

Solvable Criterion for the Contextuality of any Prepare-and-Measure Scenario

Victor Gitton and Mischa P. Woods

Institute for Theoretical Physics, ETH Zürich, Switzerland

Starting from arbitrary sets of quantum states and measurements, referred to as the prepare-and-measure scenario, an operationally noncontextual ontological model of the quantum statistics associated with the prepare-and-measure scenario is constructed. The operationally noncontextual ontological model coincides with standard Spekkens noncontextual ontological models for tomographically complete scenarios, while covering the non-tomographically complete case with a new notion of a reduced space, which we motivate following the guiding principles of noncontextuality. A mathematical criterion, called *unit separability*, is formulated as the relevant classicality criterion — the name is inspired by the usual notion of quantum state separability. Using this criterion, we derive a new upper bound on the cardinality of the ontic space. Then, we recast the unit separability criterion as a (possibly infinite) set of linear constraints, from which we obtain two separate hierarchies of algorithmic tests to witness the non-classicality or certify the classicality of a scenario. Finally, we reformulate our results in the framework of generalized probabilistic theories and discuss the implications for simplex-embeddability in such theories.

Contents

1	Introduction	1
2	Operationally noncontextual ontological model	2
3	Unit separability and cardinality bounds	8
4	Algorithmic formulation, witnesses and certifiers	12
5	Connections with generalized probabilistic theories	16
6	Conclusion	17
	References	18
A	Review of convex analysis	20

B	Presentation of the operationally noncontextual ontological model	20
C	Unit separability and cardinality bounds	29
D	Alternative reduced spaces	35
E	Algorithmic formulation, witnesses and certifiers	38
F	Connections with generalized probabilistic theories	46

1 Introduction

1.1 Background: previous notions of classicality

Studying the non-classicality of quantum mechanics is a field that originated from the collective effort of the scientific community to obtain meaningful interpretations of the ontologically opaque yet undoubtedly successful theory of quantum mechanics. One of the early influential works highlighting how quantum mechanics departs significantly from classical mechanics was that of Einstein, Podolsky and Rosen [1]: there, it was brought to light that local realism, a natural notion of classicality, is in conflict with the quantum description of nature. Realism means that one posits the existence of a hidden state that should describe the actual physics behind the scenes, the *ontic* (actual) state of the system. Local realism means that the ontic state cannot be updated from a spacelike-separated spacetime region. This notion was further studied and turned into an experimentally verifiable *no-go theorem* by Bell [2]: the no-go theorem states that quantum mechanics cannot be described by a local hidden variable model. For the perspective of this manuscript, it is important to notice that this notion of classicality only applies to spacelike-separated systems, whereas a single quantum system is not eligible to be tested via the prism of local realism.

A natural generalization of local realism is that of Kochen-Specker noncontextuality, where the associated classical model is called noncontextual hidden variable model. This notion of classicality assumes that at the ontic state level, the outcome statistics of one measurement are 1) statistically independent from the outcome statistics of any other commuting measurement and 2) non-varying with respect to

changing the jointly-measured commuting measurement. This notion was formalized and shown to be inconsistent with quantum mechanics by Kochen and Specker [3]. Only commuting measurements may be tested through the prism of Kochen-Specker noncontextuality but possibly on a single quantum system, which was not the case with local realism.

The work of Spekkens [4] lead to a new notion of noncontextuality that subsumes Bell’s local hidden variables and Kochen-Specker noncontextuality. The assumption of realism is similar to that of the previously mentioned notions of classicality, but the scope of noncontextuality is more universal. The first step towards formulating an assumption of noncontextuality is to formulate a notion of operational equivalence, such as e.g. the operational equivalence of an electron spin- $\frac{1}{2}$ degree of freedom and a photon polarization degree of freedom as two implementations of a qubit. The corresponding assumption of noncontextuality is to posit that operationally equivalent procedures have an identical representation at the level of the ontic model. In [4], several no-go theorems are presented to show the incompatibility of quantum mechanics with respect to Spekkens noncontextuality. Quantum procedures may be eligible for testing their classicality with respect to Spekkens noncontextuality irrespective of the existence of commuting measurements. Furthermore, the incompatibility of quantum mechanics and Spekkens noncontextuality has known links with computational efficiency of quantum protocols, see e.g. [5, 6], which supplement the existing links between computational efficiency and violations of Kochen-Specker noncontextuality (a special case of Spekkens noncontextuality), see e.g. [7, 8].

1.2 The objective notion of classicality

The present work aims at obtaining a notion of classicality that is applicable to an arbitrary prepare-and-measure scenario and that provides an answer to the question of whether the scenario is classical or not with respect to that notion of classicality. The prepare-and-measure scenario may consist of all states and measurements allowed by quantum mechanics within a given Hilbert space, but it can also consist of strict subsets of these: this would be interesting if for instance one has an apparatus that only allows to produce certain types of states or perform certain types of measurements. Then, one could answer the question of whether this specific apparatus has a classical description or not. Alternatively, one can associate to a given quantum protocol a corresponding prepare-and-measure scenario that only features the states and measurements relevant for the protocol. For instance, the set of states of the scenario could be special types of multi-qubit states of a quantum computer that are relevant for a given algorithm. Then, assessing the classicality of the prepare-and-measure scenario asso-

ciated to the protocol is an indirect way of assessing the classicality of the protocol itself. This assessment may help identify resources that are most useful for efficient protocols.

Local realism and Kochen-Specker noncontextuality are well-motivated and widely useful notions of classicality, but they do not quite fulfill the above requirement of being applicable to arbitrary prepare-and-measure scenarios. Indeed, local realism specializes to local measurements on spacelike separated systems, and Kochen-Specker noncontextuality specializes to commuting measurements. On the other hand, the universality of Spekkens’ notion of noncontextuality makes it a promising basis for the formulation of our objective notion of classicality.

1.3 Content overview

Section 2 will formalize the quantum prepare-and-measure scenarios under consideration, motivate Spekkens noncontextual ontological models, and define the adjustments thereof that we posit in the case of non-tomographically complete scenarios — the resulting *operationally noncontextual ontological model* that we will consider is specified in definition 6. In particular, section 2.4 will discuss the motivation for our operationally noncontextual ontological models. Then, in section 3, the existence of an operationally noncontextual ontological model for a given prepare-and-measure scenario is turned into the unit separability criterion in theorem 2. An explicit example of the application of this criterion is given in section 3.3. This criterion allows one to extract theoretical properties of the operationally noncontextual ontological model, such as the ontic space cardinality bounds of theorem 3. Furthermore, an algorithmic formulation that evaluates the criterion for a given scenario is presented in section 4. In section 5, parallel independent work treating generalized probabilistic theories is discussed and connected to the content of this manuscript.

2 Operationally noncontextual ontological model

2.1 Prepare-and-measure scenarios

Let \mathcal{H} be a finite dimensional Hilbert space corresponding to the quantum system. The set of Hermitian matrices acting on \mathcal{H} is denoted $\mathcal{L}(\mathcal{H})$. $\mathcal{L}(\mathcal{H})$ has the structure of a real inner product space of dimension $\dim(\mathcal{L}(\mathcal{H})) = \dim(\mathcal{H})^2$: its inner product, often referred to as the Hilbert-Schmidt inner product, is defined by $\langle a, b \rangle_{\mathcal{L}(\mathcal{H})} := \text{Tr}_{\mathcal{H}}[ab]$ for all $a, b \in \mathcal{L}(\mathcal{H})$. The set of density matrices, i.e., positive semi-definite, trace-one hermitian matrices acting on \mathcal{H} , is denoted $\mathcal{S}(\mathcal{H})$. The set of quantum effects, i.e., positive semi-

definite matrices E acting on \mathcal{H} such that $\mathbb{1}_{\mathcal{H}} - E$ is also positive semi-definite, is denoted $\mathcal{E}(\mathcal{H})$.¹

Definition 1 (State space). *Let $\mathfrak{s} \subseteq \mathcal{S}(\mathcal{H})$ be a nonempty subset of states that is convex.*

Physically, any non-convex set S_1 of density matrices together with the possibility of taking classical probabilistic mixtures leads to a set of states S_2 that is the convex hull of S_1 , and hence S_2 is convex.

Definition 2 (Effect space). *Let $\mathfrak{e} \subseteq \mathcal{E}(\mathcal{H})$ be a subset of effects such that*

- (i) \mathfrak{e} is convex;
- (ii) $0_{\mathcal{H}}, \mathbb{1}_{\mathcal{H}} \in \mathfrak{e}$;
- (iii) if $E \in \mathfrak{e}$, then there exists a completion $\{E_k \in \mathfrak{e}\}_k$ such that $E + \sum_k E_k = \mathbb{1}_{\mathcal{H}}$.

The convexity requirement (i) for \mathfrak{e} is motivated by allowing classical probabilistic mixtures of different measurements, see appendix B.1 for an explicit example. The requirement (ii) comes from the fact that the trivial effects should always be allowed and faithfully be represented in the ontological model. The requirement (iii) reflects the fact that in any practical application, the effects in \mathfrak{e} will come from complete POVM sets. Proposition 11 will formalize the fact that including or not incoherent coarse-grainings of measurements in \mathfrak{e} does not make a difference for our purposes. Furthermore, proposition 8 will formalize the fact that distinct but operationally equivalent quantum descriptions of a prepare-and-measure scenario are equivalent as far as the operationally non-contextual ontological model is concerned.

The pair $(\mathfrak{s}, \mathfrak{e})$ is referred to as being an instance of a quantum prepare-and-measure scenario, or just a scenario for brevity. Recall the motivation for taking $(\mathfrak{s}, \mathfrak{e})$ to be strict subsets of all states and effects on the Hilbert space in section 1.2.

2.2 Reduced space

Since we are primarily concerned with quantum protocols that involve preparing a given state $\rho \in \mathfrak{s}$ and measuring it once with a complete set of effects where each effect E belongs to \mathfrak{e} , the experimental predictions of quantum mechanics for such protocols are entirely encoded in the probabilities $\langle \rho, E \rangle_{\mathcal{L}(\mathcal{H})}$ for all $\rho \in \mathfrak{s}$ and $E \in \mathfrak{e}$.² Let us introduce the *reduced space*, which is meant to capture the minimal amount of information needed to reconstruct these operational statistics. For any set $X \subseteq \mathcal{L}(\mathcal{H})$, we denote the linear span of its elements as $\text{span}(X) \subseteq \mathcal{L}(\mathcal{H})$, which

¹We will use throughout the notation where $\mathbb{1}_{\mathcal{V}}$ (resp. $0_{\mathcal{V}}$) is the identity (resp. zero operator) on any vector space \mathcal{V} .

²One can go beyond this setting by including post-measurement states in the set \mathfrak{s} which is a minimalistic way of accounting for multiple consecutive measurements.

is the minimal vector subspace that contains X . For any $a \in \mathcal{L}(\mathcal{H})$, the projection of a over any vector subspace $\mathcal{V} \subseteq \mathcal{L}(\mathcal{H})$ equipped with an orthonormal basis $\{v_i \in \mathcal{V}\}_i$ is denoted $P_{\mathcal{V}}(a) := \sum_i \langle v_i, a \rangle_{\mathcal{L}(\mathcal{H})} v_i$. The projection of a set $X \subseteq \mathcal{L}(\mathcal{H})$ over \mathcal{V} is denoted $P_{\mathcal{V}}(X) := \{P_{\mathcal{V}}(x) : x \in X\}$.

Definition 3 (Reduced space). *Let*

$$\mathcal{R} := P_{\text{span}(\mathfrak{e})}(\text{span}(\mathfrak{s})) \quad (2.1)$$

be the reduced space associated with the scenario $(\mathfrak{s}, \mathfrak{e})$. $\mathcal{R} \subseteq \mathcal{L}(\mathcal{H})$ is a vector space that we equip with the inner product inherited from $\mathcal{L}(\mathcal{H})$.

Note that $\dim(\mathcal{R}) \leq \dim(\mathcal{L}(\mathcal{H})) = \dim(\mathcal{H})^2$. The main property of the reduced space is the following. See appendix B.2 for a proof.

Proposition 4. *For all $\rho \in \mathfrak{s}$, for all $E \in \mathfrak{e}$,*

$$\langle \rho, E \rangle_{\mathcal{L}(\mathcal{H})} = \langle P_{\mathcal{R}}(\rho), P_{\mathcal{R}}(E) \rangle_{\mathcal{R}}. \quad (2.2)$$

Proposition 4 shows that we can in fact restrict the analysis of the probabilities associated to $(\mathfrak{s}, \mathfrak{e})$ to the analysis of all probabilities $\langle \bar{\rho}, \bar{E} \rangle_{\mathcal{R}}$ for all $\bar{\rho} \in P_{\mathcal{R}}(\mathfrak{s})$ and for all $\bar{E} \in P_{\mathcal{R}}(\mathfrak{e})$.

Since the reduced space will play an important role in the definition of the operationally noncontextual ontological model (section 2.3), it is useful to distinguish the case where the reduced space projection is irrelevant. The following definition formalizes this case — note that the wording “tomographically complete” could be used differently in other contexts.

Definition 5 (Tomographically complete scenarios). *A tomographically complete scenario $(\mathfrak{s}, \mathfrak{e})$ is one where it holds that*

$$\text{span}(\mathfrak{s}) = \text{span}(\mathfrak{e}) \subseteq \mathcal{L}(\mathcal{H}). \quad (2.3)$$

In such scenarios, the projection of \mathfrak{s} and \mathfrak{e} over the reduced space is trivial, so that for our purposes we may replace $P_{\mathcal{R}}$ with $\mathbb{1}_{\mathcal{L}(\mathcal{H})}$ — specifically, it holds that $P_{\mathcal{R}}|_{\text{span}(\mathfrak{s})} = \mathbb{1}_{\text{span}(\mathfrak{s})}$ and $P_{\mathcal{R}}|_{\text{span}(\mathfrak{e})} = \mathbb{1}_{\text{span}(\mathfrak{e})}$. Furthermore, definition 5 puts states and effects on an equal footing: for instance, a scenario in which $\text{span}(\mathfrak{s}) \subset \text{span}(\mathfrak{e})$ could be turned into a tomographically complete scenario by either adding some states to $\text{span}(\mathfrak{s})$ or removing some effects from $\text{span}(\mathfrak{e})$. Also, it is readily apparent that tomographic completeness is a weaker requirement than having $(\mathfrak{s}, \mathfrak{e})$ consist of all states and effects on the Hilbert space.

2.3 Definition of the operationally noncontextual ontological model

We now motivate the construction of the operationally noncontextual ontological model that we are considering for the scenario $(\mathfrak{s}, \mathfrak{e})$. This model is

largely based on the mathematical description and notation introduced in [4]. We will motivate the conditions that we impose in the non-tomographically complete case as we go along, and we will further argue in favor of these in section 2.4. While this work was in development, a similar notion of noncontextuality was considered in [9] — see section 5.2 for differences and similarities in the results.

2.3.1 Ontological model

We introduce the notion of an ontic state space, or ontic space for short, denoted Λ . An ontic state $\lambda \in \Lambda$ is meant to describe a classical state of the system, so that Λ can be thought of as a classical phase space that will be assigned to the quantum setup.

Let \mathbf{P} denote a preparation procedure, i.e., a set of operational instructions that fully specify the steps one needs to take to obtain the same preparation. The first idea of the ontological model is to associate to each preparation procedure \mathbf{P} a classical probability distribution, i.e., normalizable and nonnegative, over Λ . We refer to these probability distributions as the ontic state distributions. The ontic state distribution gives the probability $\Pr[\lambda|\mathbf{P}]$ that the system is in the ontic state λ after having been prepared by the preparation procedure \mathbf{P} .

Let \mathbf{M} be a measurement procedure with outcomes labeled by k . We denote by \mathbf{M}_k the event that the outcome k occurred when the measurement procedure \mathbf{M} was carried out. Any operational detail should be included in the specification of \mathbf{M} . In the ontological model, the measurements will be represented as classical probability distributions over the outcomes k ; but these probability distributions, referred to as the response functions, will not depend on the quantum states directly. Instead, the response functions will “read off” the value of a given ontic state λ to produce the outcome statistics. The response function is thus represented by the conditional probabilities $\Pr[\mathbf{M}_k|\lambda]$. The actual outcome statistics, given a preparation \mathbf{P} and an event \mathbf{M}_k , will be the outcome statistics $\Pr[\mathbf{M}_k|\lambda]$ averaged over the probability that the system was in the ontic state λ , which is specified by the ontic state distribution $\Pr[\lambda|\mathbf{P}]$:

$$\Pr[\mathbf{M}_k|\mathbf{P}] = \int_{\Lambda} d\lambda \Pr[\lambda|\mathbf{P}] \Pr[\mathbf{M}_k|\lambda]. \quad (2.4)$$

2.3.2 Noncontextual state representation

In complete generality, the probability $\Pr[\lambda|\mathbf{P}]$ could depend on any detail of the preparation procedure \mathbf{P} . This is not very satisfactory: we know from quantum mechanics that all possible measurement statistics are uniquely determined from the density matrix $\rho(\mathbf{P})$ associated to the preparation procedure \mathbf{P} .

The standard assumption of noncontextuality that would prevail here was introduced by Spekkens in [4].

There, it is justified that any detail of the preparation procedure \mathbf{P} which is not reflected in the density matrix $\rho(\mathbf{P})$ is part of the context. The corresponding assumption of noncontextuality is that the noncontextual ontic state distribution only depends on the density matrix $\rho(\mathbf{P})$: thus, we make the replacement

$$\Pr[\lambda|\mathbf{P}] \rightarrow \Pr[\lambda|\rho(\mathbf{P})]. \quad (2.5)$$

For example, in the case of a mixed quantum state, the ontic state distribution associated with that quantum state does not depend on which ensemble decomposition the mixed state may have originated from. Another example is the case where a mixed state originated from the partial trace of a pure entangled state on a larger Hilbert space: the ontic state distribution does not distinguish among the different purifications.

In the setup considered here, the only states available are in the set \mathfrak{s} , so that it would be reasonable to require that there exists a valid ontic state distribution $\Pr[\lambda|\rho]$ for any $\rho \in \mathfrak{s}$, without requiring anything else for the other quantum states in $\mathcal{S}(\mathcal{H}) \setminus \mathfrak{s}$. However, in the non-tomographically complete case (definition 5), we argue that this is still too permissive given that the only measurements available are those taken out of the set \mathfrak{e} . Indeed, it is clear from proposition 4 that any detail of the preparation procedure that is reflected in $\rho \in \mathfrak{s}$ but that is not reflected in the reduced density matrix $P_{\mathcal{R}}(\rho)$ will not be resolved by the available measurement resource \mathfrak{e} and is thus part of a context. Our notion of noncontextuality is that the ontic state distribution only depends on $P_{\mathcal{R}}(\rho)$, i.e., we make the further replacement

$$\Pr[\lambda|\rho] \rightarrow \Pr[\lambda|P_{\mathcal{R}}(\rho)]. \quad (2.6)$$

The motivation for this requirement will be discussed in section 2.4. The conventional label for the ontic state distribution is μ [4]: for all $\lambda \in \Lambda$, for all $\rho \in \mathfrak{s}$,

$$\mu(P_{\mathcal{R}}(\rho), \lambda) := \Pr[\lambda|P_{\mathcal{R}}(\rho)]. \quad (2.7)$$

This means that μ has the following domain:

$$\mu : P_{\mathcal{R}}(\mathfrak{s}) \times \Lambda \rightarrow \mathbb{R}. \quad (2.8a)$$

The normalization and nonnegativity of the probability distributions read

$$\forall \bar{\rho} \in P_{\mathcal{R}}(\mathfrak{s}) : \int_{\Lambda} d\lambda \mu(\bar{\rho}, \lambda) = 1, \quad (2.8b)$$

$$\forall \lambda \in \Lambda, \forall \bar{\rho} \in P_{\mathcal{R}}(\mathfrak{s}) : \mu(\bar{\rho}, \lambda) \geq 0. \quad (2.8c)$$

It is also reasonable to require that the ontic state distribution mapping represents classical probabilistic mixtures of quantum states by classical probabilistic mixtures of ontic states. This is formulated as a convex-linearity requirement of the form:

$$\begin{aligned} \forall \lambda \in \Lambda, \forall p \in [0, 1], \forall \bar{\rho}_1, \bar{\rho}_2 \in P_{\mathcal{R}}(\mathfrak{s}) : \\ \mu(p\bar{\rho}_1 + (1-p)\bar{\rho}_2, \lambda) \\ = p\mu(\bar{\rho}_1, \lambda) + (1-p)\mu(\bar{\rho}_2, \lambda). \end{aligned} \quad (2.8d)$$

2.3.3 Noncontextual measurement representation

As previously stated, the response function distribution $\Pr[\mathbf{M}_k|\lambda]$ could in principle depend on all operational details of \mathbf{M} . The notion of noncontextuality that would prevail here [4] would be that the response function distribution does not depend on more than the POVM $\{E_k(\mathbf{M}_k)\}_k$ associated to the measurement procedure \mathbf{M} . Thus, we make the replacement

$$\Pr[\mathbf{M}_k|\lambda] \rightarrow \Pr[E_k(\mathbf{M}_k)|\lambda]. \quad (2.9)$$

This is motivated by the fact that in quantum mechanics, two distinct measurement procedures which lead to the same POVM are equivalent with respect to the statistics that are produced upon measuring any state. Equation (2.9) implies that the response function does not resolve whether a POVM originated from a coarse-graining of a finer POVM; nor does it resolve whether the POVM originated from tracing out the result of a projective measurement on a larger Hilbert space.

In our setup where all available POVM elements belong to \mathbf{e} , it is reasonable to require that there exists a valid response function $\Pr[E|\lambda]$ for all $E \in \mathbf{e}$, irrespective of what is predicted for other quantum effects in $\mathcal{E}(\mathcal{H}) \setminus \mathbf{e}$. This is however too permissive: consider distinct quantum effects $E_1, E_2 \in \mathbf{e}$. Given the set \mathbf{s} and proposition 4, it could be that $P_{\mathcal{R}}(E_1) = P_{\mathcal{R}}(E_2)$ so that the effects are indistinguishable in this setup. Thus, the part of a quantum effect $E \in \mathbf{e}$ which is not reflected in $P_{\mathcal{R}}(E)$ is part of a new kind of context, and we make the further replacement

$$\Pr[E|\lambda] \rightarrow \Pr[P_{\mathcal{R}}(E)|\lambda], \quad (2.10)$$

which is, again, nontrivial for non-tomographically complete scenarios (definition 5), and will be further commented in section 2.4. The mapping that associates a response function to each quantum effect is denoted ξ , following the notation of [4]: for all $\lambda \in \Lambda$, for all $E \in \mathbf{e}$,

$$\xi(P_{\mathcal{R}}(E), \lambda) := \Pr[\lambda|P_{\mathcal{R}}(E)]. \quad (2.11)$$

The domain of ξ is then:

$$\xi : P_{\mathcal{R}}(\mathbf{e}) \times \Lambda \rightarrow \mathbb{R}. \quad (2.12a)$$

The explicit normalization and nonnegativity are imposed as follows:

$$\begin{aligned} \forall \lambda \in \Lambda, \forall K \in \mathbb{N} \cup \{+\infty\}, \\ \forall \left\{ E_k \in \mathbf{e} : \sum_{k=1}^K E_k = \mathbb{1}_{\mathcal{H}} \right\}_{k=1}^K : \\ \sum_{k=1}^K \xi(P_{\mathcal{R}}(E_k), \lambda) = 1, \end{aligned} \quad (2.12b)$$

$$\forall \lambda \in \Lambda, \forall \bar{E} \in P_{\mathcal{R}}(\mathbf{e}) : \quad \xi(\bar{E}, \lambda) \geq 0. \quad (2.12c)$$

In addition to the properties already specified, the response function mapping should represent classical

probabilistic mixtures of quantum effects as classical probabilistic mixtures of response functions:

$$\begin{aligned} \forall \lambda \in \Lambda, \forall p \in [0, 1], \forall \bar{E}_1, \bar{E}_2 \in P_{\mathcal{R}}(\mathbf{e}) : \\ \xi(p\bar{E}_1 + (1-p)\bar{E}_2, \lambda) \\ = p\xi(\bar{E}_1, \lambda) + (1-p)\xi(\bar{E}_2, \lambda). \end{aligned} \quad (2.12d)$$

We are now able to formulate our definition of operationally noncontextual ontological model.

Definition 6 (Operationally noncontextual ontological model). *The operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) is specified as follows. Let μ be the ontic state mapping that has domain (2.8a) and that satisfies (2.8b), (2.8c), and (2.8d). Let ξ be the response function mapping that has domain (2.12a) and that satisfies (2.12b), (2.12c), and (2.12d). The operationally noncontextual ontological model is required to reproduce the statistics that quantum mechanics predicts for the available states and measurements. Using proposition 4 to write down the quantum probabilities and equation (2.4) to write down the ontological probabilities, this requirement may be formulated in the reduced space \mathcal{R} as follows:*

$$\begin{aligned} \forall \bar{\rho} \in P_{\mathcal{R}}(\mathbf{s}), \forall \bar{E} \in P_{\mathcal{R}}(\mathbf{e}) : \\ \langle \bar{\rho}, \bar{E} \rangle_{\mathcal{R}} = \int_{\Lambda} d\lambda \mu(\bar{\rho}, \lambda) \xi(\bar{E}, \lambda). \end{aligned} \quad (2.13)$$

2.4 Motivation for the reduced space prescription

In the tomographically complete case (definition 5), the definition of the operationally noncontextual ontological model, definition 6, is standard in the literature [4, 9, 10]. In the non-tomographically complete case, however, the situation is different. In order to construct a noncontextual ontological model, it is crucial to identify the operational equivalences that the model should implement as ontological equivalences. In the tomographically complete case, compelled by the success and apparent operational completeness of non-relativistic quantum mechanics in its scope of application, the standard prescription is to say that any two systems with the same quantum description should be ontologically equivalent [4].

An interesting question is to inquire about the case of a system that may be entangled in two distinct joint states with an ancilla, yet whose reduced density matrix is identical in both instances of the joint state. These two states are clearly not ‘‘globally’’ operationally equivalent, but the two states restricted to the system of interest are. Hence, noncontextuality asks for the ontological description, restricted to the system of interest, to be identical in these two cases. The operation to restrict to the system of interest is the partial trace. We argue that projecting

states and effects over the reduced space (as in definition 6) is equally well motivated as taking the partial trace of states or effects for the purpose of restricting the focus to the system of interest.

The main point of tension about this claim lies in the ontology that one attributes to the notion of a “system”, whether described by a reduced space \mathcal{R} or a Hilbert space \mathcal{H} : indeed, it seems intuitive that the Hilbert space \mathcal{H} describes a “real” system, at least in some cases, while the reduced space is just a description of an effective system. For instance, an effective qubit Hilbert space that corresponds to some 2-dimensional subspace of a larger Hilbert space, say, that of a large quantum computer, is perhaps not the best example of a “real qubit system”. Instead, one has in mind, as an example of a “real qubit”, a single spin-1/2, for instance that of a free electron.³ However, we argue that there is in fact no consistent operational way to distinguish a “real system” from an “effective” one: operationally, a system can only be defined through the specification of the set of operations that an agent is considering. Of course, certain systems have been found to be more typical or easier to isolate than others, but typicality is arguably subjective and isolation is not fundamental — one can simply think of vacuum fluctuations continuously happening around a tentatively isolated system. Ultimately, the cut between “the system” and “the rest” is nothing more but a specification of the relevant operational procedures that the agents restrict themselves to consider. In fact, we give in appendix B.3 an example construction where the reduced space projection and a partial trace in a suitably chosen tensor product structure are equivalent.

Thus, using this premise that all systems are defined only effectively through the specification of the relevant agent operations, the motivation for not allowing an ontological model to depend on the system’s surroundings can be equally well applied upon defining the system through some reduced space projection or some partial trace, depending on the type of prepare-and-measure scenario under consideration. This motivation can be justified, for instance, from desiring a minimal ontological description. Alternatively, if one embeds the system (whether a reduced space or the special case of a full Hilbert space system) within a larger system, for instance by repeating the experiment in a different lab without affecting the measurement statistics, it is natural to expect the ontological description to be invariant under these “irrelevant” modifications. Of course, one may reject this motivation both for tomographically complete and non-tomographically complete scenarios, at the expense of having to have a prescription for the extent of the information upon which a *contextual* ontological model would depend. One could

³For argument’s sake, we forget about the electron’s position and momentum.

be worried that the mathematical definition of the reduced space \mathcal{R} (definition 3) bears some arbitrariness with respect to the above motivation: in section 3.5, we describe more general reduced space constructions and the equivalence between any such constructions.

Independently of the above arguments, a minimal motivation for the reduced space prescription is the following: in the general framework of [11], the user of the framework may incorporate arbitrary sets of operational equivalences in the noncontextual ontological model, as long as they are compatible with the statistics of the prepare-and-measure scenario. The reduced space prescription essentially amounts to incorporating the maximal set of operational equivalences. When viewed from this perspective, the reduced space prescription induces a strengthening of the classicality requirements compared with a certain class of “minimal” Spekkens noncontextual ontological models when the prepare-and-measure scenario is not tomographically complete, as is formalized in the following proposition. A proof is given in appendix B.2.

Proposition 7. *We let a minimal Spekkens noncontextual ontological model for (\mathbf{s}, \mathbf{e}) be defined as in definition 6 but under the replacement of*

$$P_{\mathcal{R}} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{R} (\subseteq \mathcal{L}(\mathcal{H})) \quad (2.14a)$$

with

$$\mathbb{1}_{\mathcal{L}(\mathcal{H})} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}). \quad (2.14b)$$

In particular, the ontic state mapping μ and response function mapping ξ , in this minimal Spekkens noncontextual ontological model, have as quantum arguments $\rho \in \mathbf{s}$ and $E \in \mathbf{e}$, respectively, instead of $P_{\mathcal{R}}(\rho) \in P_{\mathcal{R}}(\mathbf{s})$ and $P_{\mathcal{R}}(E) \in P_{\mathcal{R}}(\mathbf{e})$. Then, it holds that

- (i) *any (\mathbf{s}, \mathbf{e}) that admits an operationally noncontextual ontological model (definition 6) also admits a minimal Spekkens noncontextual ontological model;*
- (ii) *there exists (\mathbf{s}, \mathbf{e}) that admits a minimal Spekkens noncontextual ontological model but no operationally noncontextual ontological model (definition 6).*

Furthermore, it is also possible to prove that two scenarios (\mathbf{s}, \mathbf{e}) and $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ that produce the same operational statistics are equivalent with respect to operationally noncontextual ontological models. This proves that it does not matter whether e.g. one chooses to use a larger Hilbert space including an ancilla so that the states are purified and/or the measurements are implemented as projective measurements in accordance with Naimark’s theorem. The proof is given in appendix B.5. Note that such a statement does not necessarily hold in standard Spekkens noncontextual ontological models [12].

Proposition 8. Let \mathcal{H} and $\tilde{\mathcal{H}}$ denote two finite dimensional Hilbert spaces. Let I be a discrete or continuous range of indices. Consider two scenarios (\mathbf{s}, \mathbf{e}) and $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ that take the form

$$\mathbf{s} = \text{conv}(\{\rho_k \in \mathcal{L}(\mathcal{H})\}_{k \in I}), \quad (2.15a)$$

$$\mathbf{e} = \text{conv}(\{E_k \in \mathcal{L}(\mathcal{H})\}_{k \in I}), \quad (2.15b)$$

and

$$\tilde{\mathbf{s}} = \text{conv}(\{\tilde{\rho}_k \in \mathcal{L}(\tilde{\mathcal{H}})\}_{k \in I}), \quad (2.16a)$$

$$\tilde{\mathbf{e}} = \text{conv}(\{\tilde{E}_k \in \mathcal{L}(\tilde{\mathcal{H}})\}_{k \in I}), \quad (2.16b)$$

both assumed to satisfy definitions 1 and 2, and that satisfy, for all $k, l \in I$,

$$\text{Tr}_{\mathcal{H}}[\rho_k E_l] = \text{Tr}_{\tilde{\mathcal{H}}}[\tilde{\rho}_k \tilde{E}_l]. \quad (2.17)$$

Then, these two scenarios define reduced spaces $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{e}))$ and $\tilde{\mathcal{R}} = P_{\text{span}(\tilde{\mathbf{e}})}(\text{span}(\tilde{\mathbf{e}}))$ that have the same dimension and the scenario (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model if and only if the scenario $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ does. In fact, if one of the two scenarios admits an operationally noncontextual ontological model with ontic space Λ , so does the other with the same ontic space Λ .

We will return to the computational equivalence of working with either version of the prepare-and-measure scenario in proposition 35.

2.5 Structure of the operationally noncontextual ontological model

Let us now use the properties of the operationally noncontextual ontological model to derive basic results related to its structure which will be useful for our later endeavors. The following proposition is proven in appendix B.4, and is motivated by the analysis of the no-go theorem developed in [10].

Proposition 9. Let $\lambda \in \Lambda$ be arbitrary. Starting from the convex-linear mappings

$$\mu(\cdot, \lambda) : P_{\mathcal{R}}(\mathbf{s}) \rightarrow \mathbb{R}_{\geq 0}, \quad (2.18a)$$

$$\xi(\cdot, \lambda) : P_{\mathcal{R}}(\mathbf{e}) \rightarrow \mathbb{R}_{\geq 0}, \quad (2.18b)$$

there exist unique linear extensions

$$\mu(\cdot, \lambda) : \mathcal{R} \rightarrow \mathbb{R}, \quad (2.19a)$$

$$\xi(\cdot, \lambda) : \mathcal{R} \rightarrow \mathbb{R}. \quad (2.19b)$$

Following [10], we may apply Riesz' representation theorem, stated in appendix B.4, for any fixed $\lambda \in \Lambda$ to obtain that there exist unique $F(\lambda) \in \mathcal{R}$, $\sigma(\lambda) \in \mathcal{R}$ such that for all $\lambda \in \Lambda$, $r \in \mathcal{R}$:

$$\mu(r, \lambda) = \langle r, F(\lambda) \rangle_{\mathcal{R}} \quad (2.20a)$$

$$\xi(r, \lambda) = \langle \sigma(\lambda), r \rangle_{\mathcal{R}}. \quad (2.20b)$$

We will express the nonnegativity requirements of (2.8c) and (2.12c) using the notion of the polar convex cone.

Definition 10 (Polar convex cone). For any real inner product space \mathcal{V} of finite dimension, for any $X \subseteq \mathcal{V}$, the polar convex cone⁴ $X^{+\mathcal{V}}$ is defined as

$$X^{+\mathcal{V}} := \{y \in \mathcal{V} : \forall x \in X, \langle x, y \rangle_{\mathcal{V}} \geq 0\}. \quad (2.21)$$

We may now formulate the following theorem which links the existence of an operationally noncontextual ontological model to the existence of specific mathematical primitives. The proof is presented in appendix B.5. Such a representation is a generalization of the frame representation of quantum mechanics introduced in [13].

Theorem 1. Given (\mathbf{s}, \mathbf{e}) that lead to the reduced space \mathcal{R} (definition 3), there exists an operationally noncontextual ontological model with ontic state space Λ (definition 6) if and only if there exist mappings F , σ with ranges

$$F : \Lambda \rightarrow P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}, \quad (2.22a)$$

$$\sigma : \Lambda \rightarrow P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}, \quad (2.22b)$$

satisfying the normalization condition

$$\forall \lambda \in \Lambda : \langle \sigma(\lambda), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1 \quad (2.23)$$

as well as the consistency requirement: for all $r, s \in \mathcal{R}$,

$$\langle r, s \rangle_{\mathcal{R}} = \int_{\Lambda} d\lambda \langle r, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), s \rangle_{\mathcal{R}}. \quad (2.24)$$

This theorem will in particular prove useful to determine the unit separability criterion in the next section. For completeness, as proven in appendix B.5, we have the alternative expressions $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = \mathcal{R} \cap (\mathbf{s}^{+\mathcal{L}(\mathcal{H})})$, and $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = \mathcal{R} \cap (\mathbf{e}^{+\mathcal{L}(\mathcal{H})})$. Furthermore, equation (2.23) is equivalent to a trace condition since $\langle \sigma(\lambda), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = \text{Tr}_{\mathcal{H}}[\sigma(\lambda)]$.

We are now in a position to easily give a prescription for including or not coarse-grained effects in the set \mathbf{e} : namely, it does not make any difference. A proof is given in appendix B.5.

Proposition 11 (Incoherent coarse-grainings). Consider any prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) . Suppose that there exist $\{E_k \in \mathbf{e}\}_{k=1}^N$ where $N \in \mathbb{N} \cup \{+\infty\}$ such that $\sum_{k=1}^N E_k \leq \mathbb{1}_{\mathcal{H}}$ and the sum $\sum_{k=1}^N E_k$ may not be in \mathbf{e} . Then, the scenario (\mathbf{s}, \mathbf{e}) and the extended scenario

$$\left(\mathbf{s}, \mathbf{e}_{\text{ext}} := \text{conv}(\mathbf{e} \cup \{\sum_{k=1}^N E_k\})\right) \quad (2.25)$$

define the same reduced space $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})) = P_{\text{span}(\mathbf{e}_{\text{ext}})}(\text{span}(\mathbf{s}))$ and same polar cone $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = P_{\mathcal{R}}(\mathbf{e}_{\text{ext}})^{+\mathcal{R}}$.

⁴See appendix A for definitions of convex and conic sets.

This proves that the scenarios (\mathbf{s}, \mathbf{e}) and $(\mathbf{s}, \mathbf{e}_{\text{ext}})$ are completely equivalent as inputs to theorem 1. In particular, any operationally noncontextual ontological model for any one of these two scenarios is valid for the other too. Furthermore, one can easily apply proposition 11 recursively to account for the inclusion of multiple coarse-grained effects.

3 Unit separability and cardinality bounds

In this section, we will derive a more powerful criterion, referred to as unit separability, for the existence of an operationally noncontextual ontological model. This criterion was inspired by the no-go theorems of [10, 13].

3.1 Mathematical preliminaries

We now introduce the mathematical tools that will be needed to state our results. The main two notions are a notion of generalized separability as well as a generalized Choi-Jamiołkowski isomorphism.

