# A histories perspective on characterizing quantum non-locality

# A histories perspective on characterizing quantum non-locality

**Fay Dowker**[1]**, Joe Henson**[1] **and Petros Wallden**[2,3]

[1] Imperial College, Blackett Laboratory, Prince Consort Road, London SW7 2BZ, UK
[2] Heriot-Watt University, Edinburgh EH14 1AS, UK
[3] University of Athens, Physics Department, Panepistimiopolis 157-71, Athens, Greece
E-mail: petros.wallden@gmail.com

## Abstract

We introduce a framework for studying non-locality and contextuality inspired by the path integral formulation of quantum theory. We prove that the existence of a strongly positive joint quantum measure—the quantum analogue of a joint probability measure—on a set of experimental probabilities implies the Navascues–Pironio–Acin (NPA) condition $Q^1$ and is implied by the stronger NPA condition $Q^{1+AB}$. A related condition is shown to be equivalent to $Q^{1+AB}$.

Keywords: non-locality, contextuality, NPA hierarchy, quantum measure, decoherence functional, path integrals, histories approach

## 1. Introduction

The phenomenon now commonly referred to as non-locality was, from the early days of quantum mechanics, central to debates on the adequacy of the standard formalism of quantum theory and other conceptual issues (see for example [1]). More recently, there has been interest in probing the exact quantitative extent of non-locality in quantum theory. For, while ordinary quantum models of experiments can violate Bell's local causality condition [2], they do not allow all sets of experimental probabilities or 'behaviours' [3–5] consistent with the weaker, operational condition, 'no superluminal signalling' [6]. A research program of 'characterizing quantum non-locality' has arisen with two closely related goals: to provide a method of

determining whether a given experimental behaviour could have been produced by an ordinary quantum model, and to discover physical or information-theoretic principles that result in constraints on possible behaviours.

Building on the pioneering work of Tsirelson, Navascues, Pironio and Acin (NPA) made substantial progress towards the first goal. NPA introduced a sequence of conditions of increasing strength, satisfaction of all of which implies the existence of an ordinary quantum model for a behaviour. Most importantly, each condition can be decided—given unlimited computing power—using standard programming techniques [7, 8]. As to the second goal, a number of proposals have been made for information-theoretic principles that result in bounds on behaviours, including 'non-triviality of information complexity' [9], 'no advantage for non-local computation' [10], 'information causality' [11], 'macroscopic locality' [12], and 'local orthogonality'/'consistent exclusivity' [13–15].

Progress on these issues could provide a more general understanding of the limitations on measures of success for quantum information processing tasks, along with new ways of bounding them. It could also throw light on the conceptual issues mentioned above. One view of the conflict between locality and quantum theory is that we should seek to uphold some essential feature of Bell's notion of local causality beyond mere no-signalling, while shedding some (at present unidentified) conceptual 'excess baggage'. Drawing an analogy with the apparent conflict between relativity of motion and constancy of the speed of light which resolved itself in the theory of relativity, the hope is that a significant gain in understanding might result from this rehabilitation of causal ideas in quantum theory: see, for example [16], for an approach to quantum theory influenced by such considerations. Developing an understanding of what restrictions on non-locality remain in quantum theory could aid this endeavour.

In the meantime, related questions have arisen more or less independently in work on the problem of quantum gravity in a path integral framework. A number of quantum gravity researchers have argued that, rather than trying to canonically quantize general relativity (GR), the best prospects for success lie in formulating quantum theory in a fundamentally relativistic way using the path integral (see e.g. [17–19]). To understand the reasoning behind this, recall that the Hilbert space formalism for quantum theory arises from canonical quantization of the Hamiltonian form of classical mechanics: in order to perform a canonical quantization of GR it must be formulated as a Hamiltonian theory based on a foliation of spacetime with spacelike hypersurfaces of constant time coordinate. General covariance implies that such a time coordinate and such hypersurfaces are not physical and the canonical quantization of this Hamiltonian form of GR leads to the infamous 'problem of time' [20, 21] or, more accurately, multiple *problems* of time [22]. These difficulties suggest that quantum gravity requires a formulation of quantum theory based fundamentally on physical, covariant, spacetime quantities. Early in the history of quantum mechanics (QM), Dirac [23] argued that the canonical theory, based as it is on a choice of time variable, is essentially non-relativistic and that the alternative is to base QM on the Lagrangian form of classical mechanics which 'can easily be expressed relativistically on account of the action functional being a relativistic invariant'. Dirac showed that this leads to the path integral or sum-over-histories which was brought to prominence by Feynman in his paper, 'A *spacetime* approach to non-relativistic quantum mechanics' (our emphasis) [24]. An approach to quantum theory based on the path integral and free from any fundamental reference to states or measurements was pioneered by

Caves, Hartle, Sorkin and others [25–29]. The many advantages for quantum gravity of a histories approach over a canonical approach have been set out by Sorkin [19].

In quantum gravity we must decide which principles from current theories to preserve and which to jettison in the search for the new theory and a principle of relativistic causality would seem to be an excellent candidate for keeping. However, the explicitly operational, no signalling condition is of no use in quantum theories of closed systems such as the whole universe as often considered in quantum gravity. A principle analogous to Bell's local causality, suitably generalized to apply in quantum theories of closed systems, which bans superluminal signalling as a special case and yet allows violations of the Bell inequalities, is what is needed [30]. Most concretely, in causal set quantum gravity Bell's local causality has been applied as a constraint on classical stochastic dynamical models of a growing discrete spacetime (causal set), with promising and suggestive results [31, 32]. If we can find a principle of *quantum Bell causality*, quantum dynamical models of discrete causal set spacetimes could then be constructed using a similar approach.

At this point, this important question in this approach to quantum gravity intersects with the programme of characterizing quantum non-locality in quantum information research in the following way. There is a trio of equivalent conditions on a behaviour in the Clauser–Horne–Shimony–Holt (CHSH) scenario: the existence of a Bell locally causal model (also known as a 'classical screening off model') of the probabilities, the satisfaction of all the CHSH (Bell) inequalities [33] and the existence of a joint probability measure (JPM) on all the possible (counterfactual) outcomes [34, 35]. This is also the case for more general experimental scenarios, with appropriate analogues, usually referred to simply as Bell inequalities, of the CHSH inequalities [36]. Thus, Bell's local causality, obedience of certain bounds on experimental probabilities and a non-contextuality condition are equivalent[4]. The path integral approach to QM reveals quantum theories to be generalized measure theories in which probability measures are generalized to *quantum measures* by weakening the Kolmogorov sum rule to an analogous rule that expresses the absence of genuine three-way interference between alternative histories for a system [29]. This immediately raises the question of what becomes of the 'Fine trio' of equivalent conditions—local causality (classical screening off), bounds on experimental probabilities and non-contextuality—when we enlarge the category of physical models we consider to ones in which dynamics is represented by quantum measures. A natural analogue of non-contextuality is the existence of a *joint quantum measure* (JQM) over all the outcomes [35]. The existence of such a measure in the CHSH scenario implies no restriction on the experimental probabilities, indeed a Popescu–Rohrlich (PR) box is possible [35, 37]. However, requiring the joint quantum measure to be *strongly positive* (SPJQM) implies the Tsirelson inequalities on the correlations [35, 37] and thus we have the beginnings of a quantum analogue of the Fine trio, with an implication of the form 'joint measure implies bounds on probabilities'. This raises the hope that this will be a guide to discovering the quantum Bell causality condition, the most physically important condition in the trio. In particular, quantum Bell causality should be a condition that implies the existence of a strongly positive joint

