a, e = i|x, z = k) = P_A(a|x), \quad (20)$$

and given an overcomplete extension (ONSEA) there are  $|\mathcal{Z}|$  pure members ensembles, where  $\mathcal{Z} = \{z\}$ , is the set of all input choices of the extending system.

The above definition of ONSEA satisfies the non-signalling condition for its both subsystems. For system  $A$ , it is by construction, and for system  $E$  it holds due to the fact that for each input output pair  $(z = k, e = i)$  of  $E$ , system  $A$  holds a behaviour  $P_E^{i,k}(a|x)$  according to Eq. (19), which gives 1 when summed over  $a$ .

We will see later that the ONSEA is actually a complete extension as it allows for ACCESS and GENERATION. However, the dimension of the extending system is larger than necessary, we therefore seek to construct a minimal complete extension<sup>4</sup>. To do so we will now introduce a representative subset of all PMEs. An ensemble  $\mathcal{E}_{\text{pure}}(P) = \{p_i, P_E^i\}_i$ , of a behaviour  $P = \sum_i p_i P_E^i$ , is called *minimal ensemble*, if any proper subset  $V' \subset V(\mathcal{E}_{\text{pure}})$ , with another choices of probabilities  $\{p'_j\}$  is not an ensemble of the behaviour  $P$ , i.e.,  $P \neq \sum_{j: P_E^j \in V'} p'_j P_E^j$ . Any minimal ensemble will be denoted as  $\mathcal{M}(P)$ , to distinguish it from an arbitrary PME that is not necessarily minimal.

Next, we will eliminate the redundant ensembles from the ONSEA, to obtain the NSEA that has the lowest number of inputs sufficient to satisfy ACCESS property. To achieve this we will show

<sup>4</sup>Minimality will be proven in Prop. 27 in the following subsection.

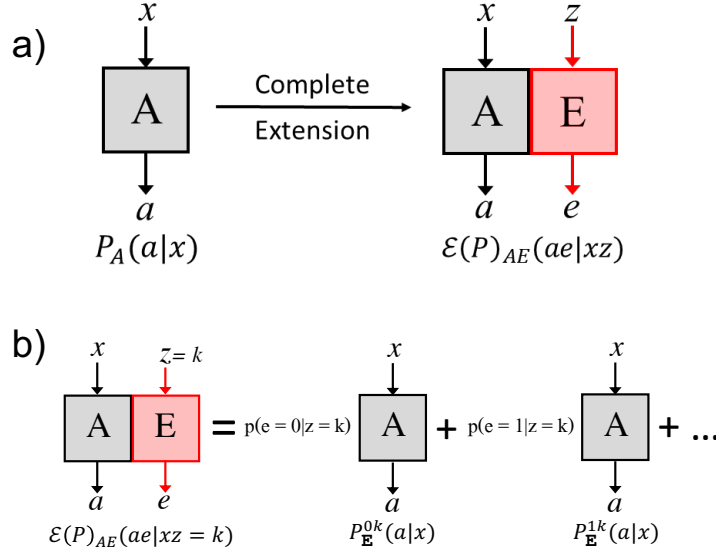


Figure 4: Panel a), schematic diagram of a NSEA of an arbitrary behaviour  $P_A$  to  $\mathcal{E}(P)_{AE}$ . Panel b), Corresponding to each input  $z = k$  in part of the extending system, the composite behaviour partitioned to a minimal ensemble  $\{p(e|z = k), P_E^{ek}(a|x)\}$  in part of  $A$ .

that having access to all minimal ensembles  $\{\mathcal{M}(P)\}$  of  $P$ , along with arbitrary randomness, one can generate any PME,  $\mathcal{E}_{pure}^{\mathcal{E}}(P)$ .

**Theorem 16.** *ACCESS to PMEs is equivalent to ACCESS to minimal ensembles.*

*Proof.* See App. A.2. □

**Remark 17.** *This theorem (Thm. 16) can be generalized to any GPT that is a discrete theory (see Definition 5).*

The above theorem motivates the following definition of *non-signalling extension with access* (NSEA), which, like the extension defined in Definition 15, also satisfies the properties of access and generation:

**Definition 18** (Non-signalling Extension with Access). *Given a behaviour  $P_A : P_A(a|x)$ , we say that a behaviour  $P_{AE} : P_{AE}(ae|xz)$  is its non-signalling extension with access extended to system  $E$  if for any input choice  $z = k$  an outcome  $e = i$  occurred in the extending system, there holds*

$$P_{AE}(a, e = i|x, z = k) = P_E^{i,k}(a|x)p(e = i|z = k) \quad (21)$$

such, that, for each  $k$ , the ensemble  $\left\{ \left( p(e = i|z = k), P_E^{i,k}(a|x) \right) \right\}_i$  is a minimal ensemble of the behaviour  $P_A$ . Moreover, corresponding to each minimal ensemble of  $P_A$ , there is exactly one input  $z = k$ , in part of the extending system which generates it.

We will represent the NSEA of an arbitrary behaviour  $P_A$  as  $\mathcal{E}(P)_{AE}$ . If there is more than one subsystem in the system  $A$  satisfying non-signalling constraints, i.e., if  $A \equiv A_1 A_2 \dots A_N$ , and none of the  $k$  systems  $\{A_{i_1}, \dots, A_{i_k}\}$  with  $I = \{i_1, \dots, i_k\}$  can signal to the remaining  $N - k$  systems  $\{A_{j_1}, \dots, A_{j_{N-k}}\}$  with  $j_l \in \{1, \dots, N\} \setminus I$ , for all  $k \in \{1, 2, \dots, N-1\}$ , then also our method of constructing NSEA holds, and it will be denoted by  $\mathcal{E}(P)_{A_1 A_2 \dots A_N E}$ . A schematic diagram of NSEA has been depicted in Fig. 4. For NSEA defined above, the inputs of the extending system correspond to the minimal ensembles of the given behaviour  $P_A$ . If there are  $N$  number of minimal ensembles, then the total number of input choices in the extending part is  $|\mathcal{Z}| = N$ .

**Proposition 19.** *For each behavior  $P_A$  of system  $A$ , its NSEA  $\mathcal{E}(P)_{AE}$  is unique up to the local relabeling on the extending system  $E$ .*

*Proof.* Let  $P_{AE}$  and  $P'_{AE}$  be two different NSEA of a behaviour  $P_A$ . First note, that  $P_A$  determines the number of inputs  $|E|$ . However the inputs of  $P_{AE}$  can be labelled in different way than that of  $P'_{AE}$ . However, due to Definition 18, each input of  $P_{AE}$  corresponds to a unique minimal ensemble of  $P_A$ . The same holds for  $P'_{AE}$ . Thus for each input  $e$  of  $P_{AE}$  there exists input  $e'$  of  $P'_{AE}$  which corresponds to the same unique minimal ensemble. Hence for this pair of inputs the distribution of outputs of  $P_{AE}$  and  $P'_{AE}$  are the same up to labeling of outputs. For this reason, the two extensions differ only by relabelling of inputs and outputs of the extending system. Thus, relabellings establish an equivalence class over different behaviors being NSEA of behavior  $P_A$ .  $\square$

Note that one can easily show that such an extension necessarily exists for any behaviour  $P_A$ . For example, one can check (see Section B.1 of the Appendix) that the NSEA of the behaviour  $P_A$ , the maximally mixed behaviour with a single binary input and single binary output, extended to system  $E$ , has the following structure:

$$P_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ & \swarrow e \searrow a & & & & \\ z & & 0 & 1 & | & 0 & 1 \\ \hline 0 & 0 & 1/2 & 0 & | & 1/2 & 0 \\ & 1 & 0 & 1/2 & | & 0 & 1/2 \\ \hline 1 & 0 & 1/2 & 0 & | & 0 & 1/2 \\ & 1 & 0 & 1/2 & | & 1/2 & 0 \end{array}, \quad (22)$$

which is nothing but the famous Popescu-Rohrlich ( $PR$ ) behaviour satisfying conditions  $x.z = a \oplus e$  with  $\oplus$  being addition modulo 2. We have thus arrived at this structure without referring to the CHSH inequality [99] (in contrast to the way in which it was done in [5]).

We are in a position to show, that having access to NSEA we can generate any PME.

**Corollary 20.** *The non-signalling extension with access (NSEA) of a behaviour  $P$  given in Definition 18, together with access to arbitrary local randomness, gives access to any pure members ensemble of a behaviour  $P$ .*

*Proof.* See App. A.3.  $\square$

From the above corollary it is clear that an arbitrary PME can be accessed from the NSEA by using a randomness generator  $\{p(k)\}$ , i.e., by using a dice(coin). To access all possible PME, one needs an access to arbitrary randomness. This can be done by setting the output ( $k$ ) of a dice with a distribution,  $p(k|z')$ , where  $z'$  is the tuning parameter, as the input of the extending party of NSEA. Here  $|k| = |\mathcal{Z}|$ , and  $|z'|$  will be equal to the possible number of PME one wants to generate. The dice can be thought of as a local behaviour with  $z'$  being the input and  $k$  as the output. Different choices of the  $z'$  can be considered as dices with different probabilities of outcome, actually led to different PMEs. Accessing all possible PME has been pictorially depicted in Fig. 5.

An explicit example of constructing an arbitrary pure members ensembles has been given in Sec B.1.2. Where we have chosen an arbitrary behaviour containing single binary input and single binary output.

**Theorem 21.** *ACCESS to PMEs is equivalent to ACCESS to all ensembles.*

*Proof.* Let us consider the access of the mixed ensemble  $\mathcal{E}_{mix}(P) = \{(p_m, P^m)\}_m$ , which is the most general ensemble, from the set of all ensembles of a given behavior  $P$ . Now each  $P^m$  lies in the same polytope as  $P$ , hence, all of them has a pure behaviour decomposition,

$$P^m = \sum_i q_i^m P_{\mathbf{E}}^i, \quad (23)$$

where  $\sum_i q_i^m = 1$ ,  $\forall m$  and  $0 \leq q_i^m \leq 1$ ,  $\forall i, m$ . Note that this decomposition is not unique, unless it is a minimal decomposition. Now  $P = \sum_m p_m P^m = \sum_{m,i} p_m q_i^m P_{\mathbf{E}}^i = \sum_i r_i P_{\mathbf{E}}^i$ , where  $r_i = \sum_m p_m q_i^m$ , implies that  $\{(r_i, P_{\mathbf{E}}^i)\}_i$  is also a PME of  $P$ . From Theorem 20, we know that by using an appropriate

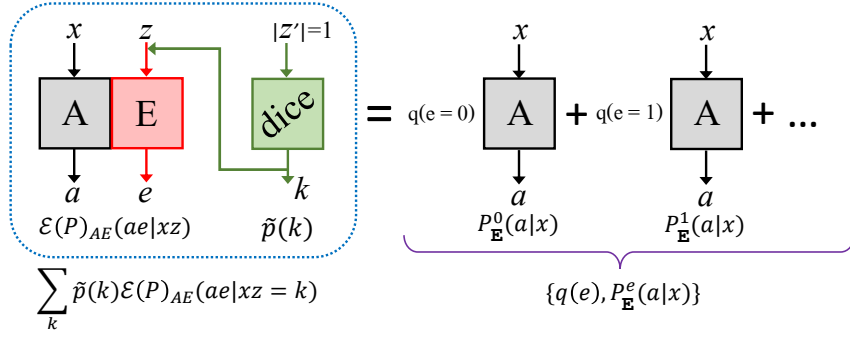


Figure 5: Schematic diagram visualizing the mixing the minimal ensembles  $\{\mathcal{M}_k\}$  of  $A$  with arbitrary randomness  $p(k)$  in part of the extending system on the NSEA of the behaviour, which is obtained from the output of a dice (a local behaviour with unary input), results an arbitrary pure members ensemble  $\{p(k), \{(p(e|z = k), P_E^{e|z=k}(a|x))\}\} = \{q(e), P_E^e(ab|xy)\}$ , where  $q(e) = \sum_k p(k)p(e|z = k)$ .

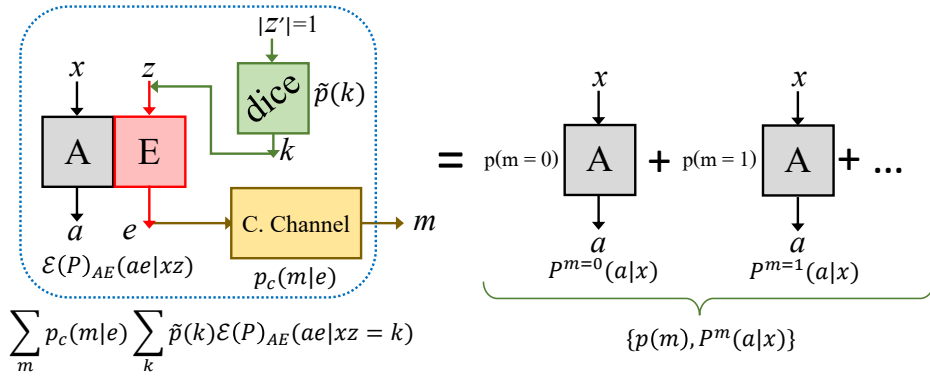


Figure 6: Explanation of Theorem 22 of accessing all possible ensembles even mixed in part of the extended system. By passing the output  $e$  of the extending party's (Eve) behaviour, through a post-processing channel PC, Eve is able to interpret the behaviour shared by Alice and Bob as an ensemble of mixed bipartite behaviours  $\{p(m), P^m(ab|xy)\}$ , where  $P^m(ab|xy)$ , are mixed behaviours and can be expanded as  $P^m(ab|xy) = \sum_e q_e^m P_E^e(ab|xy)$ . Now the post-processing channel  $p_c(m|e)$ , helps Eve to interpret the mixed behaviours as  $\sum_e p_c(m|e)r(e)P_E^e(ab|xy) = p(m)P^m(ab|xy)$ .

randomness generator in the input of the extending system  $E$  (input randomizer), such PME can be realized. To interpret the mixed ensembles, from  $\{(r_i, P_E^i)\}_i$ , the extending system send her output ( $e = i$ , of the NSEA along with an appropriate input randomizer) through a post-processing channel  $p_c(m|e)$ . The existence of this channel (24),  $p_c(m|e)$  with  $\sum_m p(m|e) = 1$ , is guaranteed by the initial decomposition of the members of the mixed ensemble into pure behaviours.

$$\sum_i p_c(m|e = i)r_i P_E^i = \sum_i \frac{p_m q_i^m}{r_i} r_i P_E^i = p_m P^m, \quad (24)$$

where  $p_c(m|e)$  take the form  $\frac{p_m q_e^m}{r_e}$ . □

**Corollary 22.** (NSEA satisfies ACCESS) *The extending system of the NSEA gives access to any possible (even mixed) ensemble of the extended behaviour.*

*Proof.* Suppose the extending system wants to access an arbitrary mixed ensemble, once she obtained the NSEA. From Corollary 20, NSEA gives access to any PME, and from Theorem 21, it is clear that access to PMEs is equivalent to access to all ensembles. Hence, NSEA gives access to any possible (even mixed) ensemble of the extended behaviour. □

From the above theorem, we observe that, corresponding to each mixed ensemble of  $P$ , there exists at least one PME of  $P$ , from which one can obtain the mixed ensemble by using an appropriate

post-processing channel. In this construction the NSEA plays an important role due to the fact that Theorem 20 certifies the existence of that particular PME. In Sec. B.1.3 of the Appendix, we exemplify Theorem 22 for a single party binary input binary output behaviour, by providing the explicit form of input randomizer and the post-processing channel. A visualization of Theorem 22, has been given in Fig. 6. Note that the input randomizer  $\tilde{p}(k)$ , which generates the required PME of P, is a dice with a unary input (trivial input),  $\tilde{p}(k|z' = \text{fix})$ . One can generate all possible ensembles by appropriately tuning the dice with the help of the dice input  $z'$  and a post-processing channel (also dependent on  $z'$ ) on part of the extending system.

We have now seen that we can construct the NSEA which is an extension with ACCESS. However, in order for this to be a complete extension we need that it also satisfies GENERATION. One can show that this is indeed the case by demonstrating that ACCESS and GENERATION are equivalent conditions in the theory of non-signalling behaviours.

**Theorem 23.** *In the theory of non-signalling behaviours ACCESS is equivalent to GENERATION.*

*Proof.* See App. A.4. □

From the above theorem, Theorem 22, and the fact that NSEAs exist for all behaviours, we obtain the main result of this section:

**Corollary 24.** *The Non-signalling Extension with Access (NSEA) has the properties of ACCESS and GENERATION, that is, it is a Non-signalling Complete Extension (NSCE). Hence, the theory of non-signalling behaviours satisfies the Complete Extension Postulate (CEP).*

## 4.2 On the dimensionality of the complete extension

In this subsection we are going to discuss the size of the non-signalling complete extension NSEA, and, what is most important, is that we are going to show that it is finite. The restriction on the size of the non-signalling complete extension (NSEA) is particularly important from point of view of non-signalling Device Independent cryptography [89, 90, 100], and interesting on its own. We finish this section with another observation, namely we show that for any behaviour  $P$ , its NSEA  $\mathcal{E}(P)$  has the lowest dimension (is minimal) among all non-signalling extension of  $P$  having the property of ACCESS.

The following Theorem holds.

**Theorem 25.** *Let be  $\mathcal{B}$  a polytope of  $n$ -partite non-signalling behaviours, with  $m_i$  inputs for parties and  $v_{ij}$  outputs respectively. Then for each  $P \in \mathcal{B}$ , there exists a non-signalling polytope  $\tilde{\mathcal{B}} \ni \mathcal{E}(P)$  which the NSEA lives in, such that:*

$$\dim \tilde{\mathcal{B}} < (\dim \mathcal{B} + 1) \times \left( \left( \binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1} \right) \dim \mathcal{B} + 1 \right), \quad (25)$$

where:

$$\dim \mathcal{B} = \prod_{i=1}^n \left( \sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}. \quad (26)$$

*Proof.* The proof is the content of Section B.2 of the Appendix. □

In Theorem 25 we were interested only in showing that the dimension of the complete extension is always finite and therefore the upper bound is very loose. For instance, the expression in equation (25) in the case of  $n = 1$ ,  $m_i = 2$ ,  $v_{ij} = 2$  yields the dimension of 339, for  $n = 2$ ,  $m_i = 2$ ,  $v_{ij} = 2$  we have c.a.  $1.2 \times 10^{54}$ , and for  $n = 2$ ,  $m_i = 3$ ,  $v_{ij} = 3$  we obtain c.a.  $1.14 \times 10^{1762}$ . All exemplified results are far above any numerical predictions. The discussion about the possible improvements in the upper bound is left to the Appendix. Importantly for us, however, is the following simple corollary of the existence of any upper bound, namely:

**Corollary 26.** *The dimension of the NSCE can always be finite.*

A tempting question to ask about NSEA is whether it has the lowest dimension amongst all NSCEs, and therefore whether it is the minimal one. Similarly to earlier in this section, we refer to the dimension of the extension as the dimension of the behaviour polytope to which it belongs. The answer to this query is positive and therefore we finish this section with the following proposition.

**Proposition 27.** *Among all non-signalling extensions of a behaviour  $P$ , having the property of ACCESS, NSEA  $\mathcal{E}(P)$  is a minimal one.*

*Proof.* Let  $P$  be a non-signalling behaviour in  $\mathcal{B}$ , and  $\mathcal{E}(P) \in \tilde{\mathcal{B}}$  be its NSEA. The dimensions of non-signalling polytopes can be determined with equation (89) as before. Suppose now, there exists another non-signalling extension  $\hat{\mathcal{E}}(P) \in \hat{\mathcal{B}}$ , having a property of ACCESS, such that  $\dim \hat{\mathcal{B}} < \dim \tilde{\mathcal{B}}$ . The fact that  $\hat{\mathcal{E}}(P)$  has the property of ACCESS implies that upon processing (possibly trivial) of inputs and outputs of the extending system with local randomness,  $\hat{\mathcal{E}}(P)$  also has access also to all minimal ensembles of  $P$ . However, the minimal ensembles of  $P$  are extremal points in the polytope of all possible ensembles (see Theorem 30) of the behaviour  $P$ , and so cannot be created via probabilistic processing of inputs and outputs in the extending system (class of operations considered in ACCESS). Therefore,  $\hat{\mathcal{E}}(P)$  having a property of ACCESS, must have for each minimal ensemble of  $P$ , an input that generates it, like  $\mathcal{E}(P)$  does. This implies that  $\dim \hat{\mathcal{B}} \geq \dim \tilde{\mathcal{B}}$ , and so proves by contradiction that NSEA  $\mathcal{E}(P)$  is the minimal extension of  $P$  having the property of ACCESS.  $\square$

## 5 Conclusions

To summarize, our main contribution is a new concept, which is the *complete extension*. We show that complete extensions are present in classical theory, quantum theory, super-selected quantum theory, the theory of non-signalling behaviours, and, moreover, any theory satisfying the purification postulate in which the product of pure states is pure. We also postulate that it may exist hypothetical beyond-quantum theories which could hyperdecohere to quantum theory. In the case of quantum and classical theory, as well as the theory of non-signalling behaviours, we can explicitly construct these complete extensions. This notion implies a number of paths for research, some of which we have exemplified in our case study on the theory of non-signalling behaviors.

The idea of the CEP sets a demarcation line in the set of results obtained on the basis of the purification postulate. It divides them into those that really require all of the purification postulate and those for which the CEP suffices. We exemplify this by considering the possibility of bit commitment showing that the no-go for it is not specific to theories with the purification postulate. This has the added benefit of giving a unified proof for the quantum and classical cases.

The CEP may also be viewed as a razor for excluding theories that can not substitute quantum theory in the future. Indeed, a theory not satisfying CEP may not be physical, as (see Theorem 24) even the theory of non-signalling behaviours satisfies it. The easiest way to obtain theories that do not satisfy CEP is to restrict state space or dynamics so that it is not possible to generate all extensions of some state. An example of a theory that does not satisfy CEP for this reason is given in [50].