3.1.1 Generalized separability

Consider the tensor product space $\mathcal{R} \otimes \mathcal{R}$ with \mathcal{R} as in definition 3. It is a real inner product vector space — its inner product is defined for product operators as follows: for all $a, b, x, y \in \mathcal{R}$,

$$\langle a \otimes b, x \otimes y \rangle_{\mathcal{R} \otimes \mathcal{R}} := \langle a, x \rangle_{\mathcal{R}} \langle b, y \rangle_{\mathcal{R}}. \quad (3.1)$$

To obtain the complete inner product, extend this expression by linearity. We now define the minimal tensor product of two sets as:

Definition 12. *Given any two sets $X, Y \subseteq \mathcal{R}$, the minimal tensor product set $X \otimes_{\text{set}} Y \subseteq \mathcal{R} \otimes \mathcal{R}$ is defined as*

$$X \otimes_{\text{set}} Y := \{x \otimes y : x \in X, y \in Y\}. \quad (3.2)$$

This allows to define the two following sets which are of primordial importance:

Definition 13 (Generalized product operators). *The set of generalized product operators is defined to be*

$$\text{Prod}(\mathbf{s}, \mathbf{e}) := P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} \otimes_{\text{set}} P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}. \quad (3.3)$$

Recall that the convex hull $\text{conv}(X)$ of a set X is the set of all convex combinations of finitely many elements of X , as defined in appendix A.

Definition 14 (Generalized separable operators). *The set of generalized separable operators is defined to be*

$$\text{Sep}(\mathbf{s}, \mathbf{e}) := \text{conv}(\text{Prod}(\mathbf{s}, \mathbf{e})). \quad (3.4)$$

Referring to the definitions introduced in appendix A, $\text{Prod}(\mathbf{s}, \mathbf{e})$ is a cone and $\text{Sep}(\mathbf{s}, \mathbf{e})$ is a convex cone. More details on the structure of $\text{Prod}(\mathbf{s}, \mathbf{e})$ and $\text{Sep}(\mathbf{s}, \mathbf{e})$ are presented in appendix C.1.

3.1.2 Choi-Jamiołkowski isomorphism

We will make use of a simple generalization of the Choi-Jamiołkowski isomorphism [14]. Let $L(\mathcal{R})$ be the space of linear maps from \mathcal{R} to \mathcal{R} . The Choi-Jamiołkowski isomorphism maps each linear map in $L(\mathcal{R})$ to an element of $\mathcal{R} \otimes \mathcal{R}$.

Definition 15. *For any $\Phi \in L(\mathcal{R})$, let $\mathbb{J}(\Phi) \in \mathcal{R} \otimes \mathcal{R}$ be the Choi-Jamiołkowski operator defined uniquely by the relations*

$$\forall r, s \in \mathcal{R} : \langle r, \Phi(s) \rangle_{\mathcal{R}} = \langle \mathbb{J}(\Phi), r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}}. \quad (3.5)$$

The proof of uniqueness, of bijectivity and explicit coordinate solutions are derived in appendix C.2. The following lemma is also proven in appendix C.2:

Lemma 16. *For any orthonormal basis of \mathcal{R} $\{R_i \in \mathcal{R}\}_{i=1}^{\dim(\mathcal{R})}$:*

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^{\dim(\mathcal{R})} R_i \otimes R_i. \quad (3.6)$$

3.2 The unit separability criterion

Starting from theorem 1, we may now derive an alternative criterion for the existence of an operationally noncontextual ontological model. First, we make an assumption for the types of ontic spaces that we are considering.

Definition 17 (Riemann integrability). *An operationally noncontextual ontological model with ontic space Λ as introduced in definition 6 is Riemann integrable if and only if there exist*

$$\Delta : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}_{\geq 0}, \quad (3.7a)$$

$$\lambda^{(\text{dis})} : \mathbb{N} \times \mathbb{N} \rightarrow \Lambda, \quad (3.7b)$$

such that

$$\begin{aligned} \forall \bar{\rho} \in P_{\mathcal{R}}(\mathbf{s}), \forall \bar{E} \in P_{\mathcal{R}}(\mathbf{e}) : & \int_{\Lambda} d\lambda \mu(\bar{\rho}, \lambda) \xi(\bar{E}, \lambda) \\ & = \lim_{N \rightarrow \infty} \sum_{k=1}^N \Delta_{N,k} \mu(\bar{\rho}, \lambda_{N,k}^{(\text{dis})}) \xi(\bar{E}, \lambda_{N,k}^{(\text{dis})}). \end{aligned} \quad (3.8)$$

N can be thought of as being a number of subsets that form a discrete partition of Λ , while k is a discrete index running over all such subsets and $\lambda^{(\text{dis})} \in \Lambda$ is a value in that subset of Λ . Riemann integrable operationally noncontextual ontological models include in particular:

- (i) operationally noncontextual ontological models equipped with discrete, finite ontic spaces Λ , which means that Λ is isomorphic to $\{1, \dots, N\}$ for some $N \in \mathbb{N}$;
- (ii) operationally noncontextual ontological models equipped with discrete, countable infinite ontic spaces Λ , which means that Λ is isomorphic to \mathbb{N} ;
- (iii) operationally noncontextual ontological models equipped with a continuous ontic space Λ isomorphic to \mathbb{R}^d for some $d \in \mathbb{N}$ such that for all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$, the real function $\mu(\bar{\rho}, \cdot)\xi(\bar{E}, \cdot) : \mathbb{R}^d \rightarrow \mathbb{R}_{\geq 0}$ is Riemann integrable. Such operationally noncontextual ontological models are reasonable physically because they may be seen as describing a system with finitely many continuous degrees of freedom such as position and momentum of finitely many particles.

Theorem 2 (Main theorem: unit separability). *The prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model (definition 17) if and only if:*

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \mathbf{Sep}(\mathbf{s}, \mathbf{e}). \quad (3.9)$$

This formulation is useful because it allows one to derive properties of the operationally noncontextual ontological model when it exists: the main theoretical application is described in section 3.4 where the cardinality of the ontic space $|\Lambda|$ is shown to be constrained by the dimension of the reduced space \mathcal{R} . Note that $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ is easy to compute, whereas $\mathbf{Sep}(\mathbf{s}, \mathbf{e})$ is harder to characterize. Still, well-known algorithmic results from convex analysis make the unit separability criterion decidable as described in section 4.

Proof overview. The complete proof is given in appendix C.3. In one direction, the goal is to show that if there exists an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) , then the ontic mappings F and σ from theorem 1 satisfy:

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \int_{\Lambda} d\lambda F(\lambda) \otimes \sigma(\lambda). \quad (3.10)$$

The assumption of Riemann integrability allows to prove that (3.10) implies $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \mathbf{Sep}(\mathbf{s}, \mathbf{e})$.

For the other direction, the idea is to show that if $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \mathbf{Sep}(\mathbf{s}, \mathbf{e})$ holds, then there exists a decomposition of the form

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^n F_i \otimes \sigma_i \quad (3.11)$$

for $F_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ and $\sigma_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ which yields a valid Riemann integrable model allowing to compute

quantum statistics as follows: for all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$,

$$\langle \bar{\rho}, \bar{E} \rangle_{\mathcal{R}} = \sum_{i=1}^n \langle \bar{\rho}, F_i \rangle_{\mathcal{R}} \langle \sigma_i, \bar{E} \rangle_{\mathcal{R}}. \quad \square \quad (3.12)$$

3.3 Explicit examples

3.3.1 Operationally noncontextual example

It is useful to pause and consider an explicit example scenario, which is taken to be the stabilizer rebit scenario as presented in [9]. We will return to the general case of solving the unit separability criterion in section 4. The Hilbert space is that of a qubit, $\mathcal{H} \simeq \mathbb{C}^2$, and we define

$$X = \{ |0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -| \}, \quad (3.13)$$

where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The state and effect spaces are

$$\mathbf{s} := \text{conv}(X), \quad (3.14a)$$

$$\mathbf{e} := \text{conv}(\{0_{\mathcal{H}}, \mathbb{1}_{\mathcal{H}}\} \cup X), \quad (3.14b)$$

Since $\text{span}(\mathbf{s}) = \text{span}(\mathbf{e})$ in this case, we have that $\mathcal{R} = \text{span}(\mathbf{s})$, and hence it is sufficient to find an orthonormal basis of $\text{span}(\mathbf{s})$, which can be taken to be made out of the following three elements:

$$R_1 := |0\rangle\langle 0|, \quad (3.15a)$$

$$R_2 := |1\rangle\langle 1|, \quad (3.15b)$$

$$R_3 := \frac{1}{\sqrt{2}}(2|+\rangle\langle +| - |0\rangle\langle 0| - |1\rangle\langle 1|). \quad (3.15c)$$

These satisfy $\langle R_i, R_j \rangle_{\mathcal{R}} \equiv \text{Tr}_{\mathcal{H}}[R_i R_j] = \delta_{ij}$ for $i, j = 1, 2, 3$. We can now express $P_{\mathcal{R}}(\mathbf{s})$ and $P_{\mathcal{R}}(\mathbf{e})$ as convex hulls of three dimensional vectors of the form $(a, b, c)^T \equiv aR_1 + bR_2 + cR_3 \in \mathcal{R}$, through the definition of the set \bar{X} composed of the four elements

$$P_{\mathcal{R}}(|0\rangle\langle 0|) = (1, 0, 0)^T, \quad (3.16a)$$

$$P_{\mathcal{R}}(|1\rangle\langle 1|) = (0, 1, 0)^T, \quad (3.16b)$$

$$P_{\mathcal{R}}(|+\rangle\langle +|) = \frac{1}{2}(1, 1, \sqrt{2})^T, \quad (3.16c)$$

$$P_{\mathcal{R}}(|-\rangle\langle -|) = \frac{1}{2}(1, 1, -\sqrt{2})^T. \quad (3.16d)$$

We then have

$$P_{\mathcal{R}}(\mathbf{s}) := \text{conv}(\bar{X}), \quad (3.17a)$$

$$P_{\mathcal{R}}(\mathbf{e}) := \text{conv}(\{P_{\mathcal{R}}(0_{\mathcal{H}}) = (0, 0, 0)^T, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) = (1, 1, 0)^T\} \cup \bar{X}). \quad (3.17b)$$

It can be verified that the polar cones are equal to the same convex cone,⁵

$$P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = \text{conv} \circ \text{coni}(\bar{Y}) \quad (3.18)$$

⁵See definition A.5 for the definition of the conic hull, denoted $\text{coni}(\cdot)$.

where $\bar{Y} = \{y_1, y_2, y_3, y_4\}$ with

$$y_1 = (\sqrt{2}, 0, 1)^T, \quad (3.19a)$$

$$y_2 = (\sqrt{2}, 0, -1)^T, \quad (3.19b)$$

$$y_3 = (0, \sqrt{2}, 1)^T, \quad (3.19c)$$

$$y_4 = (0, \sqrt{2}, -1)^T. \quad (3.19d)$$

This allows to write (see proposition 30 for more details)

$$\text{Sep}(\mathbf{s}, \mathbf{e}) = \text{conv} \circ \text{coni}(\{y_i \otimes y_j\}_{i,j=1}^4). \quad (3.20)$$

The Choi state of the identity is

$$\begin{aligned} \mathbb{J}(\mathbb{1}_{\mathcal{R}}) &= (1, 0, 0, 0, 1, 0, 0, 0, 1)^T \\ &\equiv \sum_{i=1}^3 R_i \otimes R_i. \end{aligned} \quad (3.21)$$

The unit separability criterion is then verified, i.e., $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e})$, since it can be verified explicitly that

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^4 \frac{1}{4} y_i \otimes y_i \in \text{Sep}(\mathbf{s}, \mathbf{e}). \quad (3.22)$$

This proves that this scenario admits an operationally noncontextual ontological model, in agreement to the existing literature [9]. Furthermore, from the explicit decomposition of (3.22) over four elements of $\text{Prod}(\mathbf{s}, \mathbf{e})$, one obtains the explicit operators $\{F_i = y_i/(2\sqrt{2})\}_{i=1}^4$ and $\{\sigma_i = y_i/\sqrt{2}\}_{i=1}^4$ of theorem 1 giving rise to ontic distributions over $\Lambda \simeq \{1, 2, 3, 4\}$. For instance, for the $|0\rangle\langle 0|$ state,

$$\begin{aligned} \mu(P_{\mathcal{R}}(|0\rangle\langle 0|), \lambda = 1) &= \langle F_{\lambda=1}, P_{\mathcal{R}}(|0\rangle\langle 0|) \rangle_{\mathcal{R}} \\ &= \begin{pmatrix} \frac{1}{2} \\ 0 \\ \frac{1}{2\sqrt{2}} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \frac{1}{2}. \end{aligned} \quad (3.23)$$

Similarly, one obtains the other components so that $\mu(P_{\mathcal{R}}(|0\rangle\langle 0|)) \equiv (1/2, 1/2, 0, 0)$, in agreement with [9].

3.3.2 Operationally contextual example

The simplest class of scenarios which do not admit an operationally noncontextual ontological model are those for which \mathbf{s} and \mathbf{e} are all states and effects allowed within a given finite-dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^d$. It is clear that $\text{span}(\mathbf{s}) = \text{span}(\mathbf{e}) = \mathcal{L}(\mathcal{H})$ in this case, so that $\mathcal{R} = \mathcal{L}(\mathcal{H})$. The conical hulls of \mathbf{s} and \mathbf{e} are both equal to the set $\mathcal{P}(\mathcal{H})$ of all positive semi-definite matrices within $\mathcal{L}(\mathcal{H})$. The set of semi-definite matrices, as a convex cone, is “self-dual” in the sense that its polar cone $\mathcal{P}(\mathcal{H})^{+\mathcal{L}(\mathcal{H})}$ equals $\mathcal{P}(\mathcal{H})$. Thus, we obtain that

$$P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = \mathcal{P}(\mathcal{H}), \quad (3.24)$$

and it follows that

$$\text{Sep}(\mathbf{s}, \mathbf{e}) = \text{conv}(\mathcal{P}(\mathcal{H}) \otimes_{\text{set}} \mathcal{P}(\mathcal{H})) = \mathbf{Q.Sep}, \quad (3.25)$$

where $\mathbf{Q.Sep}$ is simply the usual set of unnormalized separable bipartite quantum states. The Choi state of the identity, given $\mathcal{R} = \mathcal{L}(\mathcal{H})$, can be written as follows. First, we choose an arbitrary basis of $\mathcal{H} \simeq \mathbb{C}^d$ that we write as $\{|j\rangle\}_{j=1}^d$. Then, the orthonormal basis of $\mathcal{R} = \mathcal{L}(\mathcal{H})$ is taken to be the union of elements of the form $\{|j\rangle\langle j|\}_j$, $\{(|j\rangle\langle k| + |k\rangle\langle j|)/\sqrt{2}\}_{j \neq k}$ and $\{(i|j\rangle\langle k| - i|k\rangle\langle j|)/\sqrt{2}\}_{j \neq k}$. One then obtains, using lemma 16,

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{ij} |i\rangle\langle j| \otimes |i\rangle\langle j|. \quad (3.26)$$

Using the partial transpose criterion [15], one can see that $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ must be entangled and not separable,⁶ which proves that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \notin \mathbf{Q.Sep} = \text{Sep}(\mathbf{s}, \mathbf{e})$, and hence any such scenario does not admit an operationally noncontextual ontological model. This argument is mathematically the same as that of [13], and the same result is proven by slightly different means in [10].

3.4 Ontic space cardinality

In this section, we will show two simple bounds for the cardinality $|\Lambda|$ of the ontic state space. For our purposes, it suffices to distinguish two cases: either $|\Lambda| < \infty$ which means that Λ is a finite set consisting of $|\Lambda|$ many elements, or $|\Lambda| = \infty$ which means that Λ is countable infinite or uncountable. Then, one can show a lower and upper bound for the size of the ontic space as in the following theorem. The proof is given in appendix C.4.

Theorem 3 (Ontic space cardinality bounds). *For any (\mathbf{s}, \mathbf{e}) that admits an operationally noncontextual ontological model with ontic state space Λ , it holds that either Λ is an infinite set, or it is discrete and respects*

$$\dim(\mathcal{R}) \leq |\Lambda|. \quad (3.27)$$

Furthermore, if (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model (definition 17), then there exists an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) with discrete ontic space Λ_{\min} which satisfies

$$\dim(\mathcal{R}) \leq |\Lambda_{\min}| \leq \dim(\mathcal{R})^2 \leq \dim(\mathcal{L}(\mathcal{H}))^2. \quad (3.28)$$

Recall that $\dim(\mathcal{L}(\mathcal{H})) = \dim(\mathcal{H})^2$: the dimension of the quantum Hilbert space thus plays an important role in determining the maximal cardinality of the ontic space. Furthermore, propositions 8 and 11 together with theorem 3 prove the bounds on the number of ontic space and the minimal ontic space cardinality to be invariant under the addition of incoherent coarse-grainings in the scenario, and also under a shift of representation of the operational (quantum) primitives.

⁶The partial transpose of $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ reads in this case $\sum_{ij} |i\rangle\langle j| \otimes |j\rangle\langle i|$ which has a negative eigenvalue for the eigenvector $|1\rangle \otimes |2\rangle - |2\rangle \otimes |1\rangle$.

Addendum. After the first version of this work was released, related bounds on the number of ontic states in ontological models of operational theories were produced under different assumptions. The work of [16] is also concerned with prepare-and-measure scenarios, and our (\mathbf{s}, \mathbf{e}) would be termed a subGPT there. A stronger upper bound was given in their theorem 3 that can be written as $|\Lambda| = \dim(\mathcal{R})$: the discrepancy between this bound and that of equation (3.28) comes from additional constraints related to the uniqueness of possible classical models, effectively building up a more constrained alternative to both standard Spekkens noncontextuality and the present operationally noncontextual ontological model. In [17], the standard notion of Spekkens noncontextuality is probed in a framework that includes preparations and measurements but also features a proper treatment of transformations. One key result of their work is that including the requirement of the noncontextual ontological representation of transformations strengthens the bound of (3.28) down to, loosely speaking,⁷ $|\Lambda| = \dim(\mathcal{R})$.

3.5 Alternative reduced spaces

We have defined the reduced space in definition 3 as

$$\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})). \quad (3.29)$$

However, one could ask whether an alternative definition of the reduced space would preserve the same physical motivation for the operationally noncontextual ontological model while implying a possibly distinct notion of classicality for the prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) . Such an alternative definition could for instance be obtained from swapping the roles of \mathbf{s} and \mathbf{e} in the definition of \mathcal{R} , thus leading to a potential alternative reduced space $P_{\text{span}(\mathbf{s})}(\text{span}(\mathbf{e}))$.

In this section, we will define and motivate a generalized class of reduced spaces from which one can construct alternative operationally noncontextual ontological models, and prove that the corresponding notions of classicality are all equivalent. To start with, consider the following class of reduced spaces:

Definition 18. *An alternative reduced space is any real, finite dimensional inner product space \mathcal{R}_{alt} together with two linear maps $f, g : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{R}_{\text{alt}}$ that satisfy*

$$\forall \rho \in \mathbf{s}, \forall E \in \mathbf{e} : \quad \langle \rho, E \rangle_{\mathcal{L}(\mathcal{H})} = \langle f(\rho), g(E) \rangle_{\mathcal{R}_{\text{alt}}}, \quad (3.30a)$$

$$\text{span}(f(\mathbf{s})) = \mathcal{R}_{\text{alt}}, \quad (3.30b)$$

$$\text{span}(g(\mathbf{e})) = \mathcal{R}_{\text{alt}}. \quad (3.30c)$$

⁷There would be more to say about tomographic completeness assumptions: we use \mathcal{R} as the representative vector space by analogy with our work, but for the proper description see [17].

The fact that both maps f, g have their image in the same vector space allows one to preserve the symmetry between the treatment of states and effects. The real inner product structure of any \mathcal{R}_{alt} is a simple mathematical choice. We will return to the validity of the choice of finite dimensionality of \mathcal{R}_{alt} later. Equation (3.30c) is motivated by the fact that for any $\rho \in \mathbf{s}$, the probabilities $\{\langle \rho, E \rangle_{\mathcal{L}(\mathcal{H})} : E \in \mathbf{e}\}$ do not necessarily fully determine ρ . On the other hand, with equation (3.30c) and the non-degeneracy of the inner product at hand, the probabilities $\{\langle f(\rho), g(E) \rangle_{\mathcal{R}_{\text{alt}}} : E \in \mathbf{e}\}$ completely determine $f(\rho)$. Thus, $f(\rho)$ is a good primitive to devise a noncontextual model that only resolves the degrees of freedom that are resolved by $g(E)$. This argument can be repeated swapping each ρ, \mathbf{s} and f with E, \mathbf{e} and g respectively to motivate analogously equation (3.30b). The inner product bilinearity and equations (3.30b), (3.30c) imply that if (3.30a) is to hold then f, g have to be linear maps.

Without attempting to fully characterize the set of solutions to definition 18, we prove that while \mathcal{R} as defined in definition 3 is indeed a valid solution to definition 18, it is not the only such solution. The proof is given in appendix D.1.

Proposition 19. *The choice*

$$\mathcal{R}_{\text{alt}} := \mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})), \quad (3.31a)$$

$$f(\cdot) := P_{\mathcal{R}}(\cdot), \quad (3.31b)$$

$$g(\cdot) := P_{\mathcal{R}}(\cdot) \quad (3.31c)$$

yields a valid alternative reduced space in definition 18; and so does the swapped version

$$\mathcal{R}_{\text{alt}} := P_{\text{span}(\mathbf{s})}(\text{span}(\mathbf{e})) =: \mathcal{R}', \quad (3.32a)$$

$$f(\cdot) := P_{\mathcal{R}'}(\cdot), \quad (3.32b)$$

$$g(\cdot) := P_{\mathcal{R}'}(\cdot). \quad (3.32c)$$

We required the dimension of \mathcal{R}_{alt} to be finite: in fact, definition 18 allows to prove the following proposition, see appendix D.1 for a proof.

Proposition 20. *It holds that for any reduced space \mathcal{R}_{alt} (definition 18), $\dim(\mathcal{R}_{\text{alt}}) = \dim(\mathcal{R})$ where \mathcal{R} is defined in definition 3.*

Recall that the vector space inclusion $\mathcal{R} \subseteq \mathcal{L}(\mathcal{H})$ bounds the dimension of \mathcal{R} and thus also bounds that of any \mathcal{R}_{alt} : $\dim(\mathcal{R}_{\text{alt}}) \leq \dim(\mathcal{L}(\mathcal{H})) = \dim(\mathcal{H})^2$.

The operationally noncontextual ontological model is defined for a given alternative reduced space in appendix definition D.2 by analogy to the operationally noncontextual ontological model formulated for \mathcal{R} in definition 6. It turns out that if one were to use any alternative reduced space, one would derive equivalent results to those already obtained. In addition, a result that holds formulated in \mathcal{R} is usually equivalent to that formulated in any \mathcal{R}_{alt} . Most importantly we have the following equivalence, proven in appendix D.2:

Theorem 4 (Equivalence of reduced spaces). *Given any (\mathbf{s}, \mathbf{e}) , consider \mathcal{R} and any \mathcal{R}_{alt} constructed from (\mathbf{s}, \mathbf{e}) . There exists an operationally noncontextual ontological model with ontic space Λ constructed on \mathcal{R} (definition 6) if and only if there exists an operationally noncontextual ontological model constructed on \mathcal{R}_{alt} (appendix definition D.2) with the same ontic space Λ .*

Note that the ontic primitives of the models in \mathcal{R} and a given \mathcal{R}_{alt} may be different, in particular they may belong to distinct vector spaces; but the ontic space that underlies the operationally noncontextual ontological model is the same in either case. The implications of theorem 4 are the following:

- (i) saying that the scenario (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model is a statement which can be made regardless of which reduced space one chooses to use;
- (ii) in the case of Riemann integrable operationally noncontextual ontological models (definition 17), the generic case is that the ontic space is discrete as stated in theorem 3. Then, according to theorem 4, any choice of alternative reduced space \mathcal{R}_{alt} will yield ontic spaces of the same cardinality as those of \mathcal{R} .

Our choice to use \mathcal{R} rather than another alternative reduced space \mathcal{R}_{alt} is without significance. Some additional equivalences between the alternative reduced spaces that are relevant for the algorithmic evaluation of the unit separability criterion will be provided in section 4.5.

4 Algorithmic formulation, witnesses and certifiers

The content of this section is organized as follows. First, we describe general results from convex analysis and introduce the vertex enumeration problem in section 4.1. Then, we describe general theoretical results that hold for an arbitrary scenario (\mathbf{s}, \mathbf{e}) in section 4.2: these results help characterize the set of separable operators $\text{Sep}(\mathbf{s}, \mathbf{e})$ appearing in the unit separability criterion of theorem 2. In section 4.3, we specialize to the to-be-defined polyhedral scenarios for which the unit separability criterion can be verified exactly. In section 4.4, we show how to certify the existence or non-existence of an operationally noncontextual operational model for an arbitrary scenario (\mathbf{s}, \mathbf{e}) using the results of the polyhedral case, and discuss the convergence of the resulting hierarchies of algorithmic tests. In section 4.5, we show the equivalence between the computational complexities of the algorithm formulated in \mathcal{R} and that formulated in an alternative reduced space.

4.1 Vertex enumeration

Let us introduce some notation. A review of the main convex analysis definitions is presented in appendix A, and the proofs of the propositions of this section are presented in appendix E.1.

For any finite dimensional real inner product space \mathcal{V} , let $X \subseteq \mathcal{V}$ be an arbitrary set. The conic hull $\text{coni}(X)$ is the set of elements of the form $\lambda x \in \mathcal{V}$ for any $\lambda \in \mathbb{R}_{\geq 0}$ and $x \in X$. A convex cone $\mathcal{C} \subseteq \mathcal{V}$ is one that equals its convex hull and also its conic hull: $\mathcal{C} = \text{conv}(\mathcal{C}) = \text{coni}(\mathcal{C})$. A half-line is the conic hull of a single element of the vector space. An extremal half-line of \mathcal{C} is a half-line whose elements cannot be expressed as the average of linearly independent elements of \mathcal{C} . The set of extremal half-lines of a convex cone \mathcal{C} is denoted $\text{extr}(\mathcal{C})$.

Definition 21 (Pointed cone). *Let $\mathcal{C} \subseteq \mathcal{V}$ be a convex cone. \mathcal{C} is said to be a pointed cone if*

- (i) \mathcal{C} is closed;
- (ii) $\mathcal{C} \neq \emptyset$ and $\mathcal{C} \neq \{0\}$;
- (iii) there exists a linear function $L : \mathcal{V} \rightarrow \mathbb{R}$ such that for all $c \in \mathcal{C} \setminus \{0\}$, $L(c) > 0$.

The following proposition guarantees the representation of pointed cones as the convex hull of their extremal half-lines.

Proposition 22. *If $\mathcal{C} \subseteq \mathcal{V}$ is a pointed cone, then it holds that*

$$\mathcal{C} = \text{conv}\left(\bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C})} \mathfrak{l}\right). \quad (4.1)$$

We will also need the representation of the polar cone (definition 10) as the convex hull of its extremal half-lines: the following definition and proposition will be useful for that purpose.

Definition 23 (Spanning cone). *A convex cone $\mathcal{C} \subseteq \mathcal{V}$ is a spanning cone in \mathcal{V} if*

- (i) \mathcal{C} is closed;
- (ii) $\mathcal{C} \neq \mathcal{V}$;
- (iii) $\text{span}(\mathcal{C}) = \mathcal{V}$.

Notice that the spanning cone property depends on the vector space \mathcal{V} in which one embeds \mathcal{C} .

Proposition 24. *If $\mathcal{C} \subseteq \mathcal{V}$ is a spanning cone, then the polar cone $\mathcal{C}^{+\mathcal{V}} \subseteq \mathcal{V}$ (definition 10) is a pointed cone, which implies by proposition 22 that*

$$\mathcal{C}^{+\mathcal{V}} = \text{conv}\left(\bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\mathcal{V}})} \mathfrak{l}\right). \quad (4.2)$$

We see that if $\mathcal{C} \subseteq \mathcal{V}$ is a spanning pointed cone in \mathcal{V} , both \mathcal{C} and the polar $\mathcal{C}^{+\mathcal{V}}$ may be represented as the convex hull of their respective extremal half-lines. This defines the so-called vertex enumeration problem⁸:

⁸We follow the denomination given in the literature, e.g., in [18], but we define the vertex enumeration problem even if the cone has infinitely many extremal half-lines.

Definition 25 (Vertex enumeration problem). For $\mathcal{C} \subseteq \mathcal{V}$ a spanning pointed cone, the vertex enumeration problem consists in obtaining the extremal half-lines of $\mathcal{C}^{+\nu}$ from the extremal half-lines of \mathcal{C} . We denote the vertex enumeration map $\mathbf{V.E}_{\mathcal{V}}[\cdot]$:

$$\mathbf{V.E}_{\mathcal{V}}[\text{extr}(\mathcal{C})] := \text{extr}(\mathcal{C}^{+\nu}). \quad (4.3)$$

The last proposition that will prove useful is the following half-space representation of a convex cone, starting from its extremal half-lines⁹.

Proposition 26. A solution to the vertex enumeration problem allows one to represent a spanning pointed cone $\mathcal{C} \subseteq \mathcal{V}$ as the intersection of half-spaces:

$$\mathcal{C} = \bigcap_{l \in \mathbf{V.E}_{\mathcal{V}}[\text{extr}(\mathcal{C})]} l^{+\nu}. \quad (4.4)$$

4.2 General aspects of the algorithm

The proofs of the propositions of this section are presented in appendix E.2.

For the purpose of determining the structure of $\text{Sep}(\mathbf{s}, \mathbf{e})$, it turns out that rather than considering the convex sets $P_{\mathcal{R}}(\mathbf{e})$ and $P_{\mathcal{R}}(\mathbf{s})$, the main objects of interest are the convex cones $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$, where \bar{X} denotes the closure of X . Indeed:

Proposition 27.

$$P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = [\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))]^{+\mathcal{R}}, \quad (4.5a)$$

$$P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = [\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))]^{+\mathcal{R}}. \quad (4.5b)$$

These expressions are useful due to the fact that the vertex enumeration problem is well-defined for the sets $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$:

Proposition 28. $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ are spanning pointed cones in \mathcal{R} .

Together with proposition 27, this shows that applying the vertex enumeration map to $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ will yield the extremal half-lines of $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ and $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$:

$$\text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}) = \mathbf{V.E}_{\mathcal{R}}[\text{extr}(\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})))], \quad (4.6a)$$

$$\text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}) = \mathbf{V.E}_{\mathcal{R}}[\text{extr}(\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})))]. \quad (4.6b)$$

Knowing the extremal half-lines of $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ and $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$, the characterization of $\text{Sep}(\mathbf{s}, \mathbf{e})$ as the convex hull of its extremal half-lines is readily obtained. First consider the following proposition:

⁹A half-space $H \subseteq \mathcal{V}$ is the geometric interpretation of a homogeneous linear inequality solution set $H := \{v \in \mathcal{V} : L_H(v) \geq 0\}$ where L_H is a linear functional. By Riesz' representation theorem B.1, there exists a unique $v_H \in \mathcal{V}$ such that $L_H(\cdot) = \langle v_H, \cdot \rangle_{\mathcal{V}}$. It is then clear that $H = \{v_H\}^{+\nu} = [\text{coni}(v_H)]^{+\nu}$. Thus, the polar cone of a half-line is a half-space and conversely.

Proposition 29. $\text{Sep}(\mathbf{s}, \mathbf{e})$ is a spanning pointed cone in $\mathcal{R} \otimes \mathcal{R}$.

This proposition together with proposition 22 guarantees that we may represent $\text{Sep}(\mathbf{s}, \mathbf{e})$ as the convex hull of its extremal half-lines. If l_1 and l_2 are half-lines in \mathcal{R} , then $l_1 \otimes_{\text{set}} l_2$ is a half-line in $\mathcal{R} \otimes \mathcal{R}$. The following proposition makes explicit the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ (recall definition 12):

Proposition 30. It holds that

$$\text{extr}(\text{Sep}(\mathbf{s}, \mathbf{e})) = \left\{ l_1 \otimes_{\text{set}} l_2 : \begin{array}{l} l_1 \in \text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}), \\ l_2 \in \text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}) \end{array} \right\}. \quad (4.7)$$

Thus, knowing the extremal half-lines of $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ and $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ is equivalent to knowing the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$.

Once the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ are known, there are two options to determine whether the Choi state $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ belongs to $\text{Sep}(\mathbf{s}, \mathbf{e})$. The first, which is rather computationally intensive, is to run the vertex enumeration procedure on $\text{Sep}(\mathbf{s}, \mathbf{e})$ to go from its extremal half-line representation to its supporting half-space representation. It then suffices to verify whether $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ belongs to all the corresponding half-spaces. Alternatively, for a better computational efficiency, one may formulate the test of whether $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e})$ as a linear program, once the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ are known. An explicit such linear program computes a value t^* as follows:

$$t^* := \max_{t, \{\lambda_i \in \mathbb{R}\}_i} t \quad (4.8a)$$

$$\text{s.t. } \forall i : \lambda_i \geq t, \quad (4.8b)$$

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_i \lambda_i l_i, \quad (4.8c)$$

where i is an index for the extremal half-lines $\{l_i\}_i$ of $\text{Sep}(\mathbf{s}, \mathbf{e})$ as in proposition 30, and for each i we picked an arbitrary $0 \neq l_i \in l_i$. This linear program is always feasible, since for $t = -\infty$, it amounts to finding a real (not necessarily nonnegative) linear combination of the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ that equals $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$, which is always possible since $\text{Sep}(\mathbf{s}, \mathbf{e})$ spans $\mathcal{R} \otimes \mathcal{R}$ to which $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ belongs. If $t^* \geq 0$, then the solution coefficients $\{\lambda_i\}_i$ of the linear program are nonnegative and they yield a valid conic combination the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ that equals $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$. Overall, we that the linear program always terminates with a value t^* , and if $t^* < 0$, this proves that the scenario does not admit an operationally noncontextual ontological model, while $t^* \geq 0$ proves the scenario to admit an operationally noncontextual ontological model.¹⁰

¹⁰When $t^* < 0$, the absolute value of t^* could be used as a heuristic, scenario-dependent ‘‘measure’’ of non-classicality.

The authors were made aware of the fact that a vertex enumeration-based routine was already proposed in [19], allowing to test algorithmically the existence of a different but related classical model. One important technical difference is that the authors of [19] were looking at a reduced set of inequalities, i.e., in the light of the current results, they were looking at a lower dimensional projection of the convex cone $\text{Sep}(\mathbf{s}, \mathbf{e})$. For this reason, proposition 30 did not apply to their setup, which is why in this case it was crucial to remove the redundant elements out of the set on the right-hand side of equation (4.7) to optimize the runtime of the algorithm, as was emphasized in their work.

4.3 Solvable cases: polyhedral scenarios

In this section, we describe the case of polyhedral scenarios, for which the unit separability criterion may be evaluated in finite time.

Definition 31 (Polyhedral scenarios). *The prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) is said to be a polyhedral scenario if the convex cones $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ have finitely many extremal half-lines:*

$$N_{\mathbf{s}} := |\text{extr}(\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})))| < \infty, \quad (4.9a)$$

$$N_{\mathbf{e}} := |\text{extr}(\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})))| < \infty. \quad (4.9b)$$

A sufficient condition for (\mathbf{s}, \mathbf{e}) to form a polyhedral scenario is that \mathbf{s} is the convex hull of finitely many quantum states, and \mathbf{e} is the convex hull of finitely many quantum effects. The motivation for the name is that convex cones that are generated by finitely many extremal half-lines are special cases of the well-known polyhedral convex cones [20].

In the vertex enumeration problem, if $\mathcal{C} \subset \mathcal{V}$ is a spanning pointed cone that has finitely many extremal half-lines, i.e., if \mathcal{C} is polyhedral, then $\mathcal{C}^{+\nu}$ will have finitely many extremal half-lines, as described in, e.g., section 4.6 of [20]. Efficient algorithms to solve the vertex enumeration in this case exist in the literature such as the reverse search approach of [18].

Thus, for a polyhedral scenario (\mathbf{s}, \mathbf{e}) , the first vertex enumeration problems, i.e., those of equations (4.6), will each produce a finite number of extremal half-lines. Let there be $M_{\mathbf{s}} \in \mathbb{N}$ extremal half-lines of $P_{\mathcal{R}}(\mathbf{s})^{+\kappa}$, and $M_{\mathbf{e}} \in \mathbb{N}$ for $P_{\mathcal{R}}(\mathbf{e})^{+\kappa}$. These will form, via proposition 30, the $M_{\mathbf{s}} \cdot M_{\mathbf{e}}$ extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$. The linear program of (4.9) has $1 + M_{\mathbf{s}} \cdot M_{\mathbf{e}}$ variables and $\dim(\mathcal{R})^2 + M_{\mathbf{s}} \cdot M_{\mathbf{e}}$ constraints, which are both finite only for such polyhedral scenarios, and thus, for these polyhedral scenarios the linear program (4.9) can be implemented.

4.4 Polyhedral approximations

In the general case where (\mathbf{s}, \mathbf{e}) is not a polyhedral scenario (definition 31), or where (\mathbf{s}, \mathbf{e}) is a polyhedral

scenario but the runtime of the previous algorithm is prohibitively long due to, e.g., a large number of extremal half-lines, one may still choose any polyhedral inner or outer approximation of the relevant cones, yielding certificates for either the existence or non-existence of an operationally noncontextual ontological model as described in the following sections.

