Superposition and its connections to Uncertainty, Entanglement and the Quantum Tensor Product

Vincenzo Fiorentino[1,*] and Kuntal Sengupta[2,1,†]

[1] *Department of Mathematics, University of York, York YO10 5GH, United Kingdom*

[2] *Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

Motivation and context: Superposition is a defining feature of quantum theory, emerging from the linear structure of Hilbert spaces and serving as a fundamental resource for quantum information processing. Given a collection of perfectly distinguishable quantum states $\mathsf{D} = \{|\phi\rangle_i\}_i$, a system in state $|\psi\rangle$ is in a superposition of the elements in D if there exist non-zero complex coefficients $\{c_i\}$ such that $|\psi\rangle = \sum_i c_i|\phi\rangle_i$, with $\sum_i |c_i|^2 = 1$. Quantum states when represented as rays of a Hilbert space, however, are not observable properties of a quantum system. Therefore, it is natural to ask whether superposition can be associated to, hence expressed as, properties of empirical, observable data. In this work, we indicate that it is possible to operationally define superposition in terms of constraints on the statistics observed in an experiment in a theory-independent way that reduces precisely to the standard linear expression in the specific case of quantum theory.

Previous attempts to formalise superposition in the context of *Generalised Probabilistic Theories* (GPTs) [1, 2], which generally lack an underlying Hilbert-space structure, have introduced it not as a relational property between states (as it is usually regarded in quantum theory) but as a global, structural feature of the state and effect spaces. Moreover, these definitions either apply only to theories with infinitely many extremal states [3] or reduce simply to non-simpliciality [4].

Defining superposition in GPTs: We resolve these limitations by proposing a definition of superposition rooted in the following operational scenario. Consider a GPT system with state and effect spaces $\mathcal{S}$ and $\mathcal{E}$, respectively. Suppose an experimenter has access to a measurement device $\mathsf{M} = \{e_i\}_{i=1}^n \subseteq \mathrm{Ext}[\mathcal{E}]$ (where $\mathrm{Ext}[\mathcal{E}]$ denotes the set of extremal effects) that can perfectly distinguish some *maximal* set of extremal states $\mathsf{D} = \{s_j\}_{j=1}^n \subseteq \mathrm{Ext}[\mathcal{S}]$, i.e. $e_i(s_j) = \delta_{ij}$. Another device prepares a state r which is observed to be *probabilistic* with respect to this measurement. If this preparation device were sampling from the set $\{s_1, \ldots, s_n\}$, the observed statistics would have a classical explanation as a statistical mixture. However, if the experimenter can rule out this possibility, for instance, by certifying that r is an extremal state, then there must be a different

* vincenzo.fiorentino@york.ac.uk

† kuntal.sengupta@neel.cnrs.fr

"phenomenological explanation" for this probabilistic behaviour. In quantum theory, one would call this superposition. We abstract this insight to arbitrary GPTs: superposition is manifested in the the observed probabilistic behaviour of extremal states relative to measurements that are maximally distinguishable.

Superposition is then said to *exist* in $(\mathcal{S}, \mathcal{E})$ if the theory admits states that satisfy our definition of superposition. By imposing stronger conditions to the existence of such states, we devise three *superposition principles* that capture structural features of the quantum theory within the GPT framework. First, *complete superposition* requires that every extremal state in $\mathcal{S}$ is a superposition. Second, *uniform superposition* stipulates that every extremal state lying outside a maximal set D can be expressed as a superposition of (some of) its elements. Thirdly, we call *mutual superposition* the property according to which, whenever an extremal state r is a superposition of a set $\{s_i\}$, then each s_i must likewise be a superposition of a set containing r. They capture inequivalent structural aspects of the quantum state and effect spaces.

Results: Having set the scene by establishing the main concepts, we explore superposition in different well-known GPTs, including classical theories, Boxworld [5], Spekkens' toy theory [6] and regular polygonal models [7]. We show that the existence of superposition is not automatically implied from the non-classicality of the theory, and that the three principles do not do not form a hierarchy. There are even theories such as Boxworld that exhibit superposition but fail to satisfy any of the stronger principles. We also identify peculiar features of superposition that diverge from quantum mechanical intuition. A notable example occurs in regular n-gon models, where an extremal state can exist as a superposition of a set D while remaining perfectly discriminable from some of its members.

We then show that *preparational uncertainty* [8–10], defined informally as the absence of "common eigenstates" between two maximally distinguishable measurements, can be regarded as a special form of superposition. In fact, one can devise an analogous set of uncertainty principles and show that each implies its corresponding superposition principle. Notably, so-called *uniform uncertainty principle*, which requires every pair of maximally distinguishable measurements to exhibit uncertainty, distinguishes qubit quantum theory (for which the principle holds) from regular n-gon models (where it is violated). In a sense, imposing uniform uncertainty is tantamount to taking the $n \to \infty$ limit, thereby isolating the Bloch disk from the state spaces of regular polygonal models.

We then analyse composite GPT systems. In the special case of quantum subsystems, we show that the quantum tensor product occupies a unique position among all possible no-signaling composition rules: it is the "largest" composition satisfying all three of our superposition principles. We also identify a class of GPTs—characterised by composite systems for which the operational dimension (i.e. the cardinality of a maximal set of perfectly distinguishable states) equals the product of those of its subsystems—where superposition in a subsystem necessarily carries over to the composite system. Nevertheless, superposition can emerge as a feature of a composite system even when it is absent in its individual components (e.g. Boxworld). In this class of theories, *entanglement* can be regarded as a special form of superposition proper of multipartite systems, in agreement with the quantum mechanical idea of superposition. These results establish our operational definition of superposition as not simply a descriptive tool, but as a structural requirement that can be employed to derive properties of quantum theory and identify operational principles that might be useful to re-axiomatise quantum theory [11].

The paper is not currently available as a pre-print on `arxiv.org`; however, a self-consistent preliminary draft is provided with this submission.

---

[1] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).

[2] L. Hardy, Quantum theory from five reasonable axioms (2001), quant-ph/0101012.

[3] G. M. D'Ariano, M. Erba, and P. Perinotti, Classical theories with entanglement, Phys. Rev. A **101**, 042118 (2020).

[4] G. Aubrun, L. Lami, C. Palazuelos, and M. Plávala, Entanglement and superposition are equivalent concepts in any physical theory, Phys. Rev. Lett. **128**, 160402 (2022).

[5] A. J. Short and J. Barrett, Strong nonlocality: a trade-off between states and measurements, New Journal of Physics **12**, 033034 (2010).

[6] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110.

[7] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, Limits on nonlocal correlations from the structure of the local state space, New Journal of Physics **13**, 063024 (2011).

[8] D. Saha, M. Oszmaniec, L. Czekaj, M. Horodecki, and R. Horodecki, Operational foundations for complementarity and uncertainty relations, Phys. Rev. A **101**, 052104 (2020).

[9] R. Takakura and T. Miyadera, Preparation uncertainty implies measurement uncertainty in a class of generalized probabilistic theories, J. Math. Phys. **61**, 082203 (2020).

[10] S. Aravinda, R. Srikanth, and A. Pathak, On the origin of nonclassicality in single systems, J. Phys. A: Math. Theor. **50**, 465303 (2017).

[11] G. Chiribella and R. W. Spekkens, *Quantum theory: informational foundations and foils* (Springer, 2016).

Superposition and its connections to Uncertainty, Entanglement and the Quantum Tensor Product

Vincenzo Fiorentino[1, *] and Kuntal Sengupta[2, 1, †]

[1] *Department of Mathematics, University of York, York YO10 5GH, United Kingdom*
[2] *Univ. Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France*

Textbook quantum superposition refers to the feature that certain linear combinations of Hilbert space rays, each representing a valid state, are themselves valid states. Although superposition is commonly attributed to a phenomenological understanding of quantum theory, this notion is not operational and relies on the Hilbert space formalism of the theory. We propose an operational and theory-independent definition of superposition based on constraints on observed measurement statistics. We use this notion to formulate *superposition principles* that capture structural features of the quantum state space within the framework of generalised probabilistic theories. We show that the existence of superposition is not automatically implied from the non-classicality of a theory, and that preparational uncertainty and entanglement can be viewed as special forms of superposition. We also show that the quantum tensor product emerges as the largest composition of quantum systems that satisfy all our superposition principles.
This draft is a work in progress. The presentation of our results will be polished by the time of the conference.

I. INTRODUCTION

Superposition is a defining feature of quantum theory. It provides a non-classical explanation for interference phenomena, such as the one observed in a double-slit experiment, about which Feynman went on to say that

"In reality, it contains the only mystery" [1].

More recently, quantum superposition has been at the heart of celebrated quantum algorithms [2], kicking off the field of quantum computing. The fundamental idea behind superposition is the fact that quantum states can be represented as rays of a Hilbert space [3, 4] and that every such state can be represented as a complex linear combination of other states. The popular interpretation is that every state can be seen as an interference of these states. However, this phenomenological viewpoint relies on the Hilbert space formalism of quantum theory. Since the description of interference phenomena should be independent of the mathematical framework used to study quantum theory, it is natural to ask what would one refer to as superposition if quantum theory is not mathematically modelled by Hilbert spaces.

The main goal of our contribution is to present and justify an operational notion of superposition that is based on constraints on observed measurement statistics. By remaining theory-independent, our definition extends the notion of superposition to frameworks of theories where states are not necessarily represented by rays in a Hilbert space. Specifically, we formalise superposition within the framework of *Generalised Probabilistic Theories* (GPTs) [5–13]. There have been several attempts to capture superposition operationally [14, 15]. Our version represents another candidate that, nevertheless, resolves some of the limitations of previous versions (see Appendix D for a detailed comparison), and reduces to textbook superposition in the quantum case.

Far from being a purely descriptive excercise, an operational notion of superposition offers insights into the structure of quantum theory by analysing its relationship with other non-classical features. This approach mirrors the generalisation of entanglement and uncertainty relations from their Hilbert space-based origins to broader operational frameworks [16–20]. Investigating superposition in non-quantum contexts helps isolate the specific properties that distinguish quantum theory from other physically or operationally plausible models. In this work, we focus on the interplay between superposition, preparational uncertainty, and entanglement. We then show that our notion of superposition can be used to uniquely single out the standard quantum tensor product from the set of all no-signalling composition rules.

This paper is organised as follows. Sec. II provides a brief review of the GPT framework and introduces basic concepts. Sec. III discusses superposition in quantum theory, forming the basis for our operational definition applicable to GPTs, which is presented in Sec. IV. Within this section, we introduce *superposition principles* that capture structural features of the quantum state space related to superposition, and we explore the presence of superposition in specific GPTs. Sec. V formalises preparational uncertainty in GPTs, showing that it defines a stronger version of superposition. In Sec. VI, we show that the standard tensor product for quantum system emerges as the largest composition of quantum systems satisfying all our superposition principles. Sec. VII focuses on composite systems more generally, identifying a class of GPTs where superposition in a subsystem necessarily carries over to any composition and where en-

* vincenzo.fiorentino@york.ac.uk
† This work was partially conducted at the author's previous affiliation: the University of York. kuntal.sengupta@neel.cnrs.fr

tanglement, in agreement with quantum-mechanical intuition, can be regarded as a special type of superposition for multipartite systems.

All proofs are provided in the Appendix.

II. PRELIMINARIES

GPTs are a framework to analyse theories in which experiments can be decomposed into preparations, operations and measurements of systems. In a GPT, the set of states, i.e. the *state space*, $\mathcal{S}$, is given by by a compact convex subset of a real vector space $\mathbb{V}$. The *effect space*, $\mathcal{E}$, is a compact convex subset of the dual vector space $\mathbb{V}^*$ with the property that for every $e \in \mathcal{E}$, $e(s) \coloneqq \langle e, s\rangle \in [0,1]$ for every $s \in \mathcal{S}$, where $\langle \cdot, \cdot \rangle$ is an inner product on $\mathbb{V}$. In addition, there exists an element (*unit effect*) $u \in \mathcal{E}$, such that $u(s) = 1$ for all $s \in \mathcal{S}$ (*normalisation*) and for every $e \in \mathcal{E}$, $u - e \in \mathcal{E}$. When $\mathcal{E}$ is the largest subset of $\mathbb{V}^*$ compatible with these conditions, the GPT is said to satisfy the *no-restriction hypothesis.* A state (effect) is said to be *extremal* if it cannot be expressed as a convex combination of other states (effects). We denote by $\mathrm{Ext}[\mathcal{S}]$ and $\mathrm{Ext}[\mathcal{E}]$ the sets of extremal states and effects, respectively.