---

[4] This equivalence is the reason that any behaviour that violates a Bell inequality is commonly called 'non-local' and why phrases such as 'characterizing non-locality' have come to be used for seeking and studying conditions that limit Bell inequality-violating behaviours. We have adopted this terminology in this paper in order to make contact with much of the literature but we note here that saying that a behaviour is non-local is a loose short hand for the precise meaning that the experimental probabilities cannot arise from a Bell locally causal model.

quantum measure. One proposal for such a condition made in [35] has not, thus far, been successful in this regard.

In this paper we continue the investigation of the quantum Fine trio. We have not achieved the goal of discovering quantum Bell causality but report on progress in understanding the quantitative strength of the condition of the existence of a strongly positive joint quantum measure. We use the the machinery of the NPA hierarchy and show that the strongly positive joint quantum measure condition closely resembles conditions from the hierarchy. In particular this is facilitated by new results on the $Q^{1+AB}$ condition in [38]. This allows the set of behaviours satisfying the quantum measure condition to be positioned rather precisely within the hierarchy.

## 2. Joint measurement scenarios

We describe here the framework of *joint measurement scenarios*. It makes contact with the path integral/sum-over-histories approach to QM, is suitable for discussing and resolving questions of common interest between quantum information, foundations of QM and quantum gravity and is closely related to the formalism proposed by Liang, Spekkens and Wiseman (LSW) [39]. The framework can be conceived of purely operationally but the measurement outcomes which which it deals can also be thought of as physical events which are to be (eventually) included in the physics—quantum or otherwise.

In this conceptual framework, there is a set of devices which can perform measurements. To each measurement is associated an exhaustive set of possible, exclusive outcomes. Certain sets of measurements are compatible, that is they can be jointly performed. It may be useful to have in mind, for instance, spin measurements on a collection of atoms or polarization measurements on some photons. Let $\mathcal{M} := \left\{ M_1, M_2, \ldots M_p \right\}$ be the set of possible *basic measurements*. Let $\Xi_i$ be the set of labels for the possible outcomes for the measurement $M_i$. We will assume $\Xi_i$ has finite cardinality for all $i = 1, \ldots p$. We introduce the *non-contextuality space* (NC space), $\Xi$, an element of which specifies an outcome for each of the measurements $M_i$ even though they cannot, in general, be jointly performed:

$$\Xi := \Xi_1 \times \Xi_2 \times \ldots \times \Xi_p. \tag{1}$$

Each basic measurement, $M_i$, corresponds to a partition of $\Xi$. Each possible outcome of the measurement $M_i$ will be represented by one of the subsets in this partition, $X \in M_i$. Note that we refer to the measurement and the partition with the same symbol. For a basic measurement $M_i$, an outcome $X_i \in M_i$ is a set of the form

$$X_i = \Xi_1 \times \ldots \Xi_{i-1} \times \{a_i\} \times \Xi_{i+1} \cdots \times \Xi_p \subset \Xi, \tag{2}$$

where $a_i \in \Xi_i$. The set of all such subsets, ranging over all elements of $\Xi_i$, is the partition of $\Xi$ identified with basic measurement $M_i$.

A joint measurement is specified by a set of basic measurements $\left\{ M_{i_1}, M_{i_2}, \ldots M_{i_n} \right\}$ and the partition, $M$, corresponding to the joint measurement, is the set of all sets of the form

$$X_{i_1} \cap X_{i_2} \cap \ldots X_{i_n}, \tag{3}$$

where $X_{i_k} \in M_{i_k}$. Note that each outcome of one of the basic measurements $M_{i_k}$ is a union of outcomes for the joint measurement $M$. Also, for each joint outcome $X \in M$, and each outcome $X_{i_k}$ of basic measurement $M_{i_k}$, either $X \cap X_{i_k} = X$ or $X \cap X_{i_k} = \varnothing$.

A *joint measurement scenario*, based on the set of basic measurements $\mathcal{M}$ and NC space $\varXi$, is fixed by specifying a collection of joint measurements, to be thought of as those joint measurements which are physically possible to do. We will denote the set of all joint measurements in the scenario by $\tilde{\mathcal{M}}$. Each measurement in $\tilde{\mathcal{M}}$ corresponds to a subset of $\mathcal{M}$. All the basic measurements are elements of $\tilde{\mathcal{M}}$ and if there is a joint measurement in $\tilde{\mathcal{M}}$ of a set $K$ of basic measurements then there is a joint measurement in $\tilde{\mathcal{M}}$ of each subset of $K$. Endowing $\tilde{\mathcal{M}}$ with the corresponding order by inclusion and adding the empty measurement to $\tilde{\mathcal{M}}$ makes $\tilde{\mathcal{M}}$ a meet semilattice. A maximal joint measurement is a measurement that is maximal in this order and the basic measurements are minimal in the order if the empty measurement is removed. We say that measurement $M$ is included in measurement $M'$ if $M$ precedes $M'$ in the order. We say that a collection of measurements in $\tilde{\mathcal{M}}$ can be *jointly performed* if there is a joint measurement in $\tilde{\mathcal{M}}$ which includes them all.

Note that this formalism already includes some kind of non-contextuality assumption: an outcome, $X$, of a basic measurement, $M_i$ is dealt with by the formalism as the *same* event, no matter which joint measurement $M_i$ is included in. This would be physically most justifiable were the basic measurements all to have specified spacetime locations such that jointly performable basic measurements are in different spacetime locations. Then the non-contextuality is justified by a form of locality, namely the assumption that the basic measurement outcomes are *local events* whose description is in terms of local variables associated to the spacetime region in which the outcome occurs[5]. If the situation being modelled is supposed *not* to be of this form then the non-contextuality assumption inherent in the formalism has to be justified in some other way. We will return to these considerations later.