4.4.1 Certifying the existence of an operationally noncontextual ontological model

First, consider an outer approximation of the input cones: let $\mathcal{C}_{\mathbf{s}}^{(\text{out})}, \mathcal{C}_{\mathbf{e}}^{(\text{out})} \subseteq \mathcal{R}$ be spanning pointed cones (definitions 21 and 23) in \mathcal{R} such that

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})) \subseteq \mathcal{C}_{\mathbf{s}}^{(\text{out})}, \quad (4.10a)$$

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})) \subseteq \mathcal{C}_{\mathbf{e}}^{(\text{out})}, \quad (4.10b)$$

and such that $|\text{extr}(\mathcal{C}_{\mathbf{s}}^{(\text{out})})|, |\text{extr}(\mathcal{C}_{\mathbf{e}}^{(\text{out})})| < \infty$. Such cones always exist: let us give a constructive example. Consider the hyperspace description of $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$, $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ as in proposition 26. If one keeps a finite set of at least $\dim(\mathcal{R})$ hyperspaces, the resulting cones will be spanning pointed cones outer approximations of $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$, $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ with finitely many extremal half-lines. The algorithm described in the previous section may be run in exactly the same way as in the polyhedral case, with $\mathcal{C}_{\mathbf{s}}^{(\text{out})}$ and $\mathcal{C}_{\mathbf{e}}^{(\text{out})}$ replacing $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ as inputs to the algorithm. Let $\text{Sep}^{(\text{in})}$ be the cone that the algorithm will characterize:

$$\text{Sep}^{(\text{in})} := \text{conv} \left([\mathcal{C}_{\mathbf{s}}^{(\text{out})}]^{+\kappa} \otimes_{\text{set}} [\mathcal{C}_{\mathbf{e}}^{(\text{out})}]^{+\kappa} \right). \quad (4.11)$$

Using lemma A.7 and equations (4.10), it can be shown that

$$\text{Sep}^{(\text{in})} \subseteq \text{Sep}(\mathbf{s}, \mathbf{e}), \quad (4.12)$$

which justifies the reversed superscript of $\text{Sep}^{(\text{in})}$: outer conic approximations in (4.10) yield an inner approximation of $\text{Sep}(\mathbf{s}, \mathbf{e})$ in (4.12). Hence, if the algorithm obtains that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}^{(\text{in})}$, it also holds that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e})$, and hence this proves the scenario we started with to admit an operationally noncontextual ontological model. On the other hand, if $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \notin \text{Sep}^{(\text{in})}$, the outer approximation is inconclusive. One may then, for example, use refined polyhedral outer approximations $\mathcal{C}_{\mathbf{s}}^{(\text{out})\prime} \subset \mathcal{C}_{\mathbf{s}}^{(\text{out})}, \mathcal{C}_{\mathbf{e}}^{(\text{out})\prime} \subset \mathcal{C}_{\mathbf{e}}^{(\text{out})}$ that are subsets of the previous ones but still satisfy (4.10) to obtain a finer inner approximation of $\text{Sep}(\mathbf{s}, \mathbf{e})$, and repeat the procedure. The convergence of this hierarchy of finer approximations will be discussed shortly but let us first describe the outer approximations of $\text{Sep}(\mathbf{s}, \mathbf{e})$.

4.4.2 Certifying the non-existence of an operationally noncontextual ontological model

In parallel to attempting to certify the existence of an operationally noncontextual ontological model for

(\mathbf{s}, \mathbf{e}) by using outer approximations, one may also consider inner approximations to the input cones: choose spanning pointed cones $\mathcal{C}_{\mathbf{s}}^{(\text{in})}, \mathcal{C}_{\mathbf{e}}^{(\text{in})} \subseteq \mathcal{R}$ such that

$$\mathcal{C}_{\mathbf{s}}^{(\text{in})} \subseteq \text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})), \quad (4.13a)$$

$$\mathcal{C}_{\mathbf{e}}^{(\text{in})} \subseteq \text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})), \quad (4.13b)$$

and such that $|\text{extr}(\mathcal{C}_{\mathbf{s}}^{(\text{in})})|, |\text{extr}(\mathcal{C}_{\mathbf{e}}^{(\text{in})})| < \infty$. Such cones always exist. For example, consider the extremal half-line description of $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$, $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ as in proposition 22. By keeping a finite set of at least $\dim(\mathcal{R})$ extremal half-lines, the resulting cones will be spanning pointed cones inner approximations of $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$, $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ with finitely many extremal half-lines. The algorithm may be run in that case as well, with $\mathcal{C}_{\mathbf{s}}^{(\text{in})}, \mathcal{C}_{\mathbf{e}}^{(\text{in})}$ as inputs rather than $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$, $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$. Let $\text{Sep}^{(\text{out})}$ be the cone that the algorithm will characterize:

$$\text{Sep}^{(\text{out})} := \text{conv}\left([\mathcal{C}_{\mathbf{s}}^{(\text{in})}]^{+\mathcal{R}} \otimes_{\text{set}} [\mathcal{C}_{\mathbf{e}}^{(\text{in})}]^{+\mathcal{R}}\right). \quad (4.14)$$

Using lemma A.7 and equations (4.13), it can be shown that

$$\text{Sep}(\mathbf{s}, \mathbf{e}) \subseteq \text{Sep}^{(\text{out})}, \quad (4.15)$$

which again justifies the reversed superscripts. This time, if the algorithm concludes that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \notin \text{Sep}^{(\text{out})}$, this implies $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \notin \text{Sep}(\mathbf{s}, \mathbf{e})$ and hence the scenario does not admit an operationally noncontextual ontological model. If the algorithm concludes that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}^{(\text{out})}$, the inner approximation is inconclusive and one should use a refined inner approximation in (4.13).

4.4.3 Comments on convergence

It is important to realize that finer and finer approximations will have more and more extremal half-lines, and will yield computationally harder instances of vertex enumeration and linear programming. The procedure of repeatedly refining the inner or outer approximations will in principle converge to a definite answer provided that $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ is in an interior or exterior point of the closed (proposition C.9) convex cone $\text{Sep}(\mathbf{s}, \mathbf{e})$. An alternative approach to refining the polyhedral approximations would be to change the inner or outer approximations randomly while keeping the number of extremal half-lines fixed. This procedure would have the merit of probing more of the structure of (\mathbf{s}, \mathbf{e}) while keeping the computational complexity fixed, but there is no guarantee for the convergence of this approach. If $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ lies on the boundary of the closed (proposition C.9) cone $\text{Sep}(\mathbf{s}, \mathbf{e})$, then most inner approximations in the nonpolyhedral case will not allow to prove the scenario to admit an operationally noncontextual ontological model. However, such cases can reasonably be considered as edge cases since there will exist small deformations of the inputs (\mathbf{s}, \mathbf{e}) which will make $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$

an interior or exterior point. Hence, if one is interested in families of prepare-and-measure scenarios, it is most likely that such cases will not occur for most scenarios.

4.4.4 Connections with quantum entanglement

The present algorithm may be recast as a basic algorithm to treat the usual problem of verifying the entanglement of a given bipartite state. Let us give the key ideas to relate the two procedures. Let the convex cone of positive semi-definite matrices be $\mathcal{P}(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$. The convex cone of unnormalized separable quantum states is $\mathbf{Q.Sep} := \text{conv}(\mathcal{P}(\mathcal{H}) \otimes_{\text{set}} \mathcal{P}(\mathcal{H}))$. If a state $\Omega \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ belongs to $\mathbf{Q.Sep}$, it is said to be separable, else it is said to be entangled.

To recast the problem of determining whether Ω is entangled or not to an application of the algorithm described in the previous sections, consider the following main identifications. First, in the previous algorithm, replace \mathcal{R} with $\mathcal{L}(\mathcal{H})$. Then, the input cones $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}}))$ are both replaced with $\mathcal{P}(\mathcal{H})$. The tensor product cone $\mathbf{Q.Sep}$ is related to $(\mathcal{P}(\mathcal{H}), \mathcal{P}(\mathcal{H}))$ in the same way that $\text{Sep}(\mathbf{s}, \mathbf{e})$ is related to $(\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})), \text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})))$ (see definition 14). For this identification to work, one needs to recall the basic result that $\mathcal{P}(\mathcal{H})^{+\mathcal{L}(\mathcal{H})} = \mathcal{P}(\mathcal{H})$. Then, the state $\Omega \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ replaces $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$. Characterizing whether $\Omega \in \mathbf{Q.Sep}$ can thus be reduced to a nonpolyhedral instance of the previous algorithm, due to the infinite number of extremal half-lines of $\mathcal{P}(\mathcal{H})$: the set of extremal half-lines of $\mathcal{P}(\mathcal{H})$ is equal to the set of all half-lines $\text{coni}(|\psi\rangle\langle\psi|)$ with $|\psi\rangle \in \mathcal{H}$. Thus, it makes sense to use inner and outer approximations as described in section 4.4 to decide the separability of the state. Due to the complexity of vertex enumeration in the general case, there exist more efficient algorithms in the literature to produce entanglement witnesses such as the SDP hierarchy of [21].

4.5 Computational equivalence: changing reduced spaces and quantum primitives

4.5.1 Changing reduced spaces

In section 3.5, it was shown that the existence of an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) is a concept that is independent of whether one chooses to work with the initial reduced space \mathcal{R} (definition 3) or with any alternative reduced space \mathcal{R}_{alt} (definition 18). The previous sections suggested an algorithmic procedure to verify the existence of an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) through an evaluation of the unit separability criterion, theorem 2. One may ask whether it is simpler to execute this algorithmic procedure when working in \mathcal{R} or any other \mathcal{R}_{alt} . As stated in section 4.3, the computational complexity of the algorithm depends on 1) the dimension of the ambient vector space, but

by proposition 20 these are the same in \mathcal{R} and any \mathcal{R}_{alt} , and 2) the number of extremal half-lines of the relevant cones and their dual cones. The following propositions will prove the equivalence of number of extremal half-lines of these cones built in \mathcal{R} or any other \mathcal{R}_{alt} .

Definition 32. *Given any two finite dimensional real inner product spaces \mathcal{U} , \mathcal{V} such that $\dim(\mathcal{U}) = \dim(\mathcal{V})$, two convex cones $\mathcal{C} \subseteq \mathcal{U}$ and $\mathcal{D} \subseteq \mathcal{V}$ are said to be isomorphic, denoted $\mathcal{C} \sim \mathcal{D}$, if and only if there exists an invertible linear map $\Phi : \mathcal{U} \rightarrow \mathcal{V}$ such that*

$$\Phi(\mathcal{C}) = \mathcal{D}. \quad (4.16)$$

Applying this definition to the relevant cones in our setup, we obtain:

Proposition 33. *Choosing any alternative reduced space \mathcal{R}_{alt} with associated mappings f, g (definition 18), it holds that:*

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})) \sim \text{coni}(f(\bar{\mathbf{s}})), \quad (4.17a)$$

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})) \sim \text{coni}(g(\bar{\mathbf{e}})), \quad (4.17b)$$

$$\text{Sep}(\mathbf{s}, \mathbf{e}) \sim \text{Sep}(\mathbf{s}, \mathbf{e})_{\text{alt}}, \quad (4.17c)$$

where

$$\text{Sep}(\mathbf{s}, \mathbf{e})_{\text{alt}} := \text{conv}(f(\mathbf{s})^{+\mathcal{R}_{\text{alt}}} \otimes_{\text{set}} g(\mathbf{e})^{+\mathcal{R}_{\text{alt}}}). \quad (4.18)$$

The following proposition will allow one to assert the computational equivalence of \mathcal{R} and \mathcal{R}_{alt} :

Proposition 34. *Given any two finite dimensional real inner product spaces \mathcal{U} , \mathcal{V} such that $\dim(\mathcal{U}) = \dim(\mathcal{V})$, any two convex cones $\mathcal{C} \subseteq \mathcal{U}$ and $\mathcal{D} \subseteq \mathcal{V}$ such that $\mathcal{C} \sim \mathcal{D}$ have the following properties:*

- (i) *there is a one-to-one correspondence between the extremal half-lines of \mathcal{C} and those of \mathcal{D} ;*
- (ii) *the same holds for the extremal half-lines of the polar cones due to $\mathcal{C}^{+\mathcal{U}} \sim \mathcal{D}^{+\mathcal{V}}$.*

Propositions 33 and 34, proven in appendix E.3, prove that all the cones involved in the algorithm that verifies the unit separability criterion will yield vertex enumeration problems of the same complexity because this complexity depends primarily on the number of extremal half-lines as described in section 4.3.

4.5.2 Changing quantum primitives

We now expand on proposition 8 to show that not only are statistically equivalent scenarios equivalent for the existence of an operationally noncontextual ontological model of a certain type, but also that the computational complexity of verifying the unit separability criterion is the same for either one of the two descriptions. The proof is given in appendix E.3.

Proposition 35. *For the scenarios (\mathbf{s}, \mathbf{e}) and $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ and the corresponding reduced spaces \mathcal{R} and $\tilde{\mathcal{R}}$ defined in proposition 8, it holds that*

$$\text{coni}(P_{\mathcal{R}}(\mathbf{s})) \sim \text{coni}(P_{\tilde{\mathcal{R}}}(\tilde{\mathbf{s}})), \quad (4.19)$$

$$\text{coni}(P_{\mathcal{R}}(\mathbf{e})) \sim \text{coni}(P_{\tilde{\mathcal{R}}}(\tilde{\mathbf{e}})), \quad (4.20)$$

$$\text{Sep}(\mathbf{s}, \mathbf{e}) \sim \text{Sep}(\tilde{\mathbf{s}}, \tilde{\mathbf{e}}), \quad (4.21)$$

which proves by proposition 34 the computational equivalence of starting with either (\mathbf{s}, \mathbf{e}) or $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$.

5 Connections with generalized probabilistic theories

5.1 Generalized probabilistic reformulation

Although we formulated the operationally noncontextual ontological model of definition 6 for quantum primitives, the fact that the sets \mathbf{s} , \mathbf{e} originate from the Hilbert space of the quantum system is not crucial for the operationally noncontextual ontological model construction.

Instead, rather than considering the vector space $\mathcal{L}(\mathcal{H})$ equipped with the Hilbert-Schmidt inner product, consider any real inner product space \mathcal{V} of finite dimension:

$$\mathcal{L}(\mathcal{H}) \rightarrow \mathcal{V}. \quad (5.1a)$$

Then, replace \mathbf{s} by $\Omega \subseteq \mathcal{V}$ and \mathbf{e} by $\mathcal{E} \subseteq \mathcal{V}$, following standard notation [9, 22]:

$$\mathbf{s} \rightarrow \Omega, \quad (5.1b)$$

$$\mathbf{e} \rightarrow \mathcal{E}. \quad (5.1c)$$

The probability that an effect $E \in \mathcal{E}$ occurs upon measuring a state $\rho \in \Omega$ is given by the inner product $\langle \rho, E \rangle_{\mathcal{V}}$ by analogy with the usual Hilbert-Schmidt inner product probability rule of quantum mechanics. The properties required for Ω and \mathcal{E} that are necessary for the results of this manuscript are the following. Ω and \mathcal{E} must be nonempty, bounded convex sets such that for all $s \in \Omega$ and $e \in \mathcal{E}$: $\langle s, e \rangle_{\mathcal{V}} \geq 0$. It must be that $0 \in \mathcal{V}$ belongs to \mathcal{E} . There must exist $u \in \mathcal{E}$ such that for all $s \in \Omega$: $\langle s, u \rangle_{\mathcal{V}} = 1$ — this u replaces $\mathbb{1}_{\mathcal{H}}$:

$$\mathbb{1}_{\mathcal{H}} \rightarrow u. \quad (5.1d)$$

We also require that for all $e \in \mathcal{E}$, there exists a completion $\{e_k \in \mathcal{E}\}_k$ such that $e + \sum_k e_k = u$.

All the results of this manuscript can then easily be rederived in this generalized setting: once the prepare-and-measure scenario is defined, the derivations only rely on the axiomatic properties of the state and effect sets together with the basic real, finite-dimensional inner product space structure which are assumed both in the quantum setting and this generalized setting. As an illustration, the reduced space of definition 3 obtained under the substitution (5.1) is

$$\mathcal{R}_G := P_{\text{span}(\mathcal{E})}(\text{span}(\Omega)) \subseteq \mathcal{V}. \quad (5.2)$$

5.2 Connections with simplex-embeddability

Recently, a similar approach to the contextuality of arbitrary tomographically complete (i.e., $\text{span}(\Omega) = \text{span}(\mathcal{E}) = \mathcal{V}$) prepare-and-measure scenarios was presented in [9]. In this section, we will relate the present results to their work. For tomographically complete generalized probabilistic theories, the reduced space under the substitution (5.1) is simply the vector space \mathcal{V} , since

$$\mathcal{R}_G = P_{\text{span}(\mathcal{E})}(\text{span}(\Omega)) = P_{\mathcal{V}}(\mathcal{V}) = \mathcal{V}. \quad (5.3)$$

The definition 1 in [9] (reproduced in appendix definition F.1) of simplex-embeddability, which is formulated only for tomographically complete prepare-and-measure scenarios, turns out to yield a mathematically equivalent criterion to unit separability when the latter is restricted to such tomographically complete prepare-and-measure scenarios (see appendix F for the proof):

Proposition 36. *Any tomographically complete generalized probabilistic theory $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable in d dimensions in the sense of definition 1 of [9], if and only if the tomographically complete prepare-and-measure scenario (Ω, \mathcal{E}) admits an operationally noncontextual ontological model in the sense of definition 6 (under the substitution (5.1)) with a discrete ontic space of finite cardinality d .*

Now consider an arbitrary tomographically complete generalized probabilistic theory denoted $G := (\mathcal{V}, \Omega, \mathcal{E})$. Then, let $b(G) \in \mathbb{N}$ be such that if G is simplex-embeddable, then it is also simplex-embeddable in at most $b(G)$ dimensions. It was asked in [9] whether there existed such a bound. Proposition 36 proves as a corollary the existence of this bound:

Corollary 37. *For any tomographically complete generalized probabilistic theory $G = (\mathcal{V}, \Omega, \mathcal{E})$ for which there exists $d \in \mathbb{N}$ such that G is simplex-embeddable in d dimensions, it holds that G is also simplex-embeddable in $d_{\min} \in \mathbb{N}$ dimensions with*

$$\dim(\mathcal{V}) \leq d_{\min} \leq \dim(\mathcal{V})^2, \quad (5.4)$$

i.e., $b(G) = b(\mathcal{V}, \Omega, \mathcal{E}) = \dim(\mathcal{V})^2$.

Proof. If $G = (\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable, then by proposition 36, the prepare-and-measure scenario (Ω, \mathcal{E}) admits a finite, discrete operationally noncontextual ontological model which is a special case of Riemann integrable operationally noncontextual ontological models (definition 17 under the substitution (5.1)). By theorem 3 under the substitution (5.1), there also exists an operationally noncontextual ontological model with a minimal ontic space cardinality $|\Lambda| =: d_{\min}$ such that $\dim(\mathcal{V}) \leq d_{\min} \leq \dim(\mathcal{V})^2$ where we used (5.3) to substitute \mathcal{R} in theorem 3 with

\mathcal{V} . Again by proposition 36, this means that the generalized probabilistic theory G is simplex-embeddable in d_{\min} dimensions. \square

In [9], by leveraging arguments of [11], it was shown that if a tomographically complete generalized probabilistic theory is such that \mathcal{E} admits finitely many extremal points, then there exists such a bound $b(G)$, and the analysis of [11] also suggests that a similar bound holds if the set of states Ω has finitely many extremal points. However, this bound which is the number of extremal points of the polytope defined in the ‘‘Characterization P1 of the noncontextual measurement-assignment polytope’’ of [11], depends on the set \mathcal{E} and does not have a clear behavior as the number of extremal points of \mathcal{E} grows — it could in principle diverge. For fixed \mathcal{V} , however, the upper bound $\dim(\mathcal{V})^2$ of corollary 37 remains constant for arbitrary choice of (Ω, \mathcal{E}) , even with infinitely many extremal points.

We now turn to applying the results of [9] to the original framework of this manuscript. In [9], an argument is given about the need for so-called ‘‘dimension mismatches’’. This useful argument can be rephrased in our setup as a proof that the lower bound $\dim(\mathcal{R})$ in theorem 3 is not always tight, i.e., there exist (\mathbf{s}, \mathbf{e}) that admits a Riemann integrable operationally noncontextual ontological model with minimal ontic state space cardinality

$$d_{\min} = \dim(\mathcal{R}) + 1. \quad (5.5)$$

While not always tight, it is easy to see from the simplex-embeddability criterion of [9] that there exist (\mathbf{s}, \mathbf{e}) such that the lower bound in theorem 3 is saturated. These considerations raise the open question of whether the upper bound $\dim(\mathcal{R})^2$ in theorem 3 is tight, i.e., whether there exist (\mathbf{s}, \mathbf{e}) such that the minimal ontic space has cardinality $\dim(\mathcal{R})^2$.

6 Conclusion

After introducing the prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) and the reduced space \mathcal{R} , the operationally noncontextual ontological model was defined in theorem 1. The unit separability criterion for the existence of an operationally noncontextual ontological model was extracted in theorem 2. This criterion allowed to extract properties for the size of the ontic space Λ , with most importantly the new bound $|\Lambda| \leq \dim(\mathcal{R})^2$ in theorem 3 on page 10. The algorithmic formulation of the criterion was discussed in section 4, allowing one to numerically certify the existence or nonexistence of an operationally noncontextual ontological model for a given scenario (\mathbf{s}, \mathbf{e}) . Connections with generalized probabilistic theories were given in section 5, with most importantly the ontic space cardinality bounds translating as dimension bounds for

simplex-embeddability as in corollary 37. Future directions of research include most importantly investigating interesting quantum protocols through the prism of whether or not they admit an operationally noncontextual ontological model. In particular, linking the non-existence of an operationally noncontextual ontological model to potential quantum advantages would be most interesting.

Acknowledgments

We would like to thank Roberto Baldijão, Nuriya Nurgalieva, Martin Plávala and Elie Wolfe for fruitful discussions. M.W. acknowledges support from the Swiss National Science Foundation (SNSF) via an AMBIZIONE Fellowship (PZ00P2_179914). This work was supported as a part of NCCR QSIT, a National Centre of Competence (or Excellence) in Research, funded by the Swiss National Science Foundation (grant number 51NF40-185902).

References

- [1] A. Einstein, B. Podolsky, and N. Rosen. “Can quantum-mechanical description of physical reality be considered complete?”. *Phys. Rev.* **47**, 777–780 (1935).
- [2] J. S. Bell. “On the Einstein Podolsky Rosen paradox”. *Physics Physique Fizika* **1**, 195–200 (1964).
- [3] Simon Kochen and E. P. Specker. “The problem of hidden variables in quantum mechanics”. *Journal of Mathematics and Mechanics* **17**, 59–87 (1967). url: <http://www.jstor.org/stable/24902153>.
- [4] R. W. Spekkens. “Contextuality for preparations, transformations, and unsharp measurements”. *Phys. Rev. A* **71**, 052108 (2005).
- [5] David Schmid and Robert W. Spekkens. “Contextual advantage for state discrimination”. *Phys. Rev. X* **8**, 011015 (2018).
- [6] David R. M. Arvidsson-Shukur, Nicole Yunger Halpern, Hugo V. Lepage, Aleksander A. Lasek, Crispin H. W. Barnes, and Seth Lloyd. “Quantum advantage in postselected metrology”. *Nature Communications* **11**, 3775 (2020).
- [7] Mark Howard, Joel J. Wallman, Victor Veitch, and Joseph Emerson. “Contextuality supplies the magic for quantum computation”. *Nature* **510**, 351–355 (2014).
- [8] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. “Contextuality as a resource for models of quantum computation on qubits”. *Physical Review Letters* **119**, 120505 (2017).
- [9] David Schmid, John H. Selby, Elie Wolfe, Ravi Kunjwal, and Robert W. Spekkens. “Characterization of noncontextuality in the framework of generalized probabilistic theories”. *PRX Quantum* **2**, 010331 (2021).
- [10] Robert W. Spekkens. “Negativity and contextuality are equivalent notions of nonclassicality”. *Phys. Rev. Lett.* **101**, 020401 (2008).
- [11] David Schmid, Robert W. Spekkens, and Elie Wolfe. “All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences”. *Phys. Rev. A* **97**, 062103 (2018).
- [12] Robert W. Spekkens. “The status of determinism in proofs of the impossibility of a noncontextual model of quantum theory”. *Foundations of Physics* **44**, 1125–1155 (2014).
- [13] Christopher Ferrie and Joseph Emerson. “Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations”. *Journal of Physics A: Mathematical and Theoretical* **41**, 352001 (2008).
- [14] A. Jamiolkowski. “Linear transformations which preserve trace and positive semidefiniteness of operators”. *Reports on Mathematical Physics* **3**, 275 – 278 (1972).
- [15] Asher Peres. “Separability criterion for density matrices”. *Phys. Rev. Lett.* **77**, 1413–1415 (1996).
- [16] Farid Shahandeh. “Contextuality of general probabilistic theories”. *PRX Quantum* **2**, 010330 (2021).
- [17] David Schmid, John H. Selby, Matthew F. Pusey, and Robert W. Spekkens. “A structure theorem for generalized-noncontextual ontological models” (2020). url: arxiv.org/abs/2005.07161.
- [18] David Avis. “A revised implementation of the reverse search vertex enumeration algorithm”. Pages 177–198. Birkhäuser Basel. (2000).
- [19] Anirudh Krishna, Robert W. Spekkens, and Elie Wolfe. “Deriving robust noncontextuality inequalities from algebraic proofs of the Kochen-Specker theorem: the Peres-Mermin square”. *New Journal of Physics* **19**, 123031 (2017).
- [20] Michael J. Panik. “Fundamentals of convex analysis”. *Theory and Decision Library*. Springer, Dordrecht. (1993).
- [21] Aram W. Harrow, Anand Natarajan, and Xiaodi Wu. “An improved semidefinite programming hierarchy for testing entanglement”. *Communications in Mathematical Physics* **352**, 881–904 (2017).
- [22] Peter Janotta and Haye Hinrichsen. “Generalized probability theories: what determines the structure of quantum theory?”. *Journal of Physics A: Mathematical and Theoretical* **47**, 323001 (2014).
- [23] M. Thamban Nair and Arindama Singh. “Linear algebra”. Springer, Singapore. (2018).
- [24] Joseph Muscat. “Functional analysis”. Springer, Cham. (2014).

[25] Isaac Namioka and R. R. Phelps. “Tensor products of compact convex sets.”. *Pacific J. Math.* **31**, 469–480 (1969).

Contents

1	Introduction	1
1.1	Background: previous notions of classicality	1
1.2	The objective notion of classicality	2
1.3	Content overview	2
2	Operationally noncontextual ontological model	2
2.1	Prepare-and-measure scenarios	2
2.2	Reduced space	3
2.3	Definition of the operationally noncontextual ontological model	3
2.4	Motivation for the reduced space prescription	5
2.5	Structure of the operationally noncontextual ontological model	7
3	Unit separability and cardinality bounds	8
3.1	Mathematical preliminaries	8
3.2	The unit separability criterion	8
3.3	Explicit examples	9
3.4	Ontic space cardinality	10
3.5	Alternative reduced spaces	11
4	Algorithmic formulation, witnesses and certifiers	12
4.1	Vertex enumeration	12
4.2	General aspects of the algorithm	13
4.3	Solvable cases: polyhedral scenarios	14
4.4	Polyhedral approximations	14
4.5	Computational equivalence: changing reduced spaces and quantum primitives	15
5	Connections with generalized probabilistic theories	16
5.1	Generalized probabilistic reformulation	16
5.2	Connections with simplex-embeddability	17
6	Conclusion	17
	References	18
A	Review of convex analysis	20
B	Presentation of the operationally noncontextual ontological model	20
B.1	The prepare-and-measure scenario	20
B.2	The reduced space	20
B.3	An example correspondence between reduced space projection and partial trace	23
B.4	Linear extensions to the ontic mappings	24
B.5	Basic criterion for the existence of an operationally noncontextual ontological model	27
C	Unit separability and cardinality bounds	29
C.1	Generalized separability	29

C.2	Choi-Jamiołkowski isomorphism . . .	33
C.3	The unit separability criterion	34
C.4	Ontic space cardinality	35
D	Alternative reduced spaces	35
D.1	Construction of the alternative operationally noncontextual ontological model	35
D.2	Equivalence of reduced spaces	37
E	Algorithmic formulation, witnesses and certifiers	38
E.1	Resolution of convex cones	38
E.2	General aspects of the algorithm . . .	41
E.3	Computational equivalence: varying reduced spaces and quantum descriptions	44
F	Connections with generalized probabilistic theories	46

A Review of convex analysis

Throughout this section, we assume that \mathcal{V} is a finite dimensional real inner product space. For a more detailed review, see [20].

Definition A.1 (Convex set). *A set $X \subseteq \mathcal{V}$ is convex if and only if for all $0 \leq \lambda \leq 1$, for all $x_1, x_2 \in X$,*

$$(\lambda x_1 + (1 - \lambda)x_2) \in X. \quad (\text{A.1})$$

Definition A.2 (Convex hull). *For any set $X \subseteq \mathcal{V}$, the convex hull $\text{conv}(X)$ is defined as*

$$\text{conv}(X) := \left\{ \sum_{i=1}^n \lambda_i x_i : n \in \mathbb{N}, \lambda_i \in \mathbb{R}_{\geq 0}, \sum_{i=1}^n \lambda_i = 1, x_i \in X \right\}. \quad (\text{A.2})$$

It holds that $\text{conv}(X)$ is the smallest convex set that contains X .

Definition A.3 (Extreme points). *Let $X \subseteq \mathcal{V}$ be a convex set. $x \in X$ is an extreme point of X if and only if, given $\lambda \in]0, 1[$ and $x_1, x_2 \in X$ such that*

$$x = \lambda x_1 + (1 - \lambda)x_2, \quad (\text{A.3})$$

one necessarily has $x_1 = x_2 = x$. The set of extremal points of X is denoted $\text{ep}(X)$.

Definition A.4 (Conic set). *A set $X \subseteq \mathcal{V}$ is a cone, or a conic set, if and only if for all $\lambda \geq 0$, for all $x \in X$,*

$$\lambda x \in X. \quad (\text{A.4})$$

Definition A.5 (Conic hull). *For any set $X \subseteq \mathcal{V}$, the conic hull $\text{coni}(X)$ is defined as*

$$\text{coni}(X) := \{ \lambda x : \lambda \in \mathbb{R}_{\geq 0}, x \in X \}. \quad (\text{A.5})$$

It holds that $\text{coni}(X)$ is the smallest conic set that contains X .

Definition A.6 (Extremal half-lines, see section 4.4 in [20]). *Let $\mathcal{C} \subseteq \mathcal{V}$ be a convex cone. A vector $c_0 \in \mathcal{C}$, $c_0 \neq 0$ is an extremal direction of \mathcal{C} if and only if, for all $d_1, d_2 \in \mathcal{C}$ that satisfy*

$$c_0 = d_1 + d_2, \quad (\text{A.6})$$

d_1 and d_2 are linearly dependent. It is then easy to show that for any $\lambda \in \mathbb{R}$, $\lambda > 0$, λc_0 is also an extremal direction of \mathcal{C} . The half-line $\text{coni}(c_0)$ is said to be an extremal half-line of \mathcal{C} , i.e., all nonzero elements of the extremal half-line are extremal directions. The set of all extremal half-lines of \mathcal{C} is denoted $\text{extr}(\mathcal{C})$.

Note that $\text{extr}(\mathcal{C})$ is a set of set of points of \mathcal{C} .

Lemma A.7. *Consider two sets $X \subseteq Y \subseteq \mathcal{V}$. It holds that*

$$Y^{+\nu} \subseteq X^{+\nu}. \quad (\text{A.7})$$

Proof. Let $v \in Y^{+\nu}$. Then, consider any $x \in X$, and we will show that $\langle v, x \rangle_{\mathcal{V}} \geq 0$. But since $X \subseteq Y$, it holds that $x \in Y$. Thus by definition 10 of the polar cone and since $v \in Y^{+\nu}$, it holds that $\langle x, v \rangle_{\mathcal{V}} \geq 0$. Thus, $v \in X^{+\nu}$. \square

B Presentation of the operationally noncontextual ontological model

B.1 The prepare-and-measure scenario

Let us give an example of how convex mixtures of effects may be obtained from probabilistic mixtures. Suppose that E_1 and E_2 belong to \mathbf{e} , the set of quantum effects that are accessible in the lab, and also that $\mathbb{1}_{\mathcal{H}} - E_1$ and $\mathbb{1}_{\mathcal{H}} - E_2$ belong to \mathbf{e} .

Suppose that one associates to E_i the outcome $+1$ and to $\mathbb{1}_{\mathcal{H}} - E_i$ the outcome -1 . Then, if one measures $\{E_1, \mathbb{1}_{\mathcal{H}} - E_1\}$ with probability $\lambda \in [0, 1]$ and $\{E_2, \mathbb{1}_{\mathcal{H}} - E_2\}$ with probability $1 - \lambda$, then effectively the effect associated with obtaining outcome $+1$ is described by $E(\lambda) := \lambda E_1 + (1 - \lambda)E_2$, while the outcome -1 has associated quantum effect $\mathbb{1}_{\mathcal{H}} - E(\lambda)$. This shows that if one allows for such probabilistic mixtures, effectively the set of effects becomes convex.

B.2 The reduced space

In this section, we present general results about the reduced space that are used in the main text as well as in the following appendices.

Definition B.1. *Let \mathcal{V} be a real inner product space of finite dimension, and let $\mathcal{X} \subseteq \mathcal{V}$ be a vector subspace. Let $\{x_i \in \mathcal{X}\}_{i=1}^{\dim(\mathcal{X})}$ be an orthonormal basis*

with respect to the inner product of \mathcal{V} . Then, we define the projection over \mathcal{X} as:

$$\forall v \in \mathcal{V} : P_{\mathcal{X}}(v) := \sum_{i=1}^{\dim(\mathcal{X})} \langle x_i, v \rangle_{\mathcal{V}} x_i. \quad (\text{B.1})$$

The projection of a set $S \subseteq \mathcal{V}$ is defined as the set of projected elements of S , i.e.,

$$P_{\mathcal{X}}(S) := \{P_{\mathcal{X}}(s) : s \in S\}. \quad (\text{B.2})$$

Lemma B.2. Let \mathcal{V} be a real inner product space of finite dimension, and let $\mathcal{X} \subseteq \mathcal{V}$ be a vector subspace equipped with the inner product inherited from \mathcal{V} . Then, for all $v \in \mathcal{V}$, for all $x \in \mathcal{X}$:

$$\langle v, x \rangle_{\mathcal{V}} = \langle P_{\mathcal{X}}(v), x \rangle_{\mathcal{X}}. \quad (\text{B.3})$$

Proof. Let $\{X_i \in \mathcal{X}\}_i$ be an orthonormal basis of \mathcal{X} . Extend this basis to an orthonormal basis of \mathcal{V} of the form $\{X_i\}_i \cup \{V_j\}_j$. Due to the orthogonality relations, we have $\langle X_i, V_j \rangle_{\mathcal{V}} = 0$ for all i, j , which also implies

$$\forall x \in \mathcal{X}, \forall j : \langle x, V_j \rangle_{\mathcal{V}} = 0. \quad (\text{B.4})$$

Thus, using the completeness relation $v = \sum_i \langle X_i, v \rangle_{\mathcal{V}} X_i + \sum_j \langle V_j, v \rangle_{\mathcal{V}} V_j$, we have

$$\begin{aligned} \langle v, x \rangle_{\mathcal{V}} &= \sum_i \langle X_i, v \rangle_{\mathcal{V}} \langle X_i, x \rangle_{\mathcal{V}} + \sum_j \langle V_j, v \rangle_{\mathcal{V}} \langle V_j, x \rangle_{\mathcal{V}} \\ &= \langle \sum_i \langle X_i, v \rangle_{\mathcal{V}} X_i, x \rangle_{\mathcal{V}} \\ &= \langle P_{\mathcal{X}}(v), x \rangle_{\mathcal{V}} \\ &= \langle P_{\mathcal{X}}(v), x \rangle_{\mathcal{X}}. \end{aligned} \quad (\text{B.5})$$

□

Lemma B.3. For any real inner product space \mathcal{V} of finite dimension, and for any vector subspaces $\mathcal{X}_1 \subseteq \mathcal{X}_2 \subseteq \mathcal{V}$,

$$\forall v \in \mathcal{V} : P_{\mathcal{X}_1}(v) = P_{\mathcal{X}_1}(P_{\mathcal{X}_2}(v)). \quad (\text{B.6})$$

Proof. Choose an orthonormal basis $\{X_i\}_i$ of \mathcal{X}_1 and extend it to an orthonormal basis $\{X_i\}_i \cup \{S_j\}_j$ of \mathcal{X}_2 . Then,

$$\begin{aligned} P_{\mathcal{X}_1}(P_{\mathcal{X}_2}(v)) &= \sum_i \langle X_i, v \rangle_{\mathcal{V}} P_{\mathcal{X}_1}(X_i) + \sum_j \langle S_j, v \rangle_{\mathcal{V}} P_{\mathcal{X}_1}(S_j). \end{aligned} \quad (\text{B.7})$$

But thanks to $P_{\mathcal{X}_1}(X_i) = X_i$ for all i and $P_{\mathcal{X}_1}(S_j) = 0$ for all j , the claim follows. □

Proposition 4 is now proven as a special case of the following proposition.