We also show that an interesting “mirror” property of quantum purifications no longer holds for the case of non-signalling behaviours. That is, suppose that we have a purification  $|\psi_{AB}\rangle$  of a quantum state  $\rho_A$ , and let us define  $\sigma_B := \text{Tr}_A(|\psi_{AB}\rangle\langle\psi_{AB}|)$ . Then, any purification of  $\sigma_B$  on system  $A$ , (which necessarily exists as  $|\psi_{AB}\rangle$  is such a purification) we denote as  $|\phi_{AB}\rangle$  and call the “mirror” purification, is equal to  $|\psi_{AB}\rangle$  up to a local unitary on  $A$ . That is  $|\psi_{AB}\rangle = U_A \otimes \mathbb{1}_B |\phi_{AB}\rangle$  for some unitary  $U_A$ . In general, beyond-quantum theory, it need not be the case, as we have exemplified in the theory of non-signalling behaviours (see Sec. B.1.6 of the Appendix). A properly defined minimal distance between the complete extension and its “mirror” one, overall systems whose complete extension is not pure, can characterise to what extent a given theory departs from quantum mechanics.

Following the development of quantum cryptography, one can ask if a post-quantum theory should lead to secure communication. As it is shown in [101], this need not always be the case. In the quantum case, the system  $E$  of the purification  $\psi_{AE}$  describes the worst-case state of the knowledge of the quantum adversary that may have about a given system  $A$  in many information processing protocols

such as QKD. Similarly, in the theory of non-signalling behaviours, the system  $E$  of the NSCE  $\tau_{AE}$  represents the worst case knowledge of an adversary about system  $A$ . The minimal dimension of the system  $E$ , therefore, represents the maximal memory needed by the adversary. We, therefore, enable the study of the worst-case adversary’s capabilities in other post-quantum theories, see Ref. [51] for an application of this idea.

Significantly, one can use CEP to define composed systems in a post-quantum theory. Namely given a set of states of a single system  $A$ , one can define as valid states of a joint system  $AB$  only those that are either (i) complete extensions of the states of  $A$ , or (ii) states obtained from such complete extensions by some operations valid in this theory. We exemplify this approach by arriving at the structure of the PR box, assuming (i) in the case that  $A$  is a local system from NS with two binary inputs and two binary outputs. Together with (ii) the local transformations in NS (including e.g. relabelings of inputs and outputs), we can reach all the non-local vertices of the non-signalling polytope of behaviours with two binary inputs and two binary outputs. Given the fact, that the complete extension of a local deterministic behavior of system  $A$  is a product of two deterministic behaviors on systems  $A \otimes B$ , via operation of mixing two or more behaviors we show that the state space of non-signalling behaviours includes the whole non-signalling polytope of behaviours with two binary inputs and two binary outputs.

Interestingly, a post-quantum theory that does not satisfy the purification postulate, but rather the CEP, can have a property that there is an infinite sequence of complete extensions of complete extensions as none of them is pure. This will always hold unless the number of pure states in the theory is of cardinality continuum at least. This is indeed the case in both classical theory and the theory of non-signalling behaviours.

Verifying if CEP holds in the case of other beyond-quantum theories and studying its consequences is an important direction to follow. We formulate the hypothesis that there exists a generalised probabilistic theory with complete extensions that may naturally hyper-decohere to quantum mechanics. This is supported by demonstrating that the proof of an existing no-go result, which applies to generalised probabilistic theories with purifications, no longer holds. The confirmation of this hypothesis would open a new arena in which new physical laws and phenomena may be searched.

It is also interesting to study whether complementing CEP with a more dynamical axiom, e.g., the one linked to the Neumark extension or Stinespring dilations, can lead to a more powerful postulate. In the case of the purification postulate such a dynamical postulate can be derived, it is therefore interesting to investigate whether this is also the case for a dynamical version of the CEP, or whether this must be additionally postulated. Determining what form this dynamical postulate should take, and demonstrating that it indeed holds in the theory of non-signalling behaviours, is perhaps the most important direction for follow up research.

## Acknowledgments

MW, TD and KH acknowledge grant Sonata Bis 5 (grant number: 2015/18/E/ST2/00327) from the National Science Center. TD acknowledge Omer Sakarya for his help in numerical simulation. KH, PH and LP acknowledges ERC grant QOLAPS and Polish Ministry of Science and Higher Education Grant no. IdP2011 000361. MP acknowledges support from European Union’s Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Action OPERACQC (Grant Agreement No. 661338), and from the Foundational Questions Institute under the Physics of the Observer Programme (Grant No. FQXi-RFP-1601). RR acknowledges support from the Early Career Scheme (ECS) Grant No. 27210620, the General Research Fund (GRF) Grant No. 17211122, and the Research Impact Fund (RIF) Grant No. R7035-21. MW, TD, KH, and PH were partially supported by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (contract no. 2018/MAB/5). JHS was supported by the Foundation for Polish Science through IRAP project co-financed by EU within Smart Growth Operational Programme (contract no. 2018/MAB/5). JHS thanks Ana Belén Sainz for helpful discussions.

## Code availability

The code that supports the theoretical plots and tables within this paper is available from the corresponding author upon reasonable request.

## References

- [1] Asher Peres. “Karl Popper and the Copenhagen interpretation”. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **33**, 23–34 (2002).
- [2] Paul Adrian Dirac. “The principles of quantum mechanics (third ed.)”. *Clarendon Press Oxford*. (1948).
- [3] J. von Neumann. “Mathematische grundlagen der quantenmechanik”. *Springer, Berlin*. (1932).
- [4] Johann von Neumann. “Mathematische grundlagen der quantenmechanik”. *Springer*. (1932).
- [5] S. Popescu and D. Rohrlich. “Quantum nonlocality as an axiom”. *Found. Phys.* **24**, 379–385 (1994). url: [doi.org/10.1007/BF02058098](https://doi.org/10.1007/BF02058098).
- [6] Wolfgang Bertram. “An Essay on the Completion of Quantum Theory. I: General Setting” (2017). [arXiv:1711.08643](https://arxiv.org/abs/1711.08643).
- [7] Wolfgang Bertram. “An Essay on the Completion of Quantum Theory. II: Unitary Time Evolution” (2018). [arXiv:1807.04650](https://arxiv.org/abs/1807.04650).
- [8] Lucien Hardy. “Quantum theory from five reasonable axioms” (2001). url: [arxiv.org/abs/quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
- [9] Karol Życzkowski. “Quartic quantum theory: an extension of the standard quantum mechanics”. *J. Phys. A: Math. Theor.* **41**, 355302 (2008). url: [doi.org/10.1088/1751-8113/41/35/355302](https://doi.org/10.1088/1751-8113/41/35/355302).
- [10] Lee Smolin. “Could quantum mechanics be an approximation to another theory?” (2006). [arXiv:quant-ph/060910](https://arxiv.org/abs/quant-ph/060910).
- [11] Marshall H Stone. “On one-parameter unitary groups in hilbert space”. *Ann. Math.* **33**, 643–648 (1932). url: [doi.org/10.2307/1968538](https://doi.org/10.2307/1968538).
- [12] Andrew M Gleason. “Measures on the closed subspaces of a hilbert space”. In *The Logico-Algebraic Approach to Quantum Mechanics*. *Pages 123–133*. Springer (1975).
- [13] Lluís Masanes, Thomas D Galley, and Markus P Müller. “The measurement postulates of quantum mechanics are operationally redundant”. *Nat. Comm.* **10**, 1–6 (2019).
- [14] Borivoje Dakic and Caslav Brukner. “Quantum theory and beyond: Is entanglement special?”. *Deep Beauty: Understanding the Quantum World Through Mathematical Innovation* (ed. Halvorson, H.) (Cambridge Univ. Press, 2011) (2011).
- [15] G. Chiribella, G. M. D’Ariano, and P. Perinotti. “Probabilistic theories with purification”. *Phys. Rev. A* **81**, 062348 (2010). [arXiv:0908.1583](https://arxiv.org/abs/0908.1583).
- [16] Lucien Hardy. “Reformulating and Reconstructing Quantum Theory” (2011). [arXiv:1104.2066](https://arxiv.org/abs/1104.2066).
- [17] Rob Clifton, Jeffrey Bub, and Hans Halvorson. “Characterizing quantum theory in terms of information-theoretic constraints”. *Foundations of Physics* **33**, 1561–1591 (2003).
- [18] Philip Goyal. “Information-geometric reconstruction of quantum theory”. *Phys. Rev. A* **78**, 052120 (2008).
- [19] Lluís Masanes and Markus P Müller. “A derivation of quantum theory from physical requirements”. *New J. Phys.* **13**, 063001 (2011).
- [20] Howard Barnum, Markus P Müller, and Cozmin Ududec. “Higher-order interference and single-system postulates characterizing quantum theory”. *New Journal of Physics* **16**, 123029 (2014).



- [21] Alexander Wilce. “A Royal Road to Quantum Theory (or Thereabouts)” (2016). [arXiv:1606.09306](#).
- [22] Philipp Höhn. “Quantum theory from rules on information acquisition”. *Entropy* **19**, 98 (2017).
- [23] Agung Budiyo and Daniel Rohrlich. “Quantum mechanics as classical statistical mechanics with an ontic extension and an epistemic restriction”. *Nature Communications* **8** (2017).
- [24] John H. Selby, Carlo Maria Scandolo, and Bob Coecke. “Reconstructing quantum theory from diagrammatic postulates”. *Quantum* **5**, 445 (2021).
- [25] Sean Tull. “A categorical reconstruction of quantum theory”. *Logical Methods in Computer Science ; Volume 16* Pages Issue 1 ; 1860–5974 (2020).
- [26] John van de Wetering. “An effect-theoretic reconstruction of quantum theory”. *Compositionality* **1**, 1 (2019).
- [27] Kenji Nakahira. “Derivation of quantum theory with superselection rules”. *Physical Review A* **101** (2020).
- [28] G. Chiribella, G. M. D’Ariano, and P. Perinotti. “Informational derivation of quantum theory”. *Phys. Rev. A* **84**, 012311 (2011). [arXiv:1011.6451](#).
- [29] Ciarán M Lee and John H Selby. “Generalised phase kick-back: the structure of computational algorithms from physical principles”. *New J. Phys.* **18**, 033023 (2016). url: [doi.org/10.1088/1367-2630/18/3/033023](https://doi.org/10.1088/1367-2630/18/3/033023).
- [30] Ciarán M Lee and John H Selby. “Deriving grover’s lower bound from simple physical principles”. *New J. Phys.* **18**, 093047 (2016).
- [31] Howard Barnum, Ciarán M Lee, and John H Selby. “Oracles and query lower bounds in generalised probabilistic theories”. *Found. Phys.* **48**, 954–981 (2018).
- [32] Giulio Chiribella, Giacomo Mauro D’Ariano, and Paolo Perinotti. “Probabilistic theories with purification”. *Physical Review A* **81**, 062348 (2010).
- [33] Jamie Sikora and John Selby. “Simple proof of the impossibility of bit commitment in generalized probabilistic theories using cone programming”. *Phys. Rev. A* **97**, 042302 (2018).
- [34] Giulio Chiribella and Carlo Maria Scandolo. “Entanglement and thermodynamics in general probabilistic theories”. *New J. Phys.* **17**, 103027 (2015).
- [35] Giulio Chiribella and Carlo Maria Scandolo. “Microcanonical thermodynamics in general physical theories”. *New J. Phys.* **19**, 123043 (2017).
- [36] Howard Barnum, Ciarán M Lee, Carlo Maria Scandolo, and John H Selby. “Ruling out higher-order interference from purity principles”. *Entropy* **19**, 253 (2017).
- [37] Ciarán M Lee and John H Selby. “A no-go theorem for theories that decohere to quantum mechanics”. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **474**, 20170732 (2018).
- [38] Roman V. Buniy, Stephen D.H. Hsu, and A. Zee. “Is Hilbert space discrete?”. *Physics Letters B* **630**, 68–72 (2005).
- [39] Markus Mueller. “Does probability become fuzzy in small regions of spacetime?”. *Physics Letters B* **673**, 166–167 (2009).
- [40] T. N. Palmer. “Discretisation of the Bloch Sphere, Fractal Invariant Sets and Bell’s Theorem” (2020). [arXiv:1804.01734](#).
- [41] Bas Westerbaan and John van de Wetering. “A computer scientist’s reconstruction of quantum theory”. *J. Phys. A: Math. Theor.* **55**, 384002 (2022).
- [42] L. Hardy. “Probability theories with dynamic causal structure: a new framework for quantum gravity” (2005). [arXiv:gr-qc/0509120](#).
- [43] L. Hardy. “Towards quantum gravity: a framework for probabilistic theories with non-fixed causal structure”. *J. Phys. A* **40**, 3081–3099 (2007). [arXiv:gr-qc/0608043](#).

- [44] Ognjan Oreshkov, Fabio Costa, and Časlav Brukner. “Quantum correlations with no causal order”. *Nat. Comm.* **3**, 1–8 (2012).
- [45] Giulio Chiribella, Giacomo Mauro D’Ariano, Paolo Perinotti, and Benoit Valiron. “Quantum computations without definite causal structure”. *Phys. Rev. A* **88**, 022318 (2013).
- [46] Mateus Araújo, Adrien Feix, Miguel Navascués, and Časlav Brukner. “A purification postulate for quantum mechanics with indefinite causal order”. *Quantum* **1**, 10 (2017).
- [47] M. A. Nielsen and I. L. Chuang. “Quantum computation and quantum information”. *Cambridge University Press, Cambridge*. (2000).
- [48] J. Barrett. “Information processing in generalized probabilistic theories”. *Phys. Rev. A* **75**, 032304 (2007).
- [49] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner. “Bell nonlocality”. *Rev. Mod. Phys.* **86**, 839 (2014). [arXiv:quant-ph/1303.2849](https://arxiv.org/abs/1303.2849).
- [50] Howard Barnum, Oscar CO Dahlsten, Matthew Leifer, and Ben Toner. “Nonclassicality without entanglement enables bit commitment”. In 2008 IEEE Information Theory Workshop. *Pages 386–390*. IEEE (2008).
- [51] Marek Winczewski, Tamoghna Das, and Karol Horodecki. “Limitations on device independent secure key via squashed non-locality” (2019). [arXiv:1903.12154](https://arxiv.org/abs/1903.12154).
- [52] Martin Plávala. “General probabilistic theories: An introduction” (2021). [arXiv:2103.07469](https://arxiv.org/abs/2103.07469).
- [53] Markus Müller. “Probabilistic theories and reconstructions of quantum theory”. *SciPost Phys. Lect. Notes* **Page 28** (2021).
- [54] Ludovico Lami. “Non-classical correlations in quantum mechanics and beyond” (2018). [arXiv:1803.02902](https://arxiv.org/abs/1803.02902).
- [55] Bob Coecke. “Terminality implies non-signalling” (2014). url: [arxiv.org/abs/1405.3681v3](https://arxiv.org/abs/1405.3681v3).
- [56] Aleks Kissinger, Matty Hoban, and Bob Coecke. “Equivalence of relativistic causal structure and process terminality” (2017). url: [doi.org/10.48550/arXiv.1708.04118](https://doi.org/10.48550/arXiv.1708.04118).
- [57] Stefano Gogioso and Carlo Maria Scandolo. “Categorical probabilistic theories” (2017). url: [doi.org/10.4204/EPTCS.266.23](https://doi.org/10.4204/EPTCS.266.23).
- [58] C. Pfister and S. Wehner. “If no information gain implies no disturbance, then any discrete physical theory is classical”. *Nat. Comm.* **4**, 1851 (2013). url: [doi.org/10.1038/ncomms2821](https://doi.org/10.1038/ncomms2821).
- [59] L. Czekaj, M. Horodecki, P. Horodecki, and R. Horodecki. “Information content of systems as a physical principle”. *Phys. Rev. A* **95**, 022119 (2017).
- [60] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner. “Limits on non-local correlations from the structure of the local state space”. *New J. Phys.* **13**, 063024 (2011).
- [61] Howard Barnum and Alexander Wilce. “Ordered linear spaces and categories as frameworks for information-processing characterizations of quantum and classical theory” (2009). [arXiv:0908.2354](https://arxiv.org/abs/0908.2354).
- [62] Peter Janotta and Raymond Lal. “Generalized probabilistic theories without the no-restriction hypothesis”. *Phys. Rev. A* **87**, 052131 (2013). url: [doi.org/10.1103/PhysRevA.87.052131](https://doi.org/10.1103/PhysRevA.87.052131).
- [63] K. Kuratowski. “Introduction to set theory & topology”. *Volume 101 of International series of monographs in pure and applied mathematics*. PWN. Warsaw (1961).
- [64] Kenta Cho and Bart Jacobs. “Disintegration and bayesian inversion, both abstractly and concretely”. *Math. Struct. Comput. Sci.* **29**, 938–971 (2017). url: [doi.org/10.1017/S0960129518000488](https://doi.org/10.1017/S0960129518000488).
- [65] Manuel Blum. “Coin flipping by telephone”. In *Advances in Cryptology: A Report on CRYPTO 81*, IEEE Workshop on Communications Security. *Pages 11–15*. (1981).
- [66] Shafi Goldwasser, Silvio Micali, and Charles Rackoff. “The knowledge complexity of interactive proof systems”. *SIAM J. Comput.* **18**, 186–208 (1989).

- [67] Dominic Mayers. “Unconditionally secure quantum bit commitment is impossible”. *Phys. Rev. Lett.* **78**, 3414–3417 (1997).
- [68] Hoi-Kwong Lo and Hoi Fung Chau. “Why quantum bit commitment and ideal quantum coin tossing are impossible”. *Physica D: Nonlinear Phenomena* **120**, 177–187 (1998).
- [69] Stephen Boyd and Lieven Vandenberghe. “Convex optimization”. *Cambridge University Press*. (2004).
- [70] Sevag Gharibian, Jamie Sikora, and Sarvagya Upadhyay. “QMA variants with polynomially many provers”. *Quantum Information & Computation* **13**, 0135–0157 (2013). [arXiv:1108.0617](https://arxiv.org/abs/1108.0617).
- [71] Somshubhro Bandyopadhyay, Alessandro Cosentino, Nathaniel Johnston, Vincent Russo, John Watrous, and Nengkun Yu. “Limitations on separable measurements by convex optimization”. *IEEE Transactions on Information Theory* **61**, 3593–3604 (2015). url: [doi.org/10.1109/TIT.2015.2417755](https://doi.org/10.1109/TIT.2015.2417755).
- [72] Monique Laurent and Teresa Piovesan. “Conic approach to quantum graph parameters using linear optimization over the completely positive semidefinite cone”. *Siam J. Optim.* **25**, 2461–2493 (2015). url: [doi.org/10.1137/14097865X](https://doi.org/10.1137/14097865X).
- [73] Ashwin Nayak, Jamie Sikora, and Levent Tunçel. “A search for quantum coin-flipping protocols using optimization techniques”. *Math. Program.* **156**, 581–613 (2016). url: [doi.org/10.1007/s10107-015-0909-y](https://doi.org/10.1007/s10107-015-0909-y).
- [74] Jamie Sikora and Antonios Varvitsiotis. “Linear conic formulations for two-party correlations and values of nonlocal games”. *Math. Program.* **162**, 431–463 (2017). url: [doi.org/10.1007/s10107-016-1049-8](https://doi.org/10.1007/s10107-016-1049-8).
- [75] Samuel Fiorini, Serge Massar, Manas K Patra, and Hans Raj Tiwary. “Generalized probabilistic theories and conic extensions of polytopes”. *J. Phys. A: Math. Theor.* **48**, 025302 (2014). url: [doi.org/10.1088/1751-8113/48/2/025302](https://doi.org/10.1088/1751-8113/48/2/025302).
- [76] Anna Jenčová and Martin Plávala. “Conditions on the existence of maximally incompatible two-outcome measurements in general probabilistic theory”. *Phys. Rev. A* **96**, 022113 (2017). url: [doi.org/10.1103/PhysRevA.96.022113](https://doi.org/10.1103/PhysRevA.96.022113).
- [77] Joonwoo Bae, Dai-Gyoung Kim, and Leong-Chuan Kwek. “Structure of optimal state discrimination in generalized probabilistic theories”. *Entropy* **18**, 39 (2016). url: [doi.org/10.3390/e18020039](https://doi.org/10.3390/e18020039).
- [78] L. Lami, C. Palazuelos, and A. Winter. “Ultimate data hiding in quantum mechanics and beyond”. *Commun. Math. Phys.* **361**, 661–708 (2018).
- [79] Jamie Sikora and John H. Selby. “Impossibility of coin flipping in generalized probabilistic theories via discretizations of semi-infinite programs”. *Phys. Rev. Research* **2**, 043128 (2020).
- [80] John H Selby and Jamie Sikora. “How to make unforgeable money in generalised probabilistic theories”. *Quantum* **2**, 103 (2018). url: [doi.org/10.22331/q-2018-11-02-103](https://doi.org/10.22331/q-2018-11-02-103).
- [81] Bob Coecke, John Selby, and Sean Tull. “Two roads to classicality” (2017). url: [doi.org/10.4204/EPTCS.266.7](https://doi.org/10.4204/EPTCS.266.7).
- [82] John Selby and Bob Coecke. “Leaks: quantum, classical, intermediate and more”. *Entropy* **19**, 174 (2017). url: [doi.org/10.3390/e19040174](https://doi.org/10.3390/e19040174).
- [83] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski. “Information causality as a physical principle”. *Nature* **461**, 1101–1104 (2009). url: [doi.org/10.1038/nature08400](https://doi.org/10.1038/nature08400).
- [84] J. Barrett. “Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality”. *Phys. Rev. A* **65**, 042302 (2002). url: [doi.org/10.1103/PhysRevA.65.042302](https://doi.org/10.1103/PhysRevA.65.042302).
- [85] A. J. Short, S. Popescu, and N. Gisin. “Entanglement swapping for generalized nonlocal correlations”. *Phys. Rev. A* **73**, 012101 (2006). url: [doi.org/10.1103/PhysRevA.73.012101](https://doi.org/10.1103/PhysRevA.73.012101).