Given two state spaces $\mathcal{S}_1 \subset \mathbb{V}_1$ and $\mathcal{S}_2 \subset \mathbb{V}_2$, a bipartite state space, $\mathcal{S}_{1,2} \coloneqq \mathcal{S}_1 \boxtimes \mathcal{S}_2 \subset \mathbb{V}_1 \otimes \mathbb{V}_2$, is a collection of states identified with a composition rule $\boxtimes$ with the property that for every $s \in \mathcal{S}_{1,2}$, $u_1 \otimes \mathrm{id}_2(s) \in \mathcal{S}_{2\leqslant}$ and $\mathrm{id}_1 \otimes u_2(s) \in \mathcal{S}_{1\leqslant}$, where u_1, id_1 and u_2, id_2 are the unit effect and identity map pair on the first and second systems respectively and $\leqslant$ denotes the subnormalised state space characterised by the convex hull of a state space with the zero vector. Two composition rules of interest are the *minimal* and *maximal* tensor product compositions. The minimal composition is defined as

$$\mathcal{S}_1 \underset{\min}{\otimes} \mathcal{S}_2 \coloneqq \mathrm{ConvHull}\{s_1 \otimes s_2 | s_1 \in \mathcal{S}_1, s_2 \in \mathcal{S}_2\}.$$

The maximal $\mathcal{S}_1 \underset{\max}{\otimes} \mathcal{S}_2$ composition is the largest bipartite state space compatible with all effects in the minimal composition of the respective effect spaces and satisfying normalisation. A bipartite compostion rule $\boxtimes$ is said to be non-signalling if for any two state spaces $\mathcal{S}_1$ and $\mathcal{S}_2$, $\mathcal{S}_1 \underset{\min}{\otimes} \mathcal{S}_2 \subseteq \mathcal{S}_{1,2} \subseteq \mathcal{S}_1 \underset{\max}{\otimes} \mathcal{S}_2$.

A collection $\mathsf{M} \coloneqq \{e_j\}_j$ of effects is said to be a *measurement* iff $\sum_j e_j = u$. A collection of extremal states $\mathsf{D} \coloneqq \{s_i\}_i$ is said to be *perfectly discriminable* (PD) if and ony if there exists a measurement $\{e_j\}_j \subseteq \mathrm{Ext}[\mathcal{E}]$, such that $\langle e_j, s_i\rangle = \delta_{i,j}$, where $\delta_{i,j} = 1$ when $i = j$ and 0 otherwise. D is said to be a *maximal* PD (MPD) set iff for any other PD set of states D', $|\mathsf{D}| \geqslant |\mathsf{D}'|$ holds—$|\mathsf{D}|$ is then known as the *measurement* or *operational dimension* of the system [21, 22]. A measurement $\mathsf{M} \subseteq \mathrm{Ext}[\mathcal{E}]$ that perfectly discriminates an MPD set D and such that $|\mathsf{M}| = |\mathsf{D}|$ is said to be a *maximally distinguishing and extremal* (MDE) measurement. We denote by $\mathsf{Det}(\mathsf{M})$ the union of all MPD sets that are perfectly distinguished by M. In this work, we assume that the set of MDE measurements is *informationally complete*; that is, every state in the theory can be tomographically reconstructed using only the statistics obtained from MDEs.

Qubit quantum theory can be phrased as a GPT by identifying its state space with the set of 2×2 unit-trace positive semi-definite matrices, a compact convex subset of the real vector space, $\mathbb{H}(\mathbb{C}^2)$, of 2×2 Hermitian matrices. An effect, Π, is a 2×2 positive semi-definite matrix, such that $\mathbb{1} - \Pi$ is also positive semi-definite, where $\mathbb{1}$ is the 2×2 identity matrix. The effect space is the largest set of such matrices with $\mathbb{1}$ as the unit effect, and forms a compact convex subset of $\mathbb{H}(\mathbb{C}^2)$. The action of the effects on the states is given by the Hilbert-Schmidt inner product. Qudit quantum theory can be understood in a similar way by identifying its state space with the compact convex subset, of $d\times d$ unit-trace positive semi-definite matrices, of the real vector space of $d\times d$ Hermitian matrices, and so on. The operational dimension of a qudit is equal to d [22].

In this work we investigate a wide class of GPTs other than quantum theory and classical theory (a theory where the state space is a simplex). We consider *gbits* which are characterised by two fiducial measurements with two outcomes each $(p(0|0), p(0|1))$—the probability of the second outcome can be trivially calculated and is redundant. The vertices of the state space of gbits correspond to $(0,0)$, $(0,1)$, $(1,0)$ and $(1,1)$. System characterised by more fiducial measurements with two or more number of outcomes are called *g-systems.* The vertices of the state spaces for g-systems are obtained in a similar way as gbit state spaces. The minimal tensor product of g-systems is called Generalised Local Theory (GLT) [13]. The maximal tensor product of such systems is called Generalised Non-Signalling Theory [13] or more commonly Boxworld [23]. We also consider single-system state spaces whose geometry is given by regular polygons, providing approximations to the real qubit state space, i.e. the Bloch disc [24]. In addition, we also construct single-system state spaces corresponding to fragments of quantum theory. We call them GPT-1 and GPT-2 and outline their details in Appendix A 5.

III. SUPERPOSITION IN QUANTUM THEORY

An extremal qubit state Ψ is a rank-1 2×2 matrix and therefore $|\Psi\rangle\rangle \coloneqq |\psi\rangle\langle\psi|$, where $|\psi\rangle \in \mathbb{C}^2$ is its eigenvector. Similarly, every extremal qudit state can also be represented by a vector in $\mathbb{C}^d$. The notion of superposition in quantum theory is inherently connected to this mathematical property. Given a finite-dimensional Hilbert-space $\mathcal{H} \cong \mathbb{C}^d$, and collection of extremal qudit states $\mathsf{D} = \{|\Phi_i\rangle\rangle\}_i^{d'\leqslant d}$, an extremal qudit with state $|\Psi\rangle\rangle$ is said to be in superposition of states in D, whenever there exists a collection of complex numbers $\{c_i\}_{i=1}^{d'}$, with $c_i \notin \{0,1\}$ for any i and $\sum_{i=1}^{d'} |c_i|^2 = 1$, such that

$$|\psi\rangle = \sum_{i=1}^{d'} c_i |\phi_i\rangle \ . \tag{1}$$

In fact, for any $|\psi\rangle$, the pair of collection D and $\{c_i\}_{i=1}^{d'}$ can always be found with $d' = d$. Textbook superposition is pertained to quantum states admitting Eq. (1), often associated to the analogy of superposition in wave mechanics. However, unlike waves in wave mechanics, vectors of the form $|\psi\rangle$ cannot be observed. Therefore it is natural to ask whether the notion of superposition in quantum mechanics is an artifact of the Hilbert space formalism or one can associate observable properties to it. In this work, we indicate that it is possible to operationally define superposition as constraints on the statistics observed in an experiment, in a theory-independent way. If one then assumes that the underlying theory is quantum mechanics, Eq. (1) can be derived.

To do this, let us analyse the textbook superposition in quantum theory phrased as a GPT. Consider an experiment where one needs to perfectly discriminate states in the set $\mathsf{D} = \{|\Phi_i\rangle\rangle\}_{i=1}^{d}$. This is indeed possible whenever $\mathrm{Tr}[|\Phi_i\rangle\rangle|\Phi_j\rangle\rangle] = \delta_{i,j}$, using the measurement $\mathsf{M} = \mathsf{D}$. In addition, since M spans the vectors space of $d \times d$ Hermitian matrices, D is an MPD set and M its MDE measurement. Then, for any state $|\Psi\rangle\rangle$ such that $\mathsf{c}_i := \mathrm{Tr}[|\Phi_i\rangle\rangle|\Psi\rangle\rangle] \in (0,1)$ for all i, we can find complex numbers $\{c_i\}_i$ with the property that $|c_i|^2 = \mathsf{c}_i$ and Eq. (1) holds. Since $\{\mathsf{c}_i\}_i$ represents a probability distribution, $\sum_{i=1}^{d} |c_i|^2 = \sum_{i=1}^{d} \mathsf{c}_i = 1$. In fact, every $|\Psi\rangle\rangle \notin \mathsf{D}$ admits this. Quantum superposition, therefore, is a property of the state $|\Psi\rangle\rangle$ with respect to the MPD set D.

An operational view of the scenario above is the following: Suppose an experimenter has access to a measurement device $\mathsf{M} = \{e_i\}_{i=1}^{d}$ that can perfectly distinguish some MPD set $\mathsf{D} = \{s_j\}_{j=1}^{d}$, i.e., $e_i(s_j) = \delta_{ij}$. Another device prepares a state r which is observed to be probabilistic with respect to this measurement. If this preparation device were sampling from the set $\{s_1, \ldots, s_d\}$, the observed statistics would have a classical explanation as a statistical mixture. However, if the experimenter can rule out this possibility, for instance, by certifying that r is an extremal state, then there must be a different "phenomenological explanation" for this probabilistic behaviour. In quantum theory, one would call this superposition.

IV. SUPERPOSITION IN GPTS

The operational prescription presented above need not assume that the underlying theory is quantum mechanics; such conclusions are possible in any GPT. In particular, superposition can be defined theory-independently as a statistical property of an extremal state with respect to an MPD set. We formalise this below.

Definition IV.1 (Superposition)**.** *Let $(\mathcal{S}, \mathcal{E})$ be a pair of state and effect spaces. Let $\mathsf{D} := \{s_i\}_{i=1}^{n}$ be an MPD set and $\mathsf{M} := \{e_i\}_{i=1}^{n}$ an MDE measurement of D. Let $r \in \mathrm{Ext}[\mathcal{S}]$ be a state with the property that for some $\mathsf{M}' \subseteq \mathsf{M}$, $e_j(r) \in (0,1)$ holds for all $e_j \in \mathsf{M}'$, and $\sum_{e_j \in \mathsf{M}'} e_j(r) = 1$. Then, r is a superposition of the states in the set $\mathsf{D}' := \{s \in \mathsf{D} | e_j(s) = 1, e_j \in \mathsf{M}'\}$.*

Note first, that if we drop the requirement that D is a maximal set, then although one could infer that r is a superposition, the set of states with respect to which it is a superposition may be falsely concluded. The following qutrit example captures this: the non-maximal PD set $\mathsf{D} = \{|0\rangle, |2\rangle\}$ is perfectly discriminated by the degenerate measurement $\mathsf{M} = \{|0\rangle\rangle, \mathbb{I}_3 - |0\rangle\rangle\}$. Without the maximality requirement, a state such as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle + \gamma|2\rangle$ with $\alpha, \beta, \gamma \neq 0$ would oddly be labeled as a superposition of $|0\rangle$ and $|2\rangle$ *only*, despite its non-zero overlap with $|1\rangle$. Second, superposition is defined in this work only for extremal states. This because superposition is only attributed to extremal states in quantum theory. We discuss in Appendix D 2 a notion of superposition in the absence of extremality and its consequences.

The notion of superposition in Definition IV.1 allows us to introduce *principles* of superposition—statements applicable to theories. The principles described in Definition IV.2 capture three different structural features of the quantum-mechanical state space that are associated with superposition. *Complete superposition* generalises the property that every quantum state of a system is expressable as a superposition of some of the elements of an othogonormal basis. *Uniform superposition* captures the fact that, relative to any fixed orthogonal basis, every other state in Hilbert space is expressible as a superposition of some of its elements. Finally, *mutual superposition* formalises the relation where, if a quantum state $|\psi\rangle$ is a superposition of some elements $\{|\phi\rangle_i\}$ of an orthonormal basis, then each $|\phi\rangle_i$ must in turn be a superposition of some elements of a basis containing $|\psi\rangle$.