Proposition B.4. Let \mathcal{V} be any real inner product space of finite dimension, and let $\mathcal{X} \subseteq \mathcal{V}$ and $\mathcal{Y} \subseteq \mathcal{V}$ be vector subspaces thereof equipped with the inner product inherited from \mathcal{V} . Let

$$\mathcal{Z} := P_{\mathcal{Y}}(\mathcal{X}). \quad (\text{B.8})$$

\mathcal{Z} is a vector subspace which we equip with the inner product inherited from \mathcal{V} . Then, for all $x \in \mathcal{X}$, for all $y \in \mathcal{Y}$,

$$\langle x, y \rangle_{\mathcal{V}} = \langle P_{\mathcal{Z}}(x), P_{\mathcal{Z}}(y) \rangle_{\mathcal{Z}}. \quad (\text{B.9})$$

Proof. By lemma B.2, and due to $y \in \mathcal{Y}$,

$$\langle x, y \rangle_{\mathcal{V}} = \langle P_{\mathcal{Y}}(x), y \rangle_{\mathcal{Y}} \quad (\text{B.10})$$

But now $P_{\mathcal{Y}}(x) \in \mathcal{Z}$, so that

$$\langle x, y \rangle_{\mathcal{V}} = \langle P_{\mathcal{Y}}(x), P_{\mathcal{Z}}(y) \rangle_{\mathcal{Z}}. \quad (\text{B.11})$$

Then, thanks to lemma B.3, and using $\mathcal{Z} = P_{\mathcal{Y}}(\mathcal{X}) \subseteq \mathcal{Y}$, for any $x \in \mathcal{X}$ it holds that

$$P_{\mathcal{Z}}(x) = P_{\mathcal{Z}}(P_{\mathcal{Y}}(x)). \quad (\text{B.12})$$

But $P_{\mathcal{Y}}(x) \in \mathcal{Z}$, so that it actually holds that

$$P_{\mathcal{Z}}(x) = P_{\mathcal{Y}}(x), \quad (\text{B.13})$$

and then equation (B.11) becomes

$$\langle x, y \rangle_{\mathcal{V}} = \langle P_{\mathcal{Z}}(x), P_{\mathcal{Z}}(y) \rangle_{\mathcal{Z}}. \quad \square \quad (\text{B.14})$$

Lemma B.5. Let \mathcal{V} be any real inner product space of finite dimension. Let $\mathcal{X} \subseteq \mathcal{V}$ and $\mathcal{Y} \subseteq \mathcal{V}$ be vector subspaces. Let $\mathbf{x} \subseteq \mathcal{X}$ be any spanning set of \mathcal{X} , and let $\mathbf{y} \subseteq \mathcal{Y}$ be any spanning set of \mathcal{Y} . Let

$$\mathcal{Z} := P_{\mathcal{Y}}(\mathcal{X}). \quad (\text{B.15})$$

Then,

$$\text{span}(P_{\mathcal{Z}}(\mathbf{x})) = \mathcal{Z}, \quad (\text{B.16a})$$

$$\text{span}(P_{\mathcal{Z}}(\mathbf{y})) = \mathcal{Z}. \quad (\text{B.16b})$$

Proof. First, consider:

$$\begin{aligned} \text{span}(P_{\mathcal{Z}}(\mathbf{x})) &= P_{\mathcal{Z}}(\text{span}(\mathbf{x})) = P_{\mathcal{Z}}(\mathcal{X}) \\ &= P_{\mathcal{Z}}(P_{\mathcal{Y}}(\mathcal{X})) = P_{\mathcal{Y}}(\mathcal{X}) = \mathcal{Z}, \end{aligned} \quad (\text{B.17})$$

where we used $\mathcal{Z} \subseteq \mathcal{Y}$ and lemma B.3, to conclude $P_{\mathcal{Z}}(\mathcal{X}) = P_{\mathcal{Z}}(P_{\mathcal{Y}}(\mathcal{X}))$. Now, consider

$$\text{span}(P_{\mathcal{Z}}(\mathbf{y})) = P_{\mathcal{Z}}(\mathcal{Y}). \quad (\text{B.18})$$

Let $\{Z_i\}_i$ be an orthonormal basis of \mathcal{Z} , and extend it to an orthonormal basis $\{Z_i\}_i \cup \{Y_j\}_j$ of \mathcal{Y} (indeed, \mathcal{Z} is a subset of \mathcal{Y}). Then,

$$\begin{aligned} P_{\mathcal{Z}}(\mathcal{Y}) &= P_{\mathcal{Z}}\left(\left\{\sum_i \alpha_i Z_i + \sum_j \beta_j Y_j : \alpha_i, \beta_j \in \mathbb{R}\right\}\right) \\ &= \left\{\sum_i \alpha_i Z_i : \alpha_i \in \mathbb{R}\right\} = \mathcal{Z}. \quad \square \end{aligned} \quad (\text{B.19})$$

Lemma B.5 can be specialized as follows.

Corollary B.6. *The projected states and effects span the whole reduced space:*

$$\text{span}(P_{\mathcal{R}}(\mathbf{s})) = \mathcal{R}, \quad (\text{B.20a})$$

$$\text{span}(P_{\mathcal{R}}(\mathbf{e})) = \mathcal{R}. \quad (\text{B.20b})$$

Lemma B.7. *For all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$,*

$$\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1. \quad (\text{B.21})$$

Proof. For all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, there exists $\rho \in \mathbf{s}$ such that $\bar{\rho} = P_{\mathcal{R}}(\rho)$. Then,

$$\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = \langle P_{\mathcal{R}}(\rho), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}}. \quad (\text{B.22})$$

Using proposition 4,

$$\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = \langle \rho, \mathbb{1}_{\mathcal{H}} \rangle_{\mathcal{L}(\mathcal{H})} = \text{Tr}_{\mathcal{H}}[\rho]. \quad (\text{B.23})$$

The claim then follows from the fact that for all $\rho \in \mathbf{s} \subseteq \mathcal{S}(\mathcal{H})$, $\text{Tr}_{\mathcal{H}}[\rho] = 1$. \square

We now prove the following lemma which will be used to derive the unit separability criterion for the existence of an operationally noncontextual ontological model, theorem 2.

Lemma B.8. *For all $\sigma \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$, the trace of σ satisfies*

$$\langle \sigma, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} \geq 0 \quad (\text{B.24})$$

with equality if and only if $\sigma = 0$.

Proof. Due to $P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \in P_{\mathcal{R}}(\mathbf{e})$, for all $\sigma \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ it holds that

$$\langle \sigma, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} \geq 0. \quad (\text{B.25})$$

Now consider the case when $\sigma \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ and $\langle \sigma, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 0$. For all $E \in \mathbf{e}$, there exists $\{E_k \in \mathbf{e}\}_k$ such that $\mathbb{1}_{\mathcal{H}} = E + \sum_k E_k$, so that

$$\begin{aligned} 0 &= \langle \sigma, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} \\ &= \langle \sigma, P_{\mathcal{R}}(E) \rangle_{\mathcal{R}} + \sum_k \langle \sigma, P_{\mathcal{R}}(E_k) \rangle_{\mathcal{R}}. \end{aligned} \quad (\text{B.26})$$

However, all the terms $\langle \sigma, P_{\mathcal{R}}(E) \rangle_{\mathcal{R}}$ and $\langle \sigma, P_{\mathcal{R}}(E_k) \rangle_{\mathcal{R}}$ are nonnegative and sum to zero, which implies that they are all zero. This shows that for all $E \in \mathbf{e}$,

$$\langle \sigma, P_{\mathcal{R}}(E) \rangle_{\mathcal{R}} = 0. \quad (\text{B.27})$$

But thanks to corollary B.6 and the non-degeneracy of the inner product, this means that $\sigma = 0$. The other direction is trivial. \square

Lemma B.9. *It is impossible with the assumptions of the main text that $P_{\mathcal{R}}(\mathbf{s}) = \{0\}$ or that $P_{\mathcal{R}}(\mathbf{e}) = \{0\}$. This implies that it is impossible that $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = \{0\}$ or that $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = \{0\}$. As a corollary, the cone $\text{Prod}(\mathbf{s}, \mathbf{e})$ is never the trivial cone $\{0 \in \mathcal{R} \otimes \mathcal{R}\}$, nor is the convex cone $\text{Sep}(\mathbf{s}, \mathbf{e})$ the trivial convex set $\{0 \in \mathcal{R} \otimes \mathcal{R}\}$.*

Proof. Suppose that $P_{\mathcal{R}}(\mathbf{s}) = \{0\}$. This implies, for any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, that $\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 0$, which is a contradiction to lemma B.7.

Suppose now $P_{\mathcal{R}}(\mathbf{e}) = \{0\}$. Then, since $\mathbf{s} \neq \emptyset$ according to definition 1, choose any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$. Then, the fact that $P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \in P_{\mathcal{R}}(\mathbf{e})$ leads to $\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = \langle \bar{\rho}, 0 \rangle_{\mathcal{R}} = 0$ which again violates lemma B.7.

Then, note that $P_{\mathcal{R}}(\mathbf{e}) \subseteq P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$. Thus, if $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = \{0\}$, also $P_{\mathcal{R}}(\mathbf{e}) = \{0\}$, since $P_{\mathcal{R}}(\mathbf{e}) \neq \emptyset$ according to definition 2. But $P_{\mathcal{R}}(\mathbf{e}) = \{0\}$ has been shown to be impossible.

For the other case, note that $P_{\mathcal{R}}(\mathbf{s}) \subseteq P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$, so that if $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = \{0\}$, then also $P_{\mathcal{R}}(\mathbf{s}) = \{0\}$ since $P_{\mathcal{R}}(\mathbf{s}) \neq \emptyset$ according to definition 1. This has been shown to be impossible. \square

Proposition 7. *We let a minimal Spekkens noncontextual ontological model for (\mathbf{s}, \mathbf{e}) be defined as in definition 6 but under the replacement of*

$$P_{\mathcal{R}} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{R} (\subseteq \mathcal{L}(\mathcal{H})) \quad (\text{2.14a})$$

with

$$\mathbb{1}_{\mathcal{L}(\mathcal{H})} : \mathcal{L}(\mathcal{H}) \rightarrow \mathcal{L}(\mathcal{H}). \quad (\text{2.14b})$$

In particular, the ontic state mapping μ and response function mapping ξ , in this minimal Spekkens noncontextual ontological model, have as quantum arguments $\rho \in \mathbf{s}$ and $E \in \mathbf{e}$, respectively, instead of $P_{\mathcal{R}}(\rho) \in P_{\mathcal{R}}(\mathbf{s})$ and $P_{\mathcal{R}}(E) \in P_{\mathcal{R}}(\mathbf{e})$. Then, it holds that

- (i) *any (\mathbf{s}, \mathbf{e}) that admits an operationally noncontextual ontological model (definition 6) also admits a minimal Spekkens noncontextual ontological model;*
- (ii) *there exists (\mathbf{s}, \mathbf{e}) that admits a minimal Spekkens noncontextual ontological model but no operationally noncontextual ontological model (definition 6).*

Proof. The item (i) is trivial: any operationally noncontextual ontological model is also a valid minimal Spekkens noncontextual ontological model where the ontic primitives happen not to depend on the part $\rho - P_{\mathcal{R}}(\rho)$ of the states $\rho \in \mathbf{s}$ and the part $E - P_{\mathcal{R}}(E)$ of the effects $E \in \mathbf{e}$.

For the item (ii), we first choose any prepare-and-measure scenario $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ on some finite dimensional Hilbert space $\tilde{\mathcal{H}}$ and where $\tilde{\mathbf{s}}$ and $\tilde{\mathbf{e}}$ take the form

$$\tilde{\mathbf{s}} = \text{conv}(\{\tilde{\rho}_i \in \mathcal{L}(\tilde{\mathcal{H}})\}_{i=1}^N), \quad (\text{B.28a})$$

$$\tilde{\mathbf{e}} = \text{conv}(\{\tilde{E}_j \in \mathcal{L}(\tilde{\mathcal{H}})\}_{j \in J}) \quad (\text{B.28b})$$

for some finite $N \in \mathbb{N}$ and some discrete or continuous range J ; this $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ is assumed to satisfy definitions 1 and 2. We assume that this prepare-and-measure scenario does not admit an operationally noncontextual ontological model — one can pick such a scenario from, e.g., the example of section 3.3.2, with a sufficiently fine-grained polyhedral inner approximation (in the sense of section 4.4.2) of the state space \mathbf{s} . Let us introduce the Hilbert space

$$\mathcal{H} := \tilde{\mathcal{H}} \otimes \mathbb{C}^N. \quad (\text{B.29})$$

Despite the tensor product structure, we think of this Hilbert space as describing one single system whose Hilbert space happens to be isomorphic to such a tensor product structure. Let $\{|\phi_i\rangle\}_{i=1}^N$ be an orthonormal basis of \mathbb{C}^N . We define, for all $i = 1, \dots, N$ and $j \in J$ the states and effects

$$\rho_i := \tilde{\rho}_i \otimes |\phi_i\rangle\langle\phi_i|, \quad (\text{B.30a})$$

$$E_j := \tilde{E}_j \otimes \mathbb{1}_N. \quad (\text{B.30b})$$

The scenario (\mathbf{s}, \mathbf{e}) where

$$\mathbf{s} := \text{conv}(\{\rho_i \in \mathcal{L}(\mathcal{H})\}_{i=1}^N), \quad (\text{B.31a})$$

$$\mathbf{e} := \text{conv}(\{E_j \in \mathcal{L}(\mathcal{H})\}_{j \in J}) \quad (\text{B.31b})$$

satisfies definitions 1 and 2 and we claim that it admits a minimal Spekkens noncontextual ontological model but no operationally noncontextual ontological model. To see that it does not admit any operationally noncontextual ontological model, we first note that for all $i = 1, \dots, N$, for all $j \in J$,

$$\text{Tr}_{\mathcal{H}}[\rho_i E_j] = \text{Tr}_{\tilde{\mathcal{H}}}[\tilde{\rho}_i \tilde{E}_j]. \quad (\text{B.32})$$

Thus, proposition 8 applies here and implies that since $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ does not admit an operationally noncontextual ontological model, also (\mathbf{s}, \mathbf{e}) does not admit such a model.

However, we can construct a minimal Spekkens noncontextual ontological model for (\mathbf{s}, \mathbf{e}) . Define $\Lambda := \{1, \dots, N\}$. For all $\lambda \in \Lambda$, $\rho \in \mathbf{s}$ and all $E \in \mathbf{e}$,

$$\mu(\rho, \lambda) := \frac{\text{Tr}_{\mathcal{H}}[\rho \rho_\lambda]}{\text{Tr}_{\tilde{\mathcal{H}}}[\tilde{\rho}_\lambda^2]}, \quad (\text{B.33a})$$

$$\xi(E, \lambda) := \text{Tr}_{\mathcal{H}}[E \rho_\lambda]. \quad (\text{B.33b})$$

Since all the matrices involved are nonnegative, the nonnegativity requirements (2.8c) and (2.12c) (under the replacement $P_{\mathcal{R}} \rightarrow \mathbb{1}_{\mathcal{L}(\mathcal{H})}$ of (2.14)) are verified. By linearity of the trace, it is also clear that

the adapted versions of (2.8d) and (2.12d) are also verified. To see that the normalization of μ holds, note that for all $\rho \in \mathbf{s}$ there exist $\{p_i \in \mathbb{R}_{\geq 0}\}_{i=1}^N$ with $\sum_{i=1}^N p_i = 1$ such that $\rho = \sum_{i=1}^N p_i \rho_i$. Then, one can compute

$$\begin{aligned} \sum_{\lambda \in \Lambda} \mu(\rho, \lambda) &= \sum_{\lambda, i=1}^N p_i \frac{\text{Tr}_{\mathcal{H}}[\rho_i \rho_\lambda]}{\text{Tr}_{\tilde{\mathcal{H}}}[\tilde{\rho}_\lambda^2]} \\ &= \sum_{\lambda, i=1}^N p_i \delta_{i\lambda} = 1. \end{aligned} \quad (\text{B.34})$$

so that (2.8b) is verified, and so is clearly the normalization requirement (2.12b) (again under the same replacement) given the linearity of ξ . It remains to compute, for any $\rho = \sum_{i=1}^N p_i \rho_i$ as previously defined and $E \in \mathbf{e}$,

$$\begin{aligned} \sum_{\lambda=1}^N \mu(\rho, \lambda) \xi(E, \lambda) &= \sum_{i, \lambda=1}^N p_i \delta_{i\lambda} \text{Tr}_{\mathcal{H}}[E \rho_\lambda] \\ &= \text{Tr}_{\mathcal{H}}[E \rho], \end{aligned} \quad (\text{B.35})$$

thus reproducing the adapted version of (2.13). This concludes the proof. \square

B.3 An example correspondence between reduced space projection and partial trace

In section 2.4, we mentioned how taking a reduced space projection and a partial trace can be thought of as similar way of restricting the focus to a system. Consider a system A , with Hilbert space \mathcal{H}_A , for which we would like to build a noncontextual ontological model, and a system B , with Hilbert space \mathcal{H}_B which is an auxiliary system that can be thought of as being “close by” with respect to system A . Suppose that an agent prepares states $\{\rho_i^{(A)} \in \mathcal{L}(\mathcal{H}_A)\}_i$ on system A , but each time they do so, there is a certain influence on the auxiliary system B so that the agent effectively prepares the joint states $\{\rho_i^{(A)} \otimes \sigma_i^{(B)} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)\}_i$. The measurements of the agent are however strictly restricted to system A , and hence the effects are of the form $\{E_j^{(A)} \otimes \mathbb{1}_{\mathcal{H}_B}\}_j$.

We offer two ways of formulating a noncontextual ontological model for system A . The first is to use definition 6, with the reduced space prescription implying that the ontic state distributions and ontic response functions only depend on $\{\rho_i^{(A)}\}_i$ and $\{E_j^{(A)}\}_j$ (this will be formalized in proposition B.10). Alternatively, one may start from the basics of Spekkens noncontextual ontological models. A potential ontological model for the joint system of A and B could in principle be allowed to depend on both the quantum primitives of A and those of B . However, a Spekkens

noncontextual ontological model for system A should only depend on the reduced quantum state of A . Taking the partial trace on the system B implements this reduction, and we see that the reduced space projection and the partial trace implement the same kind (depending on whether the prepare-and-measure scenario reduced to system A is tomographically complete) of restriction for the dependency of the ontic primitives in this case. Formally, one may for instance write down the following proposition.

Proposition B.10. *Let $I, J \subseteq \mathbb{N}$. We consider a bipartite Hilbert space $\mathcal{H}^{(AB)} = \mathcal{H}^{(A)} \otimes \mathcal{H}^{(B)}$, a set of density matrix $\{\rho_i^{(AB)} \in \mathcal{L}(\mathcal{H}^{(AB)})\}_{i \in I}$ and of effects $\{E_j^{(A)} \in \mathcal{L}(\mathcal{H}^{(A)})\}_{j \in J} \supseteq \{0^{(A)}, \mathbb{1}^{(A)}\}$. We define*

$$\mathbf{s}^{(AB)} = \text{conv}\left(\{\rho_i^{(AB)}\}_{i \in I}\right), \quad (\text{B.36a})$$

$$\mathbf{e}^{(AB)} = \text{conv}\left(\{E_j^{(A)} \otimes \mathbb{1}^{(B)}\}_{j \in J}\right), \quad (\text{B.36b})$$

as well as

$$\mathbf{s}^{(A)} = \text{conv}\left(\{\text{Tr}_B[\rho_i^{(AB)}]\}_{i \in I}\right), \quad (\text{B.37a})$$

$$\mathbf{e}^{(A)} = \text{conv}\left(\{E_j^{(A)}\}_{j \in J}\right). \quad (\text{B.37b})$$

$\mathbf{e}^{(AB)}$ and $\mathbf{e}^{(A)}$ are assumed to satisfy definition 2. We further assume that

$$\text{span}\left(\mathbf{s}^{(A)}\right) = \text{span}\left(\mathbf{e}^{(A)}\right) = \mathcal{L}(\mathcal{H}). \quad (\text{B.38})$$

Then the scenario $(\mathbf{s}^{(AB)}, \mathbf{e}^{(AB)})$ admits an operationally noncontextual ontological model (definition 6) if and only if the scenario $(\mathbf{s}^{(A)}, \mathbf{e}^{(A)})$ admits a minimal Spekkens noncontextual ontological model (as defined in proposition 7).

Proof. The scenarios $(\mathbf{s}^{(AB)}, \mathbf{e}^{(AB)})$ and $(\mathbf{s}^{(A)}, \mathbf{e}^{(A)})$ satisfy the assumptions of proposition 8, and hence, there exists an operationally noncontextual ontological model for one if and only if there exists an operationally noncontextual ontological model with identical ontic space Λ for the other. Furthermore, when considering $(\mathbf{s}^{(A)}, \mathbf{e}^{(B)})$, and thanks to the assumption of equation (B.38), the projection on the reduced space $P_{\mathcal{R}}(\cdot)$ equals the identity map $\mathbb{1}_{\mathcal{L}(\mathcal{H})}(\cdot)$. Thus, a minimal Spekkens noncontextual ontological model (defined in proposition 7) is actually equivalent to an operationally noncontextual ontological for the scenario $(\mathbf{s}^{(A)}, \mathbf{e}^{(A)})$. \square

B.4 Linear extensions to the ontic mappings

We now prove proposition 9 as follows: proposition B.12 proves how the extension is built for μ , while proposition B.13 considers the extension for ξ . Both propositions B.12 and B.13 will make use of the following lemma. Note that convex and conic sets are defined in appendix A.

Lemma B.11. *Let \mathcal{V} be a real inner product space of finite dimension, and let $\mathcal{C} \subseteq \mathcal{V}$ be a convex cone such that*

$$\text{span}(\mathcal{C}) = \mathcal{V}. \quad (\text{B.39})$$

Let $f : \mathcal{C} \rightarrow \mathbb{R}$ be any function that satisfies the following two properties¹¹:

$$\forall \alpha \in \mathbb{R}_{\geq 0}, \forall c \in \mathcal{C} : f(\alpha c) = \alpha f(c), \quad (\text{B.40a})$$

$$\forall c_1, c_2 \in \mathcal{C} : f(c_1 + c_2) = f(c_1) + f(c_2). \quad (\text{B.40b})$$

Then, there exists a unique function

$$g : \mathcal{V} \rightarrow \mathbb{R} \quad (\text{B.41})$$

that is linear and that satisfies

$$g|_{\mathcal{C}} = f, \quad (\text{B.42})$$

which means that

$$\forall c \in \mathcal{C} : g(c) = f(c). \quad (\text{B.43})$$

Proof. First, we will show that f satisfies the following: for all $\alpha_i \in \mathbb{R}$, for all $c_i \in \mathcal{C}$ such that $\sum_i \alpha_i c_i \in \mathcal{C}$,

$$f\left(\sum_i \alpha_i c_i\right) = \sum_i \alpha_i f(c_i). \quad (\text{B.44})$$

To start with, let $c_1, c_2 \in \mathcal{C}$ be such that $c_1 - c_2 \in \mathcal{C}$. Then, using equation (B.40b),

$$f((c_1 - c_2) + c_2) = f(c_1 - c_2) + f(c_2), \quad (\text{B.45})$$

which is equivalent to

$$f(c_1 - c_2) = f(c_1) - f(c_2). \quad (\text{B.46})$$

Then, let $I = \{1, \dots, n\}$ for some $n \in \mathbb{N}$, let $\{\alpha_i \in \mathbb{R}\}_{i \in I}$ and $\{c_i \in \mathcal{C}\}_{i \in I}$ be such that $\sum_{i \in I} \alpha_i c_i \in \mathcal{C}$. Then, let¹²

$$I_+ := \{i \in I : \text{sgn}(\alpha_i) = +1\}, \quad (\text{B.47a})$$

$$I_0 := \{i \in I : \alpha_i = 0\}, \quad (\text{B.47b})$$

$$I_- := \{i \in I : \text{sgn}(\alpha_i) = -1\}. \quad (\text{B.47c})$$

Using these,

$$\begin{aligned} f\left(\sum_{i \in I} \alpha_i c_i\right) &= f\left(\sum_{i \in I_+} |\alpha_i| c_i - \sum_{j \in I_-} |\alpha_j| c_j\right). \end{aligned} \quad (\text{B.48})$$

Clearly, both sums $\sum_{i \in I_{\pm}} |\alpha_i| c_i$ belong to \mathcal{C} . Using equation (B.46),

$$\begin{aligned} f\left(\sum_{i \in I} \alpha_i c_i\right) &= f\left(\sum_{i \in I_+} |\alpha_i| c_i\right) - f\left(\sum_{j \in I_-} |\alpha_j| c_j\right). \end{aligned} \quad (\text{B.49})$$

¹¹Observe that $\forall c_1, c_2 \in \mathcal{C}$, $\alpha \in \mathbb{R}_{\geq 0}$, $\alpha c_1 \in \mathcal{C}$, $c_1 + c_2 \in \mathcal{C}$ since \mathcal{C} is convex and conic.

¹²The sign function $\text{sgn}(x)$ is -1 if $x < 0$, 0 if $x = 0$ and 1 if $x > 0$. The absolute value is $|x| = \text{sgn}(x)x$.

Then, using repeatedly equation (B.40b) to expand the sums, as well as equation (B.40a) to extract the positive factors, and bringing back the trivial summands $\{i \in I_0\}$, we obtain equation (B.44).

We can now easily extend f to a linear map g whose domain is \mathcal{V} . To do so, choose a basis $\{S_i \in \mathcal{C}\}_{i=1}^{\dim(\mathcal{V})}$ of \mathcal{V} . This is always possible thanks to the assumption (B.39). For any $v \in \mathcal{V}$ and $i \in I$, let $s_i(v) \in \mathbb{R}$ be the coordinate of v in the basis $\{S_i\}_i$. Of course, $s_i : \mathcal{V} \rightarrow \mathbb{R}$ is linear for each i . Then, define for all $v \in \mathcal{V}$:

$$g(v) := \sum_i s_i(v) f(S_i). \quad (\text{B.50})$$

This choice for g is unique: indeed, if $g|_{\mathcal{C}} = f$ is to hold, then in particular g has to agree with f on the basis elements $\{S_i\}_i$, but the action of a linear map on a basis completely determines its action on the whole space. Thanks to equation (B.44), it is then easy to see that indeed

$$g|_{\mathcal{C}} = f. \quad (\text{B.51})$$

Explicitly, for all $c \in \mathcal{C}$,

$$\begin{aligned} g(c) &= g\left(\sum_i s_i(c) S_i\right) = \sum_i s_i(c) f(S_i) \\ &= f\left(\sum_i s_i(c) S_i\right) = f(c). \quad \square \end{aligned} \quad (\text{B.52})$$

Proposition B.12. *Let $\lambda \in \Lambda$ be arbitrary. Starting from the convex-linear mapping*

$$\mu(\cdot, \lambda) : P_{\mathcal{R}}(\mathbf{s}) \rightarrow \mathbb{R}_{\geq 0}, \quad (\text{B.53})$$

there exists a unique linear extension

$$\mu_{\text{ext}}(\cdot, \lambda) : \mathcal{R} \rightarrow \mathbb{R}. \quad (\text{B.54})$$

Proof. Throughout this proof, we omit the fixed argument λ . Recalling $P_{\mathcal{R}}(\mathbf{s}) \subseteq \text{coni}(P_{\mathcal{R}}(\mathbf{s})) \subseteq \mathcal{R}$, let us first look for the intermediate function

$$\mu_{\text{cone}} : \text{coni}(P_{\mathcal{R}}(\mathbf{s})) \rightarrow \mathbb{R}, \quad (\text{B.55})$$

that satisfies, for all $\alpha \in \mathbb{R}_{\geq 0}$, for all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$,

$$\mu_{\text{cone}}(\alpha \bar{\rho}) = \alpha \mu(\bar{\rho}). \quad (\text{B.56})$$

Equation (B.56) is a necessary condition for linearity which can be formulated given the restriction of the domain to a conic set. Recall lemma B.7:

$$\forall \bar{\rho} \in P_{\mathcal{R}}(\mathbf{s}) : \text{Tr}_{\mathcal{H}}[\bar{\rho}] = \langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1. \quad (\text{B.57})$$

Thus, the right-hand side of equation (B.56) may be rewritten as follows when $\alpha \neq 0$:

$$\mu_{\text{cone}}(\alpha \bar{\rho}) = \text{Tr}_{\mathcal{H}}[\alpha \bar{\rho}] \mu\left(\frac{\alpha \bar{\rho}}{\text{Tr}_{\mathcal{H}}[\alpha \bar{\rho}]}\right). \quad (\text{B.58})$$

This shows that the unique choice for μ_{cone} is the following: for all $r \in \text{coni}(P_{\mathcal{R}}(\mathbf{s}))$,

$$\mu_{\text{cone}}(r) := \begin{cases} 0 & \text{if } r = 0, \\ \text{Tr}_{\mathcal{H}}[r] \mu\left(\frac{r}{\text{Tr}_{\mathcal{H}}[r]}\right) & \text{else.} \end{cases} \quad (\text{B.59})$$

One important property that μ_{cone} satisfies is that it agrees with μ when the argument is in $P_{\mathcal{R}}(\mathbf{s})$.

Let us now show that μ_{cone} satisfies the assumptions of lemma B.11. Clearly, using corollary B.6,

$$\text{span}(\text{coni}(P_{\mathcal{R}}(\mathbf{s}))) = \mathcal{R}. \quad (\text{B.60})$$

Furthermore, since \mathbf{s} is convex and $P_{\mathcal{R}}(\cdot)$ is linear, $P_{\mathcal{R}}(\mathbf{s})$ is convex. Therefore $\text{coni}(P_{\mathcal{R}}(\mathbf{s}))$ is a convex cone. This property together with equation (B.60) allows to verify equation (B.39). Let us now prove that μ_{cone} satisfies (B.40a): let $\alpha \in \mathbb{R}_{\geq 0}$, and $r \in \text{coni}(P_{\mathcal{R}}(\mathbf{s}))$. If either $\alpha = 0$ or $r = 0$, then clearly $\mu_{\text{cone}}(\alpha r) = \alpha \mu_{\text{cone}}(r)$. If both $\alpha \neq 0$ and $r \neq 0$, then

$$\begin{aligned} \mu_{\text{cone}}(\alpha r) &= \text{Tr}_{\mathcal{H}}[\alpha r] \mu\left(\frac{\alpha r}{\text{Tr}_{\mathcal{H}}[\alpha r]}\right) \\ &= \alpha \text{Tr}_{\mathcal{H}}[r] \mu\left(\frac{r}{\text{Tr}_{\mathcal{H}}[r]}\right) = \alpha \mu_{\text{cone}}(r). \end{aligned} \quad (\text{B.61})$$

Thus μ_{cone} satisfies (B.40a). Now let $r, s \in \text{coni}(P_{\mathcal{R}}(\mathbf{s}))$, and we will verify (B.40b). If $r = 0$ or $s = 0$, or both, then trivially $\mu_{\text{cone}}(r + s) = \mu_{\text{cone}}(r) + \mu_{\text{cone}}(s)$. Otherwise if $r, s \neq 0$,

$$\begin{aligned} \mu_{\text{cone}}(r + s) &= \text{Tr}_{\mathcal{H}}[r + s] \mu\left(\frac{r + s}{\text{Tr}_{\mathcal{H}}[r + s]}\right) \\ &= \text{Tr}_{\mathcal{H}}[r + s] \mu\left(\frac{r}{\text{Tr}_{\mathcal{H}}[r + s]} + \frac{s}{\text{Tr}_{\mathcal{H}}[r + s]}\right) \\ &= \text{Tr}_{\mathcal{H}}[r + s] \mu\left(p \frac{r}{\text{Tr}_{\mathcal{H}}[r]} + (1 - p) \frac{s}{\text{Tr}_{\mathcal{H}}[s]}\right) \end{aligned} \quad (\text{B.62})$$

where we defined for brevity $p = \text{Tr}_{\mathcal{H}}[r] / \text{Tr}_{\mathcal{H}}[r + s]$. By the convex-linearity of μ as in equation (2.8d), however, this becomes

$$\begin{aligned} \mu_{\text{cone}}(r + s) &= \\ \text{Tr}_{\mathcal{H}}[r + s] &\left(p \mu\left(\frac{r}{\text{Tr}_{\mathcal{H}}[r]}\right) + (1 - p) \mu\left(\frac{s}{\text{Tr}_{\mathcal{H}}[s]}\right) \right) \\ &= \text{Tr}_{\mathcal{H}}[r] \mu\left(\frac{r}{\text{Tr}_{\mathcal{H}}[r]}\right) + \text{Tr}_{\mathcal{H}}[s] \mu\left(\frac{s}{\text{Tr}_{\mathcal{H}}[s]}\right) \\ &= \mu_{\text{cone}}(r) + \mu_{\text{cone}}(s). \end{aligned} \quad (\text{B.63})$$

Thus, μ_{cone} fully satisfies the assumptions of lemma B.11. This shows that there exists a unique linear map

$$\mu_{\text{ext}} : \mathcal{R} \rightarrow \mathbb{R} \quad (\text{B.64})$$

such that

$$\mu_{\text{ext}}|_{\text{coni}(P_{\mathcal{R}}(\mathbf{s}))} = \mu_{\text{cone}}. \quad (\text{B.65})$$

But then, by subset inclusion, and because μ_{cone} extends μ ,

$$\begin{aligned}\mu_{\text{ext}}|_{P_{\mathcal{R}}(\mathbf{s})} &= \left(\mu_{\text{ext}}|_{\text{coni}(P_{\mathcal{R}}(\mathbf{s}))} \right) \Big|_{P_{\mathcal{R}}(\mathbf{s})} \\ &= \mu_{\text{cone}}|_{P_{\mathcal{R}}(\mathbf{s})} = \mu. \quad \square \quad (\text{B.66})\end{aligned}$$

Proposition B.13. *Let $\lambda \in \Lambda$ be arbitrary. Starting from the convex-linear mapping*

$$\xi(\cdot, \lambda) : P_{\mathcal{R}}(\mathbf{e}) \rightarrow \mathbb{R}_{\geq 0}, \quad (\text{B.67})$$

there exists a unique linear extension

$$\xi_{\text{ext}}(\cdot, \lambda) : \mathcal{R} \rightarrow \mathbb{R}. \quad (\text{B.68})$$

Proof. Again, we omit the fixed argument $\lambda \in \Lambda$ throughout the proof. First, note that $\xi(0) = 0$.¹³ This follows easily from (2.12b).

Then, we show the following property of ξ . If there exists $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$ and $\alpha \in \mathbb{R}_{\geq 0}$ such that also $\alpha\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$, then,

$$\xi(\alpha\bar{E}) = \alpha\xi(\bar{E}). \quad (\text{B.69})$$

Without loss of generality assume that $\alpha \leq 1$ (indeed, if $\alpha > 1$, one may simply interchange the role of $\alpha\bar{E}$ and \bar{E}). Then,

$$\begin{aligned}\xi(\alpha\bar{E}) &= \xi(\alpha\bar{E} + (1-\alpha) \cdot 0) = \alpha\xi(\bar{E}) + (1-\alpha)\xi(0) \\ &= \alpha\xi(\bar{E}), \quad (\text{B.70})\end{aligned}$$

where we used the convex linearity of ξ as in equation (2.12d).

Let us now look for the intermediate extension

$$\xi_{\text{cone}} : \text{coni}(P_{\mathcal{R}}(\mathbf{e})) \rightarrow \mathbb{R} \quad (\text{B.71})$$

that satisfies, for all $\alpha \in \mathbb{R}_{\geq 0}$, for all $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$,

$$\xi_{\text{cone}}(\alpha\bar{E}) = \alpha\xi(\bar{E}). \quad (\text{B.72})$$

Equation (B.72) is a necessary condition for linearity which can be formulated on a conic domain. The only way to define this extension is clearly the following: for all $r \in \text{coni}(P_{\mathcal{R}}(\mathbf{e}))$, by definition A.5 of the conic hull there always exist $\alpha_r \in \mathbb{R}_{\geq 0}$ and $\bar{E}_r \in P_{\mathcal{R}}(\mathbf{e})$ such that $r = \alpha_r\bar{E}_r$, and then define

$$\xi_{\text{cone}}(r) := \alpha_r\xi(\bar{E}_r). \quad (\text{B.73})$$

This definition is meaningful because it does not depend on the α_r, \bar{E}_r that one chooses. Indeed, suppose that instead of decomposing $r = \alpha_r\bar{E}_r$, one chooses instead $r = \beta_r\bar{F}_r$ where $\beta_r \in \mathbb{R}_{\geq 0}$, $\bar{F}_r \in P_{\mathcal{R}}(\mathbf{e})$. If $r = 0$ the present discussion is irrelevant, so that we

¹³ $0 \in \mathcal{R}$ is automatically in the domain of ξ since $0 \in \mathcal{L}(\mathcal{H})$ is always an allowed effect within \mathbf{e} as required in definition 2.

may assume that $\alpha_r, \beta_r > 0$. Then, the value one obtains with the alternative decomposition $r = \beta_r\bar{F}_r$ is

$$\beta_r\xi(\bar{F}_r) = \beta_r\xi\left(\frac{1}{\beta_r}r\right) = \beta_r\xi\left(\frac{\alpha_r}{\beta_r}\bar{E}_r\right). \quad (\text{B.74})$$

Using equation (B.69) applied to \bar{E}_r and $\bar{F}_r = (\alpha_r/\beta_r)\bar{E}_r$, which both belong to $P_{\mathcal{R}}(\mathbf{e})$, (B.74) becomes

$$\beta_r\xi(\bar{F}_r) = \alpha_r\xi(\bar{E}_r). \quad (\text{B.75})$$

This proves that ξ_{cone} as in (B.73) is well-defined. Also, it is clear that equation (B.72) is verified. It is then easy to see that ξ_{cone} satisfies the first assumption (B.40a) of lemma B.11:

$$\begin{aligned}\forall \alpha \in \mathbb{R}_{\geq 0}, \forall r \in \text{coni}(P_{\mathcal{R}}(\mathbf{e})) : \\ \xi_{\text{cone}}(\alpha r) = \alpha\xi_{\text{cone}}(r). \quad (\text{B.76})\end{aligned}$$

Also, thanks to corollary B.6, the span assumption (B.39) of lemma B.11 is verified in this case: indeed,

$$\text{span}(\text{coni}(P_{\mathcal{R}}(\mathbf{e}))) = \mathcal{R}. \quad (\text{B.77})$$

Let us verify the last assumption (B.40b) of lemma B.11: let $r, s \in \text{coni}(P_{\mathcal{R}}(\mathbf{e}))$, and we will show that

$$\xi_{\text{cone}}(r+s) = \xi_{\text{cone}}(r) + \xi_{\text{cone}}(s). \quad (\text{B.78})$$

Let $\alpha_r, \alpha_s \in \mathbb{R}_{\geq 0}$, $\bar{E}_r, \bar{E}_s \in P_{\mathcal{R}}(\mathbf{e})$ be such that

$$r = \alpha_r\bar{E}_r, \quad (\text{B.79a})$$

$$s = \alpha_s\bar{E}_s. \quad (\text{B.79b})$$

Then, using equation (B.76),

$$\begin{aligned}\xi_{\text{cone}}(r+s) &= \xi_{\text{cone}}(\alpha_r\bar{E}_r + \alpha_s\bar{E}_s) \\ &= (\alpha_r + \alpha_s)\xi_{\text{cone}}\left(\frac{\alpha_r}{\alpha_r + \alpha_s}\bar{E}_r + \frac{\alpha_s}{\alpha_r + \alpha_s}\bar{E}_s\right). \quad (\text{B.80})\end{aligned}$$

Note that by the convexity of \mathbf{e} and of $P_{\mathcal{R}}(\mathbf{e})$, $\left(\frac{\alpha_r}{\alpha_r + \alpha_s}\bar{E}_r + \frac{\alpha_s}{\alpha_r + \alpha_s}\bar{E}_s\right) \in P_{\mathcal{R}}(\mathbf{e})$, so that in fact

$$\begin{aligned}\xi_{\text{cone}}(r+s) \\ = (\alpha_r + \alpha_s)\xi\left(\frac{\alpha_r}{\alpha_r + \alpha_s}\bar{E}_r + \frac{\alpha_s}{\alpha_r + \alpha_s}\bar{E}_s\right). \quad (\text{B.81})\end{aligned}$$

Using the convex-linearity property (2.12d) of ξ ,

$$\begin{aligned}\xi_{\text{cone}}(r+s) \\ = \alpha_r\xi(\bar{E}_r) + \alpha_s\xi(\bar{E}_s) = \xi_{\text{cone}}(r) + \xi_{\text{cone}}(s). \quad (\text{B.82})\end{aligned}$$

Thus also the assumption (B.40b) is verified and we may apply lemma B.11 to conclude that there exists a unique linear map

$$\xi_{\text{ext}} : \mathcal{R} \rightarrow \mathbb{R} \quad (\text{B.83})$$

that agrees with ξ_{cone} on $\text{coni}(P_{\mathcal{R}}(\mathbf{e}))$, and thus also that agrees with ξ of $P_{\mathcal{R}}(\mathbf{e})$. \square

Theorem B.1 (Riesz' representation theorem, see theorem 4.47 in [23]). *Let \mathcal{V} be an arbitrary real inner product space of finite dimension. For any linear map $f : \mathcal{V} \rightarrow \mathbb{R}$, there exists a unique $F \in \mathcal{V}$ such that*

$$\forall v \in \mathcal{V} : f(v) = \langle F, v \rangle_{\mathcal{V}}. \quad (\text{B.84})$$

B.5 Basic criterion for the existence of an operationally noncontextual ontological model

We now restate and prove the basic criterion for the existence of an operationally noncontextual ontological model.