- [86] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters. “Teleporting an unknown quantum state via dual classical and einstein–podolsky–rosen channels”. *Phys. Rev. Lett.* **70**, 1895 (1993).
- [87] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert. “Event-ready detectors bell experiment via entanglement swapping”. *Phys. Rev. Lett.* **71**, 4287 (1993).
- [88] A. Acin, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani. “Device-independent security of quantum cryptography against collective attacks”. *Phys. Rev. Lett.* **98**, 230501 (2007). url: [doi.org/10.1103/PhysRevLett.98.230501](https://doi.org/10.1103/PhysRevLett.98.230501).
- [89] E. Hänggi, R. Renner, and S. Wolf. “Efficient quantum key distribution based solely on bell’s theorem”. *EUROCRYPT*Pages 216–234 (2010). [arXiv.org:0911.4171](https://arxiv.org/abs/0911.4171).
- [90] J. Barrett, L. Hardy, and A. Kent. “No signaling and quantum key distribution”. *Phys. Rev. Lett* **95**, 010503 (2005).
- [91] A. Acin, N. Gisin, and L. Masanes. “From bell’s theorem to secure quantum key distribution”. *Phys. Rev. Lett* **97**, 120405 (2006).
- [92] E. Hänggi. “Device-independent quantum key distribution”. PhD thesis. PhD Thesis, 2010. (2010). url: [doi.org/10.48550/arXiv.1012.3878](https://doi.org/10.48550/arXiv.1012.3878).
- [93] R. Colbeck and R. Renner. “Free randomness can be amplified”. *Nat. Phys.* **8**, 450–454 (2012). url: [doi.org/10.1038/nphys2300](https://doi.org/10.1038/nphys2300).
- [94] R. Gallego, L. Masanes, G. DeLaTorre, C. Dhara, L. Aolita, and A. Acin. “Full randomness from arbitrarily deterministic events”. *Nat. Comm.* **4**, 2654 (2013). url: [doi.org/10.1038/ncomms3654](https://doi.org/10.1038/ncomms3654).
- [95] P. Mironowicz, R. Gallego, and M. Pawłowski. “Amplification of arbitrarily weak randomness”. *Phys. Rev. A* **91**, 032317 (2015). url: [doi.org/10.1103/PhysRevA.91.032317](https://doi.org/10.1103/PhysRevA.91.032317).
- [96] F. G. S. L. Brandão, R. Ramanathan, A. Grudka, K. Horodecki, P. Horodecki M. Horodecki, T. Szarek, and H. Wojewódka. “Robust device-independent randomness amplification with few devices”. *Nat. Comm.* **7**, 11345 (2016). url: [doi.org/10.1038/ncomms11345](https://doi.org/10.1038/ncomms11345).
- [97] R. Ramanathan, F. G. S. L. Brandão, K. Horodecki, M. Horodecki, P. Horodecki, and H. Wojewódka. “Randomness amplification against no-signaling adversaries using two devices”. *Phys. Rev. Lett.* **117**, 230501 (2016). url: [doi.org/10.1103/PhysRevLett.117.230501](https://doi.org/10.1103/PhysRevLett.117.230501).
- [98] H. Wojewódka, F. G. S. L. Brandão, A. Grudka, M. Horodecki, K. Horodecki, P. Horodecki, M. Pawłowski, and R. Ramanathan. “Randomness amplification against no-signaling adversaries using two devices”. *IEEE Trans. Inf. Theory* **63**, 7592 (2017). url: [doi.org/10.1109/TIT.2017.2738010](https://doi.org/10.1109/TIT.2017.2738010).
- [99] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt. “Proposed experiment to test local hidden-variable theories”. *Phys. Rev. Lett.* **23**, 880–884 (1969).
- [100] Marek Winczewski, Tamoghna Das, and Karol Horodecki. “Limitations on device independent secure key via squashed non-locality” (2020). [arXiv:1903.12154](https://arxiv.org/abs/1903.12154).
- [101] P. Horodecki and R. Ramanathan. “The relativistic causality versus no-signaling paradigm for multi-party correlations”. *Nat Commun* **10**, 1701 (2019).
- [102] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts. “Non-local correlations as an information theoretic resource”. *Phys. Rev. A* **71**, 022101 (2005).
- [103] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. “Quantum entanglement”. *Rev. Mod. Phys.* **81**, 865 (2009). url: [doi.org/10.1103/RevModPhys.81.865](https://doi.org/10.1103/RevModPhys.81.865).
- [104] S. Pironio. “Lifting bell inequalities”. *Journal of Mathematical Physics* **46**, 062112 (2005). [arXiv:1210.0194](https://arxiv.org/abs/1210.0194).
- [105] A. Schrijver. “Combinatorial optimization polyhedra and efficiency”. Springer. Berlin (2003). url: [link.springer.com/book/9783540443896](https://link.springer.com/book/9783540443896).

- [106] C. Carathéodory. “Über den variabilitätsbereich der fourier’schen konstanten von positiven harmonischen funktionen”. Aus: *Rendiconti del Circolo Matematico di Palermo*. Direzione e Redazione. (1911). url: [books.google.co.in/books?id=n4SkjwEACAAJ](https://books.google.co.in/books?id=n4SkjwEACAAJ).
- [107] Günter M. Ziegler. “Lectures on polytopes”. *Springer New York*. (1995).
- [108] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu. “Bell inequalities for arbitrarily high-dimensional systems”. *Phys. Rev. Lett.* **88**, 040404 (2002). url: [doi.org/10.1103/PhysRevLett.88.040404](https://doi.org/10.1103/PhysRevLett.88.040404).
- [109] P. McMullen. “The maximum numbers of faces of a convex polytope”. *Mathematika* **17**, 179–184 (1970). arXiv:<https://londmathsoc.onlinelibrary.wiley.com/doi/pdf/10.1112/S0025579300002850>.
- [110] Khaled Elbassioni, Zvi Lotker, and Raimund Seidel. “Upper bound on the number of vertices of polyhedra with 0,1-constraint matrices”. *Information Processing Letters* **100**, 69 – 71 (2006).
- [111] Samson Abramsky and Adam Brandenburger. “The sheaf-theoretic structure of non-locality and contextuality”. *New J. Phys.* **13**, 113036 (2011). url: [doi.org/10.1088/1367-2630/13/11/113036](https://doi.org/10.1088/1367-2630/13/11/113036).
- [112] M. Araújo, M. Túlio Quintino, C. Budroni, M. Terra Cunha, and A. Cabello. “All noncontextuality inequalities for the n-cycle scenario”. *Phys. Rev. A* **88**, 022118 (2013). url: [doi.org/10.1103/PhysRevA.88.022118](https://doi.org/10.1103/PhysRevA.88.022118).
- [113] Ernst Specker. “Die logik nicht gleichzeitig entscheidbarer aussagen”. In *Ernst Specker Selecta*. Pages 175–182. Springer (1990).
- [114] Yeong-Cherng Liang, Robert W Spekkens, and Howard M Wiseman. “Specker’s parable of the overprotective seer: A road to contextuality, nonlocality and complementarity”. *Phys. Rep.* **506**, 1–39 (2011). url: [doi.org/10.1016/j.physrep.2011.05.001](https://doi.org/10.1016/j.physrep.2011.05.001).
- [115] Ravi Kunjwal, Chris Heunen, and Tobias Fritz. “Quantum realization of arbitrary joint measurability structures”. *Phys. Rev. A* **89**, 052126 (2014). url: [doi.org/10.1103/PhysRevA.88.022118](https://doi.org/10.1103/PhysRevA.88.022118).
- [116] B. Tsirelson. “Quantum generalizations of Bell’s inequality”. *Lett. Math. Phys.* **4**, 93–100 (1980). url: [doi.org/10.1007/BF00417500](https://doi.org/10.1007/BF00417500).
- [117] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik. “Quantifying Contextuality”. *Phys. Rev. Lett.* **112**, 120401 (2014). url: [doi.org/10.1103/PhysRevLett.112.120401](https://doi.org/10.1103/PhysRevLett.112.120401).

# A Proofs

## A.1 Proof of Proposition 13

For simplicity in this proof we adopt the diagrammatic notation of [15] using the conventions of [24].

Consider some state  $s$  and some purification  $\sigma^p \in \mathbf{Ext}_P[s]$  and an arbitrary extension  $\Sigma \in \mathbf{Ext}[s]$ . We want to demonstrate that we can achieve generation, which we will do by showing that there exists a transformation mapping  $\sigma^p$  to  $\Sigma$ , as  $\Sigma$  is an arbitrary extension we will therefore have demonstrated that  $\sigma^p$  is generating.

To start, let us consider another purification,  $\Sigma^p \in \mathbf{Ext}_P[\sigma]$  this time of the state  $\Sigma$ . It is clear that, as this is pure and an extension of  $s$  then it is also a purification of  $s$ . We therefore have two purifications of  $s$  and so one may be tempted to connect them via the essential uniqueness property. However, this will not always be the case, to see this let us examine the systems more carefully.  $\sigma^p$  is a state on some purifying system  $B$ :

$$\begin{array}{c} |A \quad |B \\ \hline \sigma^p \end{array} \quad (27)$$

whilst,  $\sigma$  could have been an extension on some other system  $X$  and the purification of this,  $\Sigma^p$  could involve a third system  $Y$ :

$$\begin{array}{c} |A \quad |X \quad |Y \\ \hline \Sigma^p \end{array} \quad (28)$$

Unless we are in a highly contrived scenario whereby  $B = X \otimes Y$  then we are not able to directly employ the essential uniqueness property. We therefore have to further extend our systems by composing with some extra pure states,  $\phi$  and  $\chi$ , as follows:

$$\begin{array}{c} |A \quad |B \\ \hline \sigma^p \end{array} \quad \begin{array}{c} |X \quad |Y \\ \hline \chi \end{array} \quad \text{and} \quad \begin{array}{c} |A \quad |B \quad |X \quad |Y \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \phi \end{array} \quad (29)$$

Then, thanks to our additional assumption that the parallel composite of pure states is pure, then these composite states define further purifications of the original states  $s$ . Now however, they have the same systems and hence we can use essential uniqueness to conclude that there is a reversible transformation  $T$  mapping between them, that is:

$$\begin{array}{c} |A \quad |B \quad |X \quad |Y \\ \hline T \\ \hline \begin{array}{c} |B \quad |X \quad |Y \\ \hline \sigma^p \quad \chi \end{array} \end{array} = \begin{array}{c} |A \quad |B \quad |X \quad |Y \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \phi \end{array} \quad (30)$$

Now, we can simply discard the  $B$  and  $Y$  systems on both sides of this equation to see that we can achieve generation:

$$\begin{array}{c} |A \quad |X \\ \hline \tilde{T} \\ \hline \begin{array}{c} |B \\ \hline \sigma^p \end{array} \end{array} := \begin{array}{c} |A \quad |X \\ \hline T \\ \hline \begin{array}{c} |B \quad |X \\ \hline \sigma^p \quad \chi \end{array} \end{array} = \begin{array}{c} |A \quad |B \quad |X \quad |Y \\ \hline T \\ \hline \begin{array}{c} |B \quad |X \quad |Y \\ \hline \sigma^p \quad \chi \end{array} \end{array} \quad (31)$$

$$= \begin{array}{c} |A \quad |B \quad |X \quad |Y \\ \hline \Sigma^p \end{array} \quad \begin{array}{c} \phi \end{array} = \begin{array}{c} |A \quad |X \quad |Y \\ \hline \Sigma^p \end{array} = \begin{array}{c} |A \quad |X \\ \hline \Sigma \end{array} \quad (32)$$

That is, there is a transformation  $\tilde{T}$  which maps  $\sigma^p$  to  $\Sigma$  which completes the proof.

## A.2 Proof of Theorem 16

To see this, first note that the set of all PME's,  $S_P$  is a convex set as any two PME's of  $P$ ,  $\mathcal{E}_{pure}^1(P)$ ,  $\mathcal{E}_{pure}^2(P) \in S_P$ , their convex combination<sup>5</sup>  $\lambda \mathcal{E}_{pure}^1(P) + (1-\lambda) \mathcal{E}_{pure}^2(P) \in S_P$ ,  $\forall \lambda \in (0, 1)$ . Such a mixture of PME's is defined as follows: Suppose  $\{(p_i, P_E^i)\}_i = \mathcal{E}_{pure}^1(P)$  and  $\{(q_i, P_E^i)\}_i = \mathcal{E}_{pure}^2(P)$  are two PME's of the behaviour  $P$ , where  $V(\mathcal{E}_{pure}^1)$  and  $V(\mathcal{E}_{pure}^2)$  are the set of pure behaviours corresponds to the ensembles. Now define  $V_{int} = V(\mathcal{E}_{pure}^1) \cap V(\mathcal{E}_{pure}^2)$ ,  $V_1 = V(\mathcal{E}_{pure}^1) \setminus V(\mathcal{E}_{pure}^2)$  and  $V_2 = V(\mathcal{E}_{pure}^2) \setminus V(\mathcal{E}_{pure}^1)$ . Then the convex combination of  $\mathcal{E}_{pure}^1(P)$  and  $\mathcal{E}_{pure}^2(P)$  is defined as

$$\lambda \mathcal{E}_{pure}^1(P) + (1-\lambda) \mathcal{E}_{pure}^2(P) := \{(r_i, P_E^i)\}_{P_E^i \in V(\mathcal{E}_{pure}^1) \cup V(\mathcal{E}_{pure}^2)}, \quad (33)$$

where  $r_i = \lambda p_i + (1-\lambda)q_i$ ,  $\forall P_E^i \in V_{int}$ ,  $r_i = \lambda p_i$ ,  $\forall P_E^i \in V_1$  and  $r_i = (1-\lambda)q_i$ ,  $\forall P_E^i \in V_2$ . Clearly

$$\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^1) \cup V(\mathcal{E}_{pure}^2)} r_i P_E^i = \sum_{i: P_E^i \in V_{int}} (\lambda p_i + (1-\lambda)q_i) P_E^i + \sum_{i: P_E^i \in V_1} \lambda p_i P_E^i + \sum_{i: P_E^i \in V_2} (1-\lambda)q_i P_E^i \quad (34)$$

$$= \sum_{i: P_E^i \in V_{int} \cup V_1} \lambda p_i P_E^i + \sum_{i: P_E^i \in V_{int} \cup V_2} (1-\lambda)q_i P_E^i \quad (35)$$

$$= \lambda P + (1-\lambda)P = P. \quad (36)$$

Here we use the fact that  $V_{int} \cup V_1 = V(\mathcal{E}_{pure}^1)$ ,  $V_{int} \cup V_2 = V(\mathcal{E}_{pure}^2)$  and  $\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^1)} p_i P_E^i = P$  and  $\sum_{i: P_E^i \in V(\mathcal{E}_{pure}^2)} q_i P_E^i = P$ . The above equation proves that the convex combination of two PME's of a behaviour  $P$  is also an ensemble of  $P$ , and its members are all pure. Hence,  $S_P$  forms a convex set (as we have shown) with only finite number of vertices (as we show below), and all of them are minimal ensembles, as the following lemma proves.

**Lemma 28** (Extremal  $\implies$  Minimal). *In the set of all pure members ensembles of the behaviour  $P$ , denoted by  $S_P$ , the ensembles that are extremal of  $S_P$ , are minimal.*

*Proof.* Suppose by contradiction, this is not true, i.e. there exists a PME,  $\mathcal{E}_{pure}(P) = \{(p_i, P_E^i)\}_{i=1}^n$ , with  $p_i > 0$ ,  $\forall i$ , which is extremal in  $S_P$ , but is not minimal. Then, there must exist a proper subset  $\mathcal{I} \subset V(\mathcal{E}_{pure})$  such that for all  $P_E^j \in \mathcal{I}$ , there is some other choices of probabilities  $\{q_j\}$ , which forms an ensemble  $\mathcal{M}(P) = \{(q_j, P_E^j)\}_{P_E^j \in \mathcal{I}}$ , and it is minimal. Let us now embed the distribution  $\{q_j\}_{j \in \mathcal{I}}$  which has less than  $n$  elements, to obtain new but equivalent distribution with  $n$  elements, by letting  $p'_i := q_i$ ,  $\forall P_E^i \in \mathcal{I}$  and  $p'_i := 0$ ,  $\forall P_E^i \in V(\mathcal{E}_{pure}) \setminus \mathcal{I}$ . Let us note that the minimal ensemble  $\mathcal{M}(P)$ , is now equivalent to the PME,  $\{(p'_i, P_E^i)\}_{i=1}^n$ .

Consider now an ensemble defined as:

$$\mathcal{N} = \left\{ \frac{p_i - pp'_i}{(1-p)}, P_E^i \right\}_{i=1}^n \equiv \{(r_i, P_E^i)\}_{i=1}^n, \quad (37)$$

where we define  $p = \frac{p_{min}}{p_{max}}$  with  $p_{min} = \min_i \{p_i : p_i > 0\}$  and  $p_{max} = \max_i \{p'_i\}$ . Let us first note, that  $\mathcal{N}$  is an ensemble of  $P$ . indeed, note that

$$\begin{aligned} P &= \sum_i p_i P_E^i = \sum_i (p_i - pp'_i) P_E^i + p \sum_i p'_i P_E^i, \\ &= (1-p) \sum_i r_i P_E^i + p \sum_i p'_i P_E^i, \end{aligned} \quad (38)$$

now by assumption  $\sum_i p'_i P_E^i = P$ , as  $\{(p'_i, P_E^i)\}_{i=1}^n = \mathcal{M}(P)$ . Thus:

$$P = (1-p) \sum_i r_i P_E^i + pP \quad (39)$$

which implies that  $\sum_i r_i P_E^i = P$ , i.e.  $\mathcal{N}$  is an ensemble of  $P$ , if  $0 < p < 1$ .

<sup>5</sup>Abstractly we can view PME's as probability distributions over the set of pure behaviours and hence is a convex set (in particular a simplex) the set  $S_P$  is then a subset of this simplex which we will show is closed under convex combinations.

We will argue now that the latter fact holds. Indeed, by definition,  $p$  is nonzero, resulting  $p_{min} > 0$ . To see  $p < 1$  we will prove that  $p_{min} < p'_{max}$ . If  $p_{min} \geq p'_{max}$ , then  $1 = \sum_i p_i \geq np_{min} \geq np'_{max} \geq n/|\mathcal{I}| \Rightarrow |\mathcal{I}| \geq n$ , which is a contradiction as  $\mathcal{I}$  is a proper subset of  $V(\mathcal{E}_{pure})$ . Here we use the fact that  $p'_{max} \geq 1/|\mathcal{I}|$ . Hence,  $p = \frac{p_{min}}{p'_{max}} < 1$  and  $\mathcal{N}$  is an ensemble of  $P$ , and can be denoted as  $\mathcal{N}(P)$ .

Now, we observe that by construction the PME

$$\mathcal{E}(P) = p\mathcal{M}(P) + (1-p)\mathcal{N}(P), \quad (40)$$

i.e.,  $\mathcal{E}(P)$  is a mixture of two ensembles, that are not equal to each other. This is a contradiction with assumed extremality of the ensemble  $\mathcal{E}(P)$ , since the mixture, as shown above, is non-trivial, and the assertion follows.  $\square$

The above theorem proves that all the extremal points in  $S_P$ , are minimal ensembles, i.e., there is no extremal points in  $S_P$  other than the  $\mathcal{M}(P)$ . Now we will prove the converse, that no interior point from  $S_P$  is a minimal ensemble, i.e., all minimal ensembles are also extremal. To prove it, we will need the following lemma, interesting in its own right, as it characterizes minimal ensembles as those with a unique distribution:

**Lemma 29.** *The pure members ensemble  $\mathcal{E}_{pure}(P) = \{(p_i, P_E^i)\}_i$  of a behaviour  $P$  is minimal,  $\mathcal{M}$ , iff the decomposition of this behaviour into the elements  $\{p_i : p_i > 0\}$  is unique, given by corresponding probabilities  $p_i$ .*

*Proof.* The “if” direction is trivial: if the elements  $\{p_i : p_i > 0\}$ , of the decomposition of  $\mathcal{E}_{pure}(P)$  are unique, then it is not possible to set any probability to zero. Hence, there is no proper subset  $\mathcal{I} \subset V(\mathcal{E}_{pure})$ , which forms an ensemble of  $P$ , with another choice of probabilities.

For the “only if” part, suppose  $\mathcal{M}(P) = \{(p_i, P_E^i)\}_{i=1}^m$ , is a minimal ensemble of  $P$ , we have to prove that the decomposition  $\{p_i, p_i > 0\}$ , is unique. Assume that the  $\{p_i\}$  is not unique, but being a minimal ensemble it should follow  $\sum_{i=1}^m p_i P_E^i = P$ , or in other words the set of following linear equations

$$\sum_{i=1}^m a_{ki} y_i = c_k, \quad (41)$$

for some  $k = 1, \dots, l'$ , has solution in form  $y_i = p_i$ . Here  $c_k$  are the entries of the behaviour  $P$ , for the pair  $(a, x)$ , or in other words the probability of getting  $a$ , when the input is  $x$ ,  $P(a|x) = c_k$ . Similarly the coefficients  $\{a_{ki}\}$  are the same entries for the pair  $(a, x)$  of the pure behaviours  $\{P_E^i(a|x) = a_{ki}\}$ . As the behaviour  $P$  should follow some equality constraint, and due to some internal symmetry of it, not all  $l'$  equations in Eq. (41), are linearly independent. Here, by linear independence we mean  $\sum_{k \neq k'} \lambda_k (\sum_{i=1}^m a_{ki} y_i - c_k) \neq \sum_{i=1}^m a_{k'i} y_i - c'_{k'}$ , for some  $k, k'$  and  $\lambda_k$ . Suppose, there are only  $l$  linearly independent equations. (There is also a constraint on the  $\{y_i\}$ , that  $\sum_{i=1}^m y_i = 1$  as they represent the probability of the ensemble, but we don't need to consider it separately, as the behaviour  $P$  is normalized, so Eq. (41), will take care of it.) Now the number of linearly independent equations and the number of variables can be in one of the three orders which we consider separately: 1)  $l > m$ , 2)  $l < m$  3)  $l = m$ . Notice first, that it can not be  $l > m$ , i.e., for  $l$  number of linear equation pertaining  $m$  number of variables. Otherwise there would be no solution of the set of equation:

$$\left\{ \sum_{i=1}^m a_{ki} y_i = c_k \right\}_{k=1}^l, \quad (42)$$

with variables  $y_i$  but we already have a solution, the initial one:  $y_i = p_i > 0, \forall i$ . On the other hand, if  $l < m$ , then one can always write down any set of  $l$ ,  $\{y_i\}_{i=1}^l$ , as a linear functions of the remaining  $(m-l)$   $\{y_j\}_{j=l+1}^m$ . And in that case one can always set any one (or more)  $y_i = 0$  for some  $i$ , which violates the condition of minimal ensembles. Hence we are left with  $l = m$ .