Definition IV.2 (Existence and Principles of Superposition)**.** *Let $(\mathcal{S}, \mathcal{E})$ be a state and effect pair of a GPT. Superposition exists iff some $s \in \mathrm{Ext}[\mathcal{S}]$ is a superposition, following Definition IV.1. Further, this superposition is*

- *complete iff every $s \in \mathrm{Ext}[\mathcal{S}]$ is a superposition;*
- *uniform iff for any MPD set D, every $s \in \mathrm{Ext}(\mathcal{S}) \backslash \mathsf{D}$ is a superposition of some $\mathsf{D}'_s \subseteq \mathsf{D}$;*
- *mutual iff whenever $s \in \mathrm{Ext}[\mathcal{S}]$ is a superposition of D', then every $r \in \mathsf{D}'$ is a superposition of some D'', such that $s \in \mathsf{D}''$.*

Definitions IV.2 apply to both single and composite systems. Table I summarises the superposition properties of several theories, with the mathematical derivations deferred to Appendix A.

GPT	Existence of Superposition	Superposition Principles		
		complete	uniform	mutual
Quantum	✓	✓	✓	✓
Classical	×	–	–	–
GLT	×	–	–	–
Boxworld	✓	×	×	×
Spekkens	✓	✓	✓	✓
n-gon (n odd)	✓	✓	✓	✓
n-gon (n even)	✓	✓	×	✓*
GPT-1	✓	×	✓	✓
GPT-2	✓	×	✓	×

TABLE I. Classification of superposition across various GPTs. Theories above the horizontal line include both single and composite systems, whereas for those below the line we only examine individual systems. Here, n-gon refers to regular polygons with $n > 4$ sides and ✓* denotes that the property does not hold when $n = 6$. GPT-1 and GPT-2 are models defined in Appendix A 5. The symbol – indicates that the principles are not applicable.

Classical theories, characterised by simplicial state spaces and no-restriction hypothesis, fail to exhibit any form of superposition. However, the lack of classicality is not a signature of existence of superposition. Single system gbit state spaces or their minimal tensor product composition (GLT) do not admit any form of superposition. Interestingly, Spekken's toy theory [25], introduced to motivate the classicality of interference phenomena with epistemic restrictions, admits all superposition principles. In addition, superposition can exist in composite systems even when it is is absent in the subsystems, as captured by Boxworld where, however, none of the three stronger princples hold. A more general discussion on superposition in composite systems can be found in Sec. VII.

The table also demonstrates that the three types of superposition do not form a strict hierarchy. One can construct GPTs that admit uniform and mutual superposition but does not admit complete superposition (see GPT-1 in Appendix A 5), or that only admit uniform superposition (see GPT-2 in Appendix A 5).

Furthermore, the n-gon models display a non-quantum feature of superposition as per Definition IV.1, where a state can be a superposition of a set while remaining perfectly discriminable from some of its members (see Appendix A 4). We also observe that real qubit quantum theory inherits the full suite of superposition principles from the class of odd n-gons (inner approximations of the Bloch disc). In contrast, even n-gons (outer approximations of the Bloch disc) fail to uphold uniform superposition, establishing a parity-dependent structural gap that vanishes only in the continuous limit. Whether there is a higher principle underlying such explanation is not known to the authors.

V. RELATION TO PREPARATIONAL UNCERTAINTY

Imagine a scenario, where Alice prepares a state s and sends it to Bob. Bob is allowed to randomly perform one of two measurements M or N determined by a random variable X taking values $x \in \{0, 1\}$ depending on the choice of measurements. His observation, labelled by a random variable A taking values $a \in \{0, \cdots, d-1\}$, is reported back to Alice. Alice needs to guess Bob's outcome. Quantum theory allows for MDE measurements M and N, such that for no choice of prepared state, Alice can perfectly guess, on average. In particular, the distribution of Alice's guess admits:

$$\begin{aligned} &\mathrm{P}(A) = p(x=0)\mathrm{P}(A|x=0) + p(x=1)\mathrm{P}(A|x=1) < 1 \\ &\implies \mathrm{H}_{\min}(A) := -\log_2\left(\mathrm{P}(A)\right) > 0\,, \end{aligned} \tag{2}$$

where $\mathrm{H}_{\min}(A)$ is the min-entropy of Alice's guessing probability and the implication is known as an *entropic uncertainty relation* [26]. This phenomenon is known as *preparational uncertainty*[1]. The essence of this relies on the fact that there are no states that are deterministic with respect to both M and N. This can be phrased theory-independently as follows:

Definition V.1 (Preparational Uncertainty)**.** *Two MDEs* M *and* N *exhibit preparational uncertainty iff* $\mathrm{Det}(\mathsf{M}) \cap \mathrm{Det}(\mathsf{N})) = \emptyset$.

Definition V.1 introduces preparational uncertainty as a relational property for pairs of MDEs. A GPT system will then exhibit preparational uncertainty if it contains at least a pair of MDE with such property.

Similarly to our treatment of superposition, we can introduce principles that generalise different structural features found in quantum theory that are related to preparational uncertainty. *Complete uncertainty* denotes the property that, for every non-degenerate PVM of an arbitrary quantum system, one can find another non-degenerate PVM with no shared eigenstates. *Uniform uncertainty* abstracts the exclusive property of qubit quantum theory where any pair of non-degenerate PVMs shares no common eigenstates. Unlike Definition IV.2, we only establish two principles here.

Definition V.2 (Existence and Principles of Preparational Uncertainty)**.** *Let* $(\mathcal{S}, \mathcal{E})$ *be a state and effect pair of a GPT. Preparational uncertainty exists iff some pair of MDEs* $\mathsf{M}, \mathsf{N} \subseteq \mathrm{Ext}[\mathcal{E}]$ *exhibits preparational uncertainty, following Definition V.1. Further, this uncertainty is*

[1] Preparational uncertainty [26–29] is not to be confused with measurement uncertainty [30, 31], which pertains to the impossibility to measure two observables jointly. Both these forms of uncertainty have been discussed extensively in the literature in the context of operational theories [16–20].

- *<u>complete</u> iff for every MDE* M *there exists an MDE* N *such that* M *and* N *exhibit preparational uncertainty*
- *<u>uniform</u> iff every pair of MDEs* $\mathsf{M}, \mathsf{N} \subseteq \mathrm{Ext}[\mathcal{E}]$ *exhibits preparational uncertainty*

Table II summarises these properties for the GPTs considered; detailed derivations are provided in Appendix A.

GPT	**Existence of Uncertainty**	**Uncertainty Principles**	
		complete	uniform
Quantum ($d = 2$)	✓	✓	✓
Quantum ($d > 2$)	✓	✓	×
Classical	×	–	–
GLT	×	–	–
Boxworld	×	–	–
Spekkens	✓	✓	✓
n-gon ($n \leqslant 6$)	×	–	–
n-gon ($n > 6$)	✓	✓	×
GPT-1	×	–	–
GPT-2	×	–	–

TABLE II. Classification of preparational uncertainty in various GPTs. Theories above the horizontal line include both single and composite systems, whereas for those below the line we only examine individual systems. GPT-1 and GPT-2 are defined in Appendix A 5. The symbol – indicates that the principles are not applicable.

A comparison with Table I illustrates that superposition and preparational uncertainty are *not* operationally equivalent. For instance, we mentioned that uniform uncertainty only holds for individual qubits ($d = 2$), where each projector belongs to a unique PVM. Furthermore, in contrast to the superposition principles, the uncertainty principles defined here form a hierarchy: uniform uncertainty necessarily implies complete uncertainty.

Preparational uncertainty is, in fact, a strictly stronger notion than superposition. Theorem 1 below formalises this by establishing relationships of implication between the corresponding principles in Definitions IV.2 and V.2. The last part of the proof relies on the additional operational assumption of *outcome sharpness of MDEs* (OS): for every effect e in an MDE there exists a unique extremal state s such that $e(s) = 1$.

Theorem 1. *Let* $(\mathcal{S}, \mathcal{E})$ *be a state and effect space pair of a GPT. Then,*

(i) *if preparational uncertainty exists, superposition exists;*

(ii) *if* $(\mathcal{S}, \mathcal{E})$ *admits complete uncertainty, it also admits complete superposition;*

(iii) *if* $(\mathcal{S}, \mathcal{E})$ *admits uniform uncertainty and (OS), it also admits uniform superposition.*

The connection between superposition and uncertainty is made even more clear when outcome sharpness of MDEs is assumed from the outset. In such a class of GPTs, in fact, each uncertainty principle can be redefined entirely in terms of superposition states as follows (see Appendix B 2 for proofs):

Lemma 1. *Let* $(\mathcal{S}, \mathcal{E})$ *be a state and effect space pair admitting outcome sharpness of MDEs. Then the following conditions respectively imply existence, completeness and uniformity of preparational uncertainty:*

(i) *there exists MPD sets,* D, D'*, such that every* $s \in \mathsf{D}$ *is a superposition of* D'*;*

(ii) *for every MPD set* D*, there exists an MPD set* D' *such that all* $s \in \mathsf{D}$ *are superpositions of* D'*;*

(iii) *for any pair of MPD sets* D, D'*, all* $s \in \mathsf{D}$ *are superpositions of* D' *and vice versa.*

Theories satisfying (OS) include classical and quantum systems (for all d), Spekkens' model, and the toy theories GPT-1 and GPT-2. Notably, qubit quantum theory—specifically, the real qubit state space obtained in the $n \to \infty$ limit of regular n-gon state spaces—can be uniquely isolated from the class of polygonal theories by requiring uniform uncertainty. In this class of models, preparational uncertainty arises only beyond the threshold $n = 6$, wherein complete uncertainty is satisfied; however, uniform uncertainty only emerges in the quantum limit.

VI. SINGLING OUT THE QUANTUM TENSOR PRODUCT

The study of GPTs has been principally motivated by the goal of identifying reasonable principles to re-axiomatise quantum theory [32]. Information theoretic principles have been proposed since that try to single out the set of quantum correlations [33–38]. Motivated by these, we investigate whether any of the superposition principles stated in Definition IV.2 deem useful for this purpose. Here, we do not explore the connection between superposition and the set of correlations realisable in a causal structure [2]. Rather, we investigate, assuming all subsystems are characterised by quantum state and effect spaces, what the largest composition is such that the superposition principles carry over. We show that the quantum tensor product is the largest composition to admit mutual superposition. We formalise this below.

[2] A causal structure is a collection of variables arranged as nodes of a directed acyclic graph where some of the nodes are labelled as observed. It represents the causal relations among the variables.

Theorem 2. *Let $(\mathcal{S}_{\mathcal{Q}1}, \mathcal{E}_{\mathcal{Q}1})$ and $(\mathcal{S}_{\mathcal{Q}2}, \mathcal{E}_{\mathcal{Q}2})$ be the state and effect spaces describing two quantum systems with operational dimensions d_1 and d_2, respectively. Then, the set of unit-trace positive semi-definite matrices $\mathcal{S}_{\mathcal{Q}1,2} \subset \mathbb{H}(\mathbb{C}_1^d \otimes \mathbb{C}_2^d)$ is the largest state-space for the bipartite system which, with a maximal effect space, satisfies mutual superposition.*

This result provides an operational way for characterising the quantum composition rule, in alternative to the requirement that the composite system satisfies *strong self-duality*[3]. Similar arguments for the remaining superposition principles do not hold: as shown in Lemma 2 below, the maximal tensor product of quantum systems satisfies both complete and uniform superposition.

Theorem 2 allows us to make the following, stronger statement.

Corollary 1. *The quantum tensor product is the largest composition rule for quantum systems that satisfies all superposition principles of Definition IV.2.*

VII. SUPERPOSITION FROM SUBSYSTEMS TO ARBITRARY COMPOSITIONS

We investigate whether the presence of superposition in arbitrary subsystems carries over to any of their compositions. Conversely, one might also ask whether the presence of superposition in a composite system implies it in its subsystems. Regarding the latter, we have already seen that the bipartite state space of Boxworld exhibits superposition, whereas single gbit systems do not. Consequently, superposition can emerge as a feature of a composite system even when it is absent in its individual components. We now focus on the former question, considering bipartite systems for simplicity, though the results hold for more parties.