For each measurement $M \in \tilde{\mathcal{M}}$, there is a set of outcomes which are the elements of the partition $M$. We will refer to these as the fine-grained outcomes of $M$. We consider also all *coarse-grained outcomes i.e.* all unions of the fine-grained outcomes of $M$. For each $M \in \tilde{\mathcal{M}}$ the set of all coarse-grained outcomes together with the empty outcome form the (Boolean) sigma algebra, $\mathfrak{A}_M$, generated by the fine-grained outcomes. We will use the term sigma algebra even though the case in hand is finite because one might want to extend the formalism to infinite scenarios. The union of the algebras of outcomes over all measurements in a scenario will be denoted $\tilde{C} := \bigcup_{M \in \tilde{\mathcal{M}}} \mathfrak{A}_M$ and the union of the algebras over all basic measurements will be denoted $C := \bigcup_{i=1}^{p} \mathfrak{A}_{M_i}$.

Both $\tilde{C}$ and $C$ are subsets of the power set, $2^{\varXi}$ of $\varXi$ but neither are sigma algebras: in general they are not closed under union or intersection. While $\tilde{C}$ is a poset, ordered by inclusion, it is not a lattice. However, $\tilde{C}$ is an orthocomplemented poset, meaning for each

---

[5] Note that this assumption does not require that these local variables exhaust all there is to say in physics—there may also be non-local physical variables—all it requires is that there *are* local variables and local events and that the outcomes are local events [40].

outcome $X \in \tilde{C}$ there exists $\bar{X} \in \tilde{C}$ such that: (i) $\bar{\bar{X}} = X$; (ii) $X \subset Y$ implies $\bar{Y} \subset \bar{X}$; (iii) $X \subset Y$ and $\bar{X} \subset Y$ implies $Y = \Xi$.

The NC space $\Xi$ together with the set $\tilde{\mathcal{M}}$ of all (doable) measurements specifies the joint measurement scenario $\mathcal{S} = \{\Xi, \tilde{\mathcal{M}}\}$. This framework is neither the most restrictive nor the most general that one might want to consider[6]. For example, one can imagine putting further constraints on the set of measurements, $\tilde{\mathcal{M}}$, such as, if all pairs of measurements in a set of measurements $K \subset \mathcal{M}$ are jointly performable, then $K$ is jointly performable [14, 15, 41]. One could consider more general NC spaces and/or more general partitions of the NC space as measurements. For example, one could consider scenarios in which some measurement choices can depend on the outcomes of other measurements, something we will investigate in section 5.1.

An *experimental behaviour* is a joint measurement scenario $(\Xi, \tilde{\mathcal{M}})$ together with a probability measure $P(\cdot | M)$ on the sigma algebra $\mathfrak{A}_M$, for each $M \in \tilde{\mathcal{M}}$. The probabilities in an experimental behaviour are required to be consistent *i.e*, if an outcome $X \in \tilde{C}$ is an element of $\mathfrak{A}_M$ and of $\mathfrak{A}_{M'}$ then $P(X|M) = P(X|M')$. A behaviour is therefore equivalent to a consistent *probability function $P(\cdot)$* whose domain is the set of outcomes $\tilde{C}$. $P$ is not a probability *measure* because the set of outcomes $\tilde{C}$ is not a sigma algebra, however it is a measure when restricted to any algebra of outcomes $\mathfrak{A}_M$ for a measurement $M$. $P(\cdot)$ can be referred to as the (set of) *experimental probabilities*.

The consistency of the probability function is a further non-contextuality assumption beyond the one mentioned above. There it was stressed that the formalism treats the outcome, $X$, of a measurement, $M$, as the same event no matter what measurement, $M'$, $M$ may be included in. Now we are assuming that the probability of $X$ is the same no matter what joint measurement $M$ is included in. This would be most strongly, physically justified if the basic measurements have specific spacetime locations and those locations are such that any jointly performable basic measurements are spacelike separated from each other. Then the consistency of the probability function would be justified by relativistic causality: it is a no-signalling condition. Again, if the situation being modelled does not have this underlying spacetime structure, then the assumption of the consistency of the probability function would have to be argued for in some other way.

A key question in the analysis of contextuality and non-locality takes the form, in this framework, of whether a given behaviour $(\Xi, \tilde{\mathcal{M}}, P)$ can be consistently extended in some way to a larger subset of $2^\Xi$ than $\tilde{C}$.

---

[6]  In the Fritz–Leverrier–Sainz (FLS) framework [42], as in some earlier treatments [13, 43], the basic structure is a set of outcomes along with a set of subsets of the outcomes representing measurements. While the NC space could be defined as the set of 'deterministic probabilistic models' (as defined by FLS) for an FLS scenario, measurements and their outcomes cannot always be expressed in terms of this set as they are above (it can be empty for valid FLS scenarios). There are more other differences in approach, most importantly that the FLS framework is not defined in terms of basic and joint measurements. Instead, the no-signalling and the commutation requirements for quantum models are enforced in non-locality scenarios by adding extra measurements. While this gives a simple and powerful formalism, without adding more structure it is difficult to define *e.g.* the original NPA set $Q^1$, or to discuss different versions of our quantum measure-based conditions.

**Example.** *Bell scenarios*. To make contact with the familiar special case of non-locality setups we will need some additional concepts.

*Composition* of a set of scenarios $\mathcal{T} = \left\{ \mathcal{S}_1, \mathcal{S}_1, ..., \mathcal{S}_n \right\}$ into one $\mathcal{S}(\mathcal{T}) = \left\{ \Xi_{\mathcal{T}}, \tilde{\mathcal{M}}_{\mathcal{T}} \right\}$ can be defined if we are modelling a situation in which the choice of measurements on one system does not affect what can be done with the others: a particularly natural assumption if the systems are spacelike separated for the period of measurement. The NC space in this case is

$$\Xi_{\mathcal{T}} = \Xi_{\mathcal{S}_1} \times \Xi_{\mathcal{S}_2} \times ... \times \Xi_{\mathcal{S}_n} \tag{4}$$

where $\Xi_{\mathcal{S}_i}$ is the NC space for scenario $\mathcal{S}_i$. For each outcome $X$ of a measurement $M$ in a scenario $\mathcal{S}_i \in \mathcal{T}$ there is a set $X_{\mathcal{T}} \in \Xi_{\mathcal{T}}$ defined as

$$X_{\mathcal{T}} = \Xi_{\mathcal{S}_1} \times ... \times \Xi_{\mathcal{S}_{i-1}} \times X \times \Xi_{\mathcal{S}_{i+1}} \times ... \times \Xi_{\mathcal{S}_n}, \tag{5}$$

and the set

$$M_{\mathcal{T}} = \left\{ X_{\mathcal{T}} \right\}_{X \in M} \tag{6}$$

is a new measurement partition $M_{\mathcal{T}}$ on $\Xi_{\mathcal{T}}$. The set of measurements $\tilde{\mathcal{M}}_{\mathcal{T}}$ contains all of these measurements for all members of $\mathcal{T}$. Because the operations on one system do not affect the possibilities of measurement elsewhere, any set of these measurements all of which come from different scenarios in $\mathcal{T}$ can be jointly performed, and there are no further measurements in $\tilde{\mathcal{M}}_{\mathcal{T}}$.