Theorem 1. *Given (\mathbf{s}, \mathbf{e}) that lead to the reduced space \mathcal{R} (definition 3), there exists an operationally noncontextual ontological model with ontic state space Λ (definition 6) if and only if there exist mappings F , σ with ranges*

$$F : \Lambda \rightarrow P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}, \quad (2.22a)$$

$$\sigma : \Lambda \rightarrow P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}, \quad (2.22b)$$

satisfying the normalization condition

$$\forall \lambda \in \Lambda : \langle \sigma(\lambda), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1 \quad (2.23)$$

as well as the consistency requirement: for all $r, s \in \mathcal{R}$,

$$\langle r, s \rangle_{\mathcal{R}} = \int_{\Lambda} d\lambda \langle r, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), s \rangle_{\mathcal{R}}. \quad (2.24)$$

Proof. We start from definition 6 of the operationally noncontextual ontological model, bearing in mind the extended ontic state mapping and extended ontic response function mapping introduced in proposition 9, as well as their representation as scalar products in equation (2.20). By construction, the desired convexity requirements in equations (2.8d) and (2.12d) are automatically verified as a special case of the linearity of the scalar products in (2.20). Let us constrain the mappings $F : \Lambda \rightarrow \mathcal{R}$ and $\sigma : \Lambda \rightarrow \mathcal{R}$ by imposing the relevant nonnegativity constraints (2.8c) and (2.12c):

$$\forall \lambda \in \Lambda, \forall \bar{\rho} \in P_{\mathcal{R}}(\mathbf{s}) : \langle \bar{\rho}, F(\lambda) \rangle_{\mathcal{R}} \geq 0, \quad (\text{B.85a})$$

$$\forall \lambda \in \Lambda, \forall \bar{E} \in P_{\mathcal{R}}(\mathbf{e}) : \langle \sigma(\lambda), \bar{E} \rangle_{\mathcal{R}} \geq 0. \quad (\text{B.85b})$$

Using the definition 10 of the polar cone, this is equivalent to

$$\forall \lambda \in \Lambda : F(\lambda) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}, \quad (\text{B.86a})$$

$$\forall \lambda \in \Lambda : \sigma(\lambda) \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}. \quad (\text{B.86b})$$

This proves that the nonnegativity of the ontic primitives (2.8c) and (2.12c) is equivalent to the ranges of F and σ as in equations (2.22).

The consistency requirement (2.13) in the definition 6 of the operationally noncontextual ontological model reads:

$$\begin{aligned} \forall \bar{\rho} \in P_{\mathcal{R}}(\mathbf{s}), \forall \bar{E} \in P_{\mathcal{R}}(\mathbf{e}) : \\ \langle \bar{\rho}, \bar{E} \rangle_{\mathcal{R}} = \int_{\Lambda} d\lambda \langle \bar{\rho}, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), \bar{E} \rangle_{\mathcal{R}}. \end{aligned} \quad (\text{B.87})$$

Due to $\text{span}(P_{\mathcal{R}}(\mathbf{s})) = \text{span}(P_{\mathcal{R}}(\mathbf{e})) = \mathcal{R}$ (proven in corollary B.6), it is clear that (B.87) implies, and is implied by, the consistency requirement (2.24) of theorem 1.

Let us now show that the normalization of σ as in equation (2.23) is implied by the definition 6 of the operationally noncontextual ontological model. This is easy to see: starting from the normalization (2.12b), we have in particular that $\xi(P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}), \lambda) = 1$ for all $\lambda \in \Lambda$. This translates as equation (2.23).

Let us now prove that the normalization of σ as in equation (2.23) implies the full normalization of the ontic response function (2.12b):

$$\begin{aligned} \forall \lambda \in \Lambda, \forall K \in \mathbb{N} \cup \{+\infty\}, \\ \forall \left\{ E_k \in \mathbf{e} : \sum_{k=1}^K E_k = \mathbb{1}_{\mathcal{H}} \right\} : \\ \sum_{k=1}^K \xi(P_{\mathcal{R}}(E_k), \lambda) = \sum_{k=1}^K \langle \sigma(\lambda), P_{\mathcal{R}}(E_k) \rangle_{\mathcal{R}} \\ = \langle \sigma(\lambda), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1. \end{aligned} \quad (\text{B.88})$$

The normalization of the ontic states as in equation (2.8b) reads: for any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$,

$$\begin{aligned} \int_{\Lambda} d\lambda \mu(\bar{\rho}, \lambda) &= \int_{\Lambda} d\lambda \langle \bar{\rho}, F(\lambda) \rangle_{\mathcal{R}} \\ &= \int_{\Lambda} d\lambda \langle \bar{\rho}, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} \\ &= \langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1. \end{aligned} \quad (\text{B.89})$$

We used first the normalization (2.23) of σ , then the consistency requirement (2.24) and finally lemma B.7 to conclude.

Overall, we have shown that definition 6 implies the structure of theorem 1, and that the latter suffices to recover a valid operationally noncontextual ontological model as in definition 6. \square

The following general lemma proves the alternative expressions $P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} = \mathcal{R} \cap \mathbf{s}^{+\mathcal{L}(\mathcal{H})}$ and $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = \mathcal{R} \cap \mathbf{e}^{+\mathcal{L}(\mathcal{H})}$.

Lemma B.14. *Let \mathcal{V} be a finite dimensional real inner product space. Let $X \subseteq \mathcal{V}$ be any set. Let $\mathcal{U} \subseteq \mathcal{V}$ be a vector subspace of \mathcal{V} equipped with the inner product inherited from \mathcal{V} . It holds that:*

$$P_{\mathcal{U}}(X)^{+\mathcal{U}} = \mathcal{U} \cap X^{+\mathcal{V}}. \quad (\text{B.90})$$

Proof. Let us prove that $P_{\mathcal{U}}(X)^{+u} \subseteq \mathcal{U} \cap X^{+\nu}$. Let $u \in P_{\mathcal{U}}(X)^{+u}$. Then, $u \in \mathcal{U}$ so it suffices to verify $u \in X^{+\nu}$. For all $x \in X$, using lemma B.2,

$$\langle u, x \rangle_{\mathcal{V}} = \langle u, P_{\mathcal{U}}(x) \rangle_{\mathcal{U}} \geq 0, \quad (\text{B.91})$$

where we used $u \in P_{\mathcal{U}}(X)^{+u}$ to conclude. Thus, it holds that $u \in \mathcal{U} \cap X^{+\nu}$.

Let us now prove that $\mathcal{U} \cap X^{+\nu} \subseteq P_{\mathcal{U}}(X)^{+u}$. Let $u' \in \mathcal{U} \cap X^{+\nu}$. For all $\bar{x} \in P_{\mathcal{U}}(X)$, choose $x \in X$ such that $\bar{x} = P_{\mathcal{U}}(x)$. Then, using lemma B.2,

$$\langle u', \bar{x} \rangle_{\mathcal{U}} = \langle u', x \rangle_{\mathcal{V}} \geq 0, \quad (\text{B.92})$$

where we used $u' \in X^{+\nu}$ and $x \in X$ to conclude. \square

Proposition 8. *Let \mathcal{H} and $\tilde{\mathcal{H}}$ denote two finite dimensional Hilbert spaces. Let I be a discrete or continuous range of indices. Consider two scenarios (\mathbf{s}, \mathbf{e}) and $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ that take the form*

$$\mathbf{s} = \text{conv}(\{\rho_k \in \mathcal{L}(\mathcal{H})\}_{k \in I}), \quad (\text{2.15a})$$

$$\mathbf{e} = \text{conv}(\{E_k \in \mathcal{L}(\mathcal{H})\}_{k \in I}), \quad (\text{2.15b})$$

and

$$\tilde{\mathbf{s}} = \text{conv}(\{\tilde{\rho}_k \in \mathcal{L}(\tilde{\mathcal{H}})\}_{k \in I}), \quad (\text{2.16a})$$

$$\tilde{\mathbf{e}} = \text{conv}(\{\tilde{E}_k \in \mathcal{L}(\tilde{\mathcal{H}})\}_{k \in I}), \quad (\text{2.16b})$$

both assumed to satisfy definitions 1 and 2, and that satisfy, for all $k, l \in I$,

$$\text{Tr}_{\mathcal{H}}[\rho_k E_l] = \text{Tr}_{\tilde{\mathcal{H}}}[\tilde{\rho}_k \tilde{E}_l]. \quad (\text{2.17})$$

Then, these two scenarios define reduced spaces $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{e}))$ and $\tilde{\mathcal{R}} = P_{\text{span}(\tilde{\mathbf{e}})}(\text{span}(\tilde{\mathbf{s}}))$ that have the same dimension and the scenario (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model if and only if the scenario $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ does. In fact, if one of the two scenarios admits an operationally noncontextual ontological model with ontic space Λ , so does the other with the same ontic space Λ .

Proof. Define

$$\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})) \text{ and } \tilde{\mathcal{R}} = P_{\text{span}(\tilde{\mathbf{e}})}(\text{span}(\tilde{\mathbf{s}})).$$

Equation (2.17) together with proposition 4 proves that for all $k, l \in I$,

$$\langle P_{\mathcal{R}}(\rho_k), P_{\mathcal{R}}(E_l) \rangle_{\mathcal{R}} = \langle P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k), P_{\tilde{\mathcal{R}}}(\tilde{E}_l) \rangle_{\tilde{\mathcal{R}}}. \quad (\text{B.93})$$

To prove that \mathcal{R} and $\tilde{\mathcal{R}}$ have the same dimension, we use that $\{P_{\mathcal{R}}(\rho_k)\}_{k \in I}$ spans \mathcal{R} and $\{P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k)\}_{k \in I}$ spans $\tilde{\mathcal{R}}$ (corollary B.6) and we prove that, for any $J \subseteq I$ (it suffices to take J a discrete, finite set), $\{P_{\mathcal{R}}(\rho_k)\}_{k \in J}$ is linearly dependent in \mathcal{R} if and only if

$\{P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k)\}_{k \in J}$ is linearly dependent in $\tilde{\mathcal{R}}$: if there exist $\{\alpha_k \in \mathbb{R}\}_{k \in J}$ not all zero such that

$$\sum_{k \in J} \alpha_k P_{\mathcal{R}}(\rho_k) = 0, \quad (\text{B.94})$$

then for all $l \in I$,

$$\begin{aligned} 0 &= \sum_{k \in J} \alpha_k \langle P_{\mathcal{R}}(\rho_k), P_{\mathcal{R}}(E_l) \rangle_{\mathcal{R}} \\ &\stackrel{(\text{B.93})}{=} \left\langle \sum_{k \in J} \alpha_k P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k), P_{\tilde{\mathcal{R}}}(\tilde{E}_l) \right\rangle_{\tilde{\mathcal{R}}}. \end{aligned} \quad (\text{B.95})$$

By corollary B.6, this implies that also $\sum_{k \in J} \alpha_k P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k) = 0$ which proves the ‘‘only if’’ direction; the symmetry in the definition of the two scenarios implies the converse direction. This in turn proves that bases of \mathcal{R} and $\tilde{\mathcal{R}}$ have the same cardinality: if there were a basis of \mathcal{R} of the form $\{P_{\mathcal{R}}(\rho_k)\}_{k \in J_0}$ with J_0 a discrete set consisting of $d \in \mathbb{N}$ elements and a basis of $\tilde{\mathcal{R}}$ of the form $\{P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k)\}_{k \in \tilde{J}_0}$ with \tilde{J}_0 a discrete set of $\tilde{d} \in \mathbb{N}$ elements with $\tilde{d} > d$, then we would obtain a contradiction because these \tilde{d} elements would have to be linearly dependent.

Now suppose that the scenario (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model with ontic space Λ . This is equivalent to the requirement that there exist $\{F(\lambda) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}\}_{\lambda \in \Lambda}$ and $\{\sigma_{\lambda} \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}\}_{\lambda \in \Lambda}$ as in theorem 1. Using corollary B.6, it holds that there exist two sets of indices J and K , each with $\dim(\mathcal{R})$ many indices, and real numbers $\{f_j(\lambda) \in \mathbb{R}\}_{j \in J, \lambda \in \Lambda}$, $\{s_j(\lambda) \in \mathbb{R}\}_{j \in K, \lambda \in \Lambda}$ such that

$$F(\lambda) = \sum_{j \in J} f_j(\lambda) P_{\mathcal{R}}(E_j), \quad (\text{B.96a})$$

$$\sigma(\lambda) = \sum_{j \in K} s_j(\lambda) P_{\mathcal{R}}(\rho_j). \quad (\text{B.96b})$$

To define the operationally noncontextual ontological model for the scenario $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$, it suffices to define

$$\tilde{F}(\lambda) := \sum_{j \in J} f_j(\lambda) P_{\tilde{\mathcal{R}}}(\tilde{E}_j) \in \tilde{\mathcal{R}}, \quad (\text{B.97a})$$

$$\tilde{\sigma}(\lambda) := \sum_{j \in K} s_j(\lambda) P_{\tilde{\mathcal{R}}}(\tilde{\rho}_j) \in \tilde{\mathcal{R}}. \quad (\text{B.97b})$$

Equation (B.93) implies by linearity that for all $k \in I$ and for all $\lambda \in \Lambda$,

$$\langle P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k), \tilde{F}(\lambda) \rangle_{\tilde{\mathcal{R}}} = \langle P_{\mathcal{R}}(\rho_k), F(\lambda) \rangle_{\mathcal{R}}, \quad (\text{B.98a})$$

$$\langle \tilde{\sigma}(\lambda), P_{\tilde{\mathcal{R}}}(\tilde{E}_k) \rangle_{\tilde{\mathcal{R}}} = \langle \sigma(\lambda), P_{\mathcal{R}}(E_k) \rangle_{\mathcal{R}}. \quad (\text{B.98b})$$

From there on, it is trivial to see that the primitives $\{\tilde{F}(\lambda)\}_{\lambda \in \Lambda}$ and $\{\tilde{\sigma}(\lambda)\}_{\lambda \in \Lambda}$ satisfy the constraints of theorem 1, and the symmetry in the definition of (\mathbf{s}, \mathbf{e}) versus $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ allows one to conclude that the converse direction also holds. \square

Proposition 11 (Incoherent coarse-grainings). *Consider any prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) . Suppose that there exist $\{E_k \in \mathbf{e}\}_{k=1}^N$ where $N \in \mathbb{N} \cup \{+\infty\}$ such that $\sum_{k=1}^N E_k \leq \mathbb{1}_{\mathcal{H}}$ and the sum $\sum_{k=1}^N E_k$ may not be in \mathbf{e} . Then, the scenario (\mathbf{s}, \mathbf{e}) and the extended scenario*

$$\left(\mathbf{s}, \mathbf{e}_{\text{ext}} := \text{conv}(\mathbf{e} \cup \{\sum_{k=1}^N E_k\})\right) \quad (2.25)$$

define the same reduced space $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})) = P_{\text{span}(\mathbf{e}_{\text{ext}})}(\text{span}(\mathbf{s}))$ and same polar cone $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = P_{\mathcal{R}}(\mathbf{e}_{\text{ext}})^{+\mathcal{R}}$.

Proof. First, it is clear that $\text{span}(\mathbf{e}) = \text{span}(\mathbf{e}_{\text{ext}})$ so that the equality of the reduced spaces $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s}))$ is clear. We now have to prove that $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} = P_{\mathcal{R}}(\mathbf{e}_{\text{ext}})^{+\mathcal{R}}$. Since $\mathbf{e} \subseteq \mathbf{e}_{\text{ext}}$, by lemma A.7, it holds that $P_{\mathcal{R}}(\mathbf{e}_{\text{ext}})^{+\mathcal{R}} \subseteq P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$. It remains to prove the reverse inclusion, which simply follows from the fact that if $\sigma \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$, then also $\langle \sigma, P_{\mathcal{R}}(\sum_{k=1}^N E_k) \rangle_{\mathcal{R}} \geq 0$ by linearity. \square

C Unit separability and cardinality bounds

C.1 Generalized separability

C.1.1 Review of elementary analysis

Let us first state some elementary results about convergence, sequences and closed sets. A more complete description can be found in [24] for example. Let \mathcal{V} be a finite dimensional real inner product space. \mathcal{V} is a complete normed space equipped with the norm induced by the inner product:

$$\forall v \in \mathcal{V} : \|v\|_{\mathcal{V}} := \sqrt{\langle v, v \rangle_{\mathcal{V}}}. \quad (C.1)$$

A sequence $(v_k \in \mathcal{V})_{k \in \mathbb{N}}$ is convergent if and only if there exists $v^* \in \mathcal{V}$ such that

$$\lim_{k \rightarrow \infty} v_k = v^*, \quad (C.2)$$

which is a short hand notation to state that

$$\lim_{k \rightarrow \infty} \|v_k - v^*\|_{\mathcal{V}} = 0. \quad (C.3)$$

Note that as a special case of the definition of a continuous function [24], any function $f : \mathcal{V} \rightarrow \mathbb{R}$ that is continuous has the property that for any convergent sequence $(v_k \in \mathcal{V})_{k \in \mathbb{N}}$, it holds that

$$\lim_{k \rightarrow \infty} f(v_k) = f(\lim_{k \rightarrow \infty} v_k). \quad (C.4)$$

We state without proof the following lemmas. Their proofs are either simple exercises or stated explicitly in [24].

Lemma C.1. *The norm $\|\cdot\|_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}$ is continuous, and for every fixed $v_0 \in \mathcal{V}$, the scalar products $\langle \cdot, v_0 \rangle_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}$ and $\langle v_0, \cdot \rangle_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}$ are also continuous.*

Lemma C.2. *Any subset $X \subseteq \mathcal{V}$ is closed if and only if, for any sequence $(x_k \in X)_{k \in \mathbb{N}}$ that converges to $x^* \in \mathcal{V}$, the limit x^* belongs to X .*

Lemma C.3. *Any convergent sequence $(v_k \in \mathcal{V})_{k \in \mathbb{N}}$ is also a bounded sequence. This means that there exists a finite constant $C \in \mathbb{R}$ such that*

$$\forall k \in \mathbb{N} : \|v_k\|_{\mathcal{V}} \leq C. \quad (C.5)$$

Lemma C.4. *If a sequence $(v_k \in \mathcal{V})_{k \in \mathbb{N}}$ converges, then for any subsequence defined by the strictly increasing set of indices $\{k_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N}$, it holds that*

$$\lim_{l \rightarrow \infty} v_{k_l} = \lim_{k \rightarrow \infty} v_k. \quad (C.6)$$

Lemma C.5. *Let $N \in \mathbb{N}$. For each $n = 1, \dots, N$, let $(v_k^{(n)} \in \mathcal{V})_{k \in \mathbb{N}}$ be a sequence that converges to $V^{(n)} \in \mathcal{V}$. Then, it holds that*

$$\lim_{k \rightarrow \infty} \sum_{n=1}^N v_k^{(n)} = \sum_{n=1}^N V^{(n)}. \quad (C.7)$$

Lemma C.6. *Let $N \in \mathbb{N}$. For each $n = 1, \dots, N$, let $\mathcal{V}^{(n)}$ be an arbitrary real inner product space of finite dimension. Let $(v_k^{(n)} \in \mathcal{V}^{(n)})_{k \in \mathbb{N}}$ be a real sequence that converges to $V^{(n)} \in \mathcal{V}^{(n)}$. Then, it holds that the limit of the tensor product equals the tensor product of the limits:*

$$\lim_{k \rightarrow \infty} \bigotimes_{n=1}^N v_k^{(n)} = \bigotimes_{n=1}^N V^{(n)}. \quad (C.8)$$

Proof overview. The first thing to show is that the limit of the product of two convergent sequences in \mathbb{R} is equal to the product of the limits of the sequences. Then, generalize to any number of real sequences by recursion. Finally, expand the tensor products in any basis of the underlying vector spaces and apply the result derived for the real sequence case. \square

Theorem C.1 (Bolzano-Weierstrass theorem). *Any bounded sequence $(v_k \in \mathcal{V})_{k \in \mathbb{N}}$, where \mathcal{V} is any finite-dimensional real inner product space, admits a convergent subsequence $(v_{k_l})_{l \in \mathbb{N}}$. Specifically, if the sequence $(v_k)_{k \in \mathbb{N}}$ satisfies, for some constant $C \in \mathbb{R}$ independent of k ,*

$$\forall k \in \mathbb{N} : \|v_k\|_{\mathcal{V}} \leq C, \quad (C.9)$$

then there exists a strictly increasing subset of indices, denoted $\{k_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N}$, and there exists $v^ \in \mathcal{V}$ such that the subsequence $(v_{k_l})_{l \in \mathbb{N}}$ converges to v^* :*

$$\lim_{l \rightarrow \infty} v_{k_l} = v^*. \quad (C.10)$$

Proof. We specialized the more general theorem 6.21 in [24] according to the needs of the present matter. \square

C.1.2 Generalized product operators

Referring to the definition 13 of the generalized product state set $\text{Prod}(\mathbf{s}, \mathbf{e})$, let us first verify the following lemma.

Lemma C.7. $\text{Prod}(\mathbf{s}, \mathbf{e})$ is a closed set.

Proof. Consider any sequence $(d_k \in \text{Prod}(\mathbf{s}, \mathbf{e}))_{k \in \mathbb{N}}$. By definition of $\text{Prod}(\mathbf{s}, \mathbf{e})$, there exist sequences

$$\left(a_k \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} \right)_{k \in \mathbb{N}}, \quad (\text{C.11a})$$

$$\left(b_k \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} \right)_{k \in \mathbb{N}}, \quad (\text{C.11b})$$

such that $d_k = a_k \otimes b_k$ for all $k \in \mathbb{N}$. Then, for all $k \in \mathbb{N}$, there always exist $m_k \in \mathcal{R}$ and $n_k \in \mathcal{R}$ such that

$$a_k = \|a_k\|_{\mathcal{R}} m_k, \quad \|m_k\|_{\mathcal{R}} = 1, \quad (\text{C.12a})$$

$$b_k = \|b_k\|_{\mathcal{R}} n_k, \quad \|n_k\|_{\mathcal{R}} = 1. \quad (\text{C.12b})$$

Now suppose that this sequence $(d_k)_{k \in \mathbb{N}}$ is convergent and converges to $d^* \in \mathcal{R} \otimes \mathcal{R}$. We want to show that $d^* \in \text{Prod}(\mathbf{s}, \mathbf{e})$. We know

$$d^* = \lim_{k \rightarrow \infty} d_k = \lim_{k \rightarrow \infty} (\|a_k\|_{\mathcal{R}} \cdot \|b_k\|_{\mathcal{R}})(m_k \otimes n_k). \quad (\text{C.13})$$

The norm $\|\cdot\|_{\mathcal{R} \otimes \mathcal{R}}$ being continuous according to lemma C.1, it holds that

$$\begin{aligned} \|d^*\|_{\mathcal{R} \otimes \mathcal{R}} &= \lim_{k \rightarrow \infty} \|d_k\|_{\mathcal{R} \otimes \mathcal{R}} \\ &= \lim_{k \rightarrow \infty} (\|a_k\|_{\mathcal{R}} \cdot \|b_k\|_{\mathcal{R}}), \end{aligned} \quad (\text{C.14})$$

where we used $\|m_k \otimes n_k\|_{\mathcal{R} \otimes \mathcal{R}} = \|m_k\|_{\mathcal{R}} \cdot \|n_k\|_{\mathcal{R}} = 1$ according to (C.12). This shows that the real sequence $(\|a_k\|_{\mathcal{R}} \cdot \|b_k\|_{\mathcal{R}} \in \mathbb{R})_{k \in \mathbb{N}}$ converges to $\|d^*\|_{\mathcal{R} \otimes \mathcal{R}}$. Next, consider the sequence $(m_k \in \mathcal{R})_{k \in \mathbb{N}}$: it is bounded in norm thanks to its normalization (C.12a). By the Bolzano-Weierstrass theorem C.1, we may extract a convergent subsequence with indices $\{j_l \in \mathbb{N}\}_{l \in \mathbb{N}} \subseteq \mathbb{N}$. The corresponding limit is denoted $m^* \in \mathcal{R}$, that is,

$$\lim_{l \rightarrow \infty} m_{j_l} = m^*. \quad (\text{C.15})$$

The sequence $(n_{j_l} \in \mathcal{R})_{l \in \mathbb{N}}$ is also bounded from (C.12b) so by the Bolzano-Weierstrass theorem C.1 we can further extract a convergent subsequence with indices

$$\{k_l \in \mathbb{N}\}_{l \in \mathbb{N}} \subseteq \{j_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N}, \quad (\text{C.16})$$

and we denote the corresponding limit $n^* \in \mathcal{R}$. With this further refinement of indices, both subsequences

$(m_{k_l})_{l \in \mathbb{N}}$ (using lemma C.4) and $(n_{k_l})_{l \in \mathbb{N}}$ converge in \mathcal{R} , i.e.,

$$\lim_{l \rightarrow \infty} m_{k_l} = m^*, \quad (\text{C.17a})$$

$$\lim_{l \rightarrow \infty} n_{k_l} = n^*. \quad (\text{C.17b})$$

Using lemma C.6 applied to $\mathbb{R} \otimes \mathcal{R} \otimes \mathcal{R}$ to commute the limit and the product, and using lemma C.4 for the subsequence $(d_{k_l})_{l \in \mathbb{N}}$, we obtain:

$$\begin{aligned} d^* &= \lim_{l \rightarrow \infty} d_{k_l} \\ &= \left(\lim_{l \rightarrow \infty} \|a_{k_l}\|_{\mathcal{R}} \cdot \|b_{k_l}\|_{\mathcal{R}} \right) \left(\lim_{l \rightarrow \infty} m_{k_l} \right) \otimes \left(\lim_{l \rightarrow \infty} n_{k_l} \right) \\ &= \|d^*\|_{\mathcal{R} \otimes \mathcal{R}} (m^* \otimes n^*). \end{aligned} \quad (\text{C.18})$$

To show $d^* \in \text{Prod}(\mathbf{s}, \mathbf{e})$, it only remains to show that

$$m^* \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}, \quad (\text{C.19a})$$

$$n^* \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}. \quad (\text{C.19b})$$

First, note that for all $l \in \mathbb{N}$, due to $a_{k_l} \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$, it holds that for all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$:

$$0 \leq \langle a_{k_l}, \bar{\rho} \rangle_{\mathcal{R}} = \|a_{k_l}\|_{\mathcal{R}} \langle m_{k_l}, \bar{\rho} \rangle_{\mathcal{R}}. \quad (\text{C.20})$$

For the indices $\{l \in \mathbb{N} : \|a_{k_l}\|_{\mathcal{R}} > 0\}$, equation (C.20) implies that $m_{k_l} \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$. For the remaining indices $\{l \in \mathbb{N} : \|a_{k_l}\|_{\mathcal{R}} = 0\}$, we can make an arbitrary choice in (C.12a) when we write $0 = a_{k_l} = 0 \cdot m_{k_l}$: choose any normalized $m_{k_l} \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ for these indices. This is always possible thanks to lemma B.9. This shows that for all $l \in \mathbb{N}$, $m_{k_l} \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$. Using lemma C.1, we may commute the scalar product with the limit to obtain, for all $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$,

$$\langle m^*, \bar{\rho} \rangle_{\mathcal{R}} = \lim_{l \rightarrow \infty} \langle m_{k_l}, \bar{\rho} \rangle_{\mathcal{R}} \geq 0. \quad (\text{C.21})$$

To conclude, we used that $\mathbb{R}_{\geq 0}$ is a closed interval of \mathbb{R} . This proves that $m^* \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$. By an entirely analogous reasoning we obtain that $n^* \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$. This proves that for any converging sequence $(d_k \in \text{Prod}(\mathbf{s}, \mathbf{e}))_{k \in \mathbb{N}}$, we have

$$\lim_{k \rightarrow \infty} d_k \in \text{Prod}(\mathbf{s}, \mathbf{e}), \quad (\text{C.22})$$

which according to lemma C.2 proves that $\text{Prod}(\mathbf{s}, \mathbf{e})$ is closed. \square

C.1.3 Generalized separable operators

Let us first prove the following proposition.

Proposition C.8 (Specialized Carathéodory's theorem for convex cones). *For all $\Omega \in \text{Sep}(\mathbf{s}, \mathbf{e})$, there exist $n \in \{1, \dots, \dim(\mathcal{R})^2\}$ and families*

$$\{F_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}\}_{i=1}^n, \quad (\text{C.23a})$$

$$\{\sigma_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}\}_{i=1}^n, \quad (\text{C.23b})$$

which satisfy

$$\Omega = \sum_{i=1}^n F_i \otimes \sigma_i. \quad (\text{C.24})$$

Proof. This proposition is the content of Carathéodory's theorem for convex cones as presented in theorem 4.3.2 in [20]. For completeness, we present a proof with the notation adapted to the context of this manuscript.

Suppose that there exists $\Omega \in \text{Sep}(\mathbf{s}, \mathbf{e})$ for which the shortest convex decomposition over $\text{Prod}(\mathbf{s}, \mathbf{e})$ is of length $n \geq \dim(\mathcal{R})^2 + 1$:

$$\Omega = \sum_{i=1}^n F_i \otimes \sigma_i. \quad (\text{C.25})$$

Because the space $\mathcal{R} \otimes \mathcal{R}$ is of dimension $\dim(\mathcal{R})^2$, any family of $n \geq \dim(\mathcal{R})^2 + 1$ elements of $\mathcal{R} \otimes \mathcal{R}$ has to be linearly dependent: this is the case of the set $\{F_i \otimes \sigma_i \in \mathcal{R} \otimes \mathcal{R}\}_{i=1}^n$. This implies that there exist scalars $\{\alpha_i \in \mathbb{R}\}_{i=1}^n$ not all zero such that

$$\sum_{i=1}^n \alpha_i (F_i \otimes \sigma_i) = 0. \quad (\text{C.26})$$

Suppose that for all $i = 1, \dots, n$: $\alpha_i \leq 0$. Because not all α_i are zero, there must exist i such that $\alpha_i < 0$. In that case, replace all α_i by their opposite $-\alpha_i$ so that there now exists i such that $\alpha_i > 0$.

Thus, without loss of generality, there must exist i such that $\alpha_i > 0$. We now can assert that $\max_j \alpha_j > 0$. Now, consider the following alternative decomposition of Ω where we subtracted a multiple of 0 in the form of (C.26) from the initial decomposition (C.25):

$$\begin{aligned} \Omega &= \sum_{i=1}^n (F_i \otimes \sigma_i) - \frac{1}{\max_j \alpha_j} \sum_{i=1}^n \alpha_i (F_i \otimes \sigma_i) \\ &= \sum_{i=1}^n \left(1 - \frac{\alpha_i}{\max_j \alpha_j}\right) (F_i \otimes \sigma_i). \end{aligned} \quad (\text{C.27})$$

Define

$$\theta_i := 1 - \frac{\alpha_i}{\max_j \alpha_j}. \quad (\text{C.28})$$

For all $i = 1, \dots, n$ we have $\theta_i \geq 0$. Now clearly, for j_0 such that $\max_j \alpha_j = \alpha_{j_0}$, we have that $\theta_{j_0} = 0$ which means we can rewrite Ω as a shorter positive linear combination of elements of $\text{Prod}(\mathbf{s}, \mathbf{e})$:

$$\sum_{i \in \{1, \dots, n\} \setminus \{j_0\}} \theta_i (F_i \otimes \sigma_i). \quad (\text{C.29})$$

This yields the contradiction, and we conclude that any element of $\text{Sep}(\mathbf{s}, \mathbf{e})$ can be written as a convex combination of at most $\dim(\mathcal{R})^2$ elements of $\text{Prod}(\mathbf{s}, \mathbf{e})$. \square

Proposition C.9. $\text{Sep}(\mathbf{s}, \mathbf{e})$ is a closed convex cone.

Note that this is not entirely obvious. $\text{Sep}(\mathbf{s}, \mathbf{e})$ is the convex hull of a closed *unbounded* set, namely, $\text{Prod}(\mathbf{s}, \mathbf{e})$. In general, the convex hull of a closed unbounded set may not be closed.

Proof. Consider any converging sequence $(\Omega_k \in \text{Sep}(\mathbf{s}, \mathbf{e}))_{k \in \mathbb{N}}$ with limit $\Omega^* \in \mathcal{R} \otimes \mathcal{R}$. Let $I = \{1, \dots, \dim(\mathcal{R})^2\}$. Note that by proposition C.8, for all $k \in \mathbb{N}$, there exists a decomposition of Ω_k as $\dim(\mathcal{R})^2$ elements of $\text{Prod}(\mathbf{s}, \mathbf{e})$ ¹⁴ which we write as $\{d_i^{(k)} \in \text{Prod}(\mathbf{s}, \mathbf{e})\}_{i \in I}$:

$$\Omega_k = \sum_{i \in I} d_i^{(k)}. \quad (\text{C.30})$$

Let us prove that for all $i \in I$, the sequence $(d_i^{(k)})_{k \in \mathbb{N}}$ is bounded. We have to show that there exists a finite upper bound $\lambda_i \in \mathbb{R}$ independent of k such that

$$\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}} \leq \lambda_i. \quad (\text{C.31})$$

Let $\{R_m \in \mathcal{R} \otimes \mathcal{R}\}_{m \in I}$ be an orthonormal basis of $\mathcal{R} \otimes \mathcal{R}$. Thanks to corollary B.6, we can also pick $\dim(\mathcal{R})^2$ elements of the form

$$\{\bar{\rho}_p \otimes \bar{E}_p : \bar{\rho}_p \in P_{\mathcal{R}}(\mathbf{s}), \bar{E}_p \in P_{\mathcal{R}}(\mathbf{e})\}_{p \in I} \quad (\text{C.32})$$

to obtain a basis of $\mathcal{R} \otimes \mathcal{R}$, although this basis will in general not be an orthonormal one. The two bases $\{R_m\}_{m \in I}$ and $\{\bar{\rho}_p \otimes \bar{E}_p\}_{p \in I}$ are related by an invertible change of basis: there exists a $\dim(\mathcal{R})^2 \times \dim(\mathcal{R})^2$ real, invertible matrix Q with components $\{Q_{mp} \in \mathbb{R}\}_{m, p \in I}$ such that:

$$\forall m \in I : R_m = \sum_{p \in I} Q_{mp} (\bar{\rho}_p \otimes \bar{E}_p), \quad (\text{C.33a})$$

$$\forall p \in I : \bar{\rho}_p \otimes \bar{E}_p = \sum_{m \in I} Q_{pm}^{-1} R_m. \quad (\text{C.33b})$$

Expanding the norm in the orthonormal basis $\{R_m\}_{m \in I}$, it holds that

$$\begin{aligned} &\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 \\ &= \left\langle d_i^{(k)}, \sum_{m \in I} \langle R_m, d_i^{(k)} \rangle_{\mathcal{R} \otimes \mathcal{R}} R_m \right\rangle_{\mathcal{R} \otimes \mathcal{R}} \\ &= \sum_{m \in I} \langle d_i^{(k)}, R_m \rangle_{\mathcal{R} \otimes \mathcal{R}}^2 \\ &= \sum_{m \in I} \left(\sum_{p \in I} Q_{mp} \langle d_i^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2. \end{aligned} \quad (\text{C.34})$$

¹⁴One may have to pad shorter decompositions with 0 $\in \text{Prod}(\mathbf{s}, \mathbf{e})$.