In this case, we have the same number of linearly independent equation as the number of variables, and in that case the matrix  $A = [a_{ki}]$  is non-singular and invertible, which gives a unique solution of  $y_i = p_i > 0$  for all  $i$ .  $\square$



We can pass now to prove the extremality of minimal ensembles:

**Lemma 30** (Minimal  $\implies$  Extremal). *For a behaviour  $P$ , all of its minimal ensembles  $\mathcal{M}(P)$  are extremal in the set  $S_P$  of all ensembles of a behaviour  $P$ .*

*Proof.* Suppose by contradiction, that  $\mathcal{M}(P)$  is not extremal. Then, there exist pure members ensembles  $\mathcal{E}_1(P)$  and  $\mathcal{E}_2(P)$  such that:

$$\mathcal{M}(P) = \lambda \mathcal{E}_1(P) + (1 - \lambda) \mathcal{E}_2(P) \quad (43)$$

for some  $0 < \lambda < 1$ . By the above equality,  $V(\mathcal{E}_1) \subseteq V(\mathcal{M})$  and  $V(\mathcal{E}_2) \subseteq V(\mathcal{M})$ . But by minimality of  $\mathcal{M}$ ,  $V(\mathcal{E}_1)$  can not be proper subset of  $V(\mathcal{M})$ , as there are no weights that together with any proper subset of  $V(\mathcal{M})$  form an ensemble of  $P$ . Thus  $V(\mathcal{E}_1) = V(\mathcal{M})$  and for similar reason  $V(\mathcal{E}_2) = V(\mathcal{M})$ . It would mean, that there is an ensemble (let us focus on  $\mathcal{E}_1$ ) which has different distribution, but the same set of members. It would mean that the distribution of  $\mathcal{M}$  is not unique: there is another one which together with the same set of members yields an ensemble of  $P$ . This however is not possible, since by Lemma 29, any minimal ensemble has unique distribution. This proves desired contradiction, hence the assertion follows.  $\square$

From these two Lemmas 28 and 30, we obtain, that the set  $S_P$  of all pure member ensembles (PMEs) of  $P$  is a convex hull of the set of minimal ensembles  $\mathcal{M}(P)$ . And for any behaviour  $P$ , the set of minimal ensembles is finite, as there are finite number of pure behaviours<sup>6</sup> and corresponding to Lemma 29 the decomposition of the  $p_i$  in minimal ensembles are unique, implies  $S_P$  forms a convex polytope.

### A.3 Proof of Corollary 20

*Proof.* Note that according to the Definition 18, if  $\mathcal{E}(P)_{AE}(ae|xz)$  is the NSEA of the given behaviour  $P_A(a|x)$ , then the only ensembles realized for different choices of input  $z$  of the extending party  $E$  are the minimal ensembles  $\mathcal{M}(P)$ . If there are  $\{\mathcal{M}_i(P)\}_{i=1}^N$ ,  $N$  numbers of such minimal ensembles of  $P_A$ , then there should be  $|\mathcal{Z}| = N$ , number of distinct inputs in part of the extending system. Due to Lemma 28 all the extremal points in  $S_P$  are minimal ensembles. From the latter, one can generate any pure members ensemble  $\mathcal{E}(P)$  by properly mixing the minimal ensembles by using appropriate distribution  $\{\tilde{p}(k)\}_{k=1}^N$ , with  $\sum_{k=1}^N \tilde{p}(k) = 1$ . For an arbitrary  $\mathcal{E}(P) \in S_P$ , the Lemma 28, certifies the existence of at least one such  $\{\tilde{p}(k)\}_{k=1}^N$ , which will generate it, hence,  $\mathcal{E}(P) = \sum_{k=1}^N \tilde{p}(k) \mathcal{M}_k(P)$ . Each  $\mathcal{M}_k(P) = \{(p(e = i|z = k), P_{\mathbf{E}}^{ik}(a|x))\}_i$ , has been obtained from  $\mathcal{E}(P)_{AE}$ , by setting the input  $z = k$  of the extending party. If the inputs are now chosen probabilistically according to the distribution  $\{\tilde{p}(k)\}_{k=1}^N$ , and registering the output  $e = i$ , then

$$\mathcal{E}(P) = \sum_{k=1}^N \tilde{p}(k) \left\{ \left( p(e = i|z = k), P_{\mathbf{E}}^{ik}(a|x) \right) \right\} \quad (44)$$

$$= \left\{ \left( q(i), P_{\mathbf{E}}^i(a|x) \right) \right\}_{P_{\mathbf{E}}^i \in \cup_{k=1}^N V(\mathcal{M}_k)}, \quad (45)$$

where  $q(i)$  is the probability of getting the pure behaviour  $P_{\mathbf{E}}^i$ , as given in Lemma 28.  $\square$

### A.4 Proof of Theorem 23

**Lemma 31.** (ACCESS  $\implies$  GENERATION) *Access to all ensembles implies access to arbitrary extensions of the extended system.*

<sup>6</sup>The cardinality of the set of minimal ensembles is bounded by the cardinality of the set of all subsets of pure behaviours which is finite if the set of pure behaviours is finite. See section 4.2 for an explicit tighter upper bound.

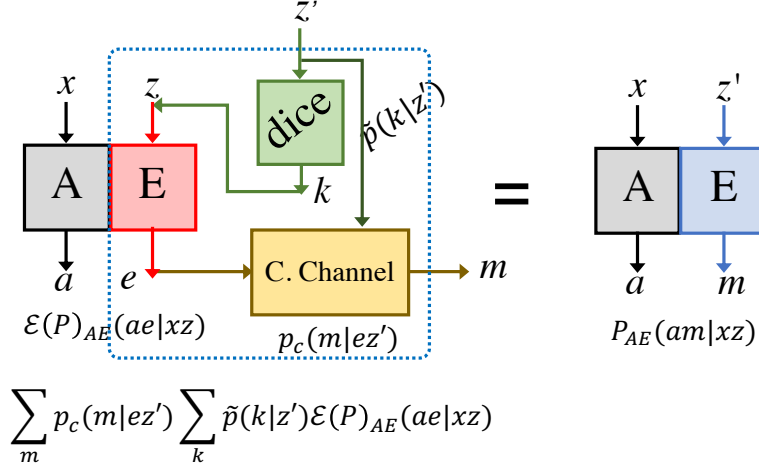


Figure 7: Pictorial depiction of Theorem 24.  $E$  holds the additional interfaces of the NSEA of behaviour  $P_A$ . She use an input randomizer (a dice)  $\tilde{p}(k|z')$ , which can be tuned by the parameter  $z'$ , and a classical post-processing channel  $p_c(m|ez')$ , which is also  $z'$  dependent. The dice, the extending system and the post-processing channel together form the new system in part of  $E$ . The input of the dice  $z'$  and the output of the channel  $m$ , can be considered as the input and output of the new system.  $E$  only choose some value of input  $z'$ , generates required randomness and also appropriate conditional probability distribution, resulting different set of mixed ensembles.

*Proof.* We first follow [102] (See Eq. (35) there) and observe that for any extension  $P \equiv P(am|xz')$  of  $P(a|x)$  there is

$$P(am|xz') = P(a|xz'm)P(m|xz') \quad (46)$$

$$= P(a|xz'm)P(m|z') \quad (47)$$

$$\equiv \{P^{ij}(a|x)P(m=i|z'=j)\}_{i,j} \quad (48)$$

In the first of the above equalities we use Bayes rule, and in the second the non-signalling from  $A$  to  $E$ . Such obtained equality implies that every bipartite behaviour can be viewed from perspective of the system  $E$  as having access to  $|z'|$  ensembles of the form  $\mathcal{E}_j(P) = \{(P(m=i|z'=j), P^{ij}(a|x))\}_i$ . We argue now, that from NSEA one can generate any of these ensembles. Thanks to Theorem 22, for  $|z'|$  inputs  $z'=j$  there exists a dice  $D_j$  with probability distribution  $\tilde{p}(k|z'=j)$  and a classical post-processing channel  $C_j$ , described with a conditional distribution  $p_c(m|e, z'=j)$ , such that when applied on system  $E$  of the NSEA  $\mathcal{E}(P)(ae|xz)$  they generate the ensemble  $\mathcal{E}_j$ . Hence from a collection of  $\{D_j\}_{j=1}^{|z'|}$  and  $\{C_j\}_{j=1}^{|z'|}$  one can build via appropriate “wiring” a behaviour  $P'$  which upon input  $z'=j$  performs according to a dice  $D_j$  and further post-process the output  $e$  (and that of  $D_j$ ) through  $C_j$  to give the final output  $m$  (see Fig. 7). In this way one assures that a new behaviour  $P'$  has access to all the ensembles  $\mathcal{E}_j$  (and no other), forming (thanks to Eq. (48) above) an extension equivalent to  $P$  up to relabelling of inputs in the extending system  $E$ .  $\square$

From the above theorem it is clear that the NSEA  $\mathcal{E}(P)_{AE}$ , of  $P_A$  together with access of arbitrary randomness (input randomizer followed by a classical post-processing channel) can generate any collection of ensembles (Theorem 22) even mixed. If we consider the whole setup as a single behaviour, as depicted in Fig. 7 (with the blue rectangular part is in possession of  $E$ ), then it becomes a proper arbitrary extension generating a particular collection of ensembles. Note that it can be easily verified that it fulfills all properties of non-signalling behaviour.

**Lemma 32.** *GENERATION*  $\implies$  *ACCESS*

*Proof.* The proof for any GPT is given in Proposition 10. For the sake of completeness, we show it now using solely arguments from the non-signalling theory. Namely NSEA is a particular extension, hence access to arbitrary extension implies access to NSEA. This further via Theorem 22 implies access to any (possibly mixed) ensemble.  $\square$

## B Explicit examples in the theory of no-signaling behaviours

### B.1 Complete extensions of binary input-output behaviours

In section 2.1, we proved that the theory of non-signaling behaviours (NS) does not possess the property of purification as there is not a pure extension of every behaviour. However, there do exist some behaviours for which the NSEA is an extremal, i.e., pure, behaviour. In other words, whilst the purification postulate fails as not every behaviour has a purification, there nonetheless do exist purifications of certain behaviours.

Here we are going to show an example of this. Namely, that if one considers a maximally mixed behaviour with a single binary input and a single binary output, then it can be extended to an extremal (pure) behaviour in a higher dimensional state space<sup>7</sup> This pure behaviour is the maximally non-local behaviour equivalent to the Popescu-Rohrlich box [5] (up to proper labeling on the extending system) defined as

$$P^{\text{PR}}(a, e|x, z) = \begin{cases} \frac{1}{2} & \text{for } a \oplus e = x \cdot z, \\ 0 & \text{otherwise.} \end{cases} \quad (49)$$

where  $a, e, x, z \in \{0, 1\}$ . This is a close analogue to the fact that in quantum theory: a purification of the maximally mixed state  $\frac{\mathbb{1}_d}{d}$  is the maximally entangled Bell state (up to local isometry) [103].

We will now prove this result. Note that the maximally mixed behaviour with a single binary input and single binary output is given by

$$P_A^m(a|x) = \begin{array}{c|cc} \begin{array}{c} \backslash a^x \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 1/2 \\ 1/2 \end{array} & \begin{array}{c} 1 \\ 1/2 \\ 1/2 \end{array} \\ \hline & & \end{array} \quad (50)$$

here  $x$  being the input and  $a$  being the output of the behaviour on system  $A$ . This maximally mixed behaviour lies in the ‘‘center’’ of the polytope (of the set of behaviours with a single binary input and a single binary output), the extremal points (vertices) of the polytope are

$$P_E^0 = \begin{array}{c|cc} \begin{array}{c} \backslash a^x \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 1 \\ 0 \end{array} & \begin{array}{c} 1 \\ 1 \\ 0 \end{array} \\ \hline & & \end{array}, \quad P_E^1 = \begin{array}{c|cc} \begin{array}{c} \backslash a^x \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \\ 1 \end{array} & \begin{array}{c} 1 \\ 0 \\ 1 \end{array} \\ \hline & & \end{array}, \quad P_E^2 = \begin{array}{c|cc} \begin{array}{c} \backslash a^x \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 1 \\ 0 \end{array} & \begin{array}{c} 1 \\ 0 \\ 1 \end{array} \\ \hline & & \end{array}, \quad P_E^3 = \begin{array}{c|cc} \begin{array}{c} \backslash a^x \\ 0 \\ 1 \end{array} & \begin{array}{c} 0 \\ 0 \\ 1 \end{array} & \begin{array}{c} 1 \\ 1 \\ 0 \end{array} \\ \hline & & \end{array}. \quad (51)$$

These pure behaviours are deterministic behaviours as  $P_E^i(a|x) = \delta_{a, g^i(x)}$ , for some function  $g^i : \{0, 1\} \rightarrow \{0, 1\}$ . The polytope has been depicted in Fig. 8, with a blue square, the corners of the square are the extremal (pure) behaviours  $\{P_E^i\}$ , represented by the black solid circles. The center point of the polytope is the maximally mixed behaviour,  $P_A^m$ , which is depicted by a white solid circle. Any point inside the polytope can be expanded as a convex combination of the vertices. It is easy to see that, as  $P_A^m$  lies at the intersection of the two diagonals of the square (behaviour polytope), it can be expanded in terms of vertex pairs  $\{P_E^0, P_E^1\}$  and  $\{P_E^2, P_E^3\}$  with equal probabilities. In particular these form the two minimal ensembles of  $P_A^m$  namely

$$\mathcal{M}_0(P_A^m) = \{(1/2, P_E^0); (1/2, P_E^1)\} = \{p(i|0), P_E^{i0}\}, \quad (52)$$

$$\mathcal{M}_1(P_A^m) = \{(1/2, P_E^2); (1/2, P_E^3)\} = \{p(i|1), P_E^{i1}\}, \quad (53)$$

and there are no other minimal ensembles. Now from the Definition 18, of NSEA to the system  $E$ , the above two minimal ensembles are obtained in part of system  $A$ , for two different measurement choices on  $E$ . We choose that the first ensemble is obtained by setting the input  $z = 0$ , and the second one

<sup>7</sup>Throughout the paper we use the following notation for behaviours,

- $P$  – *italic* represent any generic behaviour.
- $P$  – normal font represent a particular example of a behaviour.

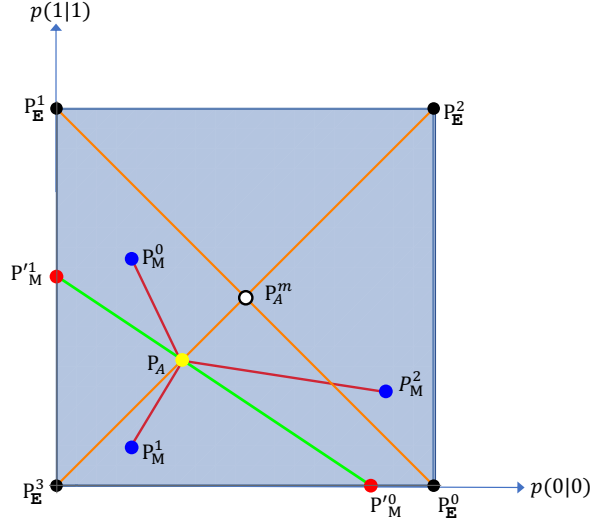


Figure 8: Polytope of the set of single party binary input output behaviours. Here the behaviour of consideration given in Eq. (56) is the yellow bullet. The white bullet is the maximally mixed behaviour. The black bullets are the deterministic, pure or extremal behaviours as given in Eq. (51). The required behaviour can be decomposed as linear combination of the deterministic behaviours as given in Eqs. (57) and (58). It can also be expanded as convex combination of mixed behaviours which are the red and blue bullets, these form the mixed ensemble of  $P_A$ .

is obtained by  $z = 1$ . Labeling the member behaviours of each ensemble with the output  $e$ , i.e.,

$z$	$e$	behaviour on $A$	Probability
0	0	$P_E^0$	1/2
	1	$P_E^1$	1/2
1	0	$P_E^2$	1/2
	1	$P_E^3$	1/2

(54)

we finally obtain the non-signaling complete extension (NSCE) of the maximally mixed behaviour that presents the same behaviour as the PR box in equation (49):

$$P_{AE}^{\text{PR}}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & \\ z & e & a & & \\ \hline 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ & 1 & 0 & 1/2 & 0 & 1/2 \\ \hline 1 & 0 & 1/2 & 0 & 0 & 1/2 \\ & 1 & 0 & 1/2 & 1/2 & 0 \end{array} \quad (55)$$

It is easy to check that if  $x \cdot z = 0$ , the resulting probability distribution on  $a$  and  $e$  is perfectly correlated, while for  $x \cdot z = 1$  it is perfectly anti-correlated (i.e.,  $e = 0$  implies  $a = 1$  and  $e = 1$  implies  $a = 0$ ). Thus in a sense we have derived a PR box solely from the principle of the complete extension (CEP). It is easy to see, that by negating  $z$  and  $e$ , one can instead obtain another maximally non-local bipartite behaviour. Similarly, all of the other maximally non-local behaviours (i.e., all non-local vertices of the polytope of two party binary input and binary output behaviours) can be obtained by proper relabeling of the  $z$  and  $e$ .

Note that the PR box is a vertex in the polytope of two party binary input and binary output behaviours [102]. Hence, we have the following conclusion:

**Corollary 33.** *The PR box is a purification of a maximally mixed behaviour with a single binary input and a single binary output.*

We have therefore constructed the PR box without any reference to CHSH inequality [99] (in contrast to how it was originally in [5]), as the non-signaling complete extension (NSCE) of a maximally

mixed behaviour (50). One could argue that the PR box is present in the theory of non-signaling behaviours from the very beginning. Whilst this is true, our derivation is based solely on the Definition 18, rather than relying on the structure of the polytope of non-signaling behaviours.

**Remark 34.** *It is tempting to say that the non-signaling complete extension (NSCE) of two party maximally mixed binary input and binary output behaviours is a tensor product of the PR boxes. It is however not the case. This is due to the fact, that one of the valid ensembles of a maximally mixed state  $\frac{\mathbb{1}_{AB}}{4} = \frac{1}{2}PR_{AB} + \frac{1}{2}\overline{PR}_{AB}$ , where  $\overline{PR}_{AB}(ab|xy) = \frac{1}{2}\delta_{a\oplus b, xy\oplus 1}$  is a non-local behaviour supported on the orthogonal subspace to that of the support of PR. Since this ensemble is clearly minimal, having 2 members, in definition of NSCE there should be the input  $z$  which allows the owner of extending system to collapse the system  $AB$  into one of these maximally non-local behaviours (each with probability half). Suppose now, by contradiction that the NSCE is of the form  $PR_{AX_A} \otimes PR_{BX_B}$ . It is then clear to see, that in such a behaviour none of direct measurements (choosing the inputs) has outcome behaviour on  $AB$  of the form expected by measurement of demanded input  $z$ . However one should consider some other possible ways of measuring system  $X_A X_B$  e.g. via wiring. Yet there is no such action on systems  $X_A X_B$ , simulating joint outcomes of  $z$ , since that would lead to the so called non-locality swapping, which is proven to be impossible in Refs. [84, 85].*

By virtue of Theorem 7, getting a pure behaviour in the higher dimensional state space through the construction of non-signaling complete extension (NSCE) is not always possible for any generic behaviour. Indeed, if we choose any other behaviour in the polytope of single binary input and single binary output, except for the maximally mixed and the four vertices behaviours, then its NSCE is not a vertex. For example, let us consider the following behaviour

$$P_A = \begin{array}{c|cc} & x & \\ \hline & 0 & 1 \\ \hline a & 0 & 1/3 & 2/3 \\ \hline & 1 & 2/3 & 1/3 \end{array} \quad (56)$$

which lies in the polytope of single party binary input and binary output behaviours, and it is represented by the yellow point in Fig. 8. Each behaviour can be expanded in terms of the pure behaviours of the polytope, hence

$$P_A = x_0 P_E^0 + x_1 P_E^1 + x_2 P_E^2 + x_3 P_E^3, \quad (57)$$

where  $x_i \geq 0, \forall i$ , and  $\sum_{i=0}^3 x_i = 1$ . The general solutions of Eq. (57) is:

$$\begin{cases} x_0 = \frac{1}{3}(2 - 3x_3), \\ x_1 = \frac{1}{3}(2 - 3x_3), \\ x_2 = \frac{1}{3}(3x_3 - 1), \\ x_3 = x_3 \end{cases}, \quad \frac{1}{3} \leq x_3 \leq \frac{2}{3} \quad (58)$$

To construct the minimal ensembles of behaviour  $P_A$ , we have to find out the set of decomposition over the pure points  $\{P_E^i\}$ , such that any proper subset of each choice can not be the ensemble of  $P_A$  with another set of probabilities. This implies that we have to find those solutions of Eq. (58), where the minimal number of  $x_i$ 's are nonzero. There are two of such choices, given by

$$\left\{ \begin{array}{l} x_0 = \frac{1}{3}, \\ x_1 = \frac{1}{3}, \\ x_2 = 0, \\ x_3 = \frac{1}{3} \end{array} \right\}, \quad \left\{ \begin{array}{l} x_0 = 0, \\ x_1 = 0, \\ x_2 = \frac{1}{3}, \\ x_3 = \frac{2}{3} \end{array} \right\}, \quad (59)$$

which, hence, form the two minimal ensembles of  $P_A$ :

$$\mathcal{M}_0(P_A) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^3)\} = \{p(i|0), P_E^{i0}\}, \quad (60)$$

$$\mathcal{M}_1(P_A) = \{(1/3, P_E^2); (2/3, P_E^3)\} = \{p(i|1), P_E^{i1}\}. \quad (61)$$

Here we label the ensembles by  $\{0, 1\}$ , according to the inputs  $z$  and the members as  $P^{00} = P_E^0$ ,  $P^{10} = P_E^1$ ,  $P^{20} = P_E^3$  and  $P^{01} = P_E^2$ ,  $P^{11} = P_E^3$ . Finally then, the NSCE of  $P_A$  to system  $E$  is given by

$$P_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & & 0 & & 1 \\ & e & a & & & \\ \hline 0 & 0 & 1/3 & 0 & 1/3 & 0 \\ & 1 & 0 & 1/3 & 0 & 1/3 \\ & 2 & 0 & 1/3 & 1/3 & 0 \\ \hline 1 & 0 & 1/3 & 0 & 0 & 1/3 \\ & 1 & 0 & 2/3 & 2/3 & 0 \end{array} \quad (62)$$

which lies in the polytope of two party behaviours, in which one party has a binary input and output, whereas the other party has a binary input but a ternary output. Moreover, as we will prove in the following section (Sec. B.1.1), this extended behaviour is not pure in the respective behaviour polytope.