*Bell scenarios* are a special type of joint measurement scenario suitable for the study of non-locality. First, consider a simple measurement scenario with $m$ mutually incompatible measurements, each with $d$ outcomes. Composing $n$ of these 'wings' as above gives a Bell scenario, specified by the triple $(n, m, d)$. Each 'local' outcome for each wing $X_{a_i x_i}$ is thus uniquely specified by two integer indices, $x_i$ for the measurement 'setting' choice and $a_i$ for the outcome, where $i$ labels the wing. Standard notation refers to such outcomes simply as $(a_i|x_i)$. Similarly, 'global' outcomes of a joint measurement on all systems are refereed to as $(a_1, a_2, ..., a_n \mid x_1, x_2, ..., x_n)$. The CHSH scenario, specified by $(2, 2, 2)$, is the simplest interesting Bell scenario. It is the most well-studied example, and also the one to which some previous results exclusively apply.

For Bell scenarios, consistency of the probabilities of outcomes over different measurements is equivalent to no signalling: the probabilities for outcomes of a measurement will be independent of any measurement choices made at spacelike separation.

## 3. Conditions on behaviours

### 3.1. The JPM condition, non-contextuality & Bell's local causality

A behaviour gives probabilities for the outcomes of doable experiments. Let us now consider a *joint probability measure* $P_J$ on the sigma algebra, $\mathfrak{A}$, generated by $\tilde{C}$ and define the set of behaviours consistent with the existence of such a measure. In our case, $\mathfrak{A} = 2^{\Xi}$ since $\Xi$ is finite.

**Definition 1.** (JPM) A behaviour $\left(\varXi, \tilde{\mathcal{M}}, P\right)$ is in the set *JPM* if there exists a joint probability measure $P_J$ on the sigma algebra $2^{\varXi}$ such that

$$P\left(X\right) = P_J\left(X\right), \quad \forall\, X \in \tilde{C}\,. \tag{7}$$

Since $2^{\varXi}$ is finite, $P_J$ is fixed by its values for the atoms of the algebra $2^{\varXi}$, *i.e.* the singleton sets corresponding to the elements of $\varXi$ itself: $P_J(\{\gamma\})$, $\gamma \in \varXi$. So, if a behaviour is in JPM then $P\left(X\right) = \sum_{\gamma \in X} P_J\left(\{\gamma\}\right)$. *JPM* is therefore a non-contextuality condition. Each measurement outcome could be said to be determined by some pre-existing physical property which is not affected by, but merely revealed by the particular measurement performed. The measure $P_J$ can be interpreted as a measure over these properties. In the language of stochastic processes, outcomes are just some of the *events*, subsets of the 'sample space' $\varXi$, *all* of which are now covered by one probability measure.

For Bell scenarios, the *JPM* condition is equivalent to the existence of a Bell locally causal model [2] of the probabilities. For the CHSH setup this was proved by Fine [34] (see [36] for a generalization to all Bell scenarios, and further references). The JPM condition for Bell scenarios is also equivalent to the satisfaction of a set of (scenario dependent) linear inequalities between experimental probabilities, generically known as 'Bell inequalities' [36]. The equivalence between non-contextuality, Bell's local causality and constraints on experimental behaviours inspires our proposals at the quantum level, where we seek analogues of these conditions.

### 3.2. Ordinary quantum behaviours

The set of behaviours $Q$ that can be derived from ordinary quantum models is of obvious interest. In the following, it will be useful to recall that the expression $X \in M$ identifies $X$ as a fine-grained outcome of the measurement $M$ ($X$ is a subset of $\varXi$ in the measurement partition $M$) while $X \in \mathfrak{A}_M$ refers to any outcome in the Boolean sigma algebra generated by the fine-grained outcomes.

**Definition 2.** (Q) A behaviour $\left(\varXi, \tilde{\mathcal{M}}, P\right)$ is in the set $Q$ if there exists an ordinary quantum model for the behaviour, that is, a Hilbert space $\mathcal{H}$, a pure state $|\psi\rangle \in \mathcal{H}$, and a projection operator $E^X$ on $\mathcal{H}$ for each fine-grained outcome, $X$, of each basic measurement such that

- for each basic measurement $M_i \in \mathcal{M}$,

$$\sum_{X \in M_i} E^X = \mathbf{1} \tag{8}$$

  where $\mathbf{1}$ is the identity on $\mathcal{H}$;
- if basic measurements $M_i, M_j \in \mathcal{M}$ can be jointly performed, then

$$\left[E^X, E^Y\right] = 0 \quad \forall\; X \in M_i, Y \in M_j; \tag{9}$$

- If outcome $X$ of a joint measurement of the set of basic measurements $\left\{ M_{i_1}, M_{i_2}, \ldots M_{i_n} \right\}$ is $X = \cap_{k=1}^{n} X_{i_k}$ where $X_{i_k} \in M_{i_k}$ then
$$\langle \psi | E^{X_{i_1}} E^{X_{i_2}} \ldots E^{X_{i_n}} | \psi \rangle = P(X). \tag{10}$$

It follows that, for all basic measurements $M_i \in \mathcal{M}$, the projectors defining an ordinary quantum model satisfy:
$$E^X E^Y = \delta_{XY} E^X \quad \forall \ X, Y \in M_i. \tag{11}$$

Since, from (8) we have that $\sum_{Y \in M} E^X E^Y E^X = E^X$ for all $X \in M$, from which it follows that $E^X E^Y E^X = 0$ for all $X \neq Y$, because all these terms are positive operators.