Then, using the triangle inequality for the absolute value:

$$\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 \leq \sum_{m \in I} \left(\sum_{p \in I} |Q_{mp}| \langle d_i^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2, \quad (\text{C.35})$$

where we used that due to $d_i^{(k)} \in \text{Prod}(\mathbf{s}, \mathbf{e})$,

$$\forall i, p \in I, \forall k \in \mathbb{N} : \langle d_i^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \geq 0, \quad (\text{C.36})$$

which allowed us to remove the absolute value off of these scalar products in (C.35). Then, let

$$\chi_1 := \sum_{m \in I} \left(\max_{p \in I} |Q_{mp}| \right)^2 \in \mathbb{R}_{\geq 0}. \quad (\text{C.37})$$

The upper bound (C.35) becomes

$$\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 \leq \chi_1 \left(\sum_{p \in I} \langle d_i^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2. \quad (\text{C.38})$$

Due to (C.36), $\forall i, p \in I, \forall k \in \mathbb{N}$:

$$\langle d_i^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \leq \sum_{j \in I} \langle d_j^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}}. \quad (\text{C.39})$$

This allows us to upper-bound equation (C.38) as

$$\begin{aligned} \|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 &\leq \chi_1 \left(\sum_{j, p \in I} \langle d_j^{(k)}, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2 \\ &= \chi_1 \left(\sum_{p \in I} \langle \Omega_k, \bar{\rho}_p \otimes \bar{E}_p \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2 \\ &= \chi_1 \left(\sum_{p, m \in I} Q_{pm}^{-1} \langle \Omega_k, R_m \rangle_{\mathcal{R} \otimes \mathcal{R}} \right)^2 \\ &\leq \chi_1 \left(\sum_{p, m \in I} |Q_{pm}^{-1}| |\langle \Omega_k, R_m \rangle_{\mathcal{R} \otimes \mathcal{R}}| \right)^2. \end{aligned} \quad (\text{C.40})$$

Let

$$\chi_2 := \left(\max_{m \in I} \sum_{p \in I} |Q_{pm}^{-1}| \right)^2 \in \mathbb{R}_{\geq 0}. \quad (\text{C.41})$$

Then, the bound becomes

$$\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 \leq \chi_1 \chi_2 \left(\sum_{m \in I} |\langle \Omega_k, R_m \rangle_{\mathcal{R} \otimes \mathcal{R}}| \right)^2. \quad (\text{C.42})$$

Note that for all $m \in I$:

$$\langle \Omega_k, R_m \rangle_{\mathcal{R} \otimes \mathcal{R}}^2 \leq \sum_{n \in I} \langle \Omega_k, R_n \rangle_{\mathcal{R} \otimes \mathcal{R}}^2 = \|\Omega_k\|_{\mathcal{R} \otimes \mathcal{R}}^2. \quad (\text{C.43})$$

The bound (C.42) becomes

$$\begin{aligned} \|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 &\leq \chi_1 \chi_2 \left(\sum_{m \in I} \|\Omega_k\|_{\mathcal{R} \otimes \mathcal{R}} \right)^2 \\ &= \chi_1 \chi_2 |I|^2 \|\Omega_k\|_{\mathcal{R} \otimes \mathcal{R}}^2 \\ &= \chi_1 \chi_2 \dim(\mathcal{R})^4 \|\Omega_k\|_{\mathcal{R} \otimes \mathcal{R}}^2. \end{aligned} \quad (\text{C.44})$$

The sequence $(\Omega_k)_{k \in \mathbb{N}}$ converges, so by lemma C.3, it is a bounded sequence: there exists $C \in \mathbb{R}_{\geq 0}$ such that for all $k \in \mathbb{N}$, $\|\Omega_k\|_{\mathcal{R} \otimes \mathcal{R}} \leq C$. We have shown that for all $i \in I$, for all $k \in \mathbb{N}$,

$$\|d_i^{(k)}\|_{\mathcal{R} \otimes \mathcal{R}}^2 \leq \chi_1 \chi_2 \dim(\mathcal{R})^4 C^2. \quad (\text{C.45})$$

We may now apply the Bolzano-Weierstrass theorem C.1 to the bounded sequence $(d_i^{(k)})_{k \in \mathbb{N}}$ to extract a first set of strictly increasing indices $\{a_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N}$ such that the induced subsequence of $(d_1^{(a_l)})_{l \in \mathbb{N}}$ converges. Then, consider the subsequence $(d_2^{(a_l)})_{l \in \mathbb{N}}$. Using (C.45), it is bounded as well, so that there exists a subset of strictly increasing indices

$$\{b_l\}_{l \in \mathbb{N}} \subseteq \{a_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N} \quad (\text{C.46})$$

such that the subsequence $(d_2^{(b_l)})_{l \in \mathbb{N}}$ converges. By lemma C.4, the subsequence $(d_1^{(b_l)})_{l \in \mathbb{N}}$ converges to the same limit as $(d_1^{(a_l)})_{l \in \mathbb{N}}$. Repeat this procedure to obtain a new set of strictly increasing indices

$$\{c_l\}_{l \in \mathbb{N}} \subseteq \{b_l\}_{l \in \mathbb{N}} \subseteq \{a_l\}_{l \in \mathbb{N}} \subseteq \mathbb{N} \quad (\text{C.47})$$

so that the sequences $(d_1^{(c_l)})_{l \in \mathbb{N}}$, $(d_2^{(c_l)})_{l \in \mathbb{N}}$ and $(d_3^{(c_l)})_{l \in \mathbb{N}}$ converge, etc., and after $\dim(\mathcal{R})^2$ steps, the process stops. We denote the final set of strictly increasing indices $\{k_l\}_{l \in \mathbb{N}}$, and we denote the limits as

$$\forall i \in I : d_i^* := \lim_{l \rightarrow \infty} d_i^{(k_l)} \in \text{Prod}(\mathbf{s}, \mathbf{e}), \quad (\text{C.48})$$

where we used lemma C.7 to conclude that the limits lie in $\text{Prod}(\mathbf{s}, \mathbf{e})$. Note that the freedom in choosing the convergent subsequences from the bounded sequences is irrelevant: in any case, using lemma C.4, the induced subsequence $(\Omega_{k_l})_{l \in \mathbb{N}}$ converges to Ω^* . Then, using lemma C.5 to commute the sum and the limit,

$$\begin{aligned} \Omega^* &= \lim_{k \rightarrow \infty} \Omega_k = \lim_{l \rightarrow \infty} \Omega_{k_l} = \lim_{l \rightarrow \infty} \sum_{i \in I} d_i^{(k_l)} \\ &= \sum_{i \in I} d_i^*. \end{aligned} \quad (\text{C.49})$$

Thanks to equation (C.48), this proves $\Omega^* \in \text{Sep}(\mathbf{s}, \mathbf{e})$, and because the sequence $(\Omega_k)_{k \in \mathbb{N}}$ was arbitrary in $\text{Sep}(\mathbf{s}, \mathbf{e})$, this proves that $\text{Sep}(\mathbf{s}, \mathbf{e})$ is a closed set. \square

C.2 Choi-Jamiołkowski isomorphism

We now prove the consistency of the definition 15 of the Choi-Jamiołkowski isomorphism. We restrict to the study of linear maps from \mathcal{R} to \mathcal{R} , i.e., maps in $L(\mathcal{R})$, but these results hold equally well should one replace \mathcal{R} with any real inner product space of finite dimension.

Lemma C.10. *For any $\Phi \in L(\mathcal{R})$, if the Choi-Jamiołkowski operator $\mathbb{J}(\Phi)$ exists, then it is unique.*

Proof. Suppose that there exist two operators $\mathbb{J}(\Phi), \tilde{\mathbb{J}}(\Phi) \in \mathcal{R} \otimes \mathcal{R}$ which satisfy the requirement of definition 15. We will show that $\mathbb{J}(\Phi) = \tilde{\mathbb{J}}(\Phi)$. Indeed, by equation (3.5), for any $r, s \in \mathcal{R}$,

$$\begin{aligned} \langle \mathbb{J}(\Phi) - \tilde{\mathbb{J}}(\Phi), r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}} &= \langle r, \Phi(s) \rangle_{\mathcal{R}} - \langle r, \Phi(s) \rangle_{\mathcal{R}} \\ &= 0. \end{aligned} \quad (\text{C.50})$$

This being valid for any $r, s \in \mathcal{R} \otimes \mathcal{R}$, by the non-degeneracy of the scalar product we obtain

$$\mathbb{J}(\Phi) = \tilde{\mathbb{J}}(\Phi). \quad (\text{C.51})$$

This proves that for any $\Phi \in L(\mathcal{R})$, the Choi-Jamiołkowski operator $\mathbb{J}(\Phi)$ is unique. \square

Lemma C.11. *The definition 15 is consistent in that for all $\Phi \in L(\mathcal{R})$, $\mathbb{J}(\Phi)$ exists and is unique. Given an orthonormal basis $\{R_i \in \mathcal{R}\}_{i=1}^{\dim(\mathcal{R})}$ of \mathcal{R} , it is given by*

$$\mathbb{J}(\Phi) = \sum_{i=1}^{\dim(\mathcal{R})} \Phi(R_i) \otimes R_i. \quad (\text{C.52})$$

Proof. The existence may be proven as follows. One can always expand $\mathbb{J}(\Phi) \in \mathcal{R} \otimes \mathcal{R}$ in the basis $\{R_i \otimes R_j\}_{i,j=1}^{\dim(\mathcal{R})}$ of $\mathcal{R} \otimes \mathcal{R}$, with coefficients $\{j(\Phi)_{i,j} \in \mathbb{R}\}_{i,j=1}^{\dim(\mathcal{R})}$:

$$\mathbb{J}(\Phi) = \sum_{i,j} j(\Phi)_{i,j} R_i \otimes R_j, \quad (\text{C.53a})$$

$$j(\Phi)_{i,j} = \langle \mathbb{J}(\Phi), R_i \otimes R_j \rangle_{\mathcal{R} \otimes \mathcal{R}}. \quad (\text{C.53b})$$

Then, using equation (3.5), for all $i, j = 1, \dots, \dim(\mathcal{R})$,

$$j(\Phi)_{i,j} = \langle R_i, \Phi(R_j) \rangle_{\mathcal{R}}. \quad (\text{C.54})$$

Inserting this result into (C.53a),

$$\mathbb{J}(\Phi) = \sum_j \left(\sum_i \langle R_i, \Phi(R_j) \rangle_{\mathcal{R}} R_i \right) \otimes R_j. \quad (\text{C.55})$$

One recognizes the completeness relation for the basis $\{R_i\}_i$:

$$\forall r \in \mathcal{R} : r = \sum_{i=1}^{\dim(\mathcal{R})} \langle R_i, r \rangle_{\mathcal{R}} R_i. \quad (\text{C.56})$$

Inserting this result in equation (C.55), equation (C.52) is readily obtained. This proves the existence of $\mathbb{J}(\mathcal{R})$ for any $\Phi \in L(\mathcal{R})$, and the uniqueness follows from lemma C.10. \square

Lemma C.11 proves the expression for $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ in lemma 16 as a corollary.

Lemma C.12. *The Choi-Jamiołkowski mapping in definition 15 is indeed an isomorphism. The inverse mapping, for any $\Omega \in \mathcal{R} \otimes \mathcal{R}$, is denoted $\mathbb{J}^{-1}[\Omega] \in L(\mathcal{R})$ and is defined by the relations:*

$$\forall r, s \in \mathcal{R} : \langle r, \mathbb{J}^{-1}[\Omega](s) \rangle_{\mathcal{R}} = \langle \Omega, r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}}. \quad (\text{C.57})$$

The injectivity of $\mathbb{J}(\cdot)$ is particularly interesting: for any two functions $\Phi_1, \Phi_2 \in L(\mathcal{R})$,

$$\mathbb{J}(\Phi_1) = \mathbb{J}(\Phi_2) \implies \Phi_1 = \Phi_2. \quad (\text{C.58})$$

Proof. It suffices to prove that

$$\forall \Phi \in L(\mathcal{R}) : \mathbb{J}^{-1}[\mathbb{J}(\Phi)] = \Phi, \quad (\text{C.59a})$$

$$\forall \Omega \in \mathcal{R} \otimes \mathcal{R} : \mathbb{J}(\mathbb{J}^{-1}[\Omega]) = \Omega, \quad (\text{C.59b})$$

which follows easily from the relations (3.5) and (C.57). \square

The following lemma will prove useful in appendix C.3.

Lemma C.13. *Let $\Phi \in L(\mathcal{R})$. Suppose that there exists n_{Φ} , and $a_i, b_i \in \mathcal{R}$ for $i = 1, \dots, n_{\Phi}$ such that*

$$\mathbb{J}(\Phi) = \sum_{i=1}^{n_{\Phi}} a_i \otimes b_i. \quad (\text{C.60})$$

Then, the dimension $\text{rank}(\Phi)$ of the image vector space of Φ satisfies

$$\text{rank}(\Phi) \leq n_{\Phi}. \quad (\text{C.61})$$

Proof. Let $\{R_i \in \mathcal{R}\}_{i=1}^{\dim(\mathcal{R})}$ be an orthonormal basis of \mathcal{R} . Then, the linear map $\Phi \in L(\mathcal{R})$ may be represented as a real matrix in this basis. We will be using the completeness relation of \mathcal{R} in the form (C.56):

$$\begin{aligned}
\Phi(r) &= \sum_{k=1}^{\dim(\mathcal{R})} \langle R_k, \Phi(r) \rangle_{\mathcal{R}} R_k \\
&= \sum_{k=1}^{\dim(\mathcal{R})} \langle \mathbb{J}(\Phi), R_k \otimes r \rangle_{\mathcal{R} \otimes \mathcal{R}} R_k \\
&= \sum_{k=1}^{\dim(\mathcal{R})} \sum_{i=1}^{n_\Phi} \langle a_i, R_k \rangle_{\mathcal{R}} \langle b_i, r \rangle_{\mathcal{R}} R_k \\
&= \sum_{i=1}^{n_\Phi} \langle b_i, r \rangle_{\mathcal{R}} a_i. \tag{C.62}
\end{aligned}$$

Clearly, this shows that the image vector subspace $\text{Im}(\Phi) := \{\Phi(r) : r \in \mathcal{R}\} \subseteq \mathcal{R}$ satisfies

$$\text{Im}(\Phi) \subseteq \text{span}(\{a_i\}_{i=1}^{n_\Phi}), \tag{C.63}$$

which implies that the dimensions respect

$$\begin{aligned}
\text{rank}(\Phi) &:= \dim(\text{Im}(\Phi)) \\
&\leq \dim(\text{span}(\{a_i\}_{i=1}^{n_\Phi})) \leq n_\Phi. \quad \square \tag{C.64}
\end{aligned}$$

C.3 The unit separability criterion

Theorem 2 (Main theorem: unit separability). *The prepare-and-measure scenario (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model (definition 17) if and only if:*

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e}). \tag{3.9}$$

Proof. Suppose that there exists a Riemann integrable operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) , so that equations (2.22), (2.23) and (2.24) of theorem 1 are verified. It is easy to show that the definition 17 of a Riemann integrable operationally noncontextual ontological model implies that the right-hand side of the consistency requirement (2.24) can be rewritten as:

$$\begin{aligned}
&\int_{\Lambda} d\lambda \langle r, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), s \rangle_{\mathcal{R}} \\
&= \lim_{N \rightarrow \infty} \sum_{k=1}^N \Delta_{N,k} \left\langle r, F(\lambda_{N,k}^{(\text{dis})}) \right\rangle_{\mathcal{R}} \left\langle \sigma(\lambda_{N,k}^{(\text{dis})}), s \right\rangle_{\mathcal{R}}. \tag{C.65}
\end{aligned}$$

Rewriting the right-hand side in tensor product form using the scalar product property (3.1):

$$\begin{aligned}
&\int_{\Lambda} d\lambda \langle r, F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), s \rangle_{\mathcal{R}} = \\
&\lim_{N \rightarrow \infty} \left\langle \sum_{k=1}^N \Delta_{N,k} F(\lambda_{N,k}^{(\text{dis})}) \otimes \sigma(\lambda_{N,k}^{(\text{dis})}), r \otimes s \right\rangle_{\mathcal{R} \otimes \mathcal{R}}. \tag{C.66}
\end{aligned}$$

Define for all $N \in \mathbb{N}$:

$$A_N := \sum_{k=1}^N \Delta_{N,k} F(\lambda_{N,k}^{(\text{dis})}) \otimes \sigma(\lambda_{N,k}^{(\text{dis})}). \tag{C.67}$$

Then, the left-hand side of equation (2.24) can be rewritten using the identity map and the defining property of the Choi-Jamiołkowski isomorphism as in (3.5):

$$\langle r, s \rangle_{\mathcal{R}} = \langle r, \mathbb{1}_{\mathcal{R}}(s) \rangle_{\mathcal{R}} = \langle \mathbb{J}(\mathbb{1}_{\mathcal{R}}), r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}}. \tag{C.68}$$

Thus, equation (2.24) is equivalent to: $\forall r, s \in \mathcal{R}$,

$$\langle \mathbb{J}(\mathbb{1}_{\mathcal{R}}), r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}} = \lim_{N \rightarrow \infty} \langle A_N, r \otimes s \rangle_{\mathcal{R} \otimes \mathcal{R}}. \tag{C.69}$$

Equation (C.69) shows in particular that all the components of A_N converge to the components of $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$, which proves the convergence of the sequence $(A_N)_{N \in \mathbb{N}}$:

$$\lim_{N \rightarrow \infty} A_N = \mathbb{J}(\mathbb{1}_{\mathcal{R}}). \tag{C.70}$$

By the definition of A_N in (C.67), the nonnegativity of $\Delta_{N,k}$ as in (3.7a), the domains of F , σ as in (2.22) and the definition 14 of $\text{Sep}(\mathbf{s}, \mathbf{e})$:

$$\forall N \in \mathbb{N} : A_N \in \text{Sep}(\mathbf{s}, \mathbf{e}). \tag{C.71}$$

Proposition C.9 proved that $\text{Sep}(\mathbf{s}, \mathbf{e})$ is a closed set within $\mathcal{R} \otimes \mathcal{R}$, so it holds that

$$\lim_{N \rightarrow \infty} A_N \in \text{Sep}(\mathbf{s}, \mathbf{e}). \tag{C.72}$$

and thus also, by (C.70), that

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e}). \tag{C.73}$$

We have proven that if the scenario (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model of the form of definitions 6 and 17, then $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e})$.

Let us consider the other direction: suppose that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \text{Sep}(\mathbf{s}, \mathbf{e})$. By proposition C.8, there exist $n \leq \dim(\mathcal{R})^2$ and

$$\{\tilde{F}_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}\}_{i=1}^n, \tag{C.74a}$$

$$\{\tilde{\sigma}_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}\}_{i=1}^n, \tag{C.74b}$$

such that

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^n \tilde{F}_i \otimes \tilde{\sigma}_i. \tag{C.75}$$

If we assume that any zero element in the decomposition (C.75) has been removed, then for all $i = 1, \dots, n$ we may assume $\tilde{\sigma}_i \neq 0$ which also implies according to lemma B.8 that

$$\forall i = 1, \dots, n : \langle \tilde{\sigma}_i, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} > 0. \tag{C.76}$$

Let, for all $i = 1, \dots, n$:

$$F_i := \langle \tilde{\sigma}_i, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} \tilde{F}_i, \tag{C.77a}$$

$$\sigma_i := (\langle \tilde{\sigma}_i, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}})^{-1} \tilde{\sigma}_i. \tag{C.77b}$$

We now show that the F_i 's and σ_i 's of equations (C.77) constitute a valid operationally noncontextual ontological model as framed in theorem 1. First, the nonnegativity requirements of equation (2.22) are verified thanks to equations (C.74), (C.76) and (C.77):

$$F_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}, \quad (\text{C.78a})$$

$$\sigma_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}. \quad (\text{C.78b})$$

The normalization in (2.23) is verified:

$$\forall i = 1, \dots, n : \langle \sigma_i, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1 \quad (\text{C.79})$$

as can be seen from equation (C.77b). Finally, the reproduction of quantum statistics in equation (2.24) is still verified: indeed, from equations (C.75) and (C.77),

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^n F_i \otimes \sigma_i, \quad (\text{C.80})$$

which in turns implies $\forall r, s \in \mathcal{R}$:

$$\langle r, s \rangle_{\mathcal{R}} = \sum_{i=1}^n \langle r, F_i \rangle_{\mathcal{R}} \langle \sigma_i, s \rangle_{\mathcal{R}}. \quad (\text{C.81})$$

It is easy to see that such a model is Riemann integrable in the sense of definition 17.

We have thus shown that $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \mathbf{Sep}(\mathbf{s}, \mathbf{e})$ if and only if the scenario (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model. \square

C.4 Ontic space cardinality

We now prove theorem 3.

Theorem 3 (Ontic space cardinality bounds). *For any (\mathbf{s}, \mathbf{e}) that admits an operationally noncontextual ontological model with ontic state space Λ , it holds that either Λ is an infinite set, or it is discrete and respects*

$$\dim(\mathcal{R}) \leq |\Lambda|. \quad (\text{3.27})$$

Furthermore, if (\mathbf{s}, \mathbf{e}) admits a Riemann integrable operationally noncontextual ontological model (definition 17), then there exists an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) with discrete ontic space Λ_{\min} which satisfies

$$\dim(\mathcal{R}) \leq |\Lambda_{\min}| \leq \dim(\mathcal{R})^2 \leq \dim(\mathcal{L}(\mathcal{H}))^2. \quad (\text{3.28})$$

Proof. Suppose that (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model with discrete ontic space $\Lambda = \{1, \dots, |\Lambda|\}$. Note that such a model is

automatically Riemann integrable according to definition 17: the primitives Δ and $\lambda^{(\text{dis})}$ take the form

$$\Delta_{N,k} = \begin{cases} 1 & \text{if } k \leq |\Lambda|, \\ 0 & \text{else,} \end{cases} \quad (\text{C.82a})$$

$$\lambda_{N,k}^{(\text{dis})} = \min(k, |\Lambda|). \quad (\text{C.82b})$$

Building upon the proof of the unit separability criterion, theorem 2, we see that equations (C.67) and (C.70) taken together in the case when Λ is discrete imply

$$\mathbb{J}(\mathbb{1}_{\mathcal{R}}) = \sum_{i=1}^{|\Lambda|} F_i \otimes \sigma_i. \quad (\text{C.83})$$

Equation (C.83) together with lemma C.13 imply

$$|\Lambda| \geq \text{rank}(\mathbb{1}_{\mathcal{R}}) = \dim(\mathcal{R}). \quad (\text{C.84})$$

This proves the first part of theorem 3.

We now prove the second part. If there exists a Riemann integrable operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) , then $\mathbb{J}(\mathbb{1}_{\mathcal{R}}) \in \mathbf{Sep}(\mathbf{s}, \mathbf{e})$ by the unit separability criterion theorem 2. By proposition C.8, there exists a decomposition of $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ over n elements of $\mathbf{Prod}(\mathbf{s}, \mathbf{e})$ where $n \leq \dim(\mathcal{R})^2$. Assume that n is minimal, i.e., that this decomposition of $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ over elements of $\mathbf{Prod}(\mathbf{s}, \mathbf{e})$ is the shortest one. In particular this implies that there are no zero elements in the decomposition so that we are in the case considered in equation (C.77) in the proof of theorem 2. This decomposition allows one to construct a valid operationally noncontextual ontological model of cardinality $|\Lambda| = n$ (where $n \leq \dim(\mathcal{R})^2$) as demonstrated in, e.g., equation (C.81). This cardinality is minimal: if there was an ontic space of cardinality $n' < n$, then equation (C.83) would yield a decomposition of $\mathbb{J}(\mathbb{1}_{\mathcal{R}})$ over $n' < n$ elements of $\mathbf{Prod}(\mathbf{s}, \mathbf{e})$ whereas n was assumed minimal. The already proven first part of theorem 3 also proves that $n \geq \dim(\mathcal{R})$. Finally, the fact that $\dim(\mathcal{R})^2 \leq \dim(\mathcal{L}(\mathcal{H}))^2$ follows easily from the fact that $\mathcal{R} \subseteq \mathcal{L}(\mathcal{H})$. \square

D Alternative reduced spaces

D.1 Construction of the alternative operationally noncontextual ontological model

Let us restate and prove proposition 19:

Proposition 19. *The choice*

$$\mathcal{R}_{\text{alt}} := \mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s})), \quad (\text{3.31a})$$

$$f(\cdot) := P_{\mathcal{R}}(\cdot), \quad (\text{3.31b})$$

$$g(\cdot) := P_{\mathcal{R}}(\cdot) \quad (\text{3.31c})$$

yields a valid alternative reduced space in definition 18; and so does the swapped version

$$\mathcal{R}_{\text{alt}} := P_{\text{span}(\mathbf{s})}(\text{span}(\mathbf{e})) =: \mathcal{R}', \quad (3.32a)$$

$$f(\cdot) := P_{\mathcal{R}'}(\cdot), \quad (3.32b)$$

$$g(\cdot) := P_{\mathcal{R}'}(\cdot). \quad (3.32c)$$

Proof. First off, the choice $\mathcal{R} = P_{\text{span}(\mathbf{e})}(\text{span}(\mathbf{s}))$ together with $f, g = P_{\mathcal{R}}$ satisfies (3.30a) by virtue of proposition 4, and satisfies equations (3.30b), (3.30c) by virtue of corollary B.6. Thus, this choice fits in as a special case of definition 18.

Now consider the choice $\mathcal{R}' = P_{\text{span}(\mathbf{s})}(\text{span}(\mathbf{e}))$ together with $f, g = P_{\mathcal{R}'}$. It satisfies (3.30a) as a corollary of proposition B.4, and equations (3.30b) and (3.30c) are verified as an application of lemma B.5. \square

We will make use of the following lemma.

Lemma D.1. *Consider any choice of \mathcal{R}_{alt} , f and g as in definition 18. For all $s \in \text{span}(\mathbf{s})$, for all $e \in \text{span}(\mathbf{e})$, it holds that*

$$\langle P_{\mathcal{R}}(s), P_{\mathcal{R}}(e) \rangle_{\mathcal{R}} = \langle f(s), g(e) \rangle_{\mathcal{R}_{\text{alt}}}. \quad (D.1)$$

Proof. It suffices to extend by linearity proposition 4 and (3.30a) of definition 18:

$$\langle P_{\mathcal{R}}(s), P_{\mathcal{R}}(e) \rangle_{\mathcal{R}} = \langle s, e \rangle_{\mathcal{L}(\mathcal{H})}, \quad (D.2a)$$

$$\langle f(s), g(e) \rangle_{\mathcal{R}_{\text{alt}}} = \langle s, e \rangle_{\mathcal{L}(\mathcal{H})}. \quad \square$$

We now restate and prove the equality between the dimensions of the alternative reduced spaces.

Proposition 20. *It holds that for any reduced space \mathcal{R}_{alt} (definition 18), $\dim(\mathcal{R}_{\text{alt}}) = \dim(\mathcal{R})$ where \mathcal{R} is defined in definition 3.*

Proof. We will first prove the existence of an invertible linear map between the two vector spaces \mathcal{R}_{alt} and \mathcal{R} . Then, theorem 2.35 in [23] allows to conclude that $\dim(\mathcal{R}_{\text{alt}}) = \dim(\mathcal{R})$.

Let $d := \dim(\mathcal{R})$ and $d_{\text{alt}} := \dim(\mathcal{R}_{\text{alt}})$ (which is finite by definition 18). Let $\{T_i \in \mathcal{R}_{\text{alt}}\}_{i=1}^{d_{\text{alt}}}$ be an orthonormal basis of \mathcal{R}_{alt} . By equation (3.30b), it is possible to find for all $i = 1, \dots, d_{\text{alt}}$ an element $s_i \in \text{span}(\mathbf{s})$ such that $T_i = f(s_i)$.

Now let $\{R_j \in \mathcal{R}\}_{j=1}^d$ be an orthonormal basis of \mathcal{R} . Then, using corollary B.6, choose for all $j = 1, \dots, d$ elements $u_j \in \text{span}(\mathbf{s})$ such that $R_j = P_{\mathcal{R}}(u_j)$.

Let $\Phi : \mathcal{R} \rightarrow \mathcal{R}_{\text{alt}}$ and $\phi : \mathcal{R}_{\text{alt}} \rightarrow \mathcal{R}$ be defined by: for all $r \in \mathcal{R}$, for all $t \in \mathcal{R}_{\text{alt}}$,

$$\Phi(r) = \sum_{i=1}^{d_{\text{alt}}} \langle P_{\mathcal{R}}(s_i), r \rangle_{\mathcal{R}} T_i, \quad (D.3)$$

$$\phi(t) = \sum_{j=1}^d \langle f(u_j), t \rangle_{\mathcal{R}_{\text{alt}}} R_j. \quad (D.4)$$

The main property of Φ is that, for any $e \in \text{span}(\mathbf{e})$, $\Phi(P_{\mathcal{R}}(e)) = g(e)$. Indeed, using lemma D.1,

$$\begin{aligned} \Phi(P_{\mathcal{R}}(e)) &= \sum_{i=1}^{d_{\text{alt}}} \langle P_{\mathcal{R}}(s_i), P_{\mathcal{R}}(e) \rangle_{\mathcal{R}} T_i \\ &= \sum_{i=1}^{d_{\text{alt}}} \langle f(s_i), g(e) \rangle_{\mathcal{R}_{\text{alt}}} T_i \\ &= \sum_{i=1}^{d_{\text{alt}}} \langle T_i, g(e) \rangle_{\mathcal{R}_{\text{alt}}} T_i = g(e). \end{aligned} \quad (D.5)$$

In the last line, we used the completeness relation of \mathcal{R}_{alt} . Similarly, for any $e \in \text{span}(\mathbf{e})$, it holds that $\phi(g(e)) = P_{\mathcal{R}}(e)$. Indeed:

$$\begin{aligned} \phi(g(e)) &= \sum_{j=1}^d \langle f(u_j), g(e) \rangle_{\mathcal{R}} R_j \\ &= \sum_{j=1}^d \langle P_{\mathcal{R}}(u_j), P_{\mathcal{R}}(e) \rangle_{\mathcal{R}} R_j \\ &= \sum_{j=1}^d \langle R_j, P_{\mathcal{R}}(e) \rangle_{\mathcal{R}} R_j = P_{\mathcal{R}}(e), \end{aligned} \quad (D.6)$$

where we used the completeness relation of \mathcal{R} .

Let us now compute $\phi(\Phi(r))$ for any $r \in \mathcal{R}$. By corollary B.6, there must exist $e \in \text{span}(\mathbf{e})$ such that $r = P_{\mathcal{R}}(e)$. Using (D.5) and (D.6):

$$\begin{aligned} \phi(\Phi(r)) &= \phi(\Phi(P_{\mathcal{R}}(e))) = \phi(g(e)) = P_{\mathcal{R}}(e) \\ &= r. \end{aligned} \quad (D.7)$$

Similarly, for any $t \in \mathcal{R}_{\text{alt}}$, there exists by equation (3.30c) $e' \in \text{span}(\mathbf{e})$ such that $t = g(e')$. Then, using (D.5) and (D.6) again:

$$\begin{aligned} \Phi(\phi(t)) &= \Phi(\phi(g(e'))) = \Phi(P_{\mathcal{R}}(e')) = g(e') \\ &= t. \end{aligned} \quad (D.8)$$

Thus we have proven that $\phi = \Phi^{-1}$, and hence $\Phi : \mathcal{R} \rightarrow \mathcal{R}_{\text{alt}}$ is an invertible linear map. The claim follows from the fact that two vector spaces are isomorphic if and only if they have the same dimension as proven, e.g., in theorem 2.35 of [23]. \square

Let us now define the operationally noncontextual ontological model on any alternative reduced space \mathcal{R}_{alt} .

Definition D.2. The operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) on a given alternative reduced space \mathcal{R}_{alt} is specified as follows. \mathcal{R}_{alt} and the mappings f and g are the primitives of definition 18. Let Λ_{alt} be the ontic space. Let μ_{alt} be the ontic state mapping that has domain

$$\mu_{\text{alt}} : f(\mathbf{s}) \times \Lambda_{\text{alt}} \rightarrow \mathbb{R} \quad (\text{D.9})$$

and that satisfies

$$\forall \bar{\rho} \in f(\mathbf{s}) : \int_{\Lambda_{\text{alt}}} d\lambda \mu_{\text{alt}}(\bar{\rho}, \lambda) = 1, \quad (\text{D.10a})$$

$$\forall \lambda \in \Lambda_{\text{alt}}, \forall \bar{\rho} \in f(\mathbf{s}) : \mu_{\text{alt}}(\bar{\rho}, \lambda) \geq 0, \quad (\text{D.10b})$$

$$\forall \lambda \in \Lambda_{\text{alt}}, \forall p \in [0, 1], \forall \bar{\rho}_1, \bar{\rho}_2 \in f(\mathbf{s}) :$$

$$\begin{aligned} \mu_{\text{alt}}(p\bar{\rho}_1 + (1-p)\bar{\rho}_2, \lambda) \\ = p\mu_{\text{alt}}(\bar{\rho}_1, \lambda) + (1-p)\mu_{\text{alt}}(\bar{\rho}_2, \lambda). \end{aligned} \quad (\text{D.10c})$$

Let ξ_{alt} be the response function mapping that has domain

$$\xi_{\text{alt}} : g(\mathbf{e}) \times \Lambda \rightarrow \mathbb{R} \quad (\text{D.11a})$$

and that satisfies

$$\forall \lambda \in \Lambda_{\text{alt}}, \forall K \in \mathbb{N} \cup \{+\infty\},$$

$$\forall \left\{ E_k \in \mathbf{e} : \sum_{k=1}^K E_k = \mathbb{1}_{\mathcal{H}} \right\}_{k=1}^K : \sum_{k=1}^K \xi_{\text{alt}}(g(E_k), \lambda) = 1, \quad (\text{D.11b})$$

$$\forall \lambda \in \Lambda_{\text{alt}}, \forall \bar{E} \in g(\mathbf{e}) : \xi_{\text{alt}}(\bar{E}, \lambda) \geq 0, \quad (\text{D.11c})$$

$$\forall \lambda \in \Lambda_{\text{alt}}, \forall p \in [0, 1], \forall \bar{E}_1, \bar{E}_2 \in g(\mathbf{e}) :$$

$$\begin{aligned} \xi_{\text{alt}}(p\bar{E}_1 + (1-p)\bar{E}_2, \lambda) \\ = p\xi_{\text{alt}}(\bar{E}_1, \lambda) + (1-p)\xi_{\text{alt}}(\bar{E}_2, \lambda). \end{aligned} \quad (\text{D.11d})$$

The operationally noncontextual ontological model is required to reproduce the statistics that quantum mechanics predicts for the available states and measurements — this is formulated using (3.30a):

$$\forall \bar{\rho} \in f(\mathbf{s}), \forall \bar{E} \in g(\mathbf{e}) :$$

$$\langle \bar{\rho}, \bar{E} \rangle_{\mathcal{R}_{\text{alt}}} = \int_{\Lambda_{\text{alt}}} d\lambda \mu_{\text{alt}}(\bar{\rho}, \lambda) \xi_{\text{alt}}(\bar{E}, \lambda). \quad (\text{D.12})$$

Lemma D.3. For any reduced space \mathcal{R}_{alt} (definition 18), for all $\bar{\rho} \in f(\mathbf{s})$,

$$\langle \bar{\rho}, g(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}_{\text{alt}}} = 1. \quad (\text{D.13})$$

Proof. Simply note that there must exist $\rho \in \mathbf{s}$ such that $\bar{\rho} = f(\rho)$, and then by virtue of (3.30a),

$$\langle \bar{\rho}, g(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}_{\text{alt}}} = \langle \rho, \mathbb{1}_{\mathcal{H}} \rangle_{\mathcal{L}(\mathcal{H})} = 1. \quad \square \quad (\text{D.14})$$

Proposition D.4. Let $\lambda \in \Lambda_{\text{alt}}$ be arbitrary. Starting from the convex-linear mappings

$$\mu_{\text{alt}}(\cdot, \lambda) : f(\mathbf{s}) \rightarrow \mathbb{R}, \quad (\text{D.15a})$$

$$\xi_{\text{alt}}(\cdot, \lambda) : g(\mathbf{e}) \rightarrow \mathbb{R}, \quad (\text{D.15b})$$

there exist unique linear extensions

$$\mu_{\text{alt}}(\cdot, \lambda) : \mathcal{R}_{\text{alt}} \rightarrow \mathbb{R}, \quad (\text{D.16a})$$

$$\xi_{\text{alt}}(\cdot, \lambda) : \mathcal{R}_{\text{alt}} \rightarrow \mathbb{R}. \quad (\text{D.16b})$$

Proof. The proof is the same as those of proposition B.12 and proposition B.13: the same constructions apply in this case. The results which needed to be verified are the span assumptions (3.30b) and (3.30c) as well as lemma D.3 for the case of μ_{alt} : replace $\text{Tr}_{\mathcal{H}}[\bar{\rho}]$ with $\langle \bar{\rho}, g(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}_{\text{alt}}}$ in the proof of proposition B.12. \square

Theorem D.1 (Basic criterion for the existence of an operationally noncontextual ontological model on \mathcal{R}_{alt}). Given (\mathbf{s}, \mathbf{e}) that lead to an alternative reduced space \mathcal{R}_{alt} with associated mappings f, g (definition 18), there exists an operationally noncontextual ontological model on \mathcal{R}_{alt} with ontic space Λ_{alt} if and only if there exist mappings $F_{\text{alt}}, \sigma_{\text{alt}}$ with ranges

$$F_{\text{alt}} : \Lambda_{\text{alt}} \rightarrow f(\mathbf{s})^{+\mathcal{R}_{\text{alt}}}, \quad (\text{D.17a})$$

$$\sigma_{\text{alt}} : \Lambda_{\text{alt}} \rightarrow g(\mathbf{e})^{+\mathcal{R}_{\text{alt}}}, \quad (\text{D.17b})$$

satisfying the normalization condition

$$\forall \lambda \in \Lambda_{\text{alt}} : \langle \sigma_{\text{alt}}(\lambda), g(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}_{\text{alt}}} = 1 \quad (\text{D.18})$$

as well as the consistency requirement: $\forall r, s \in \mathcal{R}_{\text{alt}}$,

$$\langle r, s \rangle_{\mathcal{R}_{\text{alt}}} = \int_{\Lambda_{\text{alt}}} d\lambda \langle r, F_{\text{alt}}(\lambda) \rangle_{\mathcal{R}_{\text{alt}}} \langle \sigma_{\text{alt}}(\lambda), s \rangle_{\mathcal{R}_{\text{alt}}}. \quad (\text{D.19})$$

Proof. It suffices to apply Riesz' representation theorem B.1 to the linear extensions of proposition D.4 and to read off their properties from the definition D.2 of the operationally noncontextual ontological model similarly to what was done in the proof of theorem 1. \square

D.2 Equivalence of reduced spaces

Let us restate and prove theorem 4.

Theorem 4 (Equivalence of reduced spaces). *Given any (\mathbf{s}, \mathbf{e}) , consider \mathcal{R} and any \mathcal{R}_{alt} constructed from (\mathbf{s}, \mathbf{e}) . There exists an operationally noncontextual ontological model with ontic space Λ constructed on \mathcal{R} (definition 6) if and only if there exists an operationally noncontextual ontological model constructed on \mathcal{R}_{alt} (appendix definition D.2) with the same ontic space Λ .*

Proof. Suppose first that (\mathbf{s}, \mathbf{e}) admits an operationally noncontextual ontological model on \mathcal{R} in the form of definition 6 with ontic space Λ : we will show that this implies the existence of an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) on \mathcal{R}_{alt} equipped with the same ontic space Λ .