An example of Theorem 20 has been given in Sec. B.1.2, where we have shown that any pure members ensemble of  $P_A$ , given in Eq. (56), can be constructed by taking the convex combination of its two minimal ensembles. This convex combination has been simulated by using an additional randomness (a dice) at the input of the extending party. An arbitrary ensemble of  $P_A$ , can be constructed by passing the output of the extending party through a classical post-processing channel (Theorem 22) has also been exemplified in Sec. B.1.3. The members of the mixed ensembles have been depicted by the blue dots (for mixed ensemble 1) and by red dots (for mixed ensemble 2). By enumerating the mixed ensembles with the input and the member behaviours with the output of an additional system, an extension of  $P_A$  has been generated. As the mixed ensembles of  $P_A$  are arbitrarily chosen, the extension made out of it is also an arbitrary extension of  $P_A$  – given in Eq. (79) of Sec. B.1.4. Moreover, for a generic behaviour, the non-signaling complete extension (NSCE) introduces correlations (possibly non-local) between the extending and the extended system. It is quite clear that, for a pure behaviour (a vertex of a polytope), there will be no correlation due to NSCE, whereas it can inject a maximal amount of correlation when the given behaviour is maximally mixed. For  $P_A$ , we calculate the non-local correlation in Sec. B.1.5.

Another important aspect of non-signaling complete extensions (NSCE) is that, unlike quantum purifications, the NSCE of an arbitrary behaviour and its conjugate behaviour are not the same. If  $\mathcal{E}(P)_{AE}(ae|xz)$  is the NSCE of the behaviour  $P_A(a|x)$ , then the behaviour  $P_E(e|z) = \sum_a P_{AE}(ae|xz)$ , is the conjugate behaviour of  $P_A$ . The NSCE of  $P_E(e|z)$ , according to the Definition 18,  $\mathcal{E}(P)_{A'E}(a'e|x'z)$  is not equal to the  $\mathcal{E}(P)_{AE}(ae|xz)$ . In section B.1.6, we provide an explicit example in favour of this argument, for the behaviour  $P_A$ , given in Eq. (56).

### B.1.1 Example: the non-signaling complete extension that is not a vertex

In this section we will prove that the NSCE of  $P_A$  (given in equation (56)), is not a vertex in the higher dimensional behaviour polytope. Consider an arbitrary behaviour  $P_{A_1 A_2 \dots A_n}(a_1 a_2 \dots a_n | x_1 x_2 \dots x_n)$ , with  $n$  parties  $A_1 A_2 \dots A_n$  where the  $x_i$  and  $a_i$  are respectively the input and output of the  $i$ th party  $A_i$ . Suppose there are  $m_i$  possible measurement choices of  $A_i$ , i.e.,  $x_i \in \{1, 2, \dots, m_i\}$ , and corresponding to each measurement  $x_i = j$  the number of possible outcomes is  $d_{ij}$ , such that  $a_i \in \{1, 2, \dots, d_{ij}\}$ . The total number of parameters involved in defining the behaviour is given by Ref. [104] as

$$t = \prod_{i=1}^n \left( \sum_{j=1}^{m_i} d_{ij} \right). \quad (63)$$

Among these  $t$  parameters not all of them are independent as the behaviour should obey normalization and non-signaling conditions. These conditions together imply that the behaviour must lie within a particular polytope of  $\mathbb{R}^t$ . If we construct a vector  $v \in \mathbb{R}^t$ , whose entries are those  $t$  probabilities, then the polytope can be defined as

$$\mathcal{P} = \{v \mid \mathcal{A}v \leq w\} \quad (64)$$

for some  $w \in \mathbb{R}^s$ , and  $\mathcal{A}$  an  $t \times s$  matrix. Here the condition  $\mathcal{A}u \leq w$  captures all of the constraints that the probabilities need to satisfy.

Let us take an arbitrary element  $u \in \mathcal{P}$ , and suppose  $\mathcal{A}_u u \leq w_u$  are those inequality constraints among all possible constraints that are satisfied by  $u$  with equality [105], i.e.,  $\mathcal{A}_u u = w_u$ . Here,  $\mathcal{A}_u$  and  $w_u$  are sub-matrices of  $\mathcal{A}$  and  $w$  respectively. Then from Theorem 5.7 of Ref. [105],  $u$  will be a pure point of  $\mathcal{P}$  if and only if  $\text{rank}(\mathcal{A}_u) = t$ .

The behaviour  $P_{AE}$ , given in Eq. (62), belongs to a polytope which lies in a space of  $t = 20$ . We now compute the rank of  $\mathcal{A}_u$  to demonstrate that this is not a vertex of the polytope. First we count how many independent equality constraints are acting on  $P_{AE}$ . The  $t = 20$  probabilities are  $P_{AE}(00|00), P_{AE}(01|00), \dots, P_{AE}(12|11)$ , i.e.,  $P_{AE}(ae|xz)$  for  $a, x \in \{0, 1\}$  and  $e \in \{0, 1, 2\}$  for  $z = 0$  and  $e \in \{0, 1\}$  for  $z = 1$ . The equality constraints they need to satisfy are the non-signaling constraint and normalization conditions, namely

$$\sum_a P_{AE}(ae|xz) = \sum_a P_{AE}(ae|x'z), \quad \forall e, x, x', z, \quad (65)$$

$$\sum_e P_{AE}(ae|xz) = \sum_e P_{AE}(ae|xz'), \quad \forall a, x, z, z', \quad (66)$$

and

$$\sum_{ae} P_{AE}(ae|xz) = 1, \quad \forall x, z. \quad (67)$$

There are 5 from Eq. (65)<sup>8</sup>, 4 from Eq. (66)<sup>9</sup>, and another 4 from Eq. (67), totalling 13 equality constraints. Moreover, among the 40 inequality constraints of the form  $0 \leq P_{AE}(ae|xz) \leq 1$ ,  $\forall a, e, x, z$ , (2 inequalities for each of the 20 probabilities  $P_{AE}(ae|xz)$ ), only 10 of them find equality with zeroes. Note that there is no probability which finds equality with 1, as there is no entry of 1, in Eq. (62). Hence, the total number of equality constraint are 23, and the matrix  $\mathcal{A}_u$  for which  $\mathcal{A}_u u = w_u$ , is of the dimension  $23 \times 20$ . To obtain  $\text{rank}(\mathcal{A}_u)$ , we need to find out the number of linearly independent equality constraint out of the 23. The 10 equality constraint with zero, on the probabilities are all linearly independent, but among those 13 (non-signaling constraint and normalization conditions) only 9 of them are linearly independent – by choosing any 1 normalization condition, and any 8 non-signaling condition one can generate the remaining 4 conditions. Hence, the total number of the linearly independent constraints, on the behaviour  $P_{AE}$  is  $9 + 10 = 19$ , and  $\text{rank}(\mathcal{A}_u) = 19 < 20$ . Which allows us to state that behaviour  $P_{AE}$  is not a vertex (extreme point) of the given polytope.

On the other hand the NSCE of the maximally mixed behaviour, the PR box given in Eq. (55) is a pure behaviour in the polytope of two binary input output behaviours. This can be shown in the following way: the space where the PR box lives is of  $t = 16$  [102]. Amongst the 16 parameters there are 8 equality constraints with zeroes and  $4 + 4 = 8$  linearly independent equality comes from the 8 non-signaling conditions and 4 normalization condition, hence  $\text{rank}(\mathcal{A}_u) = 16$ , which exactly matches with the dimension of the space.

In the following lines we study particular properties of NSEA in a low dimensional case.

**Observation 35.** *All non-deterministic behaviours in single party, binary input, binary output scenario have two minimal ensembles. All minimal ensembles of those have either two or three members.*

*Proof.* The polytope of single party binary input binary output behaviour belongs to set of reals of dimension  $d = 2$  and it has been given in Fig. 8. Suppose,  $P$  is any arbitrary behaviour in the polytope, then from the theorem of Carathéodory [106, 107], the maximal number of pure behaviours in each minimal ensembles of  $P$  is 3. It is also clear from the figure that any 3 pure behaviours of the polytope form a triangle, and any arbitrary point inside the polytope can be inside at most two overlapping triangles. Hence, it has at most two minimal ensembles.  $\square$

<sup>8</sup> $\sum_a P_{AE}(ae|xz) = P_E(e|z)$ ,  $e \in \{0, 1, 2\}$  for  $z = 0$  and  $e \in \{0, 1\}$  for  $z = 1$ .

<sup>9</sup> $\sum_e P_{AE}(ae|xz) = P_A(a|x)$ ,  $a, x \in \{0, 1\}$

For any arbitrary behaviour inside that polytope, the minimal ensembles consists of any combination of the following number of members.

**Corollary 36.** *Among single party, binary input, binary output behaviours, only five of them have NSCE that is a purification (NSCE which is a vertex).*

*Proof.* Due to the Observation 35 we know that the non-signaling complete extensions of behaviours in considered scenario are bipartite states in one of the following polytopes. Polytope of

- (i) one binary input and one unary input with binary outputs: behaviours lying on the edges of the polytope has single minimal ensembles with only two members.
- (ii) two binary inputs, two binary outputs behaviours: when the initial behaviour has two minimal ensembles and each of the ensembles has two members.
- (iii) two binary inputs, one binary output, one binary/ternary output (depending on the corresponding input setting) behaviours: behaviours lying on any one diagonal of the polytope.
- (iv) two binary inputs, and two ternary outputs behaviours: for any arbitrary behaviour, not belongs to the above sets.

For each local vertex (deterministic box) of the listed polytopes, if we trace out the second party by summation over all of its outcomes, the result is one of the deterministic behaviours of the initial polytope. Due to Theorem 1 of [102], we know the form of all non-local vertices in (i,ii,iii) polytopes. In each case after tracing out the extending system (summation over outcomes), the result is the maximally mixed behaviour of the initial system.

As we have investigated all the vertices which were suspected of being a purifications of behaviours from single party, one binary input, one binary output scenario and in each case we obtained one of the five states we conclude there are no other behaviours (in the initial polytope) that have purification.  $\square$

### B.1.2 Example: NSCE of $P_A$ gives ACCESS to any PME of $P_A$

In Theorem 20, we state that, the extended system of the NSCE, can access any PMEs of the behaviour  $P_A$ , if it is equipped with arbitrary randomness. Any pure ensemble of  $P_A$ ,  $\mathcal{E}(P_A) = \{x_i, P_i\}_i$ , where the  $\{x_i\}$  satisfy Eq. (58), can be written down as convex combination of the minimal ensembles, which is given below

$$\mathcal{E}(P_A) = \lambda \mathcal{M}_0(P_A) + (1 - \lambda) \mathcal{M}_1(P_A), \quad (68)$$

with  $\lambda = 2 - 3x_3 \in [0, 1]$ , as  $x_3 \in [\frac{1}{3}, \frac{2}{3}]$ . If the extending party  $X$  chooses to toss a coin  $p_t$ , (binary output) and feed it to the input  $z$  of her part of the completely extended behaviour with  $p_t(0) = \lambda$ , and  $p_t(1) = 1 - \lambda$ , then the extending system has ACCESS to any pure ensemble of  $P_A$ .

### B.1.3 Example: NSCE of $P_A$ gives ACCESS to any mixed ensemble of $P_A$

In this section, we will explicitly exemplify that the extending system  $E$  can access all possible mixed ensemble  $\mathcal{E}_{mix}(P_A) = \{p_m, P_M^m\}_m$ , of an arbitrary behaviour  $P_A$ , (given in Theorem 22). Here the behaviours  $P_M^m$  are any arbitrary behaviours.

Example 1: Suppose  $X$  wants to access the following ensemble of mixed behaviours

$$\mathcal{E}_{mix}(P_A) = \left\{ \left( \frac{33}{81}, P_M^0 \right); \left( \frac{32}{81}, P_M^1 \right); \left( \frac{16}{81}, P_M^2 \right) \right\}, \quad (69)$$

in part of system  $A^{10}$ , where the mixed behaviours (the blue points in Fig. 8), are given by

$$P_M^0 = \begin{array}{c|cc} \cancel{a^x} & 0 & 1 \\ \hline 0 & 1/5 & 2/5 \\ \hline 1 & 4/5 & 3/5 \end{array}, \quad P_M^1 = \begin{array}{c|cc} \cancel{a^x} & 0 & 1 \\ \hline 0 & 1/5 & 9/10 \\ \hline 1 & 4/5 & 1/10 \end{array}, \quad P_M^2 = \begin{array}{c|cc} \cancel{a^x} & 0 & 1 \\ \hline 0 & 7/8 & 3/4 \\ \hline 1 & 1/8 & 1/4 \end{array}. \quad (70)$$

<sup>10</sup>Here we use *tilde*, on the symbol of ensemble to denote a particular ensemble among the set of ensembles.



Each of these mixed behaviours has some decompositions over the pure behaviours, which are certainly not unique, consider the following minimal one, which are

$$\mathcal{M}(P_M^0) = \left\{ \left( \frac{2}{5}, P_E^1 \right); \left( \frac{1}{5}, P_E^2 \right); \left( \frac{2}{5}, P_E^3 \right) \right\}, \quad (71)$$

$$\mathcal{M}(P_M^1) = \left\{ \left( \frac{1}{10}, P_E^0 \right); \left( \frac{1}{10}, P_E^2 \right); \left( \frac{4}{5}, P_E^3 \right) \right\}, \quad (72)$$

and

$$\mathcal{M}(P_M^2) = \left\{ \left( \frac{3}{4}, P_E^0 \right); \left( \frac{1}{8}, P_E^1 \right); \left( \frac{1}{8}, P_E^2 \right) \right\}, \quad (73)$$

Put them into Eq. (69), the mixed ensemble then turn out to be the pure one, given by

$$\mathcal{E}_{pure}(P_A) = \left\{ \left( \frac{76}{405}, P_E^0 \right); \left( \frac{76}{405}, P_E^1 \right); \left( \frac{59}{405}, P_E^2 \right); \left( \frac{194}{405}, P_E^3 \right) \right\} = \{r(e), P_E^e\} \quad (74)$$

One can check that  $\mathcal{E}_{pure}(P_A) = \frac{76}{135} \mathcal{M}_0(P_A) + \frac{59}{135} \mathcal{M}_1(P_A)$ , and  $E$  can access  $\mathcal{E}_{pure}(P_A)$  by choosing the input according to the probability distribution,  $\{p_t(0) = \frac{76}{135}, p_t(1) = \frac{59}{135}\}$ . This can be done by feeding the output of a flipped coin to the input  $z$  of her part of the NSCE,  $\mathcal{E}(P)_{AE}(ae|xz)$ .

Once the PME in part of  $A$  has been prepared, the prefixed mixed ensembles has been constructed by passing the output  $e$  through a classical channel (post-processing channel)  $P_c(m|e)$ , which is

	$m \setminus e$	0	1	2	3	
$P_c(m e) =$	0	0	33/38	33/59	33/97	
	1	4/19	0	16/59	64/97	
	2	15/19	5/38	10/59	0	

(75)

the index  $m$  is the flag in part of  $E$ , different  $m$  give the access to different mixed behaviour  $P_M^m$  with probability  $p_m$ . One can check that  $p_m = \sum_e P_c(m|e)r(e)$ , and  $P_M^m = \frac{1}{p_m} \sum_e P_c(m|e)r(e)P_E^e$ . Thus we can see that the extending system can be able to access any ensemble of behaviour  $P_A$ , by NSCE with arbitrary randomness which will mix the minimal ensembles by mixing the input  $z$ , and then gluing the output  $e$  by a conditional classical channel.

Example 2:  $P_A$  can also be expanded as another mixed ensemble  $\mathcal{E}'_{mix}(P_A) = \left\{ \left( \frac{2}{5}, P_M^0 \right); \left( \frac{3}{5}, P_M^1 \right) \right\}$ , (the red points in Fig. 8) where

$$P_M^0 = \frac{5/6}{1/6} \Big| \frac{1}{0} = \left\{ \left( \frac{5}{6}, P_E^0 \right); \left( \frac{1}{6}, P_E^3 \right) \right\}, \quad (76)$$

$$P_M^1 = \frac{0}{1} \Big| \frac{4/9}{5/9} = \left\{ \left( \frac{5}{9}, P_E^1 \right); \left( \frac{4}{9}, P_E^3 \right) \right\}. \quad (77)$$

Now the pure ensemble turn out to  $\mathcal{E}'(P_A) = \mathcal{M}_0(P_A) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^3)\} = \{r'(e), P_E^e\}$ . For this,  $E$  will chose a completely biased coin,  $p_t(0) = 1, p_t(1) = 0$ , and feed its output to the input of the NSCE, and the post-processing channel  $P_c$  is

	$m \setminus e$	0	1	2	3	
$P_c(m e) =$	0	1	0	-	1/5	
	1	0	1	-	4/5	

(78)

which will be used to post-process the output of the extending part. Here we keep the column for  $e = 2$  “blank” as there is no such incidence that the pure behaviour  $P_E^2$  occur. Clearly  $p_0 = \sum_e r'(e)P_c(m|e) = \frac{1}{3} + \frac{1}{3} \times \frac{1}{5} = \frac{2}{5}$  and  $P_M^0 = \frac{1}{p_0} \sum_e r'(e)P_c(m=0|e)P_E^e = \frac{5}{6}P_E^0 + \frac{1}{6}P_E^3$ .

### B.1.4 NSCE can generate any extension

(Example of Theorem 24). Numbering these two examples of mixed ensembles with  $z' = 0$  and  $z' = 1$ , we obtain an arbitrary extension of  $P_A$  to the behaviour  $P_{AE}(am|xz')$ . Such that  $\{p(m|z' = 0), P^{m0}(a|x)\} = \mathcal{E}_{mix}(P_A)$  and  $\{p(m|z' = 1), P^{m1}(a|x)\} = \mathcal{E}'_{mix}(P_A)$ . And the arbitrary extended behaviour is

$$P_{AE}^{\text{mix}}(am|xz') = \begin{array}{c|cc|cc} & x & & 0 & & 1 \\ & a & & & & \\ z' & m & & & & \\ \hline & & & 0 & 1 & \\ \hline 0 & & & \frac{11}{135} & \frac{44}{135} & \frac{22}{135} & \frac{11}{45} \\ & & & \frac{32}{405} & \frac{128}{405} & \frac{16}{45} & \frac{16}{405} \\ & & & \frac{14}{81} & \frac{2}{81} & \frac{4}{27} & \frac{4}{81} \\ \hline 1 & & & \frac{1}{3} & \frac{1}{15} & \frac{2}{5} & 0 \\ & & & 0 & \frac{3}{5} & \frac{4}{15} & \frac{1}{3} \end{array} \quad (79)$$

We can consider all possible extension of  $P_A \rightarrow P_{AE}$ , which will take care of all possible ensembles of  $P_A$ .

### B.1.5 Quantifying non-locality introduced in NSCE

Here we quantify the amount of non-locality introduced among the extending and the extended system in the process of the construction of the NSCE, following Definition 18.

We have observed the fact that the completely extended behaviour of the maximally mixed single input output box, has turned out to be the Popescu-Rohrlich box [5], which (under suitable pre and post processing) can violate any kind of bipartite Bell expression maximally. In that case, the maximal amount of non-locality was introduced in the process of the construction of the NSCE. On the other hand, to quantify the non-locality of the NSCE of the non-maximally mixed single input output behaviour given in Eq. (56), the NSCE is shown in Eq. (62), to have different cardinalities of outputs. To get rid of this asymmetry in the extending system of  $X$ , we can do two possible surgeries.

**Case 1:** One can add one more outputs in the purified system and calculate the Bell like inequality defined by Collins et. al. (CGLMP) [108] (the acronym is after it finds D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu). The behaviour which maximizes the CGLMP bound has the following form after a local relabeling of the inputs and outputs

$$P_{AE}(ae|xz) = \begin{array}{c|ccc|ccc} & x & & 0 & & 1 \\ & e & & & & \\ z & a & & & & \\ \hline & & & 0 & 1 & 2 & \\ \hline 0 & & & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ & & & 0 & \frac{2}{3} & 0 & \frac{2}{3} & 0 \\ & & & 0 & 0 & 0 & 0 & 0 \\ \hline 1 & & & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ & & & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ & & & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \end{array} \quad (80)$$

For this bipartite two inputs and three-output box, the CGLMP bound turns out to be 3, which is beyond the quantum limit quoted to be 2.87 in Ref. [108].

**Case 2:** Another way to calculate the non-locality of this asymmetric behaviour by following the prescription giving in Ref. [104]. It proposes to merge the extra outcomes in the following way

$$P'_{AE}(ae = 1|xz) = P_{AE}(ae = 1|xz) + P_{AE}(ae = 2|xz) \quad (81)$$

Hence the behaviour in Eq. (62) can be transformed to a bipartite binary input output box,

$$P'_{AE}(ae|xz) = \begin{array}{c|cc|cc} & x & 0 & 1 & & \\ z & e^a & & & & \\ \hline 0 & 0 & 1/3 & 0 & 1/3 & 0 \\ & 1 & 0 & 2/3 & 1/3 & 1/3 \\ \hline 1 & 0 & 1/3 & 0 & 0 & 1/3 \\ & 1 & 0 & 2/3 & 2/3 & 0 \end{array} \quad (82)$$

For this behaviour we have the well known CHSH inequality to quantify the non-locality, and it is 3.33, which is also beyond the quantum limit. However, the amount of non-locality for this NSCE is substantially less than the amount of non-locality present in a PR box. It therefore seems that it may be possible to quantify the non-locality between subsystem and its extending system as a measure of how close a behaviour is to being a vertex for NSCE in the theory of non-signaling behaviours.

Until now, we have given examples in favor of the various properties of NSCE we have discovered so far. Now we want to shed some light on another aspects of NSCE which shows a sharp disparity with the purification principle of the QT. If  $|\psi_{AE}\rangle$ , is the purification of a quantum state  $\rho_A$ , to system  $E$ , then the same pure state is also the purification of quantum state  $\rho_X = \text{tr}_A|\psi_{AE}\rangle\langle\psi_{AE}|$ . In the latter section we give an example to show that this is not the case for the NSCE. If we have a behaviour  $P_A$ , and  $P_{AE}$  is its NSCE, then we say the behaviour  $P_E = \text{tr}_A P_{AE}$  is the conjugate box. We are going now to construct the NSCE of the conjugate box.