Superposition, as defined in Definition IV.1, requires an MDE and an extremal state that is probabilistic with respect to its outcomes. To determine whether superposition carries over to a composite system, one must first characterise the MDEs of the composition. In quantum theory, the operational dimension of a composite system $(\mathcal{S}_{\mathcal{Q}1,2}, \mathcal{E}_{\mathcal{Q}1,2})$ equals the product of the operational dimensions of its components. This ensures that the tensor product $\mathsf{M}_1 \otimes \mathsf{M}_2 := \{e_i \otimes f_j\}_{ij}$ of the local MDEs $\mathsf{M}_1 = \{e_i\}_i$ and $\mathsf{M}_2 = \{f_j\}_j$ constitutes a valid MDE for the composite system. A sufficient condition for this property was provided in [22], with the minimal and maximal products of quantum systems serving as key examples. However, it does not hold in general. A notable counterexample is the minimal tensor product of the state spaces of two regular pentagons (see Appendix C).

For the class of theories where product MDEs are MDEs for the composite system, we show that the superposition principles of a subsystem—with the exception of mutual superposition (see Theorem 2)—carry over to the unrestricted composition. For uniform superposition, outcome sharpness of MDEs (OS) is also assumed.

Lemma 2. *Let $(\mathcal{S}_1, \mathcal{E}_1)$ and $(\mathcal{S}_2, \mathcal{E}_2)$ be two GPT systems where $\mathcal{E}_i$ denotes the maximal effect space of $\mathcal{S}_i$. Let $\mathcal{S}_{1,2}$ be a no-sigalling composition of these state spaces, and let $\mathcal{E}_{1,2}$ be the associated maximal effect space, such that the operational dimension of the composite system $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ equals the product of the operational dimensions of the subsystems. Then,*

(i) *If $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies the existence of superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ also satisfies the existence of superposition;*

(ii) *If $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies complete superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ also satisfies complete superposition;*

(iii) *If $(\mathcal{S}_1, \mathcal{E}_1)$, $(\mathcal{S}_2, \mathcal{E}_2)$ and $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ satisfy (OS) and $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies uniform superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ satisfies uniform superposition.*

In quantum theory, entanglement for extremal states can be regarded as a special form of superposition proper of multipartite systems: a state with Schmidt decomposition $|\psi\rangle_{AB} = \sum_i c_i |i_A i_B\rangle$ is entangled iff two or more of its Schmidt coefficients c_i are non-zero; this implies that $|\psi\rangle_{AB}$ is probabilistic with respect to a measurement in the Schmidt product basis $\{|i_A j_B\rangle\}_{i,j}$. An analogous connection between entanglement and superposition holds whenever product MDEs are MDEs of the composite system, as formalised in Lemma 3. The proof (Appendix B 5) relies on the assumption that the MDEs of a subsystem are informationally complete.

Lemma 3. *Let $(\mathcal{S}_1, \mathcal{E}_1)$ and $(\mathcal{S}_2, \mathcal{E}_2)$ be two GPT systems where $\mathcal{E}_i$ denotes the maximal effect space of $\mathcal{S}_i$. Let $\mathcal{S}_{1,2} \supset \mathcal{S}_1 \underset{\min}{\otimes} \mathcal{S}_2$ be a no-sigalling composition of these state spaces, and let $\mathcal{E}_{1,2}$ be the associated maximal effect space, such that the operational dimension of the composite system $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ equals the product of the operational dimensions of the subsystems. Then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ exhibits superposition.*

Lemma 3 establishes that, for the considered class of GPTs, the existence of entanglement—ensured by choosing a tensor product strictly larger than the minimal—implies the presence of superposition. In Boxworld, superposition is determined exclusively by entanglement, as implied by GLT not exhibiting superposition. The converse, i.e. whether superposition implies entanglement in this class of GPTs, is not true in general. For example, the minimal tensor product of quantum systems allows product states such as $|++\rangle\rangle$ and product MDEs such as $\{|00\rangle\rangle, |01\rangle\rangle, |10\rangle\rangle, |11\rangle\rangle\}$.

[3] There exists an inner product $\langle\cdot,\cdot\rangle$ such that the cone formed by the state space and pointing on the zero vector coincides with the cone formed by the effect space and pointing on the zero vector.

VIII. SUMMARY

We have introduced an operational definition of superposition (Definition IV.1) that abstracts the standard quantum mechanical notion to arbitrary GPTs. In contrast to existing approaches [14, 15], our definition: (i) characterises superposition as a relational property of an extremal state with respect to a set of perfectly distinguishable extremal states; (ii) does not necessitate an infinite number of extremal states; and (iii) constitutes a condition strictly stronger than nonclassicality.

Building upon this definition, we have devised three superposition principles (Definition IV.2) that capture distinct geometric properties of the quantum state space, analysing their validity across commonly studied GPTs. We showed that preparational uncertainty, defined for pairs of MDEs (Definitions V.1-V.2) is a stronger condition than superposition (Theorem 1).

Turning to the composition of quantum systems, we showed that the standard quantum tensor product is the largest among the no-signalling composition rules that is compatible with all three superposition principles in Definition IV.2; specifically, it is the largest compatible with mutual superposition (Theorem 2). Extending the analysis of composition beyond quantum systems, we showed that any GPT where the operational dimension of the composite system is multiplicative, superposition in a subsystem necessarily carries over to the composite system (Lemma 2). Furthermore, in this class of theories, entanglement can be fundamentally regarded as a special manifestation of superposition in multipartite systems (Lemma 3).

IX. ACKNOWLEDGEMENTS

The authors thank Marco Erba and Stefan Weigert for useful discussions. VF was supported by the Leverhulme Trust, grant number RPG-2024-201. KS work was supported by the Departmental Studentship from the Department of Mathematics, University of York and l'Agence Nationale de la Recherche (ANR) project ANR-22-CE47-001.

[1] R. P. Feynman, *The Feynman Lectures on Physics, Volume III* (California Institute of Technology, 1964).
[2] P. Shor, Algorithms for quantum computation: discrete logarithms and factoring, in *Proceedings 35th Annual Symposium on Foundations of Computer Science* (1994) pp. 124–134.
[3] P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1930).
[4] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (Springer Berlin, 1932).
[5] I. E. Segal, Postulates for general quantum mechanics, Annals of Mathematics **48**, 930 (1947).
[6] E. B. Davies and J. T. Lewis, An operational approach to quantum probability, Commun. Math. Phys. **17**, 239 (1970).
[7] G. Ludwig, Attempt of an axiomatic foundation of quantum mechanics and more general theories. III, Communications in Mathematical Physics **9**, 1 (1968).
[8] G. Dähn, Attempt of an axiomatic foundation of quantum mechanics and more general theories. IV, Commun. Math. Phys. **9**, 192 (1968).
[9] P. Stolz, Attempt of an axiomatic foundation of quantum mechanics and more general theories VI, Communications in Mathematical Physics **23**, 117 (1971).
[10] B. Mielnik, Generalized quantum mechanics, Communications in Mathematical Physics **37**, 221 (1974).
[11] R. Giles, Foundations for quantum mechanics, J. Math. Phys. **11**, 2139 (1970).
[12] S. Gudder, Convex structures and operational quantum mechanics, Commun. Math. Phys. **29**, 249 (1973).
[13] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).
[14] G. M. D'Ariano, M. Erba, and P. Perinotti, Classical theories with entanglement, Phys. Rev. A **101**, 042118 (2020).
[15] G. Aubrun, L. Lami, C. Palazuelos, and M. Plávala, Entanglement and superposition are equivalent concepts in any physical theory, Phys. Rev. Lett. **128**, 160402 (2022).
[16] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, Comparing the degrees of incompatibility inherent in probabilistic physical theories, Europhys. Lett. **103**, 10002 (2013).
[17] R. Takakura and T. Miyadera, Preparation uncertainty implies measurement uncertainty in a class of generalized probabilistic theories, J. Math. Phys. **61**, 082203 (2020).
[18] D. Saha, M. Oszmaniec, L. Czekaj, M. Horodecki, and R. Horodecki, Operational foundations for complementarity and uncertainty relations, Phys. Rev. A **101**, 052104 (2020).
[19] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: a unified framework, New J. Phys. **21**, 113053 (2019).
[20] S. Aravinda, R. Srikanth, and A. Pathak, On the origin of nonclassicality in single systems, J. Phys. A: Math. Theor. **50**, 465303 (2017).
[21] N. Brunner, M. Kaplan, A. Leverrier, and P. Skrzypczyk, Dimension of physical systems, information processing, and thermodynamics, New J. Phys. **16**, 123050 (2014).
[22] S. Sen, E. P. Lobo, R. K. Patra, S. G. Naik, A. Das Bhowmik, M. Alimuddin, and M. Banik, Timelike correlations and quantum tensor product structure, Phys. Rev. A **106**, 062406 (2022).
[23] A. J. Short and J. Barrett, Strong nonlocality: a tradeoff between states and measurements, New Journal of Physics **12**, 033034 (2010).
[24] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, Limits on nonlocal correlations from the structure of the local state space, New Journal of Physics **13**, 063024 (2011).
[25] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110.

[26] H. Maassen and J. B. M. Uffink, Generalized entropic uncertainty relations, Phys. Rev. Lett. **60**, 1103 (1988).
[27] W. Heisenberg, Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik, Z. Phys. **43**, 172 (1927).
[28] H. P. Robertson, The uncertainty principle, Phys. Rev. **34**, 163 (1929).
[29] V. Fiorentino and S. Weigert, Uncertainty relations for the support of quantum states, J. Phys. A: Math. Theor. **55**, 495305 (2022).
[30] P. Busch, P. Lahti, and R. F. Werner, Measurement uncertainty relations, J. Math. Phys. **55**, 042111 (2014).
[31] P. Busch and N. Stevens, Direct tests of measurement uncertainty relations: What it takes, Phys. Rev. Lett. **114**, 070402 (2015).
[32] L. Hardy, Quantum theory from five reasonable axioms (2001).
[33] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on nonlocality in any world in which communication complexity is not trivial, Phys. Rev. Lett. **96**, 250401 (2006).
[34] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum nonlocality and beyond: Limits from nonlocal computation, Phys. Rev. Lett. **99**, 180502 (2007).
[35] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature **461**, 1101 (2009).
[36] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Royal Society **466**, 881 (2009).
[37] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, Nature Communications **4**, 2263 (2013).
[38] M. Müller, Probabilistic theories and reconstructions of quantum theory, SciPost Physics Lecture Notes 10.21468/scipostphyslectnotes.28 (2021).
[39] A. M. Gleason, Measures on the closed subspaces of a hilbert space, in *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution* (Springer Netherlands, Dordrecht, 1975) pp. 123–133.
[40] M. F. Pusey, Stabilizer notation for spekkens' toy theory, Foundations of Physics **42**, 688 (2012).
[41] P. Janotta and R. Lal, Generalized probabilistic theories without the no-restriction hypothesis, Phys. Rev. A **87**, 052131 (2013).
[42] S. Massar and M. K. Patra, Information and communication in polygon theories, Phys. Rev. A **89**, 052124 (2014).
[43] L. O. Hansen, A. Hauge, J. Myrheim, and P. Ø. Sollid, Extremal entanglement witnesses, Int. J. Quantum Inf. **13**, 1550060 (2015).

Appendix A: Superposition and Preparational Uncertainty in GPTs – Examples

1. Quantum and Classical theories

a. Superposition We mentioned in the justification following Definition IV.2 that, by construction, quantum theory satisfies all three superposition principles introduced in Definition IV.2. Specifically, the fact that for every (extremal) unit-trace positive semi-definite matrix, there exists a basis of the Hilbert space in which it is not diagonal implies that complete superposition is satisfied. Uniform superposition follows from the outcome sharpness of MDEs (OS) (see Sec. V), according to which every MDE M of a quantum system (that is, every non-degenerate PVM) perfectly distinguishes a unique MPD set D. Consequently, every state not belonging to D will be probabilistic with respect to M. Finally, mutual superposition follows from the property that, if $|\psi\rangle = \sum_{i=1}^{d'} c_i|\phi\rangle_i$ holds with $c_i \neq 0$ for all i, then there exists an orthonormal basis of $\mathcal{H}$ including $|\psi\rangle$, e.g. $\{|\psi\rangle, |\psi\rangle_1, \ldots, |\psi\rangle_{d-1}\}$, in which the operators $|\Phi_i\rangle\rangle$ are not diagonal. Existence of superposition and all three associated principles are preserved under the standard composition of quantum system, which ensures that the state (effect) space of a composite system with, say, two subsystems with operational dimensions d_A and d_B, respectively, is isomorphic to the state (effect) space of a non-composite system with operational dimension $d_A d_B$.