The existence of an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) of \mathcal{R} is equivalent to the existence of Λ , F and σ as in theorem 1. Consider corollary B.6, which we restate in a slightly different form:

$$P_{\mathcal{R}}(\text{span}(\mathbf{s})) = \mathcal{R}, \quad (\text{D.20a})$$

$$P_{\mathcal{R}}(\text{span}(\mathbf{e})) = \mathcal{R}. \quad (\text{D.20b})$$

This implies that for any element $r \in \mathcal{R}$, there exists $s \in \text{span}(\mathbf{s})$ such that $r = P_{\mathcal{R}}(s)$; and there exists $e \in \text{span}(\mathbf{e})$ such that $r = P_{\mathcal{R}}(e)$. Applying this reasoning for every $\lambda \in \Lambda$ to $F(\lambda) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}} \subseteq \mathcal{R}$, $\sigma(\lambda) \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}} \subseteq \mathcal{R}$, there must exist mappings

$$S : \Lambda \rightarrow \text{span}(\mathbf{s}), \quad (\text{D.21a})$$

$$E : \Lambda \rightarrow \text{span}(\mathbf{e}), \quad (\text{D.21b})$$

such that for all $\lambda \in \Lambda$, the primitives from theorem 1 satisfy

$$F(\lambda) = P_{\mathcal{R}}(E(\lambda)), \quad (\text{D.22a})$$

$$\sigma(\lambda) = P_{\mathcal{R}}(S(\lambda)). \quad (\text{D.22b})$$

At this point, the existence of an operationally noncontextual ontological model on \mathcal{R} is equivalent to the existence of the mappings S and E as in (D.22). But the existence of these mappings implies the existence of a valid operationally noncontextual ontological model on \mathcal{R}_{alt} : indeed, define for all $\lambda \in \Lambda$,

$$F_{\text{alt}}(\lambda) := g(E(\lambda)), \quad (\text{D.23a})$$

$$\sigma_{\text{alt}}(\lambda) := f(S(\lambda)). \quad (\text{D.23b})$$

Let us verify that the primitives F_{alt} , σ_{alt} satisfy all the requirements of theorem D.1, starting with the verification of (D.19). Let $r, s \in \mathcal{R}_{\text{alt}}$ be arbitrary. By equations (3.30b) and (3.30c), there exist $r' \in \text{span}(\mathbf{s})$, $s' \in \text{span}(\mathbf{e})$ such that $r = f(r')$ and $s = g(s')$. Using lemma D.1:

$$\begin{aligned} \langle r, s \rangle_{\mathcal{R}_{\text{alt}}} &= \langle f(r'), g(s') \rangle_{\mathcal{R}_{\text{alt}}} \\ &= \langle P_{\mathcal{R}}(r'), P_{\mathcal{R}}(s') \rangle_{\mathcal{R}}. \end{aligned} \quad (\text{D.24})$$

Using (2.24) to expand the last term of the previous equation, the definition (D.22) of the mappings E , S and lemma D.1:

$$\begin{aligned} \langle r, s \rangle_{\mathcal{R}_{\text{alt}}} &= \int_{\Lambda} d\lambda \langle P_{\mathcal{R}}(r'), F(\lambda) \rangle_{\mathcal{R}} \langle \sigma(\lambda), P_{\mathcal{R}}(s') \rangle_{\mathcal{R}} \\ &= \int_{\Lambda} d\lambda \langle P_{\mathcal{R}}(r'), P_{\mathcal{R}}(E(\lambda)) \rangle_{\mathcal{R}} \langle P_{\mathcal{R}}(S(\lambda)), P_{\mathcal{R}}(s') \rangle_{\mathcal{R}} \\ &= \int_{\Lambda} d\lambda \langle f(r'), g(E(\lambda)) \rangle_{\mathcal{R}} \langle f(S(\lambda)), g(s') \rangle_{\mathcal{R}} \\ &= \int_{\Lambda} d\lambda \langle r, F_{\text{alt}}(\lambda) \rangle_{\mathcal{R}_{\text{alt}}} \langle \sigma_{\text{alt}}(\lambda), s \rangle_{\mathcal{R}_{\text{alt}}}. \end{aligned} \quad (\text{D.25})$$

This derivation being valid for all $r, s \in \mathcal{R}_{\text{alt}}$ proves (D.19). Equations (D.17) and (D.18) are verified similarly. This proves that if there exists an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) constructed on \mathcal{R} with ontic space Λ , then there exists an operationally noncontextual ontological model for (\mathbf{s}, \mathbf{e}) constructed on \mathcal{R}_{alt} with the same ontic space Λ .

The other direction is proven analogously. Suppose that there exists an operationally noncontextual ontological model constructed on \mathcal{R}_{alt} with ontic space Λ_{alt} . Starting with the primitives F_{alt} , σ_{alt} of theorem D.1, choose for all $\lambda \in \Lambda_{\text{alt}}$ elements $E_{\text{alt}}(\lambda) \in \text{span}(\mathbf{e})$ such that $P_{\mathcal{R}_{\text{alt}}}(E_{\text{alt}}(\lambda)) = F_{\text{alt}}(\lambda)$, and $S_{\text{alt}}(\lambda) \in \text{span}(\mathbf{s})$ such that $P_{\mathcal{R}_{\text{alt}}}(S_{\text{alt}}(\lambda)) = \sigma_{\text{alt}}(\lambda)$. This is always possible thanks to equations (3.30b) and (3.30c). Then, define for all $\lambda \in \Lambda_{\text{alt}}$:

$$F(\lambda) := P_{\mathcal{R}}(E_{\text{alt}}(\lambda)), \quad (\text{D.26a})$$

$$\sigma(\lambda) := P_{\mathcal{R}}(S_{\text{alt}}(\lambda)). \quad (\text{D.26b})$$

It is then easy to verify that these F and σ satisfy the requirements of theorem 1, and thus that there exists an operationally noncontextual ontological model on \mathcal{R} with the same ontic space Λ_{alt} . This concludes the proof. \square

E Algorithmic formulation, witnesses and certifiers

E.1 Resolution of convex cones

Throughout this section, \mathcal{V} is a finite dimensional real inner product space.

Let us recall the definition of a pointed cone, then prove proposition 22.

Definition 21 (Pointed cone). *Let $\mathcal{C} \subseteq \mathcal{V}$ be a convex cone. \mathcal{C} is said to be a pointed cone if*

- (i) \mathcal{C} is closed;
- (ii) $\mathcal{C} \neq \emptyset$ and $\mathcal{C} \neq \{0\}$;
- (iii) there exists a linear function $L : \mathcal{V} \rightarrow \mathbb{R}$ such that for all $c \in \mathcal{C} \setminus \{0\}$, $L(c) > 0$.

Lemma E.1 (Adapted from theorem 8.1.3 in [20]). *For any nonempty compact convex set $X \subseteq \mathcal{V}$, it holds that X is the convex hull of its extremal points (definition A.3):*

$$X = \text{conv}(\text{ep}(X)). \quad (\text{E.1})$$

The following lemma relates the extremal half-lines of a given pointed cone to the extremal points of a certain compact convex “slice” of the cone.

Lemma E.2. *If $\mathcal{C} \subseteq \mathcal{V}$ is a pointed cone, then let L be the linear functional of definition 21 and let*

$$H := \{v \in \mathcal{V} : L(v) = 1\}. \quad (\text{E.2})$$

It holds that $\mathcal{C} \cap H$ is a nonempty compact convex set such that:

$$\text{extr}(\mathcal{C}) = \{\text{coni}(c) : c \in \text{ep}(\mathcal{C} \cap H)\}. \quad (\text{E.3})$$

Proof. $\mathcal{C} \cap H$ is a nonempty set: indeed, by the definition 21 of pointed cones, $\mathcal{C} \setminus \{0\}$ is a nonempty set, and for any $c \in \mathcal{C} \setminus \{0\}$ which is a nonempty set by definition 21, it holds that $L(c) > 0$. But then, $c/L(c) \in \mathcal{C}$ satisfies $L(c/L(c)) = 1$, so that $c/L(c) \in \mathcal{C} \cap H$. Clearly, $\mathcal{C} \cap H$ is a closed set: the closure follows from the closure of \mathcal{C} , by definition 21 and the closure of H in (E.2). The convexity of $\mathcal{C} \cap H$ follows from the convexity of \mathcal{C} and that of H .

Let us prove that $\mathcal{C} \cap H$ is bounded. We reason by contradiction: suppose that $\mathcal{C} \cap H$ was not bounded. Following the reasoning of section 8.1 of [20], the unboundedness of the closed convex set $\mathcal{C} \cap H$ is equivalent to the existence of $d \in \mathcal{V}$, $d \neq 0$, $c \in \mathcal{C}$ such that for all $\lambda \in \mathbb{R}_{\geq 0}$, $(c + \lambda d) \in \mathcal{C} \cap H$. Clearly, this implies $L(d) = 0$. But by the convex structure of \mathcal{C} , it holds that for any $\lambda \geq 1$, the following convex combination lies in \mathcal{C} :

$$(1 - \frac{1}{\lambda})0 + \frac{1}{\lambda}(c + \lambda d) \in \mathcal{C}. \quad (\text{E.4})$$

By the closure of \mathcal{C} , the limit point belongs to \mathcal{C} :

$$d = \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda}(c + \lambda d) \in \mathcal{C}. \quad (\text{E.5})$$

Thus, d is a point of \mathcal{C} that satisfies $L(d) = 0$: this implies $d = 0$ which is a contradiction. This proves that $\mathcal{C} \cap H$ is bounded. This proves that $\mathcal{C} \cap H$ is a nonempty compact convex set.

Let us now prove (E.3). For any $\mathfrak{l} \in \text{extr}(\mathcal{C})$, there exists $c_0 \in \mathfrak{l}$ be such that $L(c_0) = 1$. We will show that $c_0 \in \text{ep}(\mathcal{C} \cap H)$. Consider any $\lambda \in]0, 1[$ and $c_1, c_2 \in \mathcal{C} \cap H$ such that

$$c_0 = \lambda c_1 + (1 - \lambda)c_2. \quad (\text{E.6})$$

Then, because $c_1, c_2 \in \mathcal{C} \cap H \subset \mathcal{C}$, and because c_0 is part of an extremal direction (definition A.6) of \mathcal{C} , this implies without loss of generality (if $c_2 = 0$ and

$c_1 \neq 0$, swap c_1 with c_2) that there exists $\alpha \in \mathbb{R}$ such that $c_1 = \alpha c_2$. Applying the linear map L to this equation and recalling $c_1, c_2 \in H$ shows that $\alpha = 1$, and thus $c_1 = c_2$. Together with (E.6), $c_1 = c_2 = c_0$. This proves that $c_0 \in \text{ep}(\mathcal{C} \cap H)$. This shows that for any $\mathfrak{l} \in \text{extr}(\mathcal{C})$, there exists $c_0 \in \text{ep}(\mathcal{C} \cap H)$ such that

$$\mathfrak{l} = \text{coni}(c_0), \quad (\text{E.7})$$

and thus it holds that

$$\text{extr}(\mathcal{C}) \subseteq \{\text{coni}(c) : c \in \text{ep}(\mathcal{C} \cap H)\}. \quad (\text{E.8})$$

Now, let $c_0 \in \text{ep}(\mathcal{C} \cap H)$, and we will show that $\text{coni}(c_0) \in \text{extr}(\mathcal{C})$. It suffices to prove that c_0 is an extreme direction of \mathcal{C} according to definition A.6. Let $d_1, d_2 \in \mathcal{C}$ be such that $c_0 = d_1 + d_2$. Note that $1 = L(d_1) + L(d_2)$ due to $c_0 \in H$, and since $d_1, d_2 \in \mathcal{C}$: $L(d_1), L(d_2) \geq 0$, which implies in particular $L(d_1), L(d_2) \in [0, 1]$. We will show that d_1 and d_2 are linearly dependent. If either d_1 or d_2 is zero this is trivial, so assume that they are both nonzero. Thus we have $L(d_1), L(d_2) \in]0, 1[$. Then, rewrite

$$c_0 = L(d_1)\frac{d_1}{L(d_1)} + (1 - L(d_1))\frac{d_2}{L(d_2)}. \quad (\text{E.9})$$

Due to $d_i/L(d_i) \in \mathcal{C} \cap H$ for $i = 1, 2$ and the fact that c_0 is an extremal point of $\mathcal{C} \cap H$, the convex combination (E.9) implies

$$\frac{d_1}{L(d_1)} = \frac{d_2}{L(d_2)}, \quad (\text{E.10})$$

which implies the linear dependence of d_1 and d_2 . Thus, c_0 is an extremal direction of \mathcal{C} and $\text{coni}(c_0) \in \text{extr}(\mathcal{C})$. Since c_0 was arbitrary in $\text{ep}(\mathcal{C} \cap H)$, this proves the reverse inclusion to (E.8). Thus, the equality

$$\text{extr}(\mathcal{C}) = \{\text{coni}(c) : c \in \text{ep}(\mathcal{C} \cap H)\} \quad (\text{E.11})$$

holds. \square

Proposition 22. *If $\mathcal{C} \subseteq \mathcal{V}$ is a pointed cone, then it holds that*

$$\mathcal{C} = \text{conv}\left(\bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C})} \mathfrak{l}\right). \quad (\text{A.1})$$

Proof. By applying lemma E.1 to the nonempty compact convex set $\mathcal{C} \cap H$ (with H defined in lemma E.2, and the nonempty compact convex set also established in lemma E.2):

$$\mathcal{C} \cap H = \text{conv}(\text{ep}(\mathcal{C} \cap H)). \quad (\text{E.12})$$

It is clear that $\mathcal{C} = \text{coni}(\mathcal{C} \cap H)$. It is also easy to show that for any set $X \subseteq \mathcal{V}$, $\text{conv}(\text{coni}(X)) = \text{coni}(\text{conv}(X))$. Thus,

$$\begin{aligned} \mathcal{C} &= \text{coni}(\mathcal{C} \cap H) = \text{coni}(\text{conv}(\text{ep}(\mathcal{C} \cap H))) \\ &= \text{conv}(\text{coni}(\text{ep}(\mathcal{C} \cap H))). \end{aligned} \quad (\text{E.13})$$

Note that $\text{coni}(\text{ep}(\mathcal{C} \cap H)) = \bigcup_{c \in \text{ep}(\mathcal{C} \cap H)} \text{coni}(c)$. By (E.3) of lemma E.2, it holds that

$$\text{coni}(\text{ep}(\mathcal{C} \cap H)) = \bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C})} \mathfrak{l}. \quad (\text{E.14})$$

Together with (E.13) this proves

$$\mathcal{C} = \text{conv}\left(\bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C})} \mathfrak{l}\right). \quad \square \quad (\text{E.15})$$

Let us restate definition 23 and prove proposition 24.

Definition 23 (Spanning cone). *A convex cone $\mathcal{C} \subseteq \mathcal{V}$ is a spanning cone in \mathcal{V} if*

- (i) \mathcal{C} is closed;
- (ii) $\mathcal{C} \neq \mathcal{V}$;
- (iii) $\text{span}(\mathcal{C}) = \mathcal{V}$.

We will need the following lemma.

Lemma E.3. *Let $\mathcal{C} \subseteq \mathcal{V}$ be a nonempty closed convex cone. Then,*

$$[\mathcal{C}^{+\nu}]^{+\nu} = \mathcal{C}. \quad (\text{E.16})$$

Proof. One direction is easy to verify directly:

$$\mathcal{C} \subseteq [\mathcal{C}^{+\nu}]^{+\nu}. \quad (\text{E.17})$$

Let us consider the other direction. Suppose there existed $v_0 \in [\mathcal{C}^{+\nu}]^{+\nu}$ such that $v_0 \notin \mathcal{C}$. Using that \mathcal{C} is closed, by theorem 3.2.2 in [20], there exists $n \in \mathcal{V}$, $\alpha \in \mathbb{R}$ such that

$$\forall c \in \mathcal{C} : \langle n, c \rangle_{\mathcal{V}} \geq \alpha, \quad (\text{E.18a})$$

$$\langle n, v_0 \rangle_{\mathcal{V}} < \alpha. \quad (\text{E.18b})$$

Because $0 \in \mathcal{C}$ for any cone, (E.18a) implies $\alpha \leq 0$. This proves that

$$\langle n, v_0 \rangle_{\mathcal{V}} < 0. \quad (\text{E.19})$$

The conic structure of \mathcal{C} and (E.18a) imply, for all $c \in \mathcal{C}$,

$$\forall \lambda \in \mathbb{R}_{\geq 0} : \langle n, \lambda c \rangle_{\mathcal{V}} \geq \alpha, \quad (\text{E.20})$$

or, for all $\lambda > 0$, $\langle n, c \rangle_{\mathcal{V}} \geq (\alpha/\lambda)$. By taking the limit $\lambda \rightarrow \infty$, this proves that $\langle n, c \rangle_{\mathcal{V}} \geq 0$, so that

$$\forall c \in \mathcal{C} : \langle n, c \rangle_{\mathcal{V}} \geq 0, \quad (\text{E.21a})$$

$$\langle n, v_0 \rangle_{\mathcal{V}} < 0. \quad (\text{E.21b})$$

Equation (E.21a) proves that $n \in \mathcal{C}^{+\nu}$. Thus, it must hold that since $v_0 \in [\mathcal{C}^{+\nu}]^{+\nu}$, $\langle n, v_0 \rangle_{\mathcal{V}} \geq 0$, which contradicts equation (E.21b). This contradiction proves that there cannot exist $v_0 \in [\mathcal{C}^{+\nu}]^{+\nu}$ and $v_0 \notin \mathcal{C}$, which proves that $[\mathcal{C}^{+\nu}]^{+\nu} \subseteq \mathcal{C}$. Thus both inclusions are proven and the claim follows. \square

Proposition 24. *If $\mathcal{C} \subseteq \mathcal{V}$ is a spanning cone, then the polar cone $\mathcal{C}^{+\nu} \subseteq \mathcal{V}$ (definition 10) is a pointed cone, which implies by proposition 22 that*

$$\mathcal{C}^{+\nu} = \text{conv}\left(\bigcup_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})} \mathfrak{l}\right). \quad (\text{4.2})$$

Proof. Let us prove that $\mathcal{C}^{+\nu}$ is a pointed cone by verifying the assumptions of definition 21. Clearly, from definition 10 of the polar cone, $\mathcal{C}^{+\nu}$ is closed. Now, let us prove that $\mathcal{C}^{+\nu} \neq \{0\}$. We reason by contradiction: if $\mathcal{C}^{+\nu} = \{0\}$, this means that $[\mathcal{C}^{+\nu}]^{+\nu} = \mathcal{V}$. But since \mathcal{C} is closed by virtue of (i) in definition 23, by lemma E.3, it holds that $[\mathcal{C}^{+\nu}]^{+\nu} = \mathcal{C}$. Thus, if $\mathcal{C}^{+\nu} = \{0\}$, then $\mathcal{C} = \mathcal{V}$, which contradicts assumption (ii) of definition 23.

It remains to verify the property (iii) of pointed cones in definition 21. By property (iii) of spanning cones in definition 23, $\text{span}(\mathcal{C}) = \mathcal{V}$, so that we may choose a basis of \mathcal{V} of the form $\{c_i \in \mathcal{C}\}_{i=1}^{\dim(\mathcal{V})}$. Consider the linear map $L : \mathcal{V} \rightarrow \mathbb{R}$ defined by

$$L(\cdot) := \sum_{i=1}^{\dim(\mathcal{V})} \langle c_i, \cdot \rangle_{\mathcal{V}}. \quad (\text{E.22})$$

Clearly, for all $d \in \mathcal{C}^{+\nu}$, $L(d) \geq 0$. If $L(d) = 0$ for some $d \in \mathcal{C}^{+\nu}$, then by the nonnegativity of each term in the sum we must have for all $i = 1, \dots, \dim(\mathcal{V})$: $\langle c_i, d \rangle_{\mathcal{V}} = 0$. Because the c_i 's span \mathcal{V} , by linearity of the scalar product it holds that $\langle v, d \rangle_{\mathcal{V}} = 0$ for all $v \in \mathcal{V}$. The non-degeneracy of the inner product proves $d = 0$, so that L satisfies indeed assumption (iii) of definition 21.

We have proven that spanning cones defined in definition 23 have polar cones which are pointed cone as defined in definition 21. \square

Proposition 26. *A solution to the vertex enumeration problem allows one to represent a spanning pointed cone $\mathcal{C} \subseteq \mathcal{V}$ as the intersection of half-spaces:*

$$\mathcal{C} = \bigcap_{\mathfrak{l} \in \mathbb{V} \cdot \mathbb{E}_{\mathcal{V}}[\text{extr}(\mathcal{C})]} \mathfrak{l}^{+\nu}. \quad (\text{4.4})$$

Proof. Since \mathcal{C} is a spanning pointed cone, by definition 21 it is closed, so that by lemma E.3,

$$\mathcal{C} = [\mathcal{C}^{+\nu}]^{+\nu} = \{v \in \mathcal{V} : \langle v, d \rangle_{\mathcal{V}} \geq 0 \forall d \in \mathcal{C}^{+\nu}\}. \quad (\text{E.23})$$

In particular, for all $v \in \mathcal{C}$, it holds that $\langle v, d \rangle_{\mathcal{V}} \geq 0$ for all $d \in \mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})$. Thus,

$$\mathcal{C} \subseteq \bigcap_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})} \mathfrak{l}^{+\nu}. \quad (\text{E.24})$$

Now, let $v \in \bigcap_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})} \mathfrak{l}^{+\nu}$. We will show that also $v \in \mathcal{C}$. By equation (E.23), it suffices to verify that

for an arbitrary $d \in \mathcal{C}^{+\nu}$, $\langle v, d \rangle_{\mathcal{V}} \geq 0$. Now, since \mathcal{C} is a spanning cone, by proposition 24, it holds that there must exist a finite number $N \in \mathbb{N}$ of elements d_i each belonging to an extremal line \mathfrak{l}_i of $\mathcal{C}^{+\nu}$ such that

$$d = \sum_{i=1}^N d_i. \quad (\text{E.25})$$

Then, $\langle v, d \rangle_{\mathcal{V}} = \sum_{i=1}^N \langle v, d_i \rangle_{\mathcal{V}}$. Due to $v \in \bigcap_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})} \mathfrak{l}^{+\nu}$, and $d_i \in \mathfrak{l}_i \in \text{extr}(\mathcal{C}^{+\nu})$, it is clear that $\langle v, d \rangle_{\mathcal{V}} \geq 0$. Thus, it holds that

$$\mathcal{C} = \bigcap_{\mathfrak{l} \in \text{extr}(\mathcal{C}^{+\nu})} \mathfrak{l}^{+\nu}. \quad (\text{E.26})$$

Because \mathcal{C} is a spanning pointed cone, the vertex enumeration problem of definition 25 is well-defined and it holds that $\text{extr}(\mathcal{C}^{+\nu}) = \text{V.E}_{\mathcal{V}}[\text{extr}(\mathcal{C})]$, which concludes the proof. \square

E.2 General aspects of the algorithm

Lemma E.4. *It holds that*

$$P_{\mathcal{R}}(\bar{\mathfrak{s}}) = \overline{P_{\mathcal{R}}(\mathfrak{s})}, \quad (\text{E.27a})$$

$$P_{\mathcal{R}}(\bar{\mathfrak{e}}) = \overline{P_{\mathcal{R}}(\mathfrak{e})}. \quad (\text{E.27b})$$

Proof. Let us prove (E.27a). Clearly, $P_{\mathcal{R}}(\bar{\mathfrak{s}}) \subseteq \overline{P_{\mathcal{R}}(\mathfrak{s})}$. Then, let $\bar{\rho}^* \in \overline{P_{\mathcal{R}}(\mathfrak{s})}$ be arbitrary. By lemma C.2, there exists a sequence $(\bar{\rho}_n \in P_{\mathcal{R}}(\mathfrak{s}))_{n \in \mathbb{N}}$ such that $\bar{\rho}^* = \lim_{n \rightarrow \infty} \bar{\rho}_n$. For all n , choose any $\rho_n \in \mathfrak{s}$ such that $P_{\mathcal{R}}(\rho_n) = \bar{\rho}_n$. Because the set \mathfrak{s} is bounded, the sequence $(\rho_n \in \mathfrak{s})_{n \in \mathbb{N}}$ is also bounded. By the Bolzano-Weierstrass theorem C.1, there exists a set of indices $\{n_k\}_k \subseteq \mathbb{N}$ and $\rho^* \in \bar{\mathfrak{s}}$ such that

$$\lim_{k \rightarrow \infty} \rho_{n_k} = \rho^*. \quad (\text{E.28})$$

Then, by lemma C.4, the subsequence $(\bar{\rho}_{n_k})_k$ must converge to $\bar{\rho}^*$, so that

$$\bar{\rho}^* = \lim_{k \rightarrow \infty} \bar{\rho}_{n_k} = \lim_{k \rightarrow \infty} P_{\mathcal{R}}(\rho_{n_k}) = P_{\mathcal{R}}(\rho^*) \in P_{\mathcal{R}}(\bar{\mathfrak{s}}). \quad (\text{E.29})$$

Note that we used the continuity of $P_{\mathcal{R}}(\cdot)$ to conclude. The proof of (E.27b) is analogous and relies on the boundedness of \mathfrak{e} . \square

Proposition 27.

$$P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}} = [\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))]^{+\mathcal{R}}, \quad (\text{4.5a})$$

$$P_{\mathcal{R}}(\mathfrak{e})^{+\mathcal{R}} = [\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))]^{+\mathcal{R}}. \quad (\text{4.5b})$$

Proof. Let us prove equation (4.5a) explicitly. Strictly speaking, we could simply use $\overline{P_{\mathcal{R}}(\mathfrak{s})}$ rather than $P_{\mathcal{R}}(\bar{\mathfrak{s}})$ in every instance it appears, but since the two sets are equal (lemma E.4), for notational purposes we prefer the use of $P_{\mathcal{R}}(\bar{\mathfrak{s}})$. First, note that the polar of a set is equal to the polar of its closure. Indeed, consider $P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}}$. One direction is clear thanks to lemma A.7: $P_{\mathcal{R}}(\bar{\mathfrak{s}})^{+\mathcal{R}} \subseteq P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}}$. Now, consider any $s \in P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}}$, and let us show that also $s \in P_{\mathcal{R}}(\bar{\mathfrak{s}})^{+\mathcal{R}}$. It suffices to show that $\langle s, \bar{\rho} \rangle_{\mathcal{R}} \geq 0$ for any element $\bar{\rho} \in P_{\mathcal{R}}(\bar{\mathfrak{s}})$. Such elements $\bar{\rho}$ can be written as the limit of a converging sequence $(\bar{\rho}_n \in P_{\mathcal{R}}(\mathfrak{s}))_{n \in \mathbb{N}}$, thanks to lemma C.2. Then, by lemma C.1 which states the continuity of the scalar product, and by the closure of the set $\mathbb{R}_{\geq 0}$, it holds that

$$\langle s, \bar{\rho} \rangle_{\mathcal{R}} = \lim_{n \rightarrow \infty} \langle s, \bar{\rho}_n \rangle_{\mathcal{R}} \in \mathbb{R}_{\geq 0}. \quad (\text{E.30})$$

Thus we have $P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}} = P_{\mathcal{R}}(\bar{\mathfrak{s}})^{+\mathcal{R}}$. Clearly, the polar cone of a set X and the polar cone to $\text{coni}(X)$ are equal, so that $P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}} = \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))^{+\mathcal{R}}$. This proves equation (4.5a), and the proof of equation (4.5b) is completely analogous. \square

The following lemmas is an intermediate step towards proving proposition 28.

Lemma E.5. *There exists a linear map $L : \mathcal{R} \rightarrow \mathbb{R}$ such that for all $s \in \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$,*

$$L(s) \geq 0 \quad (\text{E.31})$$

with equality if and only if $s = 0$.

Proof. Choose $L(\cdot) = \langle \cdot, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}}$. For all $s \in \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$, there exists $\lambda \in \mathbb{R}_{\geq 0}$ and $\bar{\rho} \in P_{\mathcal{R}}(\bar{\mathfrak{s}})$ such that $s = \lambda \bar{\rho}$. Then, $L(s) = \lambda L(\bar{\rho}) = \lambda \geq 0$ where we used lemma B.7 to assert $\langle \bar{\rho}, P_{\mathcal{R}}(\mathbb{1}_{\mathcal{H}}) \rangle_{\mathcal{R}} = 1$ (strictly speaking, if $\bar{\rho} \in P_{\mathcal{R}}(\bar{\mathfrak{s}}) \setminus P_{\mathcal{R}}(\mathfrak{s})$, one needs to consider a converging sequence of elements of $P_{\mathcal{R}}(\mathfrak{s})$ to assert that the limit also has unit trace). Also, $L(s) = 0$ implies $\lambda = 0$ so that $s = \lambda \bar{\rho} = 0$. Thus, for all $s \in \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$, $L(s) \geq 0$ with equality if and only if $s = 0$. \square

Lemma E.6. *There exists a linear map $L : \mathcal{R} \rightarrow \mathbb{R}$ such that for all $e \in \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$,*

$$L(e) \geq 0 \quad (\text{E.32})$$

with equality if and only if $e = 0$.

Proof. Choose a (non-orthonormal) basis of \mathcal{R} in the form $\{P_{\mathcal{R}}(\rho_i) : \rho_i \in \mathfrak{s}\}_{i=1}^{\dim(\mathcal{R})}$, which is always possible by corollary B.6. Then, define for all $r \in \mathcal{R}$:

$$L(r) = \sum_{i=1}^{\dim(\mathcal{R})} \langle P_{\mathcal{R}}(\rho_i), r \rangle_{\mathcal{R}}. \quad (\text{E.33})$$

Then, consider an arbitrary element of $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ written in the form $\lambda P_{\mathcal{R}}(E)$ for some $\lambda \in \mathbb{R}_{\geq 0}$ and $E \in \bar{\mathfrak{e}}$. Thanks to proposition 4,

$$L(\lambda P_{\mathcal{R}}(E)) = \lambda \sum_{i=1}^{\dim(\mathcal{R})} \langle \rho_i, E \rangle_{\mathcal{L}(\mathcal{H})} \geq 0. \quad (\text{E.34})$$

Equality implies for all $i = 1, \dots, \dim(\mathcal{R})$ that

$$0 = \lambda \langle \rho_i, E \rangle_{\mathcal{L}(\mathcal{H})} = \langle P_{\mathcal{R}}(\rho_i), \lambda P_{\mathcal{R}}(E) \rangle_{\mathcal{R}}, \quad (\text{E.35})$$

but since the set $\{P_{\mathcal{R}}(\rho_i)\}_i$ spans \mathcal{R} , this implies $\langle r, \lambda P_{\mathcal{R}}(E) \rangle_{\mathcal{R}} = 0$ for all $r \in \mathcal{R}$, which by the non-degeneracy of the inner product implies

$$\lambda P_{\mathcal{R}}(E) = 0. \quad \square \quad (\text{E.36})$$

Proposition 28. $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ are spanning pointed cones in \mathcal{R} .

Proof. Let us verify the definition 21 of pointed cones. The closure of $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ is clear, so (i) of definition 21 is verified. By lemma B.9, $P_{\mathcal{R}}(\mathfrak{s})$ and $P_{\mathcal{R}}(\mathfrak{e})$ are strict supersets of $\{0 \in \mathcal{R}\}$, which proves that also $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ are strict supersets of $\{0\}$, satisfying (ii) of definition 23. The property (iii) was proven separately in lemmas E.5 and E.6. Thus, $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ are pointed cones in \mathcal{R} .

Now we verify the definition 23 of spanning cones. Consider the property (ii) (" $\mathcal{C} \neq \mathcal{V}$ ") of definition 23: it is automatically verified thanks to the fact that (ii) (" $\mathcal{C} \neq \emptyset, \mathcal{C} \neq \{0\}$ ") of definition 21 holds. Property (iii) follows directly from the subset inclusion $P_{\mathcal{R}}(\mathfrak{s}) \subset \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $P_{\mathcal{R}}(\mathfrak{e}) \subset \text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ together with corollary B.6. This proves that $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{s}}))$ and $\text{coni}(P_{\mathcal{R}}(\bar{\mathfrak{e}}))$ are also spanning cones. \square

Proposition 29. $\text{Sep}(\mathfrak{s}, \mathfrak{e})$ is a spanning pointed cone in $\mathcal{R} \otimes \mathcal{R}$.

Proof. The closure of $\text{Sep}(\mathfrak{s}, \mathfrak{e})$ was proven in proposition C.9. By lemma B.9, $\text{Sep}(\mathfrak{s}, \mathfrak{e}) \neq \{0\} \subset \mathcal{R} \otimes \mathcal{R}$.

Then, define the linear map $L : \mathcal{R} \otimes \mathcal{R} \rightarrow \mathbb{R}$ by: for all $\Omega \in \mathcal{R} \otimes \mathcal{R}$,

$$L(\Omega) := \sum_{i,j=1}^{\dim(\mathcal{R})} \langle \bar{\rho}_i \otimes \bar{E}_j, \Omega \rangle_{\mathcal{R} \otimes \mathcal{R}}, \quad (\text{E.37})$$

where $\{\bar{\rho}_i \in P_{\mathcal{R}}(\mathfrak{s})\}_{i=1}^{\dim(\mathcal{R})}$ and $\{\bar{E}_j \in P_{\mathcal{R}}(\mathfrak{e})\}_{j=1}^{\dim(\mathcal{R})}$ are two bases of \mathcal{R} , which is possible thanks to corollary B.6. Now consider any $\Omega \in \text{Sep}(\mathfrak{s}, \mathfrak{e})$: then, by definition 14, there exists $n \in \mathbb{N}$ and $\{F_k \in P_{\mathcal{R}}(\mathfrak{s})^{+\mathcal{R}}\}_{k=1}^n, \{\sigma_k \in P_{\mathcal{R}}(\mathfrak{e})^{+\mathcal{R}}\}_{k=1}^n$ such that $\Omega = \sum_{k=1}^n F_k \otimes \sigma_k$. Then,

$$L(\Omega) = \sum_{i,j,k} \langle \bar{\rho}_i, F_k \rangle_{\mathcal{R}} \langle \bar{E}_j, \sigma_k \rangle_{\mathcal{R}}. \quad (\text{E.38})$$

Thanks to the domains of the respective elements, each scalar product is a nonnegative number. Thus $L(\Omega) \geq 0$, and equality implies for all $k = 1, \dots, n$ that

$$\forall i, j = 1, \dots, \dim(\mathcal{R}) : \langle \bar{\rho}_i \otimes \bar{E}_j, F_k \otimes \sigma_k \rangle_{\mathcal{R} \otimes \mathcal{R}} = 0. \quad (\text{E.39})$$

But $\{\bar{\rho}_i \otimes \bar{E}_j\}_{i,j=1}^{\dim(\mathcal{R})}$ is a basis of $\mathcal{R} \otimes \mathcal{R}$, so that in fact, for all $k = 1, \dots, n$:

$$\forall R \in \mathcal{R} \otimes \mathcal{R} : \langle R, F_k \otimes \sigma_k \rangle_{\mathcal{R} \otimes \mathcal{R}} = 0. \quad (\text{E.40})$$

By the non-degeneracy of the inner product, each term $F_k \otimes \sigma_k$ is zero so that $\Omega = 0$. This proves that $\text{Sep}(\mathfrak{s}, \mathfrak{e})$ is a pointed cone.

For the spanning cone aspect, (ii) of definition 23 holds thanks to (iii) of definition 21. The spanning property (iii) of definition 23 follows from the fact that the following basis of $\mathcal{R} \otimes \mathcal{R}$:

$$\{\bar{E}_i \otimes \bar{\rho}_j : \bar{E}_i \in P_{\mathcal{R}}(\mathfrak{e}), \bar{\rho}_j \in P_{\mathcal{R}}(\mathfrak{s})\}_{i,j=1}^{\dim(\mathcal{R})} \quad (\text{E.41})$$

is a subset of $\text{Sep}(\mathfrak{s}, \mathfrak{e})$. This concludes the proof. \square

In the course of the proof of proposition 30, we will need the notion of partial scalar product over a tensor product space, which took its inspiration from the partial trace familiar from quantum mechanics, but also from the insightful proof techniques of [25].

Definition E.7 (Partial scalar product). For any $a \otimes b \in \mathcal{R} \otimes \mathcal{R}$, for any $r \in \mathcal{R}$, define

$$\mathfrak{S}_1(a \otimes b, r) := \langle a, r \rangle_{\mathcal{R}} b, \quad (\text{E.42a})$$

$$\mathfrak{S}_2(a \otimes b, r) := \langle b, r \rangle_{\mathcal{R}} a, \quad (\text{E.42b})$$

and extend these definitions by linearity to obtain bilinear maps:

$$\mathfrak{S}_1(\cdot, \cdot) : (\mathcal{R} \otimes \mathcal{R}) \times \mathcal{R} \rightarrow \mathcal{R}, \quad (\text{E.43a})$$

$$\mathfrak{S}_2(\cdot, \cdot) : (\mathcal{R} \otimes \mathcal{R}) \times \mathcal{R} \rightarrow \mathcal{R}. \quad (\text{E.43b})$$

Lemma E.8. For any $\Omega \in \text{Sep}(\mathbf{s}, \mathbf{e})$, for any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, it holds that

$$\mathbf{S}_1(\Omega, \bar{\rho}) \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}. \quad (\text{E.44})$$

For any $\Omega \in \text{Sep}(\mathbf{s}, \mathbf{e})$, for any $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$, it holds that

$$\mathbf{S}_2(\Omega, \bar{E}) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}. \quad (\text{E.45})$$

Proof. Consider any $\Omega \in \text{Sep}(\mathbf{s}, \mathbf{e})$. By definition 14, there exist $n \in \mathbb{N}$ and $\{F_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}\}_{i=1}^n$, $\{\sigma_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}\}_{i=1}^n$ such that $\Omega = \sum_{i=1}^n F_i \otimes \sigma_i$. Then, by definition E.7, for any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$,

$$\mathbf{S}_1(\Omega, \bar{\rho}) = \sum_{i=1}^n \langle F_i, \bar{\rho} \rangle_{\mathcal{R}} \sigma_i. \quad (\text{E.46})$$

Since $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ is a convex cone, $\sigma_i \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$ and $\langle F_i, \bar{\rho} \rangle_{\mathcal{R}} \geq 0$, it holds that $\mathbf{S}_1(\Omega, \bar{\rho}) \in P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$. The proof that for any $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$: $\mathbf{S}_2(\Omega, \bar{E}) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ is analogous. \square

Proposition 30. It holds that

$$\text{extr}(\text{Sep}(\mathbf{s}, \mathbf{e})) = \left\{ \mathfrak{l}_1 \otimes_{\text{set}} \mathfrak{l}_2 : \mathfrak{l}_1 \in \text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}), \mathfrak{l}_2 \in \text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}) \right\}. \quad (\text{4.7})$$

Proof. It is easy to see from definition 14 of $\text{Sep}(\mathbf{s}, \mathbf{e})$ that

$$\text{Sep}(\mathbf{s}, \mathbf{e}) = \text{conv} \left(\bigcup_{\mathfrak{l}_1 \in \text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}), \mathfrak{l}_2 \in \text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}})} \mathfrak{l}_1 \otimes_{\text{set}} \mathfrak{l}_2 \right). \quad (\text{E.47})$$

This shows that the set of extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$ have to be a subset of or equal to the set

$$\left\{ \mathfrak{l}_1 \otimes_{\text{set}} \mathfrak{l}_2 : \mathfrak{l}_1 \in \text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}), \mathfrak{l}_2 \in \text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}) \right\}. \quad (\text{E.48})$$

It remains to show that the set (E.48) contains no more than the extremal half-lines of $\text{Sep}(\mathbf{s}, \mathbf{e})$, i.e., that each half-line in (E.48) is indeed an extremal half-line of $\text{Sep}(\mathbf{s}, \mathbf{e})$.