#### B.1.6 Complete extension of the conjugate box

In this section, we will find the NSCE of the conjugate behaviour of the behaviour given in Eq. (56). The conjugate box,  $P_E$  can be obtained from Eq. (62), by  $P_E(e|z) = \sum_a P_{AE}(ae|xz)$ , and it is given by

$$P_E(e|z) = \begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 1/3 & 1/3 \\ 1 & 1 & 1/3 & 2/3 \\ \hline 2 & 2 & 1/3 & 0 \end{array} \quad (83)$$

This behaviour lies in a 4 dimensional behaviour polytope whose vertices are given by

$$\begin{aligned} P_E^0 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^1 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^2 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 2 & 2 & 1 & 1 \end{array}}{2}, \\ P_E^3 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, & P_E^4 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ \hline 2 & 2 & 0 & 1 \end{array}}{2}, & P_E^5 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ \hline 2 & 2 & 0 & 0 \end{array}}{2}, \\ P_E^6 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ \hline 2 & 2 & 1 & 0 \end{array}}{2}, & P_E^7 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ \hline 2 & 2 & 1 & 0 \end{array}}{2}, & P_E^8 &= \frac{\begin{array}{c|cc|c} & e^z & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ \hline 2 & 2 & 0 & 1 \end{array}}{2}. \end{aligned} \quad (84)$$

To obtain the NSCE of this box, we need to find the minimal ensembles of  $P_E$ , which are

$$\mathcal{M}_0(P_E) = \{(1/3, P_E^0); (1/3, P_E^1); (1/3, P_E^7)\}, \quad (85)$$

$$\mathcal{M}_1(P_E) = \{(1/3, P_E^1); (1/3, P_E^3); (1/3, P_E^6)\}, \quad (86)$$

$$\mathcal{M}_2(P_E) = \{(1/3, P_E^3); (1/3, P_E^5); (1/3, P_E^7)\}. \quad (87)$$

Consider the NSCE of  $P_E$  to a system  $A'$ , as  $P_{A'E}(a'e|x'z)$ , where  $\{p(a' = i|x' = k), P_E^{ik}(e|z)\} = \mathcal{M}_k$ , the NSCE of  $P_E$  is

$$P_{A'E}(a'e|x'z) = \begin{array}{c} z \\ \begin{array}{c|ccc|ccc|ccc} & \begin{array}{c} x' \\ e \end{array} & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \\ \hline 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 1 & 1 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} & 0 \\ 2 & 2 & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \hline 1 & 0 & \frac{1}{3} & 0 & 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 1 & 1 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{array} \end{array} \quad (88)$$

One can clearly see that  $P_{AE} \neq P_{A'E}$ , for example, by noticing the mismatch of the cardinality of the inputs of the extended party of the conjugate system.

## B.2 Dimensionality of the No-signaling complete extension

In this subsection we provide a proof for Theorem 25 of the main text.

*Proof.* Let us recall, that in quantum theory (due to the Schmidt decomposition) purifications can always be found with an extending system with the same dimension as the extended system. We are interested in similarly quantifying the size of the minimal NSCE. However, there are many possible quantifiers that we could use here, for example, we can measure the size of the NSCE using the total number of its outputs, inputs and outputs, or, the dimension of its state space (vector space). The theory of non-signalling behaviours is a discrete convex theory, hence we call its state space a convex polytope, a multidimensional generalization of a convex polyhedron. In this section, we strive to upper bound the dimension of the polytope that contains NSEA of some fixed but arbitrary behaviour  $P$ . We identify the dimension of behaviour with the dimension of a polytope it belongs. Although we chose the dimension of NSEA, as a quantifier of its size, other aspects of it, like the number of its inputs and outputs, will also be discussed in the following.

From Theorem 1 of [104] we know that the dimension of a certain behaviour polytope  $\mathcal{B}$  is:

$$\dim \mathcal{B} = \prod_{i=1}^n \left( \sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad (89)$$

where  $n$  is the total number of non-signalling parties,  $m_i$  is the total number of the inputs of the  $i^{\text{th}}$  party, with  $v_{ij}$  being the number of outputs for the  $j^{\text{th}}$  input. This polytope of  $\dim \mathcal{B}$  is contained in the vector space of  $\mathbb{R}^t$ , where  $t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}$  is the total number of outputs.

Suppose now that  $P \in \mathcal{B}$ , then there exists a polytope  $\tilde{\mathcal{B}}$ , such that the NSEA  $\mathcal{E}(P) \in \tilde{\mathcal{B}}$ . The dimension of  $\tilde{\mathcal{B}}$  can be determined by:

$$\dim \tilde{\mathcal{B}} = \prod_{i=1}^{n+1} \left( \sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1 \quad (90)$$

$$= \left( \sum_{j=1}^{m_{n+1}} (v_{n+1,j} - 1) + 1 \right) (\dim \mathcal{B} + 1) - 1, \quad (91)$$

where  $n + 1$  is the index of the extending party, and the last equality is due to expansion of product with respect to  $n + 1$ -th term.

To obtain the upper bound on  $\dim \tilde{\mathcal{B}}$  it is enough now to find upper bounds on the number of inputs and outputs of the extending system. The following considerations have a qualitative character, we are much more interested in the fact that an upper bound on  $\dim \tilde{\mathcal{B}}$  exists than in its tightness. The upper bound on the number of outputs can be found via Carathéodory theorem [107], i.e.,  $v_{n+1,j} \leq \dim \mathcal{B} + 1$ , we obtain:

$$\dim \tilde{\mathcal{B}} \leq (m_{n+1} \dim \mathcal{B} + 1) (\dim \mathcal{B} + 1) - 1. \quad (92)$$

It is instructive to notice that the above inequality can be saturated for some generic behaviours. The number of inputs of the extending system is equal to the number of minimal ensembles of the behaviour  $\mathcal{P}$ . Suppose now that  $V$  is the number of vertices of the polytope  $\mathcal{B}$ . Then, using Carathéodory theorem again, each choice of  $\dim \mathcal{B} + 1$  vertices (out of  $V$ ) leads to a minimal ensemble via elimination of the vertices, therefore  $m_{n+1} \leq \binom{V}{\dim \mathcal{B} + 1}$ , and hence:

$$\dim \tilde{\mathcal{B}} < \left( \binom{V}{\dim \mathcal{B} + 1} \dim \mathcal{B} + 1 \right) (\dim \mathcal{B} + 1), \quad (93)$$

where the last term  $(-1)$  was neglected.

The characterisation of the vertices of polytopes of behaviours with two binary inputs and two binary outputs in bipartite and tripartite scenarios was provided in [102]. In general, however, finding vertices of a polytope is a hard problem known as the ‘‘vertex enumeration problem’’. Using McMullen’s Upper Bound Theorem [107, 109, 110], we can, however, still obtain an upper bound for the number of vertices in terms of  $\dim \mathcal{B}$  and  $t$ . Here we use a fact that a polytope can be defined as an intersection of a set of halfspaces. Then following [110] (up to notation) we have the subsequent statement.

Let  $A$  be an  $m \times t$  matrix of reals and let  $b \in \mathbb{R}^m$  be a real vector. Consider a polytope

$$\mathcal{P} = \{x \in \mathbb{R}^t : Ax \leq b\}, \quad (94)$$

and  $V_{\mathcal{P}}$  be the number of vertices of  $\mathcal{P}$  then:

$$V_{\mathcal{P}} \leq \binom{m-t-s}{s} + \binom{m-s-1}{t-s-1}, \quad (95)$$

where  $s = \lfloor t/2 \rfloor$ .

To achieve our goal it is enough to determine the number of rows  $m$  of matrix  $A$ . We divide now all the constraints on the non-signalling polytope into three classes, written in terms of marginal probabilities  $p_{ijk}$ :

1. Probabilistic constraints:  $0 \leq p_{ijk} \leq 1$ .
2. Normalization constraints, i.e.,  $\sum_k^{v_{ij}} p_{ijk} = 1$ .
3. non-signalling constraints (see Section 4.1).

It is easy to see that constraints of type 1 contribute to  $2t$  rows of matrix  $A$ . In the next step we observe that constraints of type 2 and 3 are linear and together reduce the dimension of the space of correlations from  $t$  to  $\dim \mathcal{B}$ , and hence can be encoded in  $t - \dim \mathcal{B}$  linearly independent rows of matrix  $A$  (irrespective of the actual number of constraints). The total number of rows is therefore upper bound with  $m \leq 3t - \dim \mathcal{B}$  (there still is some redundancy because of the relation between constraints of types 1 and 2), hence:

$$V \leq \binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1}. \quad (96)$$

Finally, we obtain

$$\begin{aligned} \dim \tilde{\mathcal{B}} &< (\dim \mathcal{B} + 1) \\ &\times \left( \left( \binom{2t - \lfloor t/2 \rfloor - \dim \mathcal{B}}{\lfloor t/2 \rfloor} + \binom{3t - \lfloor t/2 \rfloor - (\dim \mathcal{B} + 1)}{t - \lfloor t/2 \rfloor - 1} \right) \dim \mathcal{B} + 1 \right), \end{aligned} \quad (97)$$

where:

$$\dim \mathcal{B} = \prod_{i=1}^n \left( \sum_{j=1}^{m_i} (v_{ij} - 1) + 1 \right) - 1, \quad t = \prod_{i=1}^n \sum_{j=1}^{m_i} v_{ij}. \quad (98)$$

□

In Theorem 25 we were interested only in showing that the dimension of the complete extension is bounded from above, and hence our result is very loose. In particular the number of candidates for a minimal ensemble is very inefficient upper bound on the number of minimal ensembles, although obtained with simple combinatorics. Therefore, the first place for improvement is to find tighter upper bound on the number of minimal ensembles. A more rigorous treatment can show that many candidates lead to the same minimal ensemble and moreover that some of them are not valid. Furthermore, the upper bound on the number of vertices that we used is very general and does not incorporate, for example, symmetries of the non-signalling polytope.

### B.3 Complete extensions of three-cycle contextual behaviour

In the previous sections, we have considered the NSCE of an arbitrary behavior which is inherently of the form  $P_A(a|x)$ , i.e., a conditional probability distribution consist of only single set of input  $x$  and output  $a$ . The polytope of those behaviors are completely determined by the constraints probability distributions satisfy, like non-signaling and normalization condition. Here, within the theory of non-signalling behaviours (NS) we will study contextuality scenarios, and, hence, investigate NSCE of contextual behaviours. The polytope of contextual behaviors are not simply the non-signalling polytopes but rather they are termed as no-disturbance polytopes. And while considering the minimal ensembles of a contextual behavior we will focus on the pure members of the no-disturbance polytopes.

In particular we are going to focus on the three-cycle contextuality scenario [111, 112] (aka Specker's triangle [113, 114]).

First we will introduce this contextuality scenario in standard notation before showing how to recast it within NS. The three-cycle scenario consists of a triple of binary observables  $\{X_0, X_1, X_2\}$  with outcomes  $\{a, b, c\}$  valued in  $\{-1, 1\}$ , together with the constraint that these are pairwise compatible – that is, any pair  $\{X_i, X_j\}$  can be jointly measured without disturbance, but are globally incompatible – that is, there is no way to measure all three without disturbance. Such a setup can be realised in quantum theory [115] provided that the observables are not taken to correspond to projective measurements. This compatibility structure can be captured by the set of maximal contexts [111, 112]

$$\mathcal{C}_3 = \{\{X_0, X_1\}, \{X_1, X_2\}, \{X_2, X_0\}\}. \quad (99)$$

The behaviour of a particular realisation of this scenario is captured by the joint distributions over the outcomes of the compatible observables, i.e.,  $p(a, b)$ ;  $p(b, c)$ ; and  $p(c, a)$ , which can be represented in the following table:

	$a \backslash b$	+1	-1	$b \backslash c$	+1	-1	$c \backslash a$	+1	-1	
	+1			+1			+1			
	-1			-1			-1			

(100)

Any such behaviour must satisfy certain no-disturbance constraints<sup>11</sup> [111, 112], which together define

<sup>11</sup>As it should follow the no-disturbance condition like the non-signaling one in non-locality, which is

$$p(a) = \sum_{b=-1}^1 p(a, b) = \sum_{c=-1}^1 p(c, a) \quad (101)$$

$$p(b) = \sum_{a=-1}^1 p(a, b) = \sum_{c=-1}^1 p(b, c) \quad (102)$$

$$p(c) = \sum_{b=-1}^1 p(b, c) = \sum_{a=-1}^1 p(c, a) \quad (103)$$

the no-disturbance polytope. Formally defining contextuality is beyond the scope of this paper, but, for our purposes it suffices to note that a behaviour is non-contextual if it is a convex combination of the eight deterministic vertices of the no-disturbance polytope

$$\begin{aligned} \mathcal{N}_0^{\mathcal{C}} &= \frac{1}{0} \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| & \mathcal{N}_1^{\mathcal{C}} &= \frac{1}{0} \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \\ \mathcal{N}_2^{\mathcal{C}} &= \frac{0}{0} \left| \frac{1}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| & \mathcal{N}_3^{\mathcal{C}} &= \frac{0}{1} \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| & \mathcal{N}_4^{\mathcal{C}} &= \frac{0}{0} \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \\ \mathcal{N}_5^{\mathcal{C}} &= \frac{0}{1} \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| & \mathcal{N}_6^{\mathcal{C}} &= \frac{0}{0} \left| \frac{0}{1} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{0}{0} \right| \left| \frac{1}{0} \right| & \mathcal{N}_7^{\mathcal{C}} &= \frac{0}{0} \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \left| \frac{0}{1} \right| \left| \frac{0}{1} \right| \left| \frac{0}{0} \right| \end{aligned}$$

otherwise, it is a contextual behaviour. Any other vertex of the no-disturbance polytope is therefore contextual, indeed, there are four of these which are

$$\begin{aligned} \mathcal{C}_0 &= \frac{1/2}{0} \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| & \mathcal{C}_1 &= \frac{1/2}{0} \left| \frac{0}{1/2} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \\ \mathcal{C}_2 &= \frac{0}{1/2} \left| \frac{1/2}{0} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| & \mathcal{C}_3 &= \frac{0}{1/2} \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right| \left| \frac{0}{1/2} \right| \left| \frac{1/2}{0} \right|. \end{aligned}$$

Note that the dimension of the no-disturbance polytope is 6, hence, according to the Carathéodory theorem [107], the set of minimal ensembles of an arbitrary behaviour inside the polytope has at most 7 elements.

We will now focus on a particular class of behaviours within this polytope, namely those lying on an isotropic line – a mixture of a contextual vertex, say  $\mathcal{C}_0$  and the maximally mixed behaviour<sup>12</sup> which is a non-contextual behaviour. We denote an element on this isotropic line (parameterised by  $\lambda \in (0, 1)$ ) as  $P_\lambda$  which can be explicitly written as

$$\begin{aligned} P_\lambda &= (1 - \lambda)\mathcal{C}_0 + \lambda M \\ &= \frac{2-\lambda}{4} \left| \frac{\lambda}{4} \right| \left| \frac{2-\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \left| \frac{2-\lambda}{4} \right| \left| \frac{\lambda}{4} \right| \end{aligned} \quad (105)$$

The minimal ensembles of  $P_\lambda$ , have been found by applying numerical techniques to obtain the solution of a set of linear equations, and they are as follows.

For all  $\lambda \in (0, 1)$  we have the following

$$v = \left[ \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \frac{\lambda}{2} \right] \quad (106)$$

$$\mathcal{M}_1(P) = \left[ \mathcal{N}_2^{\mathcal{C}}, \mathcal{N}_5^{\mathcal{C}}, \mathcal{C}_0 \right] \quad (107)$$

$$v = \left[ \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \lambda, \frac{\lambda}{2} \right] \quad (108)$$

$$\mathcal{M}_2 = \left[ \mathcal{N}_3^{\mathcal{C}}, \mathcal{N}_4^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_1 \right] \quad (109)$$

$$\mathcal{M}_3 = \left[ \mathcal{N}_1^{\mathcal{C}}, \mathcal{N}_6^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_2 \right] \quad (110)$$

$$\mathcal{M}_4 = \left[ \mathcal{N}_0^{\mathcal{C}}, \mathcal{N}_7^{\mathcal{C}}, \mathcal{C}_0, \mathcal{C}_3 \right] \quad (111)$$

<sup>12</sup>The maximally mixed behaviour  $M$  is given by

$$M = \frac{1}{4} \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \left| \frac{1}{4} \right| \quad (104)$$

$$v = \left[ 1 - \frac{3\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4} \right] \quad (112)$$

$$\mathcal{M}_5 = \left[ \mathcal{C}_0, \mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3 \right] \quad (113)$$

$$v = \left[ \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \lambda \right] \quad (114)$$

$$\mathcal{M}_6 = \left[ \mathcal{N}^{\mathcal{C}_1}, \mathcal{N}^{\mathcal{C}_2}, \mathcal{N}^{\mathcal{C}_3}, \mathcal{N}^{\mathcal{C}_7}, \mathcal{C}_0 \right] \quad (115)$$

$$\mathcal{M}_7 = \left[ \mathcal{N}^{\mathcal{C}_0}, \mathcal{N}^{\mathcal{C}_4}, \mathcal{N}^{\mathcal{C}_5}, \mathcal{N}^{\mathcal{C}_6}, \mathcal{C}_0 \right] \quad (116)$$

In  $\lambda \in (0, \frac{2}{3})$ , we additionally have

$$v = \left[ \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, \frac{\lambda}{4}, 1 - \frac{3\lambda}{2} \right] \quad (117)$$

$$\mathcal{M}_8 = \left[ \mathcal{N}^{\mathcal{C}_0}, \mathcal{N}^{\mathcal{C}_1}, \mathcal{N}^{\mathcal{C}_3}, \mathcal{N}^{\mathcal{C}_4}, \mathcal{N}^{\mathcal{C}_6}, \mathcal{N}^{\mathcal{C}_7}, \mathcal{C}_0 \right] \quad (118)$$

Whilst in  $\lambda \in (\frac{2}{3}, 1)$ , the behaviour has another 5 minimal ensembles. However, in this range of  $\lambda$ ,  $P_\lambda$  is non-contextual (see Theorem 1 in [112]) and we are interested in exploring the case in which  $P_\lambda$  displays contextuality so we will not go into detail about these.

To study the NSEAs for contextuality scenarios we must first recast the scenario into the theory of non-signalling behaviours. In particular, rather than describing the scenario by a triple of bipartite distributions (i.e.,  $p(a,b)$ ,  $p(b,c)$ , and  $p(c,a)$ ) we will instead describe the scenario by a conditional probability distribution. We do this by associating the input  $x \in \{0, 1, 2\}$  to the maximal contexts, i.e.,  $0 \sim \{X_0, X_1\}$ ,  $1 \sim \{X_1, X_2\}$ , and  $2 \sim \{X_2, X_0\}$ ; and associating an outcome  $a' \in \{0, 1, 2, 3\}$  to the outcome of the bipartite distributions  $\{+1, -1\} \times \{+1, -1\}$ . This defines an embedding of the no-disturbance polytope into the polytope of single input output behaviours (with the relevant dimensions).

In our particular example of interest, namely  $P_\lambda$  using this embedding we can associate the behaviour of the contextuality scenario  $P_\lambda$  to the single input output behaviour

$$P'_{A\lambda} = \begin{array}{c|ccc} & a' & 0 & 1 & 2 \\ \hline & 0 & \frac{2-\lambda}{4} & \frac{2-\lambda}{4} & \frac{\lambda}{4} \\ \hline & 1 & \frac{\lambda}{4} & \frac{\lambda}{4} & \frac{2-\lambda}{4} \\ \hline & 2 & \frac{\lambda}{4} & \frac{\lambda}{4} & \frac{2-\lambda}{4} \\ \hline & 3 & \frac{2-\lambda}{4} & \frac{2-\lambda}{4} & \frac{\lambda}{4} \end{array} \quad (119)$$

Now we are in the position that we can construct the NSCE of  $P'_{A\lambda}$ , in particular, focusing on the contextual case in which  $\lambda \in (0, \frac{2}{3})$ . Note however, that we want to construct the NSCE not for the minimal ensembles of the behaviour within the single input output polytope, but for the ensembles within the embedded no-disturbance polytope. The NSCE of this will however still live inside the polytope of no signalling behaviours with a pair of inputs and outputs (of suitable dimensions).



In particular, one can show that the NSEA of  $P'_{A\lambda}$  in  $\lambda \in (0, \frac{2}{3})$  is given by

		$x$				$0$				$1$				$2$					
		$a$	$0$	$1$	$2$	$3$	$0$	$1$	$2$	$3$	$0$	$1$	$2$	$3$	$0$	$1$	$2$	$3$	
$P'_{AE\lambda}(ae xz) =$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$
		$1$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$
		$2$	$\frac{2-\lambda}{4}$	$0$	$0$	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	$0$	$0$	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	$\frac{2-\lambda}{4}$	$0$
	$1$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$
		$1$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$
		$2$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$
		$3$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$
	$2$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$
		$1$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$
		$2$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$
		$3$	$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$
	$3$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$
$1$		$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	
$2$		$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	
$3$		$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	
$4$	$0$	$\frac{4-3\lambda}{8}$	$0$	$0$	$\frac{4-3\lambda}{8}$	$\frac{4-3\lambda}{8}$	$0$	$0$	$\frac{4-3\lambda}{8}$	$0$	$\frac{4-3\lambda}{8}$	$\frac{4-3\lambda}{8}$	$0$	$0$	$\frac{4-3\lambda}{8}$	$\frac{4-3\lambda}{8}$	$0$	$0$	
	$1$	$\frac{\lambda}{8}$	$0$	$0$	$\frac{\lambda}{8}$	$0$	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	$0$	$\frac{\lambda}{8}$	$0$	$0$	$0$	$\frac{\lambda}{8}$	$0$	$0$	$0$	$\frac{\lambda}{8}$	
	$2$	$0$	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	$0$	$\frac{\lambda}{8}$	$0$	$0$	$\frac{\lambda}{8}$	$0$	$0$	$0$	$0$	$\frac{\lambda}{8}$	$0$	$0$	$0$	$\frac{\lambda}{8}$	
	$3$	$0$	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	$0$	$0$	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	$0$	$0$	$\frac{\lambda}{8}$	$\frac{\lambda}{8}$	$0$	$0$	$0$	$0$	$0$	$0$	
$5$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	
	$1$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	
	$2$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	
	$3$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	
$6$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	
	$1$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	
	$2$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	
	$3$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	
$7$	$4$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	$\frac{1-\lambda}{2}$	$0$	$\frac{1-\lambda}{2}$	$0$	$0$	$0$	$\frac{1-\lambda}{2}$	$\frac{1-\lambda}{2}$	$0$	$0$	
	$5$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	
	$6$	$0$	$0$	$0$	$\frac{\lambda}{4}$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$0$	$\frac{\lambda}{4}$	
	$6$	$\frac{2-3\lambda}{4}$	$0$	$0$	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	$0$	$0$	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	$0$	$0$	$0$	$0$	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	$\frac{2-3\lambda}{4}$	$0$	$0$

(120)

One can then observe that non-locality is a special case of contextuality in which the no-disturbance polytope coincides with a non-signalling polytope – in such a case the NSCE that we construct for the contextuality scenario will be precisely that which we would construct in the non-signalling scenario.