Classical theories do not exhibit superposition as per Definition IV.1. They are characterised by simplicial state spaces, where the vertices (i.e. the set of all allowed extremal states) form the unique MPD set in the theory. Likewise, classical composite systems, for which the minimal and maximal tensor products coincide, feature a single product MDE.

b. Preparational Uncertainty In quantum theory, any pair of MDEs $\mathsf{M} = \{P_i\}_{i=1}^d$ and $\mathsf{M} = \{Q_i\}_{j=1}^d$ such that $[P_i, Q_j] \neq 0$ for all i, j exhibit preparational uncertainty. In every Hilbert space $\mathcal{H}$ with finite $d \coloneqq \dim(\mathcal{H})$, given any MDE M, there exist infinitely many MDEs N such that (M, N) exhibits uncertainty, hence complete uncertainty (Definition V.2) is satisfied for both single and composite systems. However, for $d \geqslant 3$, a given rank-1 projector P_i belongs to infinitely many different MDEs. This implies that uniform uncertainty does not apply to quantum systems, both single and composite, with the exception of a single qubit. In $d = 2$, in fact, each P_i belongs to a unique MDE: in the terminology of Gleason [39], projective qubit measurements do not *intertwine*. Classical theories do not exhibit preparational uncertainty as each single or composite system features a single MDE.

2. GLT and Boxworld

a. Superposition In GLT, the MDEs are given by collections of ray-extremal effects that sum up to the unit effect. Since the inner product between any extremal state and a ray-extremal effect is either 0 or 1, there does not exist an extremal state which is probabilistic with respect to an MDE measurement. Consequently, GLT does not admit any principle of superposition. On the other hand, in Boxworld, where the MDEs are also formed by collection of ray-extremal effects, there exist extremal states, namely PR-boxes, that are probabilistic. We will show this with an example. We represent a state by

$$\begin{pmatrix} p(0,0|0,0) & p(0,1|0,0) & p(0,0|0,1) & p(0,1|0,1) \\ p(1,0|0,0) & p(1,1|0,0) & p(1,0|0,1) & p(1,1|0,1) \\ p(0,0|1,0) & p(0,1|1,0) & p(0,0|1,1) & p(0,1|1,1) \\ p(1,0|1,0) & p(1,1|1,0) & p(1,0|1,1) & p(1,1|1,1) \end{pmatrix}$$

where $p(a,b|x,y)$ denotes the probability of getting the outcomes a and b when fiducial measurements x and y are performed. In this representation, the PR box is given as:

$$\frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

and the MDE with respect to which it is probabilistic is

$$\left\{ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\}.$$

This MDE perfectly discriminates the following collection of extremal states:

$$\left\{ \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right\},$$

which are all separable states. Therefore the PR box is a superposition of these separable states. This is analogous to the maximally entangled state $|\Phi^+\rangle\rangle$ being a superposition of the separable states $|00\rangle\rangle$ and $|11\rangle\rangle$. However, since there is no MDE that can discriminate four PR boxes, the stronger superposition principles of Definition IV.2 are not admitted in Boxworld.

b. Preparational Uncertainty Following from the fact that the MDEs in both GLT and Boxworld can only perfectly discriminate extremal separable states, and that the inner product of such states with any ray extremal effect is either 0 or 1, we infer that neither of these theories admit existence of preparational uncertainty. Therefore, the reamaining principles are not admitted as well.

3. Spekkens' toy model

a. Superposition Spekkens' toy theory [25] was originally introduced to show that many phenomena of quantum theory can be explained via a "classical" epistemic restriction on underlying hidden variables. The model is functionally equivalent to the stabilizer formalism of a single qubit [40], and comprises six extremal (epistemic) states (see Fig. 1) and three MDE measurements (analogous to Pauli X, Y, Z measurements). Each MDE distinguishes a unique pair of extremal states.

By construction, for any extremal state s that does *not* belong to the unique MPD set associated with a measurement $\mathsf{M} = \{e_1, e_2\}$, the model satisfies the condition $e_i(s) = 1/2$ for $i = 1, 2$. This defining characteristic, known as the *equivalence-balance principle*, ensures that the epistemic states not distinguished by a measurement appear maximally uncertain with respect to its outcomes. This principle directly implies the existence of superposition within the theory and further guarantees that the model satisfies all three superposition principles in Definition IV.2.

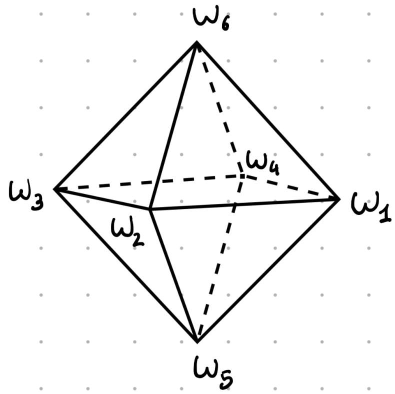

FIG. 1. [Sketch of] the state space of Spekkens' toy model, with the six extremal states labeled.

b. Preparational Uncertainty The equivalence-balance principle ensures that all three pairs of MDEs allowed in the theory exhibit uncertainty (uniform uncertainty). It follows that, along with qubit quantum theory, Spekkens' model is another example of GPT that satisfies both uncertainty principles of Definition V.2.

4. Regular n-gons

a. Superposition For $n > 4$, let $\mathcal{S}_n$ be the state space formed by the convex hull of the extremal points

$$\omega_n(i) = \begin{pmatrix} r_n \cos \frac{2\pi i}{n} \\ r_n \sin \frac{2\pi i}{n} \\ 1 \end{pmatrix},$$

where

$$r_n = \sqrt{\frac{1}{\cos(\pi/n)}},$$

which define the vertices of the regular n-gon [24, 41, 42]. The ray-extremal effects in the effect space $\mathcal{E}_n$ are

$$e_n(i) = \begin{cases} \frac{1}{2}\begin{pmatrix} r_n \cos \frac{(2i+1)\pi}{n} \\ r_n \sin \frac{(2i+1)\pi}{n} \\ 1 \end{pmatrix}, & n \text{ even}, \\ \frac{1}{1+r_n^2}\begin{pmatrix} r_n \cos \frac{2\pi i}{n} \\ r_n \sin \frac{2\pi i}{n} \\ 1 \end{pmatrix}, & n \text{ odd}, \end{cases}$$

with the unit effect being $u = (0,0,1)^\top$. When n is odd, the complements $\{u - e_n(i)\}_i$ are also extremal effects (but not ray-extremal), and the corresponding GPT $(\mathcal{S}_n, \mathcal{E}_n)$ is (strongly) self-dual. To simplify the expressions, we introduce the planar angles for effects and states, respectively:

$$\theta_j^n = \begin{cases} \frac{(2j+1)\pi}{n}, & n \text{ even} \\ \frac{2\pi j}{n}, & n \text{ odd} \end{cases}, \qquad \varphi_i^n = \frac{2\pi i}{n}.$$

Then the inner product takes the form

$$e_n(j)\big(\omega_n(i)\big) = \frac{1 + r_n^2 \cos(\vartheta_j^n - \varphi_i^n)}{D_n},$$

where

$$D_n = \begin{cases} 2, & n \text{ even}, \\ 1 + r_n^2, & n \text{ odd}. \end{cases}$$

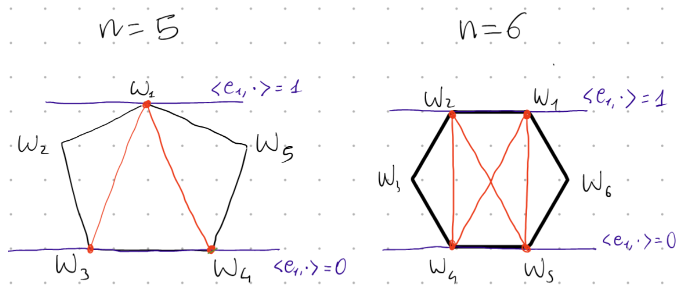

FIG. 2. [Sketches of] The pentagon ($n = 5$) and hexagon ($n = 6$) state spaces. Pairs of vertices connected by red segments constitute the MPD sets distinguished by the MDE $\mathsf{M} = \{e(1), u - e(1)\}$. Points lying on the top (resp. botom) blue line yield probability 1 (resp. 0) for the effect e_1.

For all n, the operational dimension is two. In particular, for odd n, one has $e_n(j)\big(\omega_n(i)\big) = 1$ when $\theta_j^n - \varphi_i^n = 0$—that is, when $i = j$. This means that, for each extremal state $\omega_n(i)$, there exists a unique ray-extremal effect, namely $e_n(i)$, for which it gives probability 1 (while giving probability strictly less than 1 for any other state). In other words, polygonal model with odd n satisfy *sharpness*. In addition, the equation $e_n(j)\big(\omega_n(i)\big) = 0$ is satisfied when $\cos(\vartheta_j^n - \varphi_i^n) = -\cos(\pi/n)$, i.e. when $\theta_j^n - \varphi_i^n = \pi \pm \pi/n \pmod{2\pi}$. Therefore, for each $\omega_n(i)$, there are two ray-extremal effects $e_n(j)$, where $j = i + \frac{(n\pm1)}{2} \pmod{n}$, assigning it probability zero. For example, both $\{\omega_7(2), \omega_7(6)\}$ and $\{\omega_7(2), \omega_7(5)\}$ are MPD sets that are discriminated by the MDE measurement $\{e_7(2), u - e_7(2)\}$.

Repeating the calculation for even n, we find that $e_n(j)\big(\omega_n(i)\big) = 1$ when $\theta_j^n - \varphi_i^n = \pm\pi/n$. Fixing i, the two solutions are $j = i$ and $j = i - 1 \pmod{n}$, indicating that these GPTs do not satisfy sharpness. Similarly, $e_n(j)\big(\omega_n(i)\big) = 0$ is satisfied when $j = i + n/2$ or when $j = i + (n-2)/2 \pmod{n}$. It follows that the same MDE measurement $\{e_n(j), u - e_n(j)\}$ perfectly distinguishes a total of *four* MPD sets of extremal states.

Fig. 2 depicts the state spaces of the pentagon ($n = 5$) and the hexagon ($n = 6$) models, highlighting one MDE and the maximal sets associated with it.

For any odd n, a given MDE $\mathsf{M} = \{e_n(j), u - e_n(j)\}$ has a total of three deterministic states, $\mathsf{Det}(\mathsf{M}) = \{\omega_n(j), \omega_n(j - (n+1)/2), \omega_n(j - (n-1)/2)\}$ (all sums of indices are mod n). Graphically, MPD sets always corresponds to pairs of vertices where one is an endpoint of the segment opposite the other vertex. It follows that these polygonal GPTs exhibit superposition, since they all count more than three vertices. In particular, complete superposition holds: every extremal state is probabilistic with respect to some MDE. Uniform superposition also applies: each MPD set D is distinguished by exactly two MDEs, and every extremal state not in D is probabilistic to at least one of them (Fig. 3). Moreover, given any state r in a superposition of some MPD set D, it is possible to find a extremal state t at one end of the opposing side of the polygon, such that *both* elements of D are probabilistic with respect to at least one of the two MDEs that distinguish $\{r, t\}$. Therefore, mutual superposition applies to every regular n-gon with odd n.