For any $\mathfrak{l}_1 \in \text{extr}(P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}})$, choose $F \in \mathfrak{l}_1$, $F \neq 0$. Then, for any $\mathfrak{l}_2 \in \text{extr}(P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}})$, choose $\sigma \in \mathfrak{l}_2$, $\sigma \neq 0$. Now consider any $\Omega_1, \Omega_2 \in \text{Sep}(\mathbf{s}, \mathbf{e})$ that satisfy

$$F \otimes \sigma = \Omega_1 + \Omega_2. \quad (\text{E.49})$$

Choose now any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, and take the partial scalar product defined in definition E.7:

$$\langle F, \bar{\rho} \rangle_{\mathcal{R}} \sigma = \mathbf{S}_1(\Omega_1, \bar{\rho}) + \mathbf{S}_1(\Omega_2, \bar{\rho}). \quad (\text{E.50})$$

Note that $\langle F, \bar{\rho} \rangle_{\mathcal{R}} \geq 0$ and thanks to lemma E.8, $\mathbf{S}_1(\Omega_1, \bar{\rho})$ and $\mathbf{S}_1(\Omega_2, \bar{\rho})$ belong to the cone $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$. There are two cases to consider. If $\langle F, \bar{\rho} \rangle_{\mathcal{R}} > 0$, then $\langle F, \bar{\rho} \rangle_{\mathcal{R}} \sigma$ is part of the extremal half-line \mathfrak{l}_2 of $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$. Then, the definition A.6 of extremal half-lines and equation (E.50) imply that $\mathbf{S}_1(\Omega_1, \bar{\rho})$ and $\mathbf{S}_1(\Omega_2, \bar{\rho})$ are both scalar multiples of σ :

$$\mathbf{S}_1(\Omega_1, \bar{\rho}) = f_1(\bar{\rho})\sigma, \quad (\text{E.51a})$$

$$\mathbf{S}_1(\Omega_2, \bar{\rho}) = f_2(\bar{\rho})\sigma, \quad (\text{E.51b})$$

where we allowed the scalar multiples $f_i(\bar{\rho}) \in \mathbb{R}$ to depend on $\bar{\rho}$.

This was when $\langle F, \bar{\rho} \rangle_{\mathcal{R}} > 0$. If instead, $\langle F, \bar{\rho} \rangle_{\mathcal{R}} = 0$, then (E.50) becomes $\mathbf{S}_1(\Omega_1, \bar{\rho}) + \mathbf{S}_1(\Omega_2, \bar{\rho}) = 0$. Both elements $\mathbf{S}_1(\Omega_i, \bar{\rho})$ belong to $P_{\mathcal{R}}(\mathbf{e})^{+\mathcal{R}}$, so that by taking the inner product with any $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$, we must have $\langle \mathbf{S}_1(\Omega_i, \bar{\rho}), \bar{E} \rangle_{\mathcal{R}} = 0$ for $i = 1, 2$. Since the set $P_{\mathcal{R}}(\mathbf{e})$ spans \mathcal{R} according to corollary B.6, it must be that $\mathbf{S}_1(\Omega_i, \bar{\rho}) = 0 \in \mathcal{R}$. In this case, extend the maps f_i defined in (E.51) to be 0 for such $\bar{\rho}$.

This proves that for any $\bar{\rho} \in P_{\mathcal{R}}(\mathbf{s})$, there exist maps $f_i : P_{\mathcal{R}}(\mathbf{s}) \rightarrow \mathbb{R}$ such that (E.51) still holds. The mappings f_i inherit properties from the left-hand sides of (E.51): in particular, the f_i must be convex linear over the domain $P_{\mathcal{R}}(\mathbf{s})$. By repeating the argument of proposition B.12, it holds that there exist unique linear extensions $f_i : \mathcal{R} \rightarrow \mathbb{R}$. By Riesz' representation theorem B.1, there exist $g_i, r \in \mathcal{R}$ such that $f_i(r) = \langle g_i, r \rangle_{\mathcal{R}}$ for all $r \in \mathcal{R}$. Thus, for $i = 1, 2$, we must have: for all $r, r' \in \mathcal{R}$,

$$\langle \mathbf{S}_1(\Omega_i, r), r' \rangle_{\mathcal{R}} = \langle \langle g_i, r \rangle_{\mathcal{R}} \sigma, r' \rangle_{\mathcal{R}}, \quad (\text{E.52})$$

which may be rewritten as

$$\langle \Omega_i, r \otimes r' \rangle_{\mathcal{R} \otimes \mathcal{R}} = \langle g_i \otimes \sigma, r \otimes r' \rangle_{\mathcal{R} \otimes \mathcal{R}}. \quad (\text{E.53})$$

The non-degeneracy of the inner product implies $\Omega_i = g_i \otimes \sigma$. Now, choose any $\bar{E} \in P_{\mathcal{R}}(\mathbf{e})$ such that $\langle \sigma, \bar{E} \rangle_{\mathcal{R}} \geq 0$. This is always possible, otherwise it implies that $\sigma = 0$. By lemma E.8, it must be that $\mathbf{S}_2(\Omega_i, \bar{E}) \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$, which implies that $g_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$.

We return to equation (E.49) which now reads:

$$F \otimes \sigma = g_1 \otimes \sigma + g_2 \otimes \sigma. \quad (\text{E.54})$$

Since we assumed $\sigma \neq 0$, this implies

$$F = g_1 + g_2. \quad (\text{E.55})$$

This is a decomposition of the extremal direction $F \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$ over two other directions $g_i \in P_{\mathcal{R}}(\mathbf{s})^{+\mathcal{R}}$: it must be that the g_i 's are linearly dependent, which implies the linear dependence of $\Omega_1 = g_1 \otimes \sigma$ and $\Omega_2 = g_2 \otimes \sigma$.

This proves that the direction $F \otimes \sigma$ is extremal, which proves that any half-lines in the set of equation (E.48) is an extremal half-line of $\text{Sep}(\mathbf{s}, \mathbf{e})$. This concludes the proof. \square

E.3 Computational equivalence: varying reduced spaces and quantum descriptions

Proposition 34. *Given any two finite dimensional real inner product spaces \mathcal{U}, \mathcal{V} such that $\dim(\mathcal{U}) = \dim(\mathcal{V})$, any two convex cones $\mathcal{C} \subseteq \mathcal{U}$ and $\mathcal{D} \subseteq \mathcal{V}$ such that $\mathcal{C} \sim \mathcal{D}$ have the following properties:*

- (i) *there is a one-to-one correspondence between the extremal half-lines of \mathcal{C} and those of \mathcal{D} ;*
- (ii) *the same holds for the extremal half-lines of the polar cones due to $\mathcal{C}^{+u} \sim \mathcal{D}^{+v}$.*

Proof. Since $\mathcal{C} \sim \mathcal{D}$, let $\Phi : \mathcal{U} \rightarrow \mathcal{V}$ be the invertible linear map of definition 32 that satisfies $\Phi(\mathcal{C}) = \mathcal{D}$.

(i) It suffices to prove that given an extremal direction $c \in \mathcal{C}$, $\Phi(c)$ is an extremal direction of \mathcal{D} . Take $d_1, d_2 \in \mathcal{D}$ such that

$$\Phi(c) = d_1 + d_2. \quad (\text{E.56})$$

Then, apply the inverse linear map Φ^{-1} :

$$c = \Phi^{-1}(d_1) + \Phi^{-1}(d_2). \quad (\text{E.57})$$

The fact that $c \in \mathcal{C}$ is an extremal direction of \mathcal{C} and the fact that $\Phi^{-1}(d_1), \Phi^{-1}(d_2) \in \mathcal{C}$ imply that there exist $\lambda_1, \lambda_2 \in \mathbb{R}_{\geq 0}$ such that $\Phi^{-1}(d_i) = \lambda_i c$, $i = 1, 2$; but by applying the linear map Φ this proves that $d_i = \lambda_i \Phi(c)$, and thus $\Phi(c)$ is an extremal direction of \mathcal{D} . Thus, Φ is a linear isomorphism that makes a one-to-one correspondence between extremal lines of \mathcal{C} and of \mathcal{D} .

(ii) Let us prove that $\mathcal{C}^{+u} \sim \mathcal{D}^{+v}$. It suffices to exhibit an invertible isomorphism $\Psi : \mathcal{U} \rightarrow \mathcal{V}$ such that

$$\Psi(\mathcal{C}^{+u}) = \mathcal{D}^{+v}. \quad (\text{E.58})$$

Let Ψ be the dual map to the inverse of Φ : $\Psi := (\Phi^{-1})^*$. This map is invertible and linear: its inverse is simply Φ^* . Let us verify (E.58): let $u \in \mathcal{C}^{+u}$, and we will first prove that $(\Phi^{-1})^*(u) \in \mathcal{D}^{+v}$. Indeed, for all $d \in \mathcal{D}$, the dual property reads

$$\langle (\Phi^{-1})^*(u), d \rangle_{\mathcal{V}} = \langle u, \Phi^{-1}(d) \rangle_{\mathcal{U}}. \quad (\text{E.59})$$

The properties of Φ inherited from definition 32 prove that $\Phi^{-1}(d) \in \mathcal{C}$. Due to $u \in \mathcal{C}^{+u}$ and $\Phi^{-1}(d) \in \mathcal{C}$, the right-hand side of (E.59) is nonnegative. $d \in \mathcal{D}$ was arbitrary so that indeed, for all $u \in \mathcal{C}^{+u}$,

$$\Psi(u) = (\Phi^{-1})^*(u) \in \mathcal{D}^{+v}, \quad (\text{E.60})$$

thus proving $\Psi(\mathcal{C}^{+u}) \subseteq \mathcal{D}^{+v}$. The other direction can be proven as follows. First, it is easy to verify that for all $t \in \mathcal{D}^{+v}$, $\Phi^*(t) \in \mathcal{C}^{+u}$ due to $\Phi(\mathcal{C}) = \mathcal{D}$. Using the invertibility of Φ , we obtain that $t = (\Phi^{-1})^* \circ \Phi^*(t) \in (\Phi^{-1})^*(\mathcal{C}^{+u}) \equiv \Psi(\mathcal{C}^{+u})$, which proves that also $\mathcal{D}^{+v} \subseteq \Psi(\mathcal{C}^{+u})$. \square

Lemma E.9. *For any \mathcal{R}_{alt} with associated mappings f, g (definition 18), there exist invertible linear maps $\Phi_{\mathbf{s}}, \Phi_{\mathbf{e}} : \mathcal{R} \rightarrow \mathcal{R}_{\text{alt}}$ such that for all $s \in \text{span}(\mathbf{s})$, for all $e \in \text{span}(\mathbf{e})$,*

$$\Phi_{\mathbf{s}}(P_{\mathcal{R}}(s)) = f(s), \quad (\text{E.61a})$$

$$\Phi_{\mathbf{e}}(P_{\mathcal{R}}(e)) = g(e). \quad (\text{E.61b})$$

Proof. A suitable choice for $\Phi_{\mathbf{e}}$ was given in the proof of proposition 20. We will slightly tune this construction to obtain a valid choice for $\Phi_{\mathbf{s}}$.

Let $d := \dim(\mathcal{R}) = \dim(\mathcal{R}_{\text{alt}})$ (by proposition 20). Let $\{T_i \in \mathcal{R}_{\text{alt}}\}_{i=1}^d$ be an orthonormal basis of \mathcal{R}_{alt} . By equation (3.30c), there exist elements $e_i \in \text{span}(\mathbf{e})$ such that $T_i = g(e_i)$ for all $i = 1, \dots, d$.

Now let $\{R_j \in \mathcal{R}\}_{j=1}^d$ be an orthonormal basis of \mathcal{R} . By corollary B.6, there exist elements $f_j \in \text{span}(\mathbf{s})$ such that $R_j = P_{\mathcal{R}}(f_j)$ for all $j = 1, \dots, d$.

Then, define the linear maps $\Phi_{\mathbf{s}} : \mathcal{R} \rightarrow \mathcal{R}_{\text{alt}}$ and $\phi_{\mathbf{s}} : \mathcal{R}_{\text{alt}} \rightarrow \mathcal{R}$ by: for all $r \in \mathcal{R}$, for all $t \in \mathcal{R}_{\text{alt}}$,

$$\Phi_{\mathbf{s}}(r) = \sum_{i=1}^d \langle P_{\mathcal{R}}(e_i), r \rangle_{\mathcal{R}} T_i, \quad (\text{E.62a})$$

$$\phi_{\mathbf{s}}(t) = \sum_{j=1}^d \langle g(f_j), t \rangle_{\mathcal{R}_{\text{alt}}} R_j. \quad (\text{E.62b})$$

It holds that $\Phi_{\mathbf{s}}(P_{\mathcal{R}}(s)) = f(s)$ for all $s \in \text{span}(\mathbf{s})$. Indeed, using lemma D.1 but swapping the arguments of the symmetric scalar product: for all $s \in \text{span}(\mathbf{s})$,

$$\begin{aligned} \Phi_{\mathbf{s}}(P_{\mathcal{R}}(s)) &= \sum_{i=1}^d \langle P_{\mathcal{R}}(e_i), P_{\mathcal{R}}(s) \rangle_{\mathcal{R}} T_i \\ &= \sum_{i=1}^d \langle g(e_i), f(s) \rangle_{\mathcal{R}_{\text{alt}}} T_i \\ &= \sum_{i=1}^d \langle T_i, f(s) \rangle_{\mathcal{R}_{\text{alt}}} T_i = f(s). \end{aligned} \quad (\text{E.63})$$

In the last line, we used the resolution of the identity for \mathcal{R}_{alt} in the basis $\{T_i\}_i$. Similarly, for all $s \in \text{span}(\mathbf{s})$ it holds that $\phi_{\mathbf{s}}(f(s)) = P_{\mathcal{R}}(s)$. Indeed, using again lemma D.1: for all $s \in \text{span}(\mathbf{s})$,

$$\begin{aligned} \phi_{\mathbf{s}}(f(s)) &= \sum_{j=1}^d \langle g(f_j), f(s) \rangle_{\mathcal{R}_{\text{alt}}} R_j \\ &= \sum_{j=1}^d \langle P_{\mathcal{R}}(f_j), P_{\mathcal{R}}(s) \rangle_{\mathcal{R}} R_j \\ &= \sum_{j=1}^d \langle R_j, P_{\mathcal{R}}(s) \rangle_{\mathcal{R}} R_j = P_{\mathcal{R}}(s). \end{aligned} \quad (\text{E.64})$$

We used the completeness relation of \mathcal{R} in the basis $\{R_i\}_i$. Let us now verify that $\phi_{\mathbf{s}} = \Phi_{\mathbf{s}}^{-1}$. For all $r \in \mathcal{R}$, let $s \in \text{span}(\mathbf{s})$ be such that $r = P_{\mathcal{R}}(s)$.

$$\phi_{\mathbf{s}} \circ \Phi_{\mathbf{s}}(r) = \phi_{\mathbf{s}}(\Phi_{\mathbf{s}}(P_{\mathcal{R}}(s))) = \phi_{\mathbf{s}}(f(s)) = P_{\mathcal{R}}(s) = r, \quad (\text{E.65})$$

where we used the properties (E.63) and (E.64) to conclude. This shows $\phi_{\mathbf{s}} \circ \Phi_{\mathbf{s}} = \mathbb{1}_{\mathcal{R}}$. The proof that $\Phi_{\mathbf{s}} \circ \phi_{\mathbf{s}} = \mathbb{1}_{\mathcal{R}_{\text{alt}}}$ is analogous. Thus $\phi_{\mathbf{s}} = \Phi_{\mathbf{s}}^{-1}$. \square

Lemma E.10. *Let \mathcal{U}, \mathcal{V} be two finite dimensional inner product spaces such that $\dim(\mathcal{U}) = \dim(\mathcal{V})$, and let $\mathcal{C}_1, \mathcal{C}_2 \subset \mathcal{U}$, $\mathcal{D}_1, \mathcal{D}_2 \subset \mathcal{V}$ be convex cones. If it holds that*

$$\mathcal{C}_1 \sim \mathcal{D}_1, \quad (\text{E.66})$$

$$\mathcal{C}_2 \sim \mathcal{D}_2, \quad (\text{E.67})$$

then it also holds that

$$\text{conv}(\mathcal{C}_1^{+u} \otimes_{\text{set}} \mathcal{C}_2^{+u}) \sim \text{conv}(\mathcal{D}_1^{+v} \otimes_{\text{set}} \mathcal{D}_2^{+v}). \quad (\text{E.68})$$

Proof. Proposition 34 proved that if two cones are isomorphic in the sense of definition 32, their polar cones are also isomorphic. For $i = 1, 2$, let $\Psi_i : \mathcal{U} \rightarrow \mathcal{V}$ be the invertible linear map such that

$$\Psi_i(\mathcal{C}_i^{+u}) = \mathcal{D}_i^{+v}. \quad (\text{E.69})$$

The linear map that establish (4.17c) is then simply the tensor product map $\Psi_1 \otimes \Psi_2 : \mathcal{U} \otimes \mathcal{U} \rightarrow \mathcal{V} \otimes \mathcal{V}$ that acts as follows: for any $u, u' \in \mathcal{U}$,

$$\Psi_1 \otimes \Psi_2(u \otimes u') = \Psi_1(u) \otimes \Psi_2(u'), \quad (\text{E.70})$$

and extend this definition by linearity. This linear map is invertible, and allows one to easily verify that

$$\begin{aligned} \Psi_1 \otimes \Psi_2(\text{conv}(\mathcal{C}_1^{+u} \otimes_{\text{set}} \mathcal{C}_2^{+u})) \\ = \text{conv}(\mathcal{D}_1^{+v} \otimes_{\text{set}} \mathcal{D}_2^{+v}). \end{aligned} \quad (\text{E.71})$$

which proves (4.17c). \square

Proposition 33. *Choosing any alternative reduced space \mathcal{R}_{alt} with associated mappings f, g (definition 18), it holds that:*

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{s}})) \sim \text{coni}(f(\bar{\mathbf{s}})), \quad (\text{4.17a})$$

$$\text{coni}(P_{\mathcal{R}}(\bar{\mathbf{e}})) \sim \text{coni}(g(\bar{\mathbf{e}})), \quad (\text{4.17b})$$

$$\text{Sep}(\mathbf{s}, \mathbf{e}) \sim \text{Sep}(\mathbf{s}, \mathbf{e})_{\text{alt}}, \quad (\text{4.17c})$$

where

$$\text{Sep}(\mathbf{s}, \mathbf{e})_{\text{alt}} := \text{conv}(f(\mathbf{s})^{+\mathcal{R}_{\text{alt}}} \otimes_{\text{set}} g(\mathbf{e})^{+\mathcal{R}_{\text{alt}}}). \quad (\text{4.18})$$

Proof. Due to $\text{coni}(\bar{\mathbf{s}}) \subseteq \text{span}(\mathbf{s})$, the invertible linear map $\Phi_{\mathbf{s}}$ from lemma E.9 satisfies, thanks to equation (E.61a):

$$\Phi_{\mathbf{s}}(P_{\mathcal{R}}(\text{coni}(\bar{\mathbf{s}}))) = f(\text{coni}(\bar{\mathbf{s}})). \quad (\text{E.72})$$

The fact that linear operations and conical hulls commute implies (4.17a), and the proof of (4.17b) is analogous. Lemma E.10 together with proposition 27 then allows to prove equation (4.17c). \square

Proposition 35. *For the scenarios (\mathbf{s}, \mathbf{e}) and $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$ and the corresponding reduced spaces \mathcal{R} and $\tilde{\mathcal{R}}$ defined in proposition 8, it holds that*

$$\text{coni}(P_{\mathcal{R}}(\mathbf{s})) \sim \text{coni}(P_{\tilde{\mathcal{R}}}(\tilde{\mathbf{s}})), \quad (\text{4.19})$$

$$\text{coni}(P_{\mathcal{R}}(\mathbf{e})) \sim \text{coni}(P_{\tilde{\mathcal{R}}}(\tilde{\mathbf{e}})), \quad (\text{4.20})$$

$$\text{Sep}(\mathbf{s}, \mathbf{e}) \sim \text{Sep}(\tilde{\mathbf{s}}, \tilde{\mathbf{e}}), \quad (\text{4.21})$$

which proves by proposition 34 the computational equivalence of starting with either (\mathbf{s}, \mathbf{e}) or $(\tilde{\mathbf{s}}, \tilde{\mathbf{e}})$.

Proof. Let $d = \dim(\mathcal{R}) = \dim(\tilde{\mathcal{R}})$ (proposition 8). We first construct a linear map $\Phi_{\mathbf{s}} : \mathcal{R} \rightarrow \tilde{\mathcal{R}}$ such that for all $k \in I$,

$$\Phi_{\mathbf{s}}(P_{\mathcal{R}}(\rho_k)) = P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k). \quad (\text{E.73})$$

To do so, let $\{\tilde{R}_i \in \tilde{\mathcal{R}}\}_{i=1}^d$ be an orthonormal basis of $\tilde{\mathcal{R}}$. Let $K \subseteq I$ be such that $\{P_{\tilde{\mathcal{R}}}(\tilde{E}_j)\}_{j \in K}$ is a basis of $\tilde{\mathcal{R}}$ (corollary B.6). For each i , there exist $\{\alpha_{ij} \in \mathbb{R}\}_{1 \leq i \leq d; j \in K}$ such that

$$\tilde{R}_i = \sum_{1 \leq i \leq d; j \in K} \alpha_{ij} P_{\tilde{\mathcal{R}}}(\tilde{E}_j). \quad (\text{E.74})$$

Then define for all $r \in \mathcal{R}$

$$\Phi_{\mathbf{s}}(r) = \sum_{1 \leq i \leq d; j \in K} \alpha_{ij} \langle P_{\mathcal{R}}(E_j), r \rangle_{\mathcal{R}} \tilde{R}_i. \quad (\text{E.75})$$

This map is linear and satisfies equation (E.73) (here one uses equations (B.93), (E.74) and the completeness relation of the orthonormal basis $\{\tilde{R}_i\}_i$). It is easy to mimic this construction to obtain a linear map that maps $P_{\tilde{\mathcal{R}}}(\tilde{\rho}_k)$ to $P_{\mathcal{R}}(\rho_k)$ for all $k \in I$: this map is exactly the inverse of $\Phi_{\mathbf{s}}$ since the $\{P_{\mathcal{R}}(\rho_k)\}_{k \in I}$ span \mathcal{R} (corollary B.6), which proves the invertibility of $\Phi_{\mathbf{s}}$ (alternatively, it is easy to prove its injectivity and surjectivity). Having this map at hand, it follows easily (using that linear maps commute with the conical and convex hull) that

$$\Phi_{\mathbf{s}}(\text{coni}(P_{\mathcal{R}}(\mathbf{s}))) = (\text{coni}(P_{\tilde{\mathcal{R}}}(\tilde{\mathbf{s}}))), \quad (\text{E.76})$$

which proves (4.19).

The proof of equation (4.20) is completely analogous: one starts by building a linear map $\Phi_{\mathbf{e}} : \mathcal{R} \rightarrow \tilde{\mathcal{R}}$ such that for all $k \in I$,

$$\Phi_{\mathbf{e}}(P_{\mathcal{R}}(E_j)) = P_{\tilde{\mathcal{R}}}(\tilde{E}_j). \quad (\text{E.77})$$

Then, one uses a decomposition of the form

$$\tilde{R}_i = \sum_{1 \leq i \leq d; j \in J} \beta_{ij} P_{\tilde{\mathcal{R}}}(\tilde{\rho}_j) \quad (\text{E.78})$$

instead of (E.74), with J such that $\{P_{\tilde{\mathcal{R}}}(\tilde{\rho}_j)\}_{j \in J}$ is a basis of $\tilde{\mathcal{R}}$ and $\beta_{ij} \in \mathbb{R}$. The definition of $\Phi_{\mathbf{e}}$ is then for all $r \in \mathcal{R}$,

$$\Phi_{\mathbf{e}}(r) = \sum_{1 \leq i \leq d; j \in J} \beta_{ij} \langle P_{\mathcal{R}}(\rho_j), r \rangle_{\mathcal{R}} \tilde{R}_i. \quad (\text{E.79})$$

This map is invertible, satisfies (E.77), and allows to prove (4.20).

Equation (4.21) then follows directly from lemma E.10 and from the simplification of proposition 27. \square

F Connections with generalized probabilistic theories

Let us restate the definition of a simplex-embeddable generalized probabilistic theory. We build on top of the notation of section 5, and additional notation comes from [9].

Definition F.1 (Adapted from [9]). *A tomographically complete generalized probabilistic theory $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex embeddable in d dimensions if and only if there exist:*

- (i) a d -dimensional real inner product space \mathcal{W} ;
- (ii) a simplex $\Delta_d \subset \mathcal{W}$ with d linearly independent vertices denoted $\{\delta_i \in \mathcal{W}\}_{i=1}^d$;
- (iii) a linear map $\iota : \mathcal{V} \rightarrow \mathcal{W}$ such that $\iota(\Omega) \subseteq \Delta_d$;
- (iv) a linear map $\kappa : \mathcal{V} \rightarrow \mathcal{W}$ such that for all $E \in \mathcal{E}$, for all $i = 1, \dots, d$: $\langle \kappa(E), \delta_i \rangle_{\mathcal{W}} \in [0, 1]$;

where the maps ι, κ must satisfy the consistency requirement: for all $\rho \in \Omega$, $E \in \mathcal{E}$:

$$\langle \rho, E \rangle_{\mathcal{V}} = \langle \iota(\rho), \kappa(E) \rangle_{\mathcal{W}}. \quad (\text{F.1})$$

The fact that the vertices of the simplex Δ_d are linearly independent in (ii) of definition F.1 is equivalent to the fact that their affine span does not contain the origin $0 \in \mathcal{W}$, which was the condition stated in definition 1 of [9]. Let us now prove proposition 36.

Proposition 36. *Any tomographically complete generalized probabilistic theory $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable in d dimensions in the sense of definition 1 of [9], if and only if the tomographically complete prepare-and-measure scenario (Ω, \mathcal{E}) admits an operationally noncontextual ontological model in the sense of definition 6 (under the substitution (5.1)) with a discrete ontic space of finite cardinality d .*

Proof. First, note that the existence of an operationally noncontextual ontological model as in definition 6 under the substitution (5.1) is equivalent to the criterion given in theorem 1 under the substitution (5.1). We will prove the claim using the latter rather than the former.

Suppose that the generalized probabilistic theory $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable in d dimensions as in definition F.1: we will first prove that there exists an operationally noncontextual ontological model as in theorem 1 under the substitution (5.1) with a discrete, finite ontic space of cardinality d . The set $\{\delta_i \in \mathcal{W}\}_{i=1}^d$ forms a basis of \mathcal{W} , so that there must exist coordinate functions $\lambda_i : \mathcal{V} \rightarrow \mathbb{R}$ for all $i = 1, \dots, d$ such that

$$\forall v \in \mathcal{V} : \iota(v) = \sum_{i=1}^d \lambda_i(v) \delta_i. \quad (\text{F.2})$$

By Riesz' representation theorem B.1, there must exist $\{F_i \in \mathcal{V}\}_{i=1}^d$ such that: for all $v \in \mathcal{V}$, for all $i = 1, \dots, d$,

$$\lambda_i(v) = \langle v, F_i \rangle_{\mathcal{V}} \quad (\text{F.3})$$

Consider now for any $i = 1, \dots, d$ the linear maps

$$\langle \kappa(\cdot), \delta_i \rangle_{\mathcal{W}} : \mathcal{V} \rightarrow \mathbb{R}. \quad (\text{F.4})$$

Applying Riesz' representation theorem B.1 again, there must exist $\{\sigma_i \in \mathcal{V}\}_{i=1}^d$ such that: for all $v \in \mathcal{V}$, for all $i = 1, \dots, d$,

$$\langle \kappa(v), \delta_i \rangle_{\mathcal{W}} = \langle \sigma_i, v \rangle_{\mathcal{V}}. \quad (\text{F.5})$$

The consistency requirement (F.1) of definition F.1 reads: for all $\rho \in \Omega$, for all $E \in \mathcal{E}$,

$$\begin{aligned} \langle \rho, E \rangle_{\mathcal{V}} &= \langle \iota(\rho), \kappa(E) \rangle_{\mathcal{W}} = \sum_{i=1}^d \lambda_i(\rho) \langle \delta_i, \kappa(E) \rangle_{\mathcal{W}} \\ &= \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \langle \sigma_i, E \rangle_{\mathcal{V}}. \end{aligned} \quad (\text{F.6})$$

Consider now theorem 1 under the substitution (5.1). Recall that for a tomographically complete generalized probabilistic theory, the reduced space is simply given by the whole vector space \mathcal{V} as was illustrated in (5.3). We will show that the primitives $\Lambda = \{1, \dots, d\}$, $\{F_i \in \mathcal{V}\}_{i=1}^d$ and $\{\sigma_i \in \mathcal{V}\}_{i=1}^d$ match

the requirements of theorem 1. First off, the generalized version of the consistency requirement (2.24) of theorem 1 is equivalent to (F.6) together with the fact that $\text{span}(\Omega) = \text{span}(\mathcal{E}) = \mathcal{V}$ which is the tomographic completeness assumption.

Let us now verify the positivity relations of equations (2.22). Let $\rho \in \Omega$. By definition F.1, it holds that

$$\iota(\rho) \in \Delta_d. \quad (\text{F.7})$$

The main property of a simplex such as Δ_d is that it is the convex hull of its d extremal points which are linearly independent: thus, any point $\iota(\rho) \in \Delta_d$ may be written as a convex combination

$$\iota(\rho) = \sum_{i=1}^d \lambda_i(\rho) \delta_i, \quad (\text{F.8})$$

with

$$\forall i = 1, \dots, d: \lambda_i(\rho) \geq 0, \quad (\text{F.9a})$$

$$\sum_{i=1}^d \lambda_i(\rho) = 1. \quad (\text{F.9b})$$

Because the set $\{\delta_i\}_{i=1}^d$ forms a basis of \mathcal{W} , the $\lambda_i(\rho)$'s are unique and are thus the same as in (F.2). Rewriting (F.9) with the operators F_i defined in (F.3): for all $\rho \in \Omega$,

$$\forall i = 1, \dots, d: \langle F_i, \rho \rangle_{\mathcal{V}} \geq 0, \quad (\text{F.10a})$$

$$\sum_{i=1}^d \langle F_i, \rho \rangle_{\mathcal{V}} = 1. \quad (\text{F.10b})$$

Equation (F.10a) proves that

$$F_i \in \Omega^{+\nu}, \quad (\text{F.11})$$

which corresponds to the nonnegativity requirement (2.22a) under the substitution (5.1) (again recall $\mathcal{R} = \mathcal{V}$ in this case). We now recall the property (iv) of the definition F.1 of simplex-embeddability, and rewrite it using equation (F.5): for all $E \in \mathcal{E}$, for all $i = 1, \dots, d$,

$$\langle \sigma_i, E \rangle_{\mathcal{V}} = \langle \kappa(E), \delta_i \rangle_{\mathcal{V}} \in [0, 1]. \quad (\text{F.12})$$

The fact that $\langle \sigma_i, E \rangle_{\mathcal{V}} \geq 0$ proves (2.22b) under the substitution (5.1): indeed, it holds that for all $i = 1, \dots, d$,

$$\sigma_i \in \mathcal{E}^{+\nu}. \quad (\text{F.13})$$

Equations (2.22) and (2.24) are verified, so let us verify (2.23). First, recall the defining property of the unit element $u \in \mathcal{E}$: for all $\rho \in \Omega$, $\langle \rho, u \rangle_{\mathcal{V}} = 1$. Using (F.6), for any $\rho \in \Omega$,

$$1 = \langle \rho, u \rangle_{\mathcal{V}} = \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \langle \sigma_i, u \rangle_{\mathcal{V}}. \quad (\text{F.14})$$

Since for all $i = 1, \dots, d$, it holds that $\langle \rho, F_i \rangle_{\mathcal{V}} \geq 0$ according to (F.11) and $\langle \sigma_i, u \rangle_{\mathcal{V}} \in [0, 1]$ according to (F.12), if there existed j such that $\langle \sigma_j, u \rangle_{\mathcal{V}} < 1$, then,

$$1 = \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \langle \sigma_i, u \rangle_{\mathcal{V}} < \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} = 1. \quad (\text{F.15})$$

We used (F.10b) to conclude. This yields a contradiction so that we conclude that $\langle \sigma_i, u \rangle_{\mathcal{V}} = 1$ for all $i = 1, \dots, d$ which proves (2.23) under the substitution (5.1). Thus, we conclude that if the tomographically complete generalized probabilistic theory $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable in d dimensions, then there exists an operationally noncontextual ontological model for the tomographically complete prepare-and-measure scenario (Ω, \mathcal{E}) as in theorem 1 under the substitution (5.1) with the ontic space $\Lambda = \{1, \dots, d\}$.

Consider now the other direction: suppose that the tomographically complete prepare-and-measure scenario (Ω, \mathcal{E}) admits an operationally noncontextual ontological model with discrete ontic space of cardinality d . Again, the assumption of tomographically complete (Ω, \mathcal{E}) imply $\mathcal{R} = \mathcal{V}$ as in (5.3). Let the ontic primitives of theorem 1, under the substitution (5.1) be denoted $\{F_i \in \Omega^{+\nu}\}_{i=1}^d$ and $\{\sigma_i \in \mathcal{E}^{+\nu}\}_{i=1}^d$. Now, we consider the euclidean space \mathbb{R}^d , equipped with an orthonormal basis $\{\delta_i \in \mathbb{R}^d\}_{i=1}^d$. These define the simplex $\Delta_d \subset \mathbb{R}^d$:

$$\Delta_d := \text{conv}(\{\delta_i\}_{i=1}^d). \quad (\text{F.16})$$

Now, define the linear maps $\iota, \kappa: \mathcal{V} \rightarrow \mathbb{R}^d$ by: for all $v \in \mathcal{V}$,

$$\iota(v) = \sum_{i=1}^d \langle v, F_i \rangle_{\mathcal{V}} \delta_i, \quad (\text{F.17a})$$

$$\kappa(v) = \sum_{i=1}^d \langle \sigma_i, v \rangle_{\mathcal{V}} \delta_i. \quad (\text{F.17b})$$

We will now verify first the consistency requirement (F.1), then (iii) and (iv) of definition F.1: this will prove the validity of the simplex-embedding under consideration. Using the orthonormality of the basis $\{\delta_i\}_i$: for all $\rho \in \Omega$, for all $E \in \mathcal{E}$,

$$\langle \iota(\rho), \kappa(E) \rangle_{\mathbb{R}^d} = \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \langle \sigma_i, E \rangle_{\mathcal{V}} = \langle \rho, E \rangle_{\mathcal{V}}, \quad (\text{F.18})$$

where we used the consistency requirement (2.24) of the theorem 1 for the existence of the operationally noncontextual ontological model to conclude. This proves the consistency requirement (F.1) of definition F.1. Let us now prove (iii), i.e., that for all $\rho \in \Omega$: $\iota(\rho) \in \Delta_d$. Fix the argument $\rho \in \Omega$. By the definition (F.17a), $\iota(\rho) = \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \delta_i$. The property (2.22a) stating in this case that $F_i \in \Omega^{+\nu}$ proves that: for all $i = 1, \dots, d$,

$$\lambda_i := \langle \rho, F_i \rangle_{\mathcal{V}} \geq 0. \quad (\text{F.19})$$

Furthermore, using the normalization $\langle \sigma_i, u \rangle_{\mathcal{V}} = 1$ as in (2.23) first, and then the consistency requirement (2.24):

$$\begin{aligned} \sum_{i=1}^d \lambda_i &= \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} = \sum_{i=1}^d \langle \rho, F_i \rangle_{\mathcal{V}} \langle \sigma_i, u \rangle_{\mathcal{V}} \\ &= \langle \rho, u \rangle_{\mathcal{V}} = 1. \end{aligned} \quad (\text{F.20})$$

Equations (F.19) and (F.20) prove that $\iota(\rho) = \sum_{i=1}^d \lambda_i \delta_i$ is a convex combination of the δ_i 's, and hence for any $\rho \in \Omega$ it holds that $\iota(\rho) \in \Delta_d \subseteq \mathbb{R}^d$. Hence (iii) of definition F.1 is verified.

Let us now verify (iv), i.e that for all $E \in \mathcal{E}$, for all $i = 1, \dots, d$: $\langle \kappa(E), \delta_i \rangle_{\mathbb{R}^d} \in [0, 1]$. Fix $i \in \{1, \dots, d\}$ and $E \in \mathcal{E}$. We will make use of the fact that there exists $\{E_k \in \mathcal{E}\}_k$ such that $u = E + \sum_k E_k$. By the definition (F.17b) of the mapping κ and the orthonormality of the δ_i 's: $\langle \kappa(E), \delta_i \rangle_{\mathbb{R}^d} = \langle \sigma_i, E \rangle_{\mathcal{V}}$. Using the normalization (2.23):

$$1 = \langle \sigma_i, u \rangle_{\mathcal{V}} = \langle \sigma_i, E \rangle_{\mathcal{V}} + \sum_k \langle \sigma_i, E_k \rangle_{\mathcal{V}}. \quad (\text{F.21})$$

Using that $E, E_k \in \mathcal{E}$ and using (2.22b) that states here that $\sigma_i \in \mathcal{E}^{+\nu}$, it holds that $\langle \sigma_i, E \rangle_{\mathcal{V}}, \langle \sigma_i, E_k \rangle_{\mathcal{V}} \geq 0$. Then, (F.21) implies also that $\langle \sigma_i, E \rangle_{\mathcal{V}} \leq 1$. This proves that

$$\langle \kappa(E), \delta_i \rangle_{\mathbb{R}^d} = \langle \sigma_i, E \rangle_{\mathcal{V}} \in [0, 1]. \quad (\text{F.22})$$

Thus (iv) of definition F.1 is also verified. This proves that indeed $(\mathcal{V}, \Omega, \mathcal{E})$ is simplex-embeddable in d dimensions. \square