**Lemma 3.** *A behaviour $\left( \Xi, \tilde{\mathcal{M}}, P \right)$ is in the set $Q$ if and only if there exists a Hilbert space $\mathcal{H}$, a pure state $|\psi\rangle \in \mathcal{H}$, and a map $E \colon \tilde{C} \to L$, from the set of outcomes to the lattice, $L$, of projection operators on $\mathcal{H}$ (equivalently, the lattice of subspaces of $\mathcal{H}$) such that*

- *for each measurement $M \in \tilde{\mathcal{M}}$, the image of $E$ restricted to $\mathfrak{A}_M$, $E\left( \mathfrak{A}_M \right)$, is a Boolean algebra and the restriction of $E$,*
$$E \colon \ \mathfrak{A}_M \to E\left( \mathfrak{A}_M \right),$$
*is a Boolean algebra homomorphism;*
- *the probabilities are given by*
$$\langle \psi | E\left( X \right) | \psi \rangle = P(X), \quad \forall \ X \in \tilde{C} \ . \tag{12}$$

**Proof.** If the behaviour is in $Q$ then we have projection operators for each of the fine-grained outcomes, $X_{i_k}$, of each basic measurement so define $E\left( X_{i_k} \right) := E^{X_{i_k}}$. We then define a projector for a fine-grained outcome of a joint measurement by forming the product of the projectors for the relevant basic measurement outcomes: for outcome $X$ of a joint measurement of the set of basic measurements $\left\{ M_{i_1}, M_{i_2}, \ldots M_{i_n} \right\}$ is $X = \cap_{k=1}^{n} X_{i_k}$ where $X_{i_k} \in M_{i_k}$,
$$E\left( X \right) := E^{X_{i_1}} E^{X_{i_2}} \ldots E^{X_{i_n}} \ , \tag{13}$$

where the $E^{X_{i_k}}$ all commute and so this is indeed a projector. Finally, the projector, $E(Y)$ for a coarse-grained outcome, $Y \in \mathfrak{A}_M$, is defined to be the sum of the projectors for the fine-grained outcomes of $M$ of which $Y$ is the union. This defines a map $E \colon \tilde{C} \to L$ which has the properties above, by construction.

If the map $E$ exists with the properties above, the projectors $E(X)$ for the fine-grained outcomes of the basic measurements will satisfy the conditions in definition 2. Note here that we are using the definition of Boolean algebra homomorphism which includes the condition that the unit element is mapped to the unit element. $\qquad \square$

We will refer to such a Hilbert space and map $E$ as an *ordinary quantum model* for the behaviour. Note that, throughout this paper, we assiduously refer to such quantum models as

'ordinary' in order to distinguish them from other models that, from a histories perspective, are also classified as 'quantum', in particular the quantum measure theory models introduced below.

In Bell scenarios, the condition in definition 2, equation (9), translates to the requirement that the projectors for spacelike separated measurements commute. The definition of ordinary quantum behaviours here follows the definition given in [8] in which it is not required that the Hilbert space be a tensor product. See that reference for a discussion of the fact that it is not known in general whether requiring a tensor product Hilbert space is a strictly stronger condition ('Tsirelson's problem').

A number of questions arise regarding $Q$, including: what assumptions are sufficient to determine the shape of $Q$; can we derive the condition based on simple, more-or-less physical criteria rather than by invoking the above definition, which is considerably less conceptually simple than the definition of *JPM*; does the set of allowed behaviours remain the same when an alternative framework for QM is considered, such as the path integral?

### 3.3. NPA conditions

Tsirelson pioneered the study of bounds on $Q$, which can be thought of as defining sets of behaviours containing $Q$ [3–5]. More recently, NPA introduced a hierarchy of conditions resulting in a nested sequence of sets of behaviours for bipartite Bell scenarios which contain and converge to $Q$ [7, 8]. The conditions defining these sets can be cast as semidefinite programmes which facilitates solution by computer, and allows *e.g.* estimation of violations of Bell inequalities in QM, beyond the CHSH scenario. Rather than review the whole hierarchy in detail here, two conditions that will be of particular relevance will be defined. The definitions have been generalized to joint measurement scenarios but are equivalent to the original ones given by NPA when we restrict to bipartite Bell scenarios.

The set of behaviours $Q^1$ is the first in the original NPA hierarchy. The basic observation used to define this set is that, in ordinary quantum models, the inner products between all pairs of vectors from the set $\{|\psi\rangle\} \cup \{E^X|\psi\rangle;\ X$ an outcome of a basic measurement$\}$ must obey certain conditions, including relations to the experimental probabilities. If a model is in $Q$, there must therefore exist a Hilbert space and vectors in it satisfying these conditions. The set of behaviours obeying this condition is called $Q^1$. This set of behaviours has received substantial attention in the literature. In the CHSH scenario it is equivalent to a simple bound, and the maximum violation of several Bell inequalities for $Q^1$ behaviours are known [8]. It is also equivalent to the 'macroscopic locality' condition proposed by Navascues and Wunderlich [12].

**Definition 4.** ($Q^1$) A behaviour $(\Xi, \tilde{\mathcal{M}}, P)$ is in $Q^1$ if there exists a Hilbert space $\mathcal{H}$ and vectors $\{|X\rangle\}_{X \in C}$ indexed by the coarse-grained outcomes of the *basic* measurements, that span $\mathcal{H}$, and satisfy the following conditions:

(i) if $M_i$ is a basic measurement and $X, Y \in \mathfrak{A}_{M_i}$ are disjoint outcomes, $X \cap Y = \varnothing$, then

$$|X \cup Y\rangle = |X\rangle + |Y\rangle; \tag{14}$$

(ii) if basic measurements $M_i$, $M_j \in \mathcal{M}$ are jointly performable and $X \in \mathfrak{A}_{M_i}$ and $Y \in \mathfrak{A}_{M_j}$, then

$$\langle X | Y \rangle = P(X \cap Y). \tag{15}$$

It is clear from the definition that $Q \subseteq Q^1$ and it is known that $Q^1$ is actually strictly larger than $Q$ [8].

The definition of $Q^1$ is adapted directly from the NPA definition in terms of positive matrices, and is equivalent to it for bipartite Bell scenarios. The proof is in appendix A. For the discussion below, it will be easier to work directly with this definition expressed in terms of vectors. Note that the the condition that the vectors span $\mathcal{H}$ can be dropped without altering the set $Q^1$ since if such a set of vectors exists that doesn't span the Hilbert space, then $\mathcal{H}$ can be redefined to be the span of the vectors.

The basic measurements play a key role in the definition of $Q^1$ and an obvious strengthening of the condition is to require that vectors corresponding to all outcomes of all joint measurements exist satisfying the appropriate conditions. This leads to the set of behaviours called $Q^{1+AB}$ in the NPA approach[7], and is the first in the modified version of the hierarchy introduced later by FLS [42].