For even n, a given MDE $\mathsf{M} = \{e_n(j), u - e_n(j)\} = \{e_n(j), e_n(j + n/2)\}$ has four deterministic states: $\mathsf{Det}(\mathsf{M}) = \{\omega_n(j), \omega_n(j+1), \omega_n(j - n/2), \omega_n(j - (n-2)/2)\}$ (all sums of indices are mod n). Superposition exists in these models since $n > 4$; in particular, complete superposition applies because of the symmetric properties of the state and effect spaces: as in the odd case, each extremal state is probabilistic with respect to some MDE. In contrast with the odd case, however, here there are two classes of MPD sets. In fact, any vertex forms an MPD pair with its antipodal vertex, but also with either of the two vertices adjacent to the antipodal vertex. In the first case (where the vertices are antipodal), there are two MDEs that distinguish them; whereas in the second case (where there vertices are not antipodal) there is only one MDE that distinguishes them (Fig. 4). In the first case, we find that any third extremal state is probabilistic with at least one of these MDEs—this is the same argument that justifies uniform superposition for odd n. However, for MPD pairs of the second kind, there will always exist two additional extremal states that are deterministic with respect to the unique MDE associated with it. This is true for *any* even $n > 4$: we conclude that, in general, uniform superposition does not hold in these models. A separate treatment for the two kinds of MPD sets is required to check whether mutual superposition holds. For a hexagon ($n = 6$), we can provide an invalidating example: consider an arbitrary MPD set $\{s_1, s_2\}$ of the second type; the vertex between these two extremal states, call it s_3, will be probabilistic with respect to the unique MDE associated with $\{s_1, s_2\}$. Hence, it is a superposition of the two states. However, whatever MPD set we consider that includes s_3, all of their associated MDEs will include either s_1 or s_2 among their deterministic states (Fig. 4). Thus, mutual superposition is violated for $n = 6$. The same is not true for larger even number of sides. In Fig. 5, we show this visually for the case of $n = 8$: for either kind of MPD set, the superposition states (in blue) can be joined to form MPD sets with respect to which both of the original states are superpositions.

As noted in the main text, regular n-gon GPTs exhibit a peculiar feature of superposition, which can be extrapolated

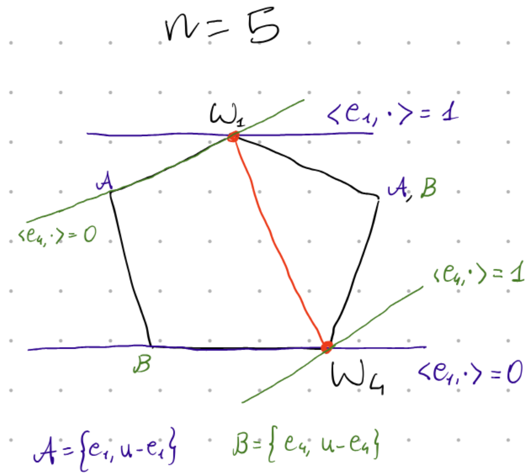

FIG. 3. [Sketch of the] regular pentagon GPT. The MPD set $\{\omega(1), \omega(4)\}$ is highlighted, along with the two MDEs A, B that discriminate it; vertices lying on the blue and green lines represent the deterministic extremal states of A and B, respectively. For each vertex not belonging to the MPD set, we indicate the MDE(s) for which they are probabilistic. Those that are probabilistic with respect to a single MDE are the extremal states that, while in a superposition of $\{\omega(1), \omega(4)\}$, form an MPD set with one of its elements.

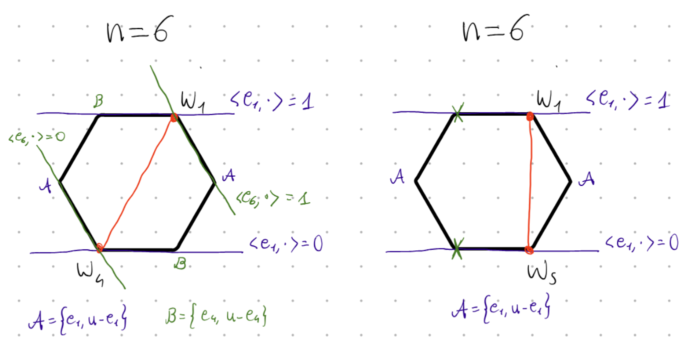

FIG. 4. [Sketches of the] regular hexagon GPT. We apply the same analysis as in Fig. 3 to the MPD sets $\{\omega(1), \omega(4)\}$ (left; antipodal vertices) and $\{\omega(1), \omega(6)\}$ (right; non-antipodal vertices). Left: similar to the pentagon case, the states that are probabilistic with respect to a single MDE are the extremal states that, while in a superposition of $\{\omega(1), \omega(4)\}$, form a MPD set with one of its elements. Right: the existence of two vertices that are not probabilistic with respect to A demonstrates that the hexagon fails to satisfy uniform superposition.

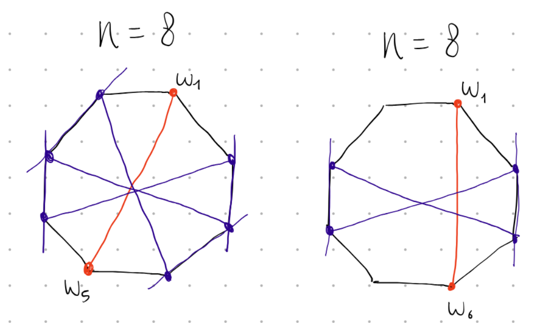

FIG. 5. [Sketches of the] regular octagon GPT: for each type of MPD set D (left: antipodal; right: non-antipodal), the states in a superposition of D (indicated by blue vertices) can be connected (see blue vertices) to form MPD pairs w.r.t. which the elements of D are in turn a superposition.

from Figs. 3-4. Regardless of the value of n and the kind of MPD set considered, one can always identify a state r that is a superposition of a pair $\{s_1, s_2\}$, yet simultaneously forms a valid MPD set with s_1 or with s_2. Such property hinges on the fact that a single MDE in a regular n-gon GPT perfectly distinguishes multiple MPD pairs—i.e. these models do not satisfy (OS) (see Sec. V), in contrast with qubit quantum theory, where (OS) holds and this phenomenon is absent.

b. Preparational Uncertainty For both even and odd n, the fact that every extremal state is deterministic with respect to multiple MDEs implies that none of the GPTs in this class satisfy uniform preparational uncertainty. Specifically, preparational uncertainty does not exist for $n = 5, 6$: for these regular polygons, the sets of deterministic states of any two MDEs always overlap (see Figure 2). As the number of sides increases, such pairs do exist (cf. the octagon in Figure 5); in particular, due to the n-fold rotational symmetry of the n-gons, each MDE will belong to at least a pair of uncertain MDEs, i.e. complete preparational uncertainty applies.

5. GPT-1 and GPT-2

a. Superposition We call GPT-1 the theory whose state space is a subset of the state space of a qutrit ($d = 3$) and given by the convex hull of the extremal qutrit states corresponding to the following vector-states,

$$S = \{|0\rangle, |\psi\rangle = a|1\rangle + b|2\rangle : \alpha, \beta \in \mathbb{C}, |a|^2 + |b|^2 = 1\} \subset \mathcal{H}_3 \,.$$

The effect space is generated by following set of rank-1 qutrit projectors

$$E = \{|0\rangle\langle 0|, |\psi\rangle\langle\psi| : |\psi\rangle \in S\}$$

along with the unit effect $u = \mathbb{I}_3$ and the zero operator. The MDEs in GPT-1 take the form

$$\mathsf{M}_\psi = \{|0\rangle\rangle, |\psi\rangle\rangle, |\psi^\perp\rangle\rangle\} \,.$$

Each $|\psi\rangle\rangle \neq |0\rangle\rangle$ is deterministic with respect to a single MDE, hence GPT-1 exhibits superposition as per Definition IV.2. However, by construction, the state $|0\rangle\rangle$ is deterministic with respect to all MDEs: it follows that complete superposition is not satisfied. At the same time, however, every state $|\phi\rangle \in S$ not belonging to a basis such as $\{|0\rangle, |\psi\rangle, |\psi^\perp\rangle\}$, is probabilistic with respect to the MDE associated with it, hence uniform superposition holds. Moreover, as GPT-1 inherits outcome sharpness of MDEs (OS) from qutrit quantum theory, $|\phi\rangle$ identifies a unique MDE, $\{|0\rangle\rangle, |\phi\rangle\rangle, |\phi^\perp\rangle\rangle\}$, with respect to which both $|\psi\rangle\rangle$ and $|\psi^\perp\rangle\rangle$ are probabilistic: that is, mutual superposition is satisfied.

We call GPT-1 the theory whose state space $\mathcal{S}$ is a subset of the state space of a qutrit ($d = 3$) and given by the convex hull of the extremal qutrit states corresponding to the following vector-states,

We call GPT-2 the theory whose state space is a subset of the state space of a qubit ($d = 2$) and given by the convex hull of theree extremal qubit states corresponding to the following vector-states

$$S = \{|0\rangle, |1\rangle, |+\rangle = (|1\rangle + b|2\rangle)/\sqrt{2}\} \subset \mathcal{H}_2$$

The extremal effects, as in GPT-1, are the corresponding rank-1 projectors, and those obtained by subtracting them to the unit effect $u = \mathbb{I}_2$. In GPT-2, there is a unique MPD set, $\mathsf{D} = \{|0\rangle\rangle, |1\rangle\rangle\}$, which is its own MDE, $\mathsf{M} = \mathsf{D}$. The only other extremal state, $|+\rangle\rangle$, is a superposition of D (i.e. uniform superposition applies). However, it does not belong to any MPD set (i.e. mutual superposition does not apply), which also means that neither element of D is a superposition (i.e. complete superposition does not apply).

b. Preparational Uncertainty In GPT-1, the extremal state $|0\rangle\rangle$ is deterministic with respect to all MDEs, hence the system does not exhibit preparational uncertainty. In GPT-2, there exist only one MPD, distinguished by a unique MDE, hence in this model there is also no preparational uncertainty.

Appendix B: Proofs of the Theorems and Lemmata

1. Uncertainty implies Superposition: Proof of Theorem 1

Theorem 1. *Let $(\mathcal{S}, \mathcal{E})$ be a state and effect space pair of a GPT. Then,*

(i) *if preparational uncertainty exists, superposition exists;*

(ii) *if* $(\mathcal{S}, \mathcal{E})$ *admits complete uncertainty, it also admits complete superposition;*

(iii) *if* $(\mathcal{S}, \mathcal{E})$ *admits uniform uncertainty and (OS), it also admits uniform superposition.*

Proof. (i). If the MDE pair $\mathsf{M}, \mathsf{N} \subseteq \mathrm{Ext}[\mathcal{E}]$ exhibit preparational uncertainty, then every $s \in \mathsf{Det}(\mathsf{N})$ is probabilistic with respect to M, hence they are superpositions of (some of) the elements of $\mathsf{Det}(\mathsf{M})$. Similarly, every $t \in \mathsf{Det}(\mathsf{M})$ is probabilistic with respect to N. Hence, $(\mathcal{S}, \mathcal{E})$ exhibits superposition.

(ii). Complete uncertainty implies that for each MDE M there exists some MDE N such that $\mathsf{Det}(\mathsf{M}) \cap \mathsf{Det}(\mathsf{N}) = \emptyset$. For every s that belongs to some MPD set D associated to an MDE M, there will exist an MDE N with respect to which it is probabilistic, hence by Definition IV.1 a superposition of (some of) the elements of $\mathsf{Det}(\mathsf{N})$. For every t that does *not* belong to any MPD set, it follows that t is probabilistic with respect *every* MDE, hence it is also a superposition.

(iii). From uniform preparational uncertainty, it follows that, for arbitrary M, every $s \in \mathrm{Ext}[\mathcal{S}] \setminus \mathsf{Det}(\mathsf{M})$ is a superposition of *every* MPD set $\mathsf{D} \subseteq \mathsf{Det}(\mathsf{M})$. If (OS) holds, then $\mathsf{D} = \mathsf{Det}(\mathsf{M})$, and the previous statement becomes equivalent to uniform superposition.