Note that as the no-disturbance polytope is a discrete theory, that theorem 7 applies and so in general the NSEAs that we construct will not be purifications. However, it would be interesting to study the cases in which they are.

## B.4 Complete extension of the behaviours lying on the isotropic line

The aim of this section is to construct explicitly the NSCE of the isotropic behaviour, that is, a mixture of the  $PR$  and anti- $PR$  box. We therefore focus on the polytope of behaviours,  $P_{AB}(ab|xy)$ , with two binary inputs  $x, y \in \{0, 1\}$ , two binary outputs  $a, b \in \{0, 1\}$ , and satisfying the non-signaling condition. There are 24 extremal (pure) behaviours (vertices) of this polytope, among which 16 are local behaviours given by

$$L_{\alpha\beta\gamma\delta}(ab|xy) = \begin{cases} 1 & \text{if } a = \alpha x \oplus b, \\ & b = \gamma y \oplus \delta \\ 0 & \text{otherwise.} \end{cases} \quad (121)$$

with  $\alpha, \beta, \gamma$  and  $\delta \in \{0, 1\}$ . And another 8 non-local behaviours, which are

$$B_{rst}(ab|xy) = \begin{cases} 1/2 & \text{if } a \oplus b = xy \oplus rx \oplus sy \oplus t \\ 0 & \text{otherwise.} \end{cases} \quad (122)$$

with  $r, s, t$  taking values either 0 or 1 [102]. The triple  $(r, s, t)$  enumerates the non-local behaviours in the polytope of behaviours with two binary inputs and two binary outputs. Hence, the isotropic behaviour can be formulated as,  $B(\eta)_{AB}(ab|xy) = \eta B_{000}(ab|xy) + (1 - \eta)B_{111}(ab|xy)$ , where  $\eta \in (0, 1)$ , and where  $B_{000}$  and  $B_{111}$  are the  $PR$  and anti- $PR$  boxes respectively. One can easily check that  $B(\frac{1}{2}) = P_{AB}^m(ab|xy)$ , is a maximally mixed behaviour in the polytope of two input two output behaviours (the one taking value  $P(ab|xy) = \frac{1}{4}$  for all  $a, b, x, y \in \{0, 1\}$ ). All the behaviours  $B(\eta)$ , for  $\eta \in (1, \frac{1}{2})$  can be transformed into to the behaviours  $B(\eta)$ , for  $\eta \in (\frac{1}{2}, 0)$  by local relabeling of inputs and outputs. Thus, in our investigation we will consider only the behaviours lying on the isotropic line from the  $PR$  box to maximally mixed behaviour, i.e., for  $B(\eta)$ , where  $\eta \in [\frac{1}{2}, 1)$ .

We know from the definition of NSCE, Definition 18, that, corresponding to each input of the extending party, there is a minimal ensemble  $\mathcal{M}(B(\eta))$ , where the members of the ensembles are enumerated by the outputs of the extending system. Hence, finding all possible minimal ensembles of a behaviour is sufficient for the construction of NSCE.

### B.4.1 Minimal ensembles for isotropic behaviours

Our aim is to find the minimal ensembles for the behaviours lying on the isotropic line, but, in this section we first focus on the minimal ensembles for a particular behaviour, the Bell Tsirelson behaviour [116]. The Bell Tsirelson behaviour, lying on the isotropic line for  $\eta = \frac{2+\sqrt{2}}{4}$ , reaches the quantum limit in violating the CHSH inequality [99]. The challenge in finding the minimal ensembles is that the Bell Tsirelson behaviour is specified by an irrational number, however, despite this, we are able to find all of its minimal ensembles analytically, due to the following observations and a theorem.

**Observation 37.** *All the pure members ensembles of the behaviours  $B(\eta)$ , in the isotropic line, contains the  $PR$  box,  $B_{000}$  as one of its member element, for all  $\eta \in (\frac{3}{4}, 1]$ .*

*Proof.* By direct inspection of the 24 vertices of the polytope of behaviours with two binary inputs and two binary outputs we observe that  $PR$ -box is the only behaviour, which has value 4 of the CHSH functional which reads

$$\beta(P) = \sum_{x,y} \sum_{(a,b) \in \text{supp}(B_{000})} P(ab|xy) \quad (123)$$

Moreover, for all other vertices the functional achieves values  $\leq 3$ . It is also clear that the functional  $\beta$  is linear (in fact it is equal to  $2\langle P|B_{000} \rangle$  where  $\langle \cdot | \cdot \rangle$  denotes Euclidean scalar product of vectors (see Lemma 4 of [117]). Let us then suppose that there exists an ensemble  $\{p_i, V_i\}_i$  of a behaviour  $P$  from the isotropic line  $(3/4, 1]$ , which does not have  $PR$  in its set of members (here  $V_i$  are the vertices of the polytope). We have then  $\beta(P) = \beta(\sum_i p_i V_i) = \sum_i p_i \beta(V_i) \leq \sum_i p_i 3 \leq 3$ . This contradicts the fact that all behaviours from the considered set satisfy  $\beta(P) > 3$  i.e. they violate CHSH inequality.  $\square$

Let  $S_{\mathcal{M}}(\eta)$  denote the set of all minimal ensembles of the isotropic behaviour  $B(\eta)$ .

**Lemma 38.** *Let  $1 > \eta' > \eta > \frac{3}{4}$ , then for any minimal ensemble from  $S_{\mathcal{M}}(\eta)$ , there exists a unique minimal ensemble of  $S_{\mathcal{M}}(\eta')$  with the same set of members.*

*Proof.* Since  $\eta, \eta' \in (\frac{3}{4}, 1)$  by observation 37, all the minimal ensembles of  $B(\eta)$  must contain  $B_{000}$  as its members elements. Now, from lemma 29, for the set  $V(\mathcal{M}_1)$  there exists unique weights  $\{p_i > 0\}_{i=0}^n$  such that

$$\sum_{V_i \in V(\mathcal{M}_1)} p_i V_i = B(\eta), \quad (124)$$

where  $V_0 = B_{000}$  and  $V_i$  for  $i \in \{1, \dots, n\}$  are vertices of the polytope of behaviours with two binary inputs and two binary outputs. Now, from  $\eta' > \eta$  there exists unique weight  $1 > q > 0$  such, that  $B(\eta') = qB(\eta) + (1 - q)B_{000}$ . Hence, for the set  $\{V_i\}_{i=0}^n$  there exist unique weights  $\{r_i\}_{i=0}^n$  given by equations  $r_0 = qp_0 + (1 - q)$  and  $r_i = p_i(1 - q)$  for  $i \in \{1, \dots, n\}$ , such that  $\sum_i r_i V_i = B(\eta')$ . Hence, by lemma 29, the ensemble  $\{r_i, V_i\}_{i=0}^n$  is minimal ensemble of  $B(\eta')$ , and the assertion follows.  $\square$

We will now adopt the notation, that  $\mathcal{V}$  is the set of the sets of members of minimal ensembles:

$$\mathcal{V}(P) := \bigcup_{\mathcal{M} \in S_{\mathcal{M}}(P)} \{V(\mathcal{M}(P))\}. \quad (125)$$

From lemma 38 we have the immediate corollary:

**Corollary 39.** *Let  $1 > \eta' > \eta > \frac{3}{4}$ , and  $B(\eta) = \eta B_{000} + (1 - \eta)B_{111}$  then, there is  $\mathcal{V}(P(\eta)) \subseteq \mathcal{V}(P(\eta'))$*

**Theorem 40.** *Let  $\frac{3}{4} < \eta' < \eta < \eta'' < 1$ . If  $\mathcal{V}(B(\eta')) = \mathcal{V}(B(\eta'')) \equiv \mathcal{V}$ , then we also have  $\mathcal{V}(\eta) = \mathcal{V}$ .*

*Proof.* Since  $\eta', \eta'' \in (\frac{3}{4}, 1)$ , from the Corollary 39, we have

$$\mathcal{V}(B(\eta')) \subseteq \mathcal{V}(B(\eta)) \subseteq \mathcal{V}(B(\eta'')) \quad (126)$$

However by assumption there is  $\mathcal{V}(B(\eta'')) \subseteq \mathcal{V}(B(\eta'))$ , hence  $\mathcal{V}(B(\eta)) \subseteq \mathcal{V}(B(\eta'))$ , and the assertion follows.  $\square$

The above theorem holds for the range of the isotropic parameter  $(\frac{3}{4}, 1)$ . Moreover, due to this theorem we are able to find out the minimal ensembles of all non-local  $B(\eta)$  for  $\eta \in (\frac{1}{2}, 1)$ . However, numerical investigate indicates that it holds true also for the  $(\frac{1}{2}, \frac{3}{4}]$ , hence we conjecture as follows:

**Conjecture 41.** *Let  $\frac{1}{2} < \eta' < \eta < \eta'' < 1$ . If  $\mathcal{V}(B(\eta')) = \mathcal{V}(B(\eta'')) \equiv \mathcal{V}$ , then we also have  $\mathcal{V}(\eta) = \mathcal{V}$ .*

The above observation and theorem help us to find out the NSEA of the Bell-Tsirelson's box. There are a total of 354 minimal ensembles of  $B(\eta)$ , for  $\eta = \frac{2+\sqrt{2}}{4}$ . Surprisingly, the NSEA of a noisy PR box leads to an extending system  $E$  embedded in a vector space of dimension 2,837 (taking into account the normalization constraints it effectively lives in a 2,483 dimensional space) and the entire NSEA, is lying in a polytope which is embedded in a vector space of dimension 45,392 taking into account the normalization and non-signaling constraints effectively lives in space of dimension 22,355. The list of these ensembles are given in Sec. B.4.3.

#### B.4.2 Dimension of the extending party for the behaviours lying on the isotropic line

In the previous section, we discussed how one can construct the NSCE of the Bell-Tsirelson box. In this section, due to the large number of minimal ensembles for a generic  $B(\eta)$ , we will discuss some of the statistics of the NSCE of more general isotropic behaviours.

For any value of  $\eta \in [\frac{1}{2}, 1)$ , we have provided an estimate of the NSCE for the bipartite behaviour  $B(\eta)$  for  $\eta \in (\frac{1}{2}, 1)$ . This is indeed an NSCE of this behaviour provided that conjecture 41 holds true. The number of elements in each of the minimal ensembles is bounded between 2 and  $\dim \mathcal{B} + 1$ , which is 9, for two input and two output binary behaviours. In the entire range of  $\eta \in (\frac{1}{2}, 1)$ , we have numerically calculated the total number of inputs (number of minimal ensembles  $|\mathcal{M}|$ ) of the

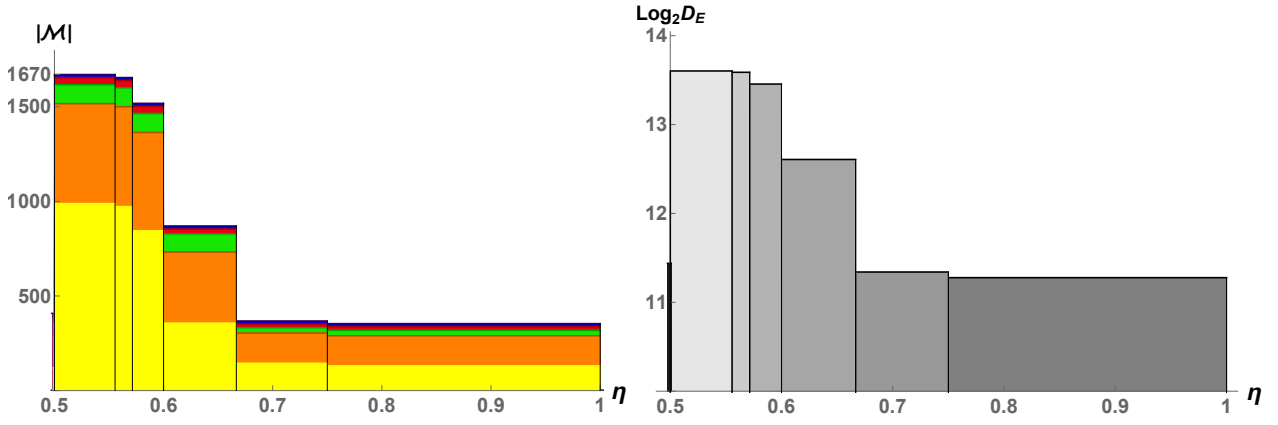


Figure 9: Panel a) Histogram plot of the total number of inputs  $|\mathcal{Z}| = |\mathcal{M}|$  of the extending system of the NSEA of the behaviours along the isotropic line. Different color in each column of the histogram shows the the number of inputs having different numbers of outputs. The height of the yellow color in each column represents the number of inputs with 9 outputs, orange represents the same for 8 outputs, green is for 7 outputs, red is for 6 outputs and so on. The inputs of the extending system for various values of outputs are given in Table 1. Panel b) Histogram plot of the memory required in the extending system to store the information about the minimal ensembles of the behaviours along the isotropic line. Here  $D_E$  represents the dimension of the extending system of NSEA. All the plots have been given from  $\eta = 0$  to  $\eta = \frac{1}{2}$ , i.e., from the PR box to the maximally mixed box.

extending system, which have an equal number of outputs  $v_j \in \{2, 3, \dots, 9\}$ . This is given in table 1. The Bell-Tsirelson box,  $\eta = \frac{2+\sqrt{2}}{4}$ , lying in the range  $\eta \in (\frac{3}{4}, 1)$ , is given in the last column of Table 1, having total number of inputs as 354. Among these 354 inputs there are 130 inputs having 9 outputs, 160 inputs having 8 outputs and so on. The minimal ensembles associated with each input of the Bell-Tsirelson box have been given in Sec. B.4.3 of the supplemental material. Here,  $\mathcal{M}$  represent the number of minimal ensembles of a behaviour and the length of the column represent the number of pure behaviours in each of the minimal ensembles.

$v_j \backslash \eta$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{5}{9}]$	$(\frac{5}{9}, \frac{4}{7}]$	$(\frac{4}{7}, \frac{3}{5}]$	$(\frac{3}{5}, \frac{2}{3}]$	$(\frac{2}{3}, \frac{3}{4}]$	$(\frac{3}{4}, 1)$
2	4	1	1	1	1	1	1
3	0	3	3	3	3	3	3
4	12	0	0	0	0	0	0
5	0	12	12	12	12	12	12
6	32	38	38	38	26	20	20
7	64	100	100	100	96	28	28
8	176	528	528	520	376	160	160
9	120	988	972	844	356	144	130
total	408	1670	1654	1518	870	368	354

Table 1: Table shows the number of inputs (number of minimal ensembles  $|\mathcal{M}|$ ), of the extending system which have same number of outputs  $v_j$ , for various values of  $v_j \in \{2, 3, \dots, 9\}$ , in the entire range of the parameter  $\eta$ .

The number of inputs of the extending system for various values of  $\eta$ , has been given in figure 9(a). The color represent the number of inputs among the total number of inputs having equal number of outputs. The yellow color stands for the set of inputs having 9 outputs, orange is for 8, green is for 7 and red is for 6 and so on. We have also plotted the memory required by the extending system, to store the information about the all possible minimal ensembles of the behaviours lying on the isotropic

line, in figure 9(b). Here  $D_E$  represent the dimension of the extending system.

$\dim \eta$	$\frac{1}{2}$	$(\frac{1}{2}, \frac{5}{9}]$	$(\frac{5}{9}, \frac{4}{7}]$	$(\frac{4}{7}, \frac{3}{5}]$	$(\frac{3}{5}, \frac{2}{3}]$	$(\frac{2}{3}, \frac{3}{4}]$	$(\frac{3}{4}, 1)$
$E$	2776	12445	12317	11237	6241	2595	2483
$NSCE$	24992	112013	110861	101141	56177	23363	22355

Table 2: Table shows the dimension of the extending system as well as the total dimension of the NSCE, of the behaviours lying on the isotropic line.

The dimension of the extending system as well as the total dimension of the NSCE, as given in Eq. (89), for various values of  $\eta \in (\frac{1}{2}, 1)$  has been enlisted in Table 2. We have observed that the dimension of the NSCE is a maximum when the behaviour is in the vicinity of the maximally mixed behaviour. The system size is considerably lower in the range of our interest i.e., when  $\eta \in (\frac{3}{4}, 1)$ , that is, when the behaviour  $B(\eta)$  starts showing non-classical features.

### B.4.3 Minimal ensembles for non-local isotropic behaviours

Here we present the minimal ensembles of the Bell-Tsirelson box, which lies on the isotropic line

$$B(\eta) = \eta B_{000} + (1 - \eta) B_{001}, \quad (127)$$

for  $\eta = \frac{2+\sqrt{2}}{4}$ . This behaviour lies in a polytope of dimension  $d = 8$  [102], and, according to the theorem of Carathéodory [106], the minimal ensembles of  $B(\eta)$ , consists of at most  $d+1$  pure behaviours. Hence, we present only those minimal ensembles having at most 9 elements.

We group the ensembles by the intervals of the parameter  $\eta$  for which they are valid minimal ensembles. They are minimal, as otherwise we would find their subsets which would be minimal and also valid for higher  $\alpha$ . The interval of the parameter  $\alpha$  is the due to the requirement that the  $B_{000}$  coefficient must be greater or equal 0.

We also group the ensembles by the vector of coefficients corresponding to the behaviours of the decomposition which summed up give the  $B(\alpha)$ .

1. Having  $\eta \in (0, 1)$ .

$$v = [\eta, 1 - \eta]$$

$$\mathcal{M}_1 = [B_{000}, B_{001}]$$

2. Having  $\eta \in (\frac{1}{4}, 1)$ .

$$v = \left[ \frac{4\eta}{3} - \frac{1}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\mathcal{M}_2 = [B_{000}, L_{0001}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1010}, L_{1100}, L_{1111}]$$

3. Having  $\eta \in (\frac{1}{3}, 1)$ .

$$v = \left[ \frac{3\eta}{2} - \frac{1}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_3 &= [\mathbf{B}_{000}, L_{0000}, L_{0001}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_4 &= [\mathbf{B}_{000}, L_{0001}, L_{0011}, L_{0100}, L_{0101}, L_{1010}, L_{1100}] \\
\mathcal{M}_5 &= [\mathbf{B}_{000}, L_{0001}, L_{0011}, L_{0110}, L_{0111}, L_{1010}, L_{1100}] \\
\mathcal{M}_6 &= [\mathbf{B}_{000}, L_{0001}, L_{0110}, L_{1000}, L_{1001}, L_{1100}, L_{1111}] \\
\mathcal{M}_7 &= [\mathbf{B}_{000}, L_{0001}, L_{0110}, L_{1010}, L_{1011}, L_{1100}, L_{1111}] \\
\mathcal{M}_8 &= [\mathbf{B}_{000}, L_{0010}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_9 &= [\mathbf{B}_{000}, L_{0011}, L_{0100}, L_{1001}, L_{1010}, L_{1100}, L_{1101}] \\
\mathcal{M}_{10} &= [\mathbf{B}_{000}, L_{0011}, L_{0100}, L_{1001}, L_{1010}, L_{1110}, L_{1111}]
\end{aligned}$$

$$v = \left[ \frac{3\eta}{2} - \frac{1}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_{11} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, L_{0001}, L_{0011}, L_{0100}, L_{0110}] \\
\mathcal{M}_{12} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{1001}, L_{1010}, L_{1100}, L_{1111}] \\
\mathcal{M}_{13} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, L_{0001}, L_{0100}, L_{1001}, L_{1100}] \\
\mathcal{M}_{14} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, L_{0011}, L_{0110}, L_{1010}, L_{1111}] \\
\mathcal{M}_{15} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0011}, L_{0110}, L_{1001}, L_{1100}] \\
\mathcal{M}_{16} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0100}, L_{1010}, L_{1111}]
\end{aligned}$$

4. Having  $\eta \in \left(\frac{2}{5}, 1\right)$ .

$$v = \left[ \frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\begin{aligned}
\mathcal{M}_{17} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0001}, L_{0011}, L_{0101}, L_{0110}, L_{1010}, L_{1011}, L_{1100}] \\
\mathcal{M}_{18} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0010}, L_{0100}, L_{0110}, L_{1000}, L_{1001}, L_{1111}]
\end{aligned}$$

$$v = \left[ \frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\begin{aligned}
\mathcal{M}_{19} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, L_{0000}, L_{0011}, L_{0100}, L_{0110}, L_{1001}, L_{1110}, L_{1111}] \\
\mathcal{M}_{20} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, L_{0001}, L_{0011}, L_{0100}, L_{0111}, L_{1010}, L_{1100}, L_{1101}]
\end{aligned}$$

$$v = \left[ \frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\begin{aligned}
\mathcal{M}_{21} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0000}, L_{0010}, L_{0100}, L_{0110}, L_{1001}, L_{1111}] \\
\mathcal{M}_{22} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0001}, L_{0011}, L_{0101}, L_{0111}, L_{1010}, L_{1100}]
\end{aligned}$$

$$v = \left[ \frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\mathcal{M}_{23} = [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0010}, L_{0110}, L_{1000}, L_{1001}, L_{1100}, L_{1111}]$$

$$v = \left[ \frac{5\eta}{3} - \frac{2}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6} \right]$$

$$\mathcal{M}_{24} = [\mathbf{B}_{000}, \mathbf{B}_{011}, L_{0011}, L_{0111}, L_{1001}, L_{1010}, L_{1100}, L_{1101}]$$







$$\begin{aligned}
\mathcal{M}_{73} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0100}] \\
\mathcal{M}_{74} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0110}] \\
\mathcal{M}_{75} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{1001}, \mathbf{L}_{1100}] \\
\mathcal{M}_{76} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{1010}, \mathbf{L}_{1111}]
\end{aligned}$$