**Definition 5.** ($Q^{1+AB}$) A behaviour $(\Xi, \tilde{\mathcal{M}}, P)$ is in $Q^{1+AB}$ if there exists a Hilbert space $\mathcal{H}$ and a set of vectors $\left\{ |X\rangle \right\}_{X \in \tilde{C}}$ indexed by the outcomes of the joint measurements, that span $\mathcal{H}$ and satisfy the following conditions:

(i) if $M$ is a measurement and $X, Y \in \mathfrak{A}_M$ are disjoint outcomes, $X \cap Y = \varnothing$, then

$$|X \cup Y\rangle = |X\rangle + |Y\rangle; \tag{16}$$

(ii) for each measurement $M \in \tilde{\mathcal{M}}$ and all outcomes $X, Y \in \mathfrak{A}_M$

$$\langle X | Y \rangle = P(X \cap Y); \tag{17}$$

(iii) if $X, Y \in \tilde{C}$ and there exists a basic measurement $M_i$ and outcomes $X', Y' \in M_i$ such that $X' \cap Y' = \varnothing$, $X \subset X'$ and $Y \subset Y'$ then

$$\langle X | Y \rangle = 0. \tag{18}$$

Crucially, condition (iii) says that two vectors corresponding to outcomes $X$ and $Y$ in $\tilde{C}$ are orthogonal if they imply disjoint outcomes for any basic measurement, *even if the two outcomes X and Y are not compatible and are not both elements of any single measurement sigma algebra* $\mathfrak{A}_M$.

The set $Q^{1+AB}$ contains $Q$, and it follows easily from the definitions that $Q^1$ contains $Q^{1+AB}$. There is strong computational evidence that these are strict containments [8]. The condition

---

[7] In the special case of the (2,m,d) Bell scenarios, the only joint measurements are joint measurements of one measurement on A's wing and one on B's wing of the experiment, hence the $AB$ in $Q^{1+AB}$.

$Q^{1+AB}$ is well-studied and has a number of remarkable properties. For instance, it can be shown to be equivalent to a bound on the Lovász number of the 'orthogonality graph' associated to the behaviour [42, 43], an important quantity for classical information theory that was defined well before the advent of quantum information. Also, all of the currently proposed information-theoretic principles that restrict non-locality can be shown to be implied by the $Q^{1+AB}$ bound [38].

Although the higher levels of the NPA hierarchy will not enter into the following discussion, note that level $n$ is equivalent to the existence of vectors with the same inner product properties as those obtained from ordinary quantum models by applying all *products* of $n$ projectors to the state $|\psi\rangle$ (the projectors corresponding to basic measurements for the NPA version, and joint measurements for the FLS version of the hierarchy).

## 4. Conditions on behaviours from the quantum measure

### 4.1. Quantum measure theory

We briefly review the histories approach to quantum theory. For more details see [29, 44–46]. In this approach quantum theory is understood as a species of generalized measure theory, a generalization of classical stochastic processes such as Brownian motion. In histories-based theories—classical, quantum or transquantum—the fundamental kinematical concept is the set of spacetime histories, $\Omega$. For a particular physical system, each history in $\Omega$ is as complete a description of the system as is conceivable in the theory, over all time of interest. The kind of elements in $\Omega$ varies from theory to theory: in $n$-particle classical or QM, a history is a set of $n$ trajectories; in a classical or quantum scalar field theory, a history is a real or complex function on spacetime; in GR, a history is a Lorentzian geometry. Discovering the appropriate set of histories $\Omega$ for a particular system is part of the business of physics. In quantum theory, conceived of as *quantum measure theory*, $\Omega$ is the set over which the integration of the path integral takes place. Even for the non-relativistic particle, the path integral remains to be defined rigorously as a genuine integral over paths [47], something that will not concern us here as we restrict ourselves to finite systems.

Once the set of histories has been settled upon, any proposition about the system is represented by a subset of $\Omega$. For example in the case of the non-relativistic particle, if $R$ is a particular spacetime region, the proposition 'the particle traverses $R$' corresponds to the set of all spacetime trajectories which intersect $R$. We follow the standard terminology of stochastic processes and refer to such subsets of $\Omega$ as *events*.

In this framework, a physical theory based on $\Omega$ is expressed as a generalized measure theory, specified by a sigma algebra, $\mathfrak{A} \subseteq 2^{\Omega}$, of events and a measure, $\mu : \mathfrak{A} \to \mathbb{R}$, where $\mu$ is a non-negative function which encodes (a combination of) the dynamics and initial condition. Sorkin has identified a hierarchy of measure theories defined by a sequence of strictly weakening conditions on the measure, $\mu$, of which the first is the Kolmogorov sum rule for probability measures. Classical stochastic theories such as Brownian motion therefore lie in the first level of the Sorkin hierarchy and classical deterministic theories are special cases.

The second level in the Sorkin hierarchy is quantum measure theory. A *quantum measure* on an event algebra $\mathfrak{A}$ is a map $\mu : \mathfrak{A} \to \mathbb{R}$ such that:

    1. for all $\alpha \in \mathfrak{A}$, $\mu(\alpha) \geqslant 0$ (*Positivity*);

2. for all mutually disjoint $\alpha, \beta, \gamma \in \mathfrak{A}$,

$$\mu(\alpha \cup \beta \cup \gamma) - \mu(\alpha \cup \beta) - \mu(\beta \cup \gamma) - \mu(\gamma \cup \alpha) + \mu(\alpha) + \mu(\beta) + \mu(\gamma) = 0 \qquad (19)$$

(*Quantum Sum Rule*);

3. $\mu(\Omega) = 1$ (*Normalization*).

The quantum sum rule expresses the lack of interference between triples of histories, just as the classical Kolmogorov rule expresses the lack of interference between pairs of histories. The higher levels of the Sorkin hierarchy generalize this to lack of interference between $k$ histories. Each level of the hierarchy contains the previous levels, for example, a classical stochastic level 1 theory satisfies the level 2 condition.

There are, broadly, two interpretational frameworks for quantum measure theory in the literature. In both frameworks, the formalism of histories, events and the decoherence functional is the complete physics of the system: histories-based quantum theory is quantum theory of a closed system, a quantum theory without external observers, agents or measuring devices. Observers, agents and measuring devices, should they exist, are described within the theory. In *decoherent or consistent histories* [26, 48, 49] probabilities are considered to be fundamental and attention is focussed on subalgebras of $\mathfrak{A}$ on which the quantum measure restricts to a probability measure. The fact that there are (infinitely) many such subalgebras must then be grappled with [50, 51]. In the *co-event* framework [52, 53], attention is focussed on events of measure zero and an interpretation sought based on the maxim 'events of measure zero do not happen'.

In all known quantum theories, the quantum measure is not given directly but via the agency of a *decoherence functional* on the event algebra. A decoherence functional is a map $D: \mathfrak{A} \times \mathfrak{A} \to \mathbb{C}$ such that [18, 27]

1. For all $\alpha, \beta \in \mathfrak{A}$, we have $D(\alpha, \beta) = D(\beta, \alpha)^*$ (*Hermiticity*).

2. For all $\alpha, \beta, \gamma \in \mathfrak{A}$ with $\beta \cap \gamma = \varnothing$, we have $D(\alpha, \beta \cup \gamma) = D(\alpha, \beta) + D(\alpha, \gamma)$ (*Bi-additivity*).