□

2. Restating Uncertainty in terms of Superposition: Proof of Lemma 1

Lemma 1. *Let* $(\mathcal{S}, \mathcal{E})$ *be a state and effect space pair admitting outcome sharpness of MDEs. Then the following conditions respectively imply existence, completeness and uniformity of preparational uncertainty:*

(i) *there exists MPD sets,* D, D'*, such that every* $s \in \mathsf{D}$ *is a superposition of* D'*;*

(ii) *for every MPD set* D*, there exists an MPD set* D' *such that all* $s \in \mathsf{D}$ *are superpositions of* D'*;*

(iii) *for any pair of MPD sets* D, D'*, all* $s \in \mathsf{D}$ *are superpositions of* D' *and vice versa.*

Proof. (i). If there exists MDEs M, N such that $\mathsf{Det}(\mathsf{M}) \cap \mathsf{Det}(\mathsf{N}) = \emptyset$, then there must exist an MPD set $\mathsf{D} \subseteq \mathsf{Det}(\mathsf{M})$ composed entirely of states in a superposition of an MPD set $\mathsf{D}' \subseteq \mathsf{Det}(\mathsf{N})$. To show the converse implication, consider (OS): we have that $\mathsf{D} = \mathsf{Det}(\mathsf{M})$ and $\mathsf{D}' = \mathsf{Det}(\mathsf{N})$. Therefore, requiring, say, D to be composed *entirely* of states in a superposition of D' implies that $\mathsf{Det}(\mathsf{M}) \cap \mathsf{Det}(\mathsf{N}) = \emptyset$, hence preparational uncertainty exists in the GPT.

(ii). Complete uncertainty implies that for all M there exists N such that $\mathsf{Det}(\mathsf{M}) \cap \mathsf{Det}(\mathsf{N}) = \emptyset$. Therefore, for all MPDs $\mathsf{D} \subseteq \mathsf{Det}(\mathsf{M})$ there exists some MPD set $\mathsf{D}' \subseteq \mathsf{Det}(\mathsf{N})$ such that each element of D' is probabilistic with respect to M, hence a superposition of the elements of D. To show the converse, we again assume (OS): the requirement that for every MPD set D there is some D' composed entirely of states in a superposition of D immediately implies that for all MDEs M there exists an MDE N such that $\mathsf{Det}(\mathsf{M}) \cap \mathsf{Det}(\mathsf{N}) = \emptyset$.

(iii). Assuming (OS), uniform uncertainty ensures that each maximal state belongs to *at most* one MPD set. Hence, for any pair of MDEs D, D', all elements of one are superpositions of all the elements of the other. To show the converse, suppose that $\mathsf{D} \cap \mathsf{D}' = \emptyset$ holds for all pairs of MPD sets $\mathsf{D} \neq \mathsf{D}'$. Due to (OS), we have that $\mathsf{D} = \mathsf{Det}(\mathsf{M})$ and $\mathsf{D}' = \mathsf{Det}(\mathsf{N})$ for some MDEs M, N. It follows that any pair of MDEs exhibit uniform uncertainty. □

3. The Quantum Tensor Product: Proof of Theorem 2

Theorem 2. *Let* $(\mathcal{S}_{\mathcal{Q}_1}, \mathcal{E}_{\mathcal{Q}_1})$ *and* $(\mathcal{S}_{\mathcal{Q}_2}, \mathcal{E}_{\mathcal{Q}_2})$ *be the state and effect spaces describing two quantum systems with operational dimensions* d_1 *and* d_2*, respectively. Then, the set of unit-trace positive semi-definite matrices* $\mathcal{S}_{\mathcal{Q}_{1,2}} \subset \mathbb{H}(\mathbb{C}_1^d \otimes \mathbb{C}_2^d)$ *is the largest state-space for the bipartite system which, with a maximal effect space, satisfies mutual superposition.*

Proof. Let s be an arbitrary extremal state in a superposition of some $\mathsf{D}' \subseteq \mathsf{D}$, with D an MPD set. Mutual superposition requires that all elements of D' are probabilistic with respect to an MDE M such that $s \in \mathsf{Det}(\mathsf{M})$. Therefore, mutual superposition implies that every extremal state is deterministic with respect to an MDE, hence it must belong to some MPD set.

For the finite-dimensional quantum state spaces $\mathcal{S}_{\mathcal{Q}_1}$ and $\mathcal{S}_{\mathcal{Q}_2}$, let $\boxtimes$ be any tensor product such that $\mathcal{S}_{\mathcal{Q}_{1,2}} \subset \mathcal{S}_{1,2} = \mathcal{S}_1 \boxtimes \mathcal{S}_2 \subseteq \mathcal{S}_1 \otimes_{\max} \mathcal{S}_2$. In other words, the states of the composite system $\mathcal{S}_{1,2}$ are represented by a convex and compact subset of the set of *positive on pure tensors* (POPT) operators[4] that is strictly larger than the set of

[4] The linear operator W_{12} on $\mathcal{H}_1 \otimes \mathcal{H}_2$ is POPT if $\mathrm{Tr}(W_{12}) = 1$ and if $\mathrm{Tr}(W_{12}(e_1 \otimes e_2)) \geqslant 0$ for all $e_1 \in \mathcal{E}_1$ and $e_2 \in \mathcal{E}_2$.

unit-trace positive semi-definite matrices. That is, they must include at least an *entanglement witness*, i.e. a POPT operator W_{12} such that $\mathrm{Tr}(W_{12}E) < 0$ for some entangled effect $E \in \mathcal{E}_{\mathcal{Q}_{1,2}}$ [41]. Without loss of generality, we can assume that E is a rank-1 projector, $E = |\Psi\rangle\rangle$, associated with some entangled vector $|\Psi\rangle$. E is not an element of the maximal effect space $\mathcal{E}_{1,2}$ associated with $\mathcal{S}_{1,2}$, as it assigns negative probability to the state $s \in \mathcal{S}_{1,2}$ represented by W_{12}. However, the quantum state s_E represented by the same projector $E = |\Psi\rangle\rangle$ corresponds to an extremal point[5] of $\mathcal{S}_{1,2} \supset \mathcal{S}_{\mathcal{Q}_{1,2}}$. Therefore, $\mathcal{S}_{1,2}$ includes an extremal quantum state s_E but $\mathcal{E}_{1,2}$ does not include the *unique* (due to the *sharpness* property of quantum theory) ray-extremal quantum effect (i.e. rank-1 projector) e_E, also represented by E, for which $e_E(s_E) = 1$. In other words, s_E is an extremal state of $\mathcal{S}_{1,2}$ that does not belong to an MPD set, as no rank-1 projector in $\mathcal{E}_{1,2} \subset \mathcal{E}_{\mathcal{Q}_{1,2}}$ features s_E among its deterministic states. We conclude that the quantum tensor product $\mathcal{S}_{\mathcal{Q}_{1,2}}$ produces the largest composite state space where every extremal state belongs to an MPD set. As $(\mathcal{S}_{\mathcal{Q}_{1,2}}, \mathcal{E}_{\mathcal{Q}_{1,2}})$ satisfies mutual superposition, it is also the "largest" tensor product compatible with mutual superposition. □

4. Superposition from Subsystems: Proof of Lemma 2

Lemma 2. *Let $(\mathcal{S}_1, \mathcal{E}_1)$ and $(\mathcal{S}_2, \mathcal{E}_2)$ be two GPT systems where $\mathcal{E}_i$ denotes the maximal effect space of $\mathcal{S}_i$. Let $\mathcal{S}_{1,2}$ be a no-sigalling composition of these state spaces, and let $\mathcal{E}_{1,2}$ be the associated maximal effect space, such that the operational dimension of the composite system $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ equals the product of the operational dimensions of the subsystems. Then,*

(i) *If $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies the existence of superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ also satisfies the existence of superposition;*

(ii) *If $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies complete superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ also satisfies complete superposition;*

(iii) *If $(\mathcal{S}_1, \mathcal{E}_1)$, $(\mathcal{S}_2, \mathcal{E}_2)$ and $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ satisfy (OS) and $(\mathcal{S}_1, \mathcal{E}_1)$ satisfies uniform superposition, then $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ satisfies uniform superposition.*

Proof. (i). If $(\mathcal{S}_1, \mathcal{E}_1)$ exhibits superposition, then there exists some $s \in \mathrm{Ext}[\mathcal{S}_1]$ and some MDE $\mathsf{M}_1 = \{e_i\}$ such that $s \notin \mathsf{Det}(\mathsf{M}_1)$. Let t be an arbitrary extremal element of $\mathcal{S}_2$ and $\mathsf{M}_2 = \{f_j\}$ be an arbitrary MDE of system 2, then $e_i \otimes f_j(s \otimes t) = e_i(s) f_j(t) < 1$ for all i and j, hence $s \otimes t \in \mathrm{Ext}[\mathcal{S}_{1,2}]$ is a superposition of (some of the elements of) any MPD set distinguished by $\mathsf{M}_1 \otimes \mathsf{M}_2$. Hence, superposition exists in the composite system $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$.

(ii). For any pair of MDEs M_1 and M_2 of systems 1 and 2, respectively, it follows that every composite state $z \in \mathrm{Ext}[\mathcal{S}_{1,2}] \backslash \mathsf{Det}(\mathsf{M}_1 \otimes \mathsf{M}_2)$ is a superposition. Instead, for any $z \in \mathsf{Det}(\mathsf{M}_1 \otimes \mathsf{M}_2)$, there exists some $e_{i*} \otimes f_{j*} \in \mathsf{M}_1 \otimes \mathsf{M}_2$ such that $(e_{i*} \otimes f_{j*})(z) = 1$. Then $e_i(z_1) = e_i \otimes u_2(z) = \sum_j (e_i \otimes f_j)(z) = \sum_j \delta_{ii*}\delta_{jj*} = \delta_{ii*}$. It follows that the marginal state z_1 is a (not necessarily extremal) state in $\mathcal{S}_1$ that is deterministic with respect to M_1; in particular, $z_1 = \sum_k p_k s_k^{i*}$ where $\{p_k\}$ is a probability distribution, $s_k^{i*} \in \mathrm{Ext}[\mathcal{S}_1]$ and $e_{i*}(s_k^{i*}) = 1$ for all k. Consider some $k*$ such that $p_{k*} \neq 0$. Due to uniform superposition for subsystem 1, it follows that for the state s_{k*}^{i*} in the decomposition of z_1 there exists some MDE N_1 such that $s_{k*}^{i*} \notin \mathsf{Det}(\mathsf{N}_1)$, i.e. $e'_l(s_{k*}^{i*}) < 1$ for all $e'_l \in \mathsf{N}_1$. But this implies that $e'_l(z_1) = \sum_k p_k e'_l(s_k^{i*}) < 1$ for all l, hence z_1 is not deterministic with respect to N_1; in particular, z_1 cannot be written as a mixture of extremal states that give probability 1 to *the same* element of N_1. Hence, for any choice of MDE N_2 of subsystem 2, we have that $z \notin \mathsf{Det}(\mathsf{N}_1 \otimes \mathsf{N}_2)$. We conclude that for every element of $\mathsf{Det}(\mathsf{M}_1 \otimes \mathsf{M}_2)$ there must exist another product MDE $\mathsf{N}_1 \otimes \mathsf{N}_2$ for which it is probabilistic. Hence, every extremal element of the composite state space $\mathcal{S}_{1,2}$ is a superposition.

(iii). This is an immediate consequence of the more general fact that, if (OS) holds in a GPT, then the existence of superposition alone implies uniform superposition. In fact, if each MDE M is associated with a unique MPD D_M, then every state $s \in \mathrm{Ext}[\mathcal{S}] \setminus \mathsf{D}_\mathsf{M}$ will be probabilistic with respect to M. Such extremal states exist if and only if superposition exists, i.e. when $\mathrm{Ext}[\mathcal{S}] \supset \mathsf{D}_\mathsf{M}$. In the context of the composite system $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$, uniform (hence existence of) superposition in $(\mathcal{S}_1, \mathcal{E}_1)$ implies existence of superposition in $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ (cf. point (i) above), and since $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ satisfies (OS) in the composite system, uniform superposition also applies.

□

[5] Extremal quantum states form a proper subset of the extremal elements of $\mathcal{S}_1 \otimes_{\max} \mathcal{S}_2$, hence they are also extremal points of the state space $\mathcal{S}_{1,2}$, where $\mathcal{S}_{\mathcal{Q}_{1,2}} \subset \mathcal{S}_{1,2} \subseteq \mathcal{S}_1 \otimes_{\max} \mathcal{S}_2$ [43].