5. Having  $\eta \in \left(\frac{3}{7}, 1\right)$ .

$$\begin{aligned}
v &= \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8} \right] \\
\mathcal{M}_{77} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{3}{8} - \frac{3\eta}{8} \right]$$

$$\begin{aligned}
\mathcal{M}_{78} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0011}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1100}] \\
\mathcal{M}_{79} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1111}]
\end{aligned}$$

$$\begin{aligned}
v &= \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4} \right] \\
\mathcal{M}_{80} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8} \right]$$

$$\begin{aligned}
\mathcal{M}_{81} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1100}, \mathbf{L}_{1101}] \\
\mathcal{M}_{82} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_{83} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1111}] \\
\mathcal{M}_{84} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1010}, \mathbf{L}_{1101}]
\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8} \right]$$

$$\begin{aligned}
\mathcal{M}_{85} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\
\mathcal{M}_{86} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8} \right]$$

$$\begin{aligned}
\mathcal{M}_{87} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1111}] \\
\mathcal{M}_{88} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1100}]
\end{aligned}$$



$$\begin{aligned}\mathcal{M}_{105} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1011}, \mathbf{L}_{1100}] \\ \mathcal{M}_{106} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1001}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]\end{aligned}$$

$$v = \left[ \frac{7\eta}{4} - \frac{3}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{3}{8} - \frac{3\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{8} - \frac{\eta}{8}, \frac{1}{8} - \frac{\eta}{8} \right]$$

$$\begin{aligned}\mathcal{M}_{107} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1101}, \mathbf{L}_{1110}] \\ \mathcal{M}_{108} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1100}, \mathbf{L}_{1111}]\end{aligned}$$

6. Having  $\eta \in \left(\frac{4}{9}, 1\right)$ .

$$v = \left[ \frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5} \right]$$

$$\begin{aligned}\mathcal{M}_{109} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1100}] \\ \mathcal{M}_{110} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1111}]\end{aligned}$$

$$v = \left[ \frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5} \right]$$

$$\begin{aligned}\mathcal{M}_{111} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1101}] \\ \mathcal{M}_{112} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1111}] \\ \mathcal{M}_{113} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1110}, \mathbf{L}_{1111}] \\ \mathcal{M}_{114} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]\end{aligned}$$

$$v = \left[ \frac{9\eta}{5} - \frac{4}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{2}{5} - \frac{2\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5}, \frac{1}{5} - \frac{\eta}{5} \right]$$

$$\begin{aligned}\mathcal{M}_{115} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1100}] \\ \mathcal{M}_{116} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]\end{aligned}$$

7. Having  $\eta \in \left(\frac{1}{2}, 1\right)$ .

$$v = \left[ 2\eta - 1, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\begin{aligned}\mathcal{M}_{117} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1110}] \\ \mathcal{M}_{118} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}] \\ \mathcal{M}_{119} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1101}] \\ \mathcal{M}_{120} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1111}]\end{aligned}$$

$$v = \left[ 2\eta - 1, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$







$$\begin{aligned}
\mathcal{M}_{205} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1101}] \\
\mathcal{M}_{206} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1111}] \\
\mathcal{M}_{207} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1111}] \\
\mathcal{M}_{208} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1010}, \mathbf{L}_{1100}, \mathbf{L}_{1101}] \\
\mathcal{M}_{209} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1010}, \mathbf{L}_{1110}, \mathbf{L}_{1111}] \\
\mathcal{M}_{210} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1101}] \\
\mathcal{M}_{211} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1101}, \mathbf{L}_{1111}] \\
\mathcal{M}_{212} &= [\mathbf{B}_{000}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ 2\eta - 1, \quad \frac{1}{2} - \frac{\eta}{2}, \quad \frac{1}{2} - \frac{\eta}{2}, \quad \frac{1}{4} - \frac{\eta}{4}, \quad \frac{1}{4} - \frac{\eta}{4}, \quad \frac{1}{4} - \frac{\eta}{4}, \quad \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\begin{aligned}
\mathcal{M}_{213} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0001}, \mathbf{L}_{0100}, \mathbf{L}_{1011}, \mathbf{L}_{1110}] \\
\mathcal{M}_{214} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{L}_{0011}, \mathbf{L}_{0110}, \mathbf{L}_{1000}, \mathbf{L}_{1101}] \\
\mathcal{M}_{215} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1110}] \\
\mathcal{M}_{216} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0100}, \mathbf{L}_{1000}, \mathbf{L}_{1101}] \\
\mathcal{M}_{217} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{L}_{0010}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1100}] \\
\mathcal{M}_{218} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{1010}, \mathbf{L}_{1111}] \\
\mathcal{M}_{219} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0101}, \mathbf{L}_{1001}, \mathbf{L}_{1100}] \\
\mathcal{M}_{220} &= [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1111}] \\
\mathcal{M}_{221} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{1001}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1110}] \\
\mathcal{M}_{222} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0111}] \\
\mathcal{M}_{223} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0101}, \mathbf{L}_{0110}] \\
\mathcal{M}_{224} &= [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{1000}, \mathbf{L}_{1010}, \mathbf{L}_{1101}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = \left[ 2\eta - 1, \quad \frac{1}{2} - \frac{\eta}{2}, \quad \frac{1}{2} - \frac{\eta}{2}, \quad \frac{1}{2} - \frac{\eta}{2}, \quad \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{225} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0100}, \mathbf{L}_{0101}] \\
\mathcal{M}_{226} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0110}, \mathbf{L}_{0111}] \\
\mathcal{M}_{227} &= [\mathbf{B}_{000}, \mathbf{L}_{0000}, \mathbf{L}_{0100}, \mathbf{L}_{1001}, \mathbf{L}_{1101}] \\
\mathcal{M}_{228} &= [\mathbf{B}_{000}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1100}] \\
\mathcal{M}_{229} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0100}, \mathbf{L}_{0101}] \\
\mathcal{M}_{230} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0110}, \mathbf{L}_{0111}] \\
\mathcal{M}_{231} &= [\mathbf{B}_{000}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1111}] \\
\mathcal{M}_{232} &= [\mathbf{B}_{000}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1110}] \\
\mathcal{M}_{233} &= [\mathbf{B}_{000}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1100}, \mathbf{L}_{1101}] \\
\mathcal{M}_{234} &= [\mathbf{B}_{000}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}, \mathbf{L}_{1111}] \\
\mathcal{M}_{235} &= [\mathbf{B}_{000}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}] \\
\mathcal{M}_{236} &= [\mathbf{B}_{000}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]
\end{aligned}$$

$$v = [2\eta - 1, \quad 1 - \eta, \quad 1 - \eta]$$

$$\begin{aligned}
\mathcal{M}_{237} &= [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{011}] \\
\mathcal{M}_{238} &= [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{101}] \\
\mathcal{M}_{239} &= [\mathbf{B}_{000}, \mathbf{B}_{110}, \mathbf{B}_{111}]
\end{aligned}$$

8. Having  $\eta \in \left(\frac{5}{9}, 1\right)$ .

$$v = \left[ \frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{240} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{241} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0100}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}]$$

$$v = \left[ \frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\mathcal{M}_{242} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{243} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{244} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0001}, \mathbf{L}_{0111}, \mathbf{L}_{1010}, \mathbf{L}_{1110}]$$

$$\mathcal{M}_{245} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0010}, \mathbf{L}_{0011}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1100}]$$

$$v = \left[ \frac{9\eta}{4} - \frac{5}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4}, \frac{1}{4} - \frac{\eta}{4} \right]$$

$$\mathcal{M}_{246} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{100}, \mathbf{B}_{111}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{0101}, \mathbf{L}_{1011}, \mathbf{L}_{1111}]$$

$$\mathcal{M}_{247} = [\mathbf{B}_{000}, \mathbf{B}_{011}, \mathbf{B}_{101}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0110}, \mathbf{L}_{0111}, \mathbf{L}_{1001}, \mathbf{L}_{1101}]$$

9. Having  $\eta \in \left(\frac{4}{7}, 1\right)$ .

$$v = \left[ \frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{248} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{111}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0100}, \mathbf{L}_{1010}, \mathbf{L}_{1011}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{249} = [\mathbf{B}_{000}, \mathbf{B}_{100}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$v = \left[ \frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{250} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{100}, \mathbf{L}_{0001}, \mathbf{L}_{0011}, \mathbf{L}_{0111}, \mathbf{L}_{1000}, \mathbf{L}_{1001}, \mathbf{L}_{1110}]$$

$$v = \left[ \frac{7\eta}{3} - \frac{4}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{6} - \frac{\eta}{6}, \frac{1}{3} - \frac{\eta}{3} \right]$$

$$\mathcal{M}_{251} = [\mathbf{B}_{000}, \mathbf{B}_{010}, \mathbf{B}_{110}, \mathbf{L}_{0000}, \mathbf{L}_{0010}, \mathbf{L}_{0110}, \mathbf{L}_{1011}, \mathbf{L}_{1100}, \mathbf{L}_{1101}]$$

$$\mathcal{M}_{252} = [\mathbf{B}_{000}, \mathbf{B}_{101}, \mathbf{B}_{111}, \mathbf{L}_{0001}, \mathbf{L}_{0010}, \mathbf{L}_{0101}, \mathbf{L}_{1000}, \mathbf{L}_{1110}, \mathbf{L}_{1111}]$$











$$\begin{aligned}
\mathcal{M}_{328} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{100}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1011}, \text{L}_{1100}] \\
\mathcal{M}_{329} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{101}, \text{L}_{0000}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1010}, \text{L}_{1101}] \\
\mathcal{M}_{330} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{110}, \text{L}_{0000}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1001}, \text{L}_{1110}] \\
\mathcal{M}_{331} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{111}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1111}] \\
\mathcal{M}_{332} &= [\text{B}_{000}, \text{B}_{011}, \text{L}_{0000}, \text{L}_{0111}, \text{L}_{1001}, \text{L}_{1010}, \text{L}_{1101}, \text{L}_{1110}] \\
\mathcal{M}_{333} &= [\text{B}_{000}, \text{B}_{011}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{1000}, \text{L}_{1011}, \text{L}_{1100}, \text{L}_{1111}] \\
\mathcal{M}_{334} &= [\text{B}_{000}, \text{B}_{100}, \text{B}_{111}, \text{L}_{0010}, \text{L}_{0100}, \text{L}_{0111}, \text{L}_{1011}, \text{L}_{1101}] \\
\mathcal{M}_{335} &= [\text{B}_{000}, \text{B}_{101}, \text{B}_{110}, \text{L}_{0000}, \text{L}_{0101}, \text{L}_{0110}, \text{L}_{1011}, \text{L}_{1101}]
\end{aligned}$$

$$v = \left[ \frac{5\eta}{2} - \frac{3}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{336} &= [\text{B}_{000}, \text{B}_{010}, \text{B}_{100}, \text{B}_{110}, \text{L}_{1011}, \text{L}_{1110}] \\
\mathcal{M}_{337} &= [\text{B}_{000}, \text{B}_{010}, \text{B}_{101}, \text{B}_{111}, \text{L}_{1000}, \text{L}_{1101}] \\
\mathcal{M}_{338} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{100}, \text{B}_{111}, \text{L}_{0010}, \text{L}_{0111}] \\
\mathcal{M}_{339} &= [\text{B}_{000}, \text{B}_{011}, \text{B}_{101}, \text{B}_{110}, \text{L}_{0000}, \text{L}_{0101}]
\end{aligned}$$

11. Having  $\eta \in \left(\frac{2}{3}, 1\right)$ .

$$v = \left[ 3\eta - 2, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{340} &= [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0001}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1110}] \\
\mathcal{M}_{341} &= [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0010}, \text{L}_{0100}, \text{L}_{0101}, \text{L}_{1011}, \text{L}_{1101}] \\
\mathcal{M}_{342} &= [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0010}, \text{L}_{0110}, \text{L}_{0111}, \text{L}_{1011}, \text{L}_{1101}] \\
\mathcal{M}_{343} &= [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1001}, \text{L}_{1101}, \text{L}_{1110}] \\
\mathcal{M}_{344} &= [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0111}, \text{L}_{1010}, \text{L}_{1011}, \text{L}_{1101}, \text{L}_{1110}] \\
\mathcal{M}_{345} &= [\text{B}_{000}, \text{L}_{0010}, \text{L}_{0011}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1110}] \\
\mathcal{M}_{346} &= [\text{B}_{000}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{1000}, \text{L}_{1011}, \text{L}_{1100}, \text{L}_{1101}] \\
\mathcal{M}_{347} &= [\text{B}_{000}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{1000}, \text{L}_{1011}, \text{L}_{1110}, \text{L}_{1111}]
\end{aligned}$$

$$v = \left[ 3\eta - 2, 1 - \eta, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\begin{aligned}
\mathcal{M}_{348} &= [\text{B}_{000}, \text{B}_{010}, \text{L}_{1000}, \text{L}_{1011}, \text{L}_{1101}, \text{L}_{1110}] \\
\mathcal{M}_{349} &= [\text{B}_{000}, \text{B}_{011}, \text{L}_{0000}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{0111}] \\
\mathcal{M}_{350} &= [\text{B}_{000}, \text{B}_{100}, \text{L}_{0010}, \text{L}_{0111}, \text{L}_{1011}, \text{L}_{1110}] \\
\mathcal{M}_{351} &= [\text{B}_{000}, \text{B}_{101}, \text{L}_{0000}, \text{L}_{0101}, \text{L}_{1000}, \text{L}_{1101}] \\
\mathcal{M}_{352} &= [\text{B}_{000}, \text{B}_{110}, \text{L}_{0000}, \text{L}_{0101}, \text{L}_{1011}, \text{L}_{1110}] \\
\mathcal{M}_{353} &= [\text{B}_{000}, \text{B}_{111}, \text{L}_{0010}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1101}]
\end{aligned}$$

12. Having  $\eta \in \left(\frac{3}{4}, 1\right)$ .

$$v = \left[ 4\eta - 3, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2}, \frac{1}{2} - \frac{\eta}{2} \right]$$

$$\mathcal{M}_{354} = [\text{B}_{000}, \text{L}_{0000}, \text{L}_{0010}, \text{L}_{0101}, \text{L}_{0111}, \text{L}_{1000}, \text{L}_{1011}, \text{L}_{1101}, \text{L}_{1110}]$$

## List of Symbols and abbreviations:

PP: Purification postulate.

CEP: Complete extension postulate.

A: Access property.

G: Generation property.

EU: Essential uniqueness.

ONSEA: Overcomplete non-signaling extension with access.

NSEA: Complete non-signaling extension with access.

NSCE: Non-signaling complete extension.

GPT: Generalised probabilistic theory.

$\mathcal{G}$ : A generalized probabilistic theory.

$\text{Syst}[\mathcal{G}]$ : Systems associated with  $\mathcal{G}$ .

$\otimes$ : Composition rule in a GPT.

$\times$ : Cartesian product.

$V_A$ : A finite dimensional real vector space associated to system  $A$ .

$V_A^*$ : Dual space of  $V_A$ .

$\mathbb{R}^\Lambda$ : Vector space of real valued functions for some finite sample space  $\Lambda$  in classical theory.

$\Omega_A$ : Convex set within the vector space  $V_A$ , represents the state space of system  $A$ .

$\Omega_A^*$ : subset of dual vectors in  $V_A^*$  which evaluate to (a subset of) the unit interval on  $\Omega_A$ .

$\mathcal{E}_A$ : Convex set within the dual space  $V_A^*$ , represents the effect space of system  $A$ .

$K_A$ : Convex cone associated with  $\Omega_A$ .

$e$ : An instance of an effect in the effect space  $\mathcal{E}_A$ .

$e(\Omega_A)$ : Image of  $e$  when its domain is restricted to  $\Omega_A$ .

$u_A$ : Unit effect on  $\Omega_A$ .

$\mathcal{T}_A^B$ : Space of transformations from system  $A$  to system  $B$ .

$M$ : Measurement, belongs to  $\mathcal{T}_A^B$ .

$\Delta_I$ : Classical systems contained in every GPT, where  $I$  denotes the set of classical (deterministic) states.

$\delta_i$ : a vertex of  $\Delta_I$ .

$\epsilon_i$ : a vertex of  $\Delta_I^*$ .

$\mathbb{1}_{\Delta_I}$ : Identity transformation for system  $\Delta_I$ .

$\omega_A$ : Arbitrary state of system  $A$ .

$\epsilon_{AB}$ : A non-signalling extension of  $\omega_A$  to system  $B$ .

$E_{\omega_A}$ : The set of all states that purify  $\omega_A$ .

$\text{Tr}_B$ : Partial trace over system  $B$ , of a composite system.

$T(T')$ : A system type in  $\text{Syst}[\mathcal{G}]$ .

$s$ : An arbitrary state in a GPT.

$s(T)$ : An arbitrary state of type  $T$ .

$\{(p_i, s_i)\}$ : An ensemble of states.

$\mathbf{Ens}[s]$ : The set of all possible ensembles for a state  $s$ .

$\mathbf{Ens}_P[s]$ : The set of all pure ensembles of  $s$ .

$\mathbf{Face}[s]$ : Face of a state  $s$ .

$\mathbf{Ext}[s]$ : The set of extensions of  $s$ .

$\mathbf{Ext}_P[s]$ : The set of all pure extensions of  $s$ .

$\mathbf{Ext}_{class}[s]$ : A class of extension of  $s$ , where the extending system is classical.

$\text{Vert}[\Delta_I]$ : The set of vertices of  $\Delta_I$ .

$\sigma^P$ : A purification of system  $s$ , belongs to  $\mathbf{Ext}_P[s]$ .

$\Sigma$ : Arbitrary extension of system  $s$ , belongs to  $\mathbf{Ext}[s]$ .

$T_{1 \rightarrow 2}$ : A reversible transformation in the extending system.

$p_A^*$ : Alice's cheating probability in integer-commitment.

$p_B^*$ : Bob's cheating probability in integer-commitment.

$\text{Tr}_B$ : Partial trace on system  $B$  of a composite quantum state  $\rho_{AB}$ .

$\text{Tr}_{\neq A_i}$ : Partial trace on all systems despite of system  $A_i$  of a composite state  $\rho_{A_1 A_2 \dots A_N}$ .

$V$ : Number of the vertices within a theory.

$\aleph_0$ : The cardinality of the set of natural numbers.

$\mathfrak{c}$ : The cardinality of the continuum.

$\mathcal{P}$ : A partition of a composite system into mutually non-signaling subsystems.

$\mathcal{S}_i^{\mathcal{P}}$ : An  $i^{\text{th}}$  system of a partition  $\mathcal{P}$ .

$\rho_A$ : Arbitrary quantum state of system  $A$ .

$|\psi_{AE}\rangle$ : A pure quantum state of the composite system A and E.

$\Theta_E$ : Quantum channel in part of system E.

$\{(\rho_i, \rho_A^i)\}$ : Ensemble of a quantum state.

$P_A$ : A behaviour of system A.

$P_{AE}$ : A composite behaviour of system A and E.

$\{p_i, P_A^i\}$ : Ensemble of the behaviour of system A.

$p_{A|X}(a|x)$ : A conditional probability distribution.

$p_{AE|XZ}(ae|xz)$ : A bipartite conditional probability distribution.

$\mathcal{X}$ : The set of all input  $\{x\}$ , of the conditional probability distribution  $p_{A|X}(a|x)$ .

$\mathcal{Z}$ : The set of all input  $\{z\}$ , of the bipartite conditional probability distribution  $p_{AE|XZ}(ae|xz)$ .

$\mathcal{A}$ : The set of all output  $\{a\}$ , of the conditional probability distribution  $p_{A|X}(a|x)$ .

$\mathcal{E}$ : The set of all output  $\{e\}$ , of the bipartite conditional probability distribution  $p_{AE|XZ}(ae|xz)$ .

$\mathcal{P}_E$ : Classical pre(post)-processing channel acting on the inputs(outputs) of system E.

$P_E$ : A behaviour which is an extremal point.

$\mathcal{E}(P)$ : An arbitrary ensemble of the behaviour  $P$ .

$V(\mathcal{E})$ : The set of members of the ensemble  $\mathcal{E}$ .

$\mathcal{E}_{pure}(P)$ : Pure members ensemble of the behaviour  $P$ .

$\mathcal{M}(P)$ : Minimal ensemble of the behaviour  $P$ .

$\mathcal{E}(P)_{AE}$ : Non-signaling Extension with Access of behaviour  $P_A$ .

$D_j$ : A dice, characterized by the probability distribution  $\tilde{p}(k|z' = j)$ .

$C_j$ : Classical post-processing channel, characterized by conditional distribution  $p_c(m|e, z' = j)$ .

$\mathcal{B}$ : A behaviour polytope.

$\dim \mathcal{B}$ : Dimension of a behaviour polytope.

$P^{PR}$ : Popescu-Rohrlich box.

$P^m$ : Maximally mixed behaviour.

$P_{A'}$ : Conjugate behaviour of behaviour  $P_A$ .

$L_{\alpha\beta\gamma\delta}$ : Local vertices of the polytope of binary input out behaviours,  $\alpha, \beta, \gamma, \delta \in \{0, 1\}$ .

$B_{rst}$ : Non-local vertices of the polytope of binary input out behaviours,  $r, s, t \in \{0, 1\}$ .

$B(\eta)$ : Arbitrary behaviour on the line joining PR box and anti-PR box.

$\beta(P)$ : Linear map on the behaviour  $P$ , computing the CHSH functional.

$\mathcal{V}(P)$ : Set of the sets of members of minimal ensembles.

$\mathcal{C}_3$ : The set of maximal contexts for the three-cycle contextuality scenario.

$\mathcal{N}_j^C$ :  $j^{\text{th}}$  non-contextual vertex of the no-disturbance polytope.

$\mathcal{C}_j$ :  $j^{\text{th}}$  contextual vertex of the no-disturbance polytope.