3. $D(\Omega, \Omega) = 1$ (*Normalization*).

4. For all $\alpha \in \mathfrak{A}$, $D(\alpha, \alpha) \geqslant 0$ (Positivity).

The existence of a quantum measure is equivalent to the existence of a decoherence functional. If $D: \mathfrak{A} \times \mathfrak{A} \to \mathbb{C}$ is a decoherence functional then the map $\mu: \mathfrak{A} \to \mathbb{R}$ defined by $\mu(\alpha) := D(\alpha, \alpha)$ satisfies the conditions of being a quantum measure. And conversely, if $\mu$ is a quantum measure, there exists a (non-unique) decoherence functional $D$ such that $\mu(\alpha) = D(\alpha, \alpha)$ [29]. Given this, we will refer to a decoherence functional also as a quantum measure. We call a triple $(\Omega, \mathfrak{A}, D)$ a quantum measure system.

It is also the case that in all quantum theories describing known physics the decoherence functional satisfies the condition of strong positivity [54]: a decoherence functional is *strongly positive* (SP) if for any finite collection of events, $\alpha_1, \ldots \alpha_N$, $\alpha_i \in \mathfrak{A}$, the $N \times N$ matrix $D(\alpha_i, \alpha_j)$ is positive semidefinite. For the special case of finite event algebras (which is the case of relevance in this paper) a decoherence functional is strongly positive if the matrix of entries of $D$ indexed by the set of atoms of the algebra is a positive matrix.

**Lemma 6.** *The product of two independent finite, strongly positive quantum measure systems is a strongly positive quantum measure system.*

**Proof.** The product decoherence functional on the atoms of the tensor product algebra is given by the Hadamard product of the two decoherence functionals of the subsystems. The Hadamard product of two positive matrices is positive. ☐

Strong positivity thus implies composability of subsystems and is related to complete positivity of quantum operations in the context of evolving systems in quantum measure theory [54]. There is therefore a good reason to adopt strong positivity as a physical condition on quantum measures, and if one were convinced by this reason, one might want to redefine quantum measures to include this condition. We will not do so in this paper as we want to discuss the consequences of adopting or not adopting the condition.

### 4.2. A quantum analogue of non-contextuality

One response to the Bell theorem, is to retreat to talk of the experimental settings and outcomes only, and to investigate operational principles that constrain the experimental probabilities in order to try to understand the bounds on $Q$. Loosely speaking, this is to start from the most general set of all behaviours and ask why the quantum constraints are so strong [6]. We propose a complementary perspective, continuing to focus on the NC space $\Xi$ for a behaviour, but from a histories perspective within the framework of quantum measure theory. The approach then becomes one of generalizing the non-contextuality condition *JPM*, asking why the quantum constraints are so weak.

The *JPM* condition for a behaviour is that there exists a joint probability measure on the full sigma algebra $\mathfrak{A}$ generated by $\tilde{C}$. Probability measures are Sorkin level one measures and since quantum measures are Sorkin level two measures this suggests the following obvious quantum analogue of JPM:

**Definition 7.** (JQM) A behaviour $(\Xi, \tilde{M}, P)$ is in the set *JQM* if there exists a joint quantum measure, $\mu_J$ on sigma algebra $2^\Xi$ such that

$$\mu_J(X) = P(X), \quad \forall\ X \in \tilde{C} \quad \text{(Experimental probabilities)}. \tag{20}$$

This condition says that the marginal quantum measure—$\mu_J$ restricted to the subalgebra $\mathfrak{A}_M$— for each measurement $M$, coincides with the behaviour's probability measure $P(\cdot | M)$ on $\mathfrak{A}_M$. *JQM* is a quantum non-contextuality condition.

We can re-express the JQM condition in terms of decoherence functionals:

**Lemma 8.** *A behaviour $(\Xi, \tilde{M}, P)$ is in the set JQM if and only if there exists a joint decoherence functional, $D_J$ over $2^\Xi$ such that*

- $D_J(X, X) = P(X), \quad \forall\ X \in \tilde{C}$ *(Experimental probabilities);*
- $D_J(X, Y) = 0, \quad \forall\ X, Y \in \mathfrak{A}_M, \ X \cap Y = \varnothing, \quad \forall\ M \in \tilde{M}$ *(Decoherence of alternative experimental outcomes).*

**Proof.** Suppose a behaviour is in *JQM* so there exists a quantum measure $\mu_J: 2^\Xi \to \mathbb{R}$ such that $\mu_J(X) = P(X)$ for all $X \in \tilde{C}$. Then, there exists a real decoherence functional $D_J$ such that $D_J(X, X) = \mu_J(X)$ for all $X \in 2^\Xi$ [29]. Any decoherence functional satisfies

$$2 \operatorname{Re}\left[D(X, Y)\right] = D(X \cup Y, X \cup Y) - D(X, X) - D(Y, Y), \tag{21}$$

for any two disjoint events, $X$ and $Y$. For any measurement, $M$ and $X, Y \in \mathfrak{A}_M$, $X \neq Y$ we also have

$$D_J(X \cup Y, X \cup Y) = P(X \cup Y), \quad D_J(X, X) = P(X) \quad \text{and} \quad D_J(Y, Y) = P(Y). \tag{22}$$

Since $D$ is real, the decoherence condition $D_J(X, Y) = 0$ follows. Note that the two conditions together are equivalent to $D_J(X, Y) = P(X \cap Y)$, $\quad \forall \ X, Y \in \mathfrak{A}_M$, $\quad \forall \ M \in \tilde{\mathcal{M}}$.

The converse is immediate on defining $\mu_J(X) = D_J(X, X)$ for all $X \in 2^\Xi$. $\qquad \square$

An obvious strengthening of the *JQM* condition is:

**Definition 9.** (SPJQM) A behaviour $\left(\Xi, \tilde{\mathcal{M}}, P\right)$ is in the set *SPJQM* if there exists a joint decoherence functional, $D_J$ over $2^\Xi$ such that

- $D_J(X, Y) = P(X \cap Y)$, $\quad \forall \ X, Y \in \mathfrak{A}_M$, $\quad \forall \ M \in \tilde{\mathcal{M}}$;
- $D_J$ is strongly positive (SP).