5. Entanglement as a form of Superposition: Proof of Lemma 3

Lemma 3. *Let* $(\mathcal{S}_1, \mathcal{E}_1)$ *and* $(\mathcal{S}_2, \mathcal{E}_2)$ *be two GPT systems where* $\mathcal{E}_i$ *denotes the maximal effect space of* $\mathcal{S}_i$*. Let* $\mathcal{S}_{1,2} \supset \mathcal{S}_1 \underset{\min}{\otimes} \mathcal{S}_2$ *be a no-sigalling composition of these state spaces, and let* $\mathcal{E}_{1,2}$ *be the associated maximal effect space, such that the operational dimension of the composite system* $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ *equals the product of the operational dimensions of the subsystems. Then* $(\mathcal{S}_{1,2}, \mathcal{E}_{1,2})$ *exhibits superposition.*

Proof. Let $z \in \mathrm{Ext}[\mathcal{S}_{1,2}]$ be entangled, and suppose the composite system does *not* exhibit superposition. This means that z is deterministic with respect to all MDEs of the composite system. In particular, we must have that $z \in \mathsf{Det}(\mathsf{M}_1 \otimes \mathsf{M}_2)$ for all product MDEs $\mathsf{M}_1 \otimes \mathsf{M}_2$. We have $(e_i \otimes f_j)(z) = 1$ for some $e_i \in \mathsf{M}_1$ and $f_j \in \mathsf{M}_2$ (and 0 for all other effects of $\mathsf{M}_1 \otimes \mathsf{M}_2$), hence $\sum_j (e_i \otimes f_j)(z) = (e_i \otimes \sum_j f_j)(z) = (e_i \otimes u_2)(z) \equiv e_i(z_1) = 1$, where z_1 is the (mixed) reduced state of z for subsystem 1 and u_2 is the unit effect of system 2. This implies that z_1 can be written as a convex combination of maximal states giving probability 1 to the local effect e_i: $z_1 = \sum_k p_k s_k$, where $p_k \neq 0$, $s_k \in \mathrm{Ext}[\mathcal{S}_1]$ and $e_i(s_k) = 1$ for all k. But since z is deterministic with respect to all product MDEs, it follows that z_1 must be deterministic with respect to all MDEs of system 1. That is, each extremal state s_k (with the non-zero weight p_k) in the decomposition of z_1 must give probability one to *exactly the same effect* in each MDE in $\mathcal{E}_1$. In other words, the states s_k have the same (deterministic) probability distributions over all MDEs of system 1. However, since MDEs form an informationally-complete set of measurements, it must follow that the s_k states are *the same state*, $s = s_k$ for all k. Hence $z_1 = s$ is a extremal state, in contradiction with the assumption that z is entangled. We conclude that z cannot be deterministic with respect to all product MDEs, hence it is a superposition of product states. □

Appendix C: Non-multiplicative Operational Dimension: Minimal Tensor Product of Regular Pentagons

Consider the bipartite system described by $(\mathcal{S}_5 \underset{\min}{\otimes} \mathcal{S}_5, \mathcal{E}_5 \underset{\max}{\otimes} \mathcal{E}_5)$, where $(\mathcal{S}_5, \mathcal{E}_5)$ denotes the state and effect spaces of a regular pentagon, cf. Appendix A 4. Each subsystem has operational dimension of two (see Figs. 2-3). Consider the following set of five product states:

$$\mathsf{D} = \{(1,1), (2,3), (3,5), (4,2), (5,4)\}\,, \quad \text{where } (i,j) \coloneqq \omega_5(i) \otimes \omega_5(j)\,.$$

We define effects $E_k \in \mathcal{E}_5 \underset{\max}{\otimes} \mathcal{E}_5$ by the following rule:

$$E_k((i,j)) = \begin{cases} 1 & \text{if } j - i = k \,(\mathrm{mod}\, 5)\,, \\ 0 & \text{otherwise}\,. \end{cases}$$

For each $k = 1, \dots, 5$, $E_k(i,j) \in \{0,1\}$ for all $(i,j) \in \mathrm{Ext}(\mathcal{S}_5 \underset{\min}{\otimes} \mathcal{S}_5)$, hence $E_k \in \mathrm{Ext}[\mathcal{E}_5 \underset{\max}{\otimes} \mathcal{E}_5]$. Moreover, $\sum_{k=1}^5 E_k(s) = 1$ for all states $s \in \mathcal{S}_5 \underset{\min}{\otimes} \mathcal{S}_5$, hence $\sum_{k=1}^5 E_k = u_{5,5}$, $u_{5,5}$ being the unit effect. The extremal measurement $\mathsf{M} = \{E_1, E_2, E_3, E_4, E_5\}$ perfectly discriminates the elements of D. It follows that the operational dimension of $(\mathcal{S}_5 \underset{\min}{\otimes} \mathcal{S}_5, \mathcal{E}_5 \underset{\max}{\otimes} \mathcal{E}_5)$ is *at least*[6] 5, hence greater than the product of the operational dimensions of the two subsystems.

Appendix D: Connections to Related Work

In this section, we compare our operational definition of superposition (Definitions IV.1-IV.2) with previous approaches presented in [14, 15].

1. Superposition in D'Ariano *et al.*, 2020

In [14], D'Ariano et al. introduce a hierarchy of three superposition principles. Similar to our construction, they consider an MPD set $\mathsf{D} = \{s_i\}_{i=1}^d$ and some measurement $\mathsf{M} = \{e_i\}_{i=1}^d$ that distinguishes them (though in their case

[6] In fact, it is equal to 5, though we won't prove this here.

M is not necessarily an MDE). Let $\mathsf{p} = \{p_i\}_{i=1}^{d}$ denote a probability distribution and r an extremal state, a GPT admits: (i) *ultraweak superposition* iff for every choice of p, there exist M and r such that

$$p_i = e_i(r) \quad \text{for all } i = 1, \ldots, d \tag{D1}$$

holds; (ii) *weak superposition* iff for every choice of p and M, there exists r such that Eq. (D1) holds; (iii) *strong superposition* iff for every choice of p, there exists r such that, for every choice of M, Eq. (D1) holds.

Superposition is not introduced as a relational property between a state and a PD set; instead, it is characterised via global structural properties of the state and effect spaces. Because the proposed definitions rely on arbitrary choices for the probability vector p, they only apply to GPTs with infinitely many extremal states. Consequently, in their approach superposition remains undefined for theories such as Spekkens' model and regular n-gons.

Using their notation, $(\mathcal{S}, \mathcal{E})$ exhibits superposition according to our Definition IV.1 if there exists a *non-dispersion-free* distribution p, an *MDE* M, and an extremal state r such that Eq. (D1) holds. Therefore, our basic notion of superposition is implied by each of their definitions; nonetheless, our principles of complete, uniform, and mutual superposition (Definition IV.2) highlight separate structural features of the state space. Crucially, both approaches share the core premise that the operational hallmark of superposition—the operational phenomenon that defines it—is the observed probabilistic behaviour of extremal states relative to measurements that are maximally distinguishable.

The authors of [14] then show that classical theories (defined by a simplicial state-space whose vertices are perfectly distinguishable) do not exhibit any of their superposition principles. Simplicial theories that are non-classical (i.e. in which the perfect distinguishability of the vertices is violated, thereby preclusing state tomography) fail to satisfy weak (and strong) superposition. However, an explicit example of a non-classical simplicial theory satisfying ultraweak superposition remains unknown.

2. Superposition in Aubrun *et al.*, 2022

The authors of [15] discuss how an operational notion of superposition in GPTs can "*stem from the comparison between the two ensembles $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$. The ensembles, in fact, when mixed with equal weights, give rise to the same density matrix; further, identifying the state unambiguously via a measurement is possible or impossible depending on whether the ensemble is revealed before or after the measurement is carried out.*".

Based on this idea, they introduce the property of *strong incompatibility* as "*a way to identify the existence of superpositions in an arbitrary GPT*". A GPT satisfies strong incompatibility if there exist (unnormalised) states $0 \neq s_0, s_1, s_+, s_-$ and (unnormalised) effects e_0, e_1, e_+, e_- such that (i) $s_0 + s_1 = s_+ + s_-$; (ii) $e_o(s_1) = e_1(s_0) = e_+(s_-) = e_-(s_+) = 0$; and (iii) $e_i + e_j$ is strictly positive, for all $i \in \{0, 1\}$, $j \in \{+, -\}$. The connection to superposition arises from the parallel between the ensembles $\{s_0, s_1\}$ and $\{s_+, s_-\}$ and the ensembles $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, where the effects e_0, e_1, e_+, e_- correspond to the rank-1 projections generated by the respective vectors.

Strong incompatibility (which is assumed to be equivalent to the existence of superposition) is shown to boil down to the non-classicality of the theory, and thus with the possibility of combining two of more of such systems with a tensor product that allows entanglement. While this notion of superposition also applies to GPTs with finitely many extremal state, it is weaker than those provided in [14] and in Definition IV.1 of this paper. For instance, the gbit system exhibits this form of superposition; to show this, pick diagonal vertices for the pairs of s_i states and equal mixtures of extremal effects for the e_i effects. Furthermore, as in [14], superposition here is *not* treated as a relational property between states, as is conventional in quantum theory.

The version of superposition introduced in [15] is implied by our own. What is more, their notion proves to be equivalent to a weakened version of Definition IV.1, in which the state r is allowed to be mixed, provided it is not a mixture of the elements of D. This weaker notion of superposition is formalised as follows:

Definition D.1 (Weak Superposition)**.** *Let $(\mathcal{S}, \mathcal{E})$ be a pair of state and effect spaces. Let $\mathsf{D} \coloneqq \{s_i\}_{i=1}^{n}$ be an MPD set and $\mathsf{M} \coloneqq \{e_i\}_{i=1}^{n}$ an MDE measurement of D. Let $r \notin \mathrm{ConvHull}[\mathsf{D}]$ be a state with the property that for some $\mathsf{M}' \subseteq \mathsf{M}$, $e_j(r) \in (0, 1)$ holds for all $e_j \in \mathsf{M}'$, and $\sum_{e_j \in \mathsf{M}'} e_j(r) = 1$. Then, r is a superposition of the states in the set $\mathsf{D}' \coloneqq \{s \in \mathsf{D} | e_j(s) = 1, e_j \in \mathsf{M}'\}$.*

Theorem 3. *Let $(\mathcal{S}, \mathcal{E})$ be a state and effect pair of a GPT. Weak superposition (as per Definition D.1) exists if and only if $(\mathcal{S}, \mathcal{E})$ is non-classical.*

Proof. If $(\mathcal{S}, \mathcal{E})$ exhibits weak superposition as per Definition D.1, then $\mathcal{S} \setminus \mathrm{ConvHull}(\mathsf{D}) \neq \emptyset$, hence it is non-classical, since in a classical GPTs, $\mathcal{S} = \mathrm{ConvHull}(\mathsf{D})$, where D is the (unique) MPD set formed by the vertices.

For the converse, suppose $\mathcal{S}$ is not a simplex. Then, for any choice of MPD $\mathsf{D} = \{s_1, \ldots, s_n\}$, there is at least one $t \in \mathrm{Ext}[\mathsf{S}] \setminus \mathsf{D}$. Fix an MDE $\mathsf{M} = \{e_i\}_{i=1}^{n}$ that distinguishes D: if t is probabilistic with respect to M, then

$(\mathcal{S}, \mathcal{E})$ exhibits superposition as per Definition IV.1, hence also the weaker form of Definition D.1. If $t \in \mathsf{Det}(\mathsf{M})$, then $e_j(t) = 1$ for exactly one $j \in \{1, \ldots, n\}$, with $e_i(t) = 0$ for all $i \neq j$. Consider the mixture $\eta_p = pt + (1-p)s_i$, with $i \neq j$. We have that $\eta_0 \in \mathrm{ConvHull}(\mathsf{D})$, $\eta_1 \notin \mathrm{ConvHull}(\mathsf{D})$ and that $e_j(\eta_p) \in (0,1)$ as long as $p \neq 0, 1$. By the hyperplane separation theorem, there exists $p' \in (0,1)$ such that $\eta_p \in \mathrm{ConvHull}(\mathsf{D})$ for all $p \leqslant p'$ and $\eta_p \notin \mathrm{ConvHull}(\mathsf{D})$ for all $p > p'$. Hence there exist non-trivial mixtures of t and s_i which are probabilistic with respect to M and lie outside $\mathrm{ConvHull}(\mathsf{D})$. Hence $(\mathcal{S}, \mathcal{E})$ exhibits weak superposition. $\square$