Definitions 7 and 9 are generalizations of sets defined for the CHSH scenario in [35, 37].
Recall that positive semidefiniteness of a Hermitian matrix is equivalent to the matrix being the inner product matrix, the *Gram matrix*, of a set of vectors spanning a Hilbert space $\mathcal{H}$. Using this, we can present an equivalent definition of the *SPJQM* set which makes manifest its connection to the NPA heirarchy of conditions:

**Lemma 10.** *A behaviour* $\left(\Xi, \tilde{\mathcal{M}}, P\right)$ *is in SPJQM if and only if there exists a Hilbert space $\mathcal{H}$ spanned by a set of vectors indexed by elements of the sigma algebra $2^\Xi$, $\left\{\left|X\right\rangle\right\}_{X \in 2^\Xi}$, such that their inner products satisfy the following conditions:*

- *for any $X \in 2^\Xi$,*
$$\left|X\right\rangle = \sum_{\gamma \in X} \left|\{\gamma\}\right\rangle; \tag{23}$$

- *for each measurement $M \in \tilde{\mathcal{M}}$,*
$$\langle X|Y \rangle = P(X \cap Y) \quad \forall \ X, Y \in \mathfrak{A}_M. \tag{24}$$

*In particular, if two outcomes, $X$ and $Y$, are disjoint then $\langle X|Y \rangle = 0$: disjoint outcomes decohere.*

Although it is less simple to state, *SPJQM* is physically more interesting than *JQM*. A theory would be problematic if composition of independent behaviours allowed by the theory

yielded a behaviour not allowed by the theory (see e.g. [14, 15, 42] for applications of this). The *JQM* condition is not composable: there exist behaviours for the CHSH scenario that are in *JQM* but such that the composition of two independent copies is not [55]. The *SPJQM* condition on the other hand respects composition due to lemma 6. It is an open question whether the *SPJQM* condition follows from the *JQM* condition plus some simple principle expressing closure under composition of independent systems.

## 5. Comparing the conditions

**Lemma 11.** *Every behaviour for the CHSH (Bell (2,2,2)) scenario is in JQM.*

For more general scenarios this is not the case, for example the composition of two independent PR boxes is not in *JQM* [55]. The fact that two independent PR boxes when composed takes us outside the space of JQM is reminiscent to the situation for LO. This is not a coincidence, since one can show that JQM with some modification implies LO.

**Corollary 12.** *Q is a proper subset of JQM.*

**Lemma 13.** *If a behaviour for the CHSH scenario is in SPJQM, then the experimental probabilities satisfy the Tsirelson–Landau–Masanes bound [4, 56, 57].*

The proofs of these two lemmas are in [35, 37].

**Lemma 14.** $Q \subseteq SPJQM$

**Proof.** Suppose a behaviour is in $Q$. Then projection operators $\left\{ E^X \right\}$ indexed by the fine grained outcomes of basic measurements and satisfying the conditions of definition 2 exist. We define a vector $\left| \{ \gamma \} \right\rangle$ for each atom $\{ \gamma \}$ of $2^{\Xi}$:

$$\left| \{ \gamma \} \right\rangle := E^{X_1} ... E^{X_p} \left| \psi \right\rangle \tag{25}$$

where $\{ \gamma \} = X_1 \cap ... X_p$ and $X_i \in M_i$. We define vectors for each $X \in \tilde{C}$ by addition:

$$\left| X \right\rangle := \sum_{\gamma \in X} \left| \{ \gamma \} \right\rangle. \tag{26}$$

These vectors satisfy the conditions of lemma 10.                                       □

On comparison of the *SPJQM* condition with the NPA conditions, similarities immediately become apparent. Both require the existence of vectors with an inner product matrix that has a certain relationship to the experimental probabilities and a number of orthogonality relations. Much of the work has been done by formulating the conditions in this way. Based on this observation, the following relations can be discovered.

**Theorem 15.** $SPJQM \subseteq Q^1$.

**Proof.** Given a behaviour in *SPJQM*, there exists a set of vectors obeying the conditions (23)–(24). To show that there exists a set of vectors with the properties (14) and (15) from definition 4 of $Q^1$ is an easy task: the *SPJQM* vectors range over all subsets of $\Xi$, while the $Q^1$ vectors range over a subset of these, so we can define each $Q^1$ vector as equal to the corresponding *SPJQM* vector. This immediately gives (14) from (23) and (15) from (24). ☐

This theorem subsumes the earlier result given in lemma 13: the Tsirelson–Landau–Masanes bound is implied by $Q^1$ [8]. It should be noted that, although the *SPJQM* vectors are indexed by outcomes of all measurements in $\tilde{\mathcal{M}}$ and not just the basic ones, one cannot use the same reasoning as above to prove that $SPJQM \subseteq Q^{1+AB}$. This is because the relations (18) in the definition of $Q^{1+AB}$ are *not* implied by the conditions in the *SPJQM* definition. The significance of these extra orthogonality relations will be expanded on later.

**Theorem 16.** $Q^{1+AB} \subseteq SPJQM$.

Miguel Navascues found this result for bipartite non-locality scenarios (private communication) using a new reformulation of $Q^{1+AB}$ [38]. The proof presented here for all joint measurement scenarios is based on his idea.

**Proof.** Consider a behaviour in $Q^{1+AB}$. By definition 5 there exists of a set of vectors $\left\{|X\rangle\right\}_{X\in\tilde{C}}$ associated to the outcomes of coarse-grained measurements, obeying the properties (16)–(18). From these vectors we will construct a set of vectors $\left\{|X\rangle\right\}_{X\in 2^{\Xi}}$ indexed by all subsets of the NC space, that obeys (23) and contains the original set $\left\{|X\rangle\right\}_{X\in\tilde{C}}$ as its restriction to $\tilde{C}$. With this established, lemma 10 shows that the behaviour is in *SPJQM*.

We first construct vectors $|\{\gamma\}\rangle$ for all singleton sets of histories, $\gamma \in \Xi$. For each basic measurement $M_i \in \mathcal{M}$, we select, arbitrarily, one outcome $X^* \in M_i$. For every other outcome $X \in M_i$, $X \neq X^*$ we define the projection operator,

$$E^X = \text{proj}\left[\mathcal{H}^X\right], \tag{27}$$

where proj [·] projects onto the given subspace, and

$$\mathcal{H}^X = \text{span}\left\{|X'\rangle : X' \in \tilde{C} \text{ s.t. } X' \subset X\right\}. \tag{28}$$

That is, $\mathcal{H}^X$ is the subspace spanned by the vectors for all measurement outcomes that imply the outcome $X$ for basic measurement $M_i$. The orthogonality relation (18) implies that $\mathcal{H}^X$ and $\mathcal{H}^Y$ are orthogonal subspaces for $X \neq Y$ and so we have $E^X E^Y = \delta_{X,Y} E^X$ for all outcomes $X, Y \in M_i$, such that $X, Y \neq X^*$. We then define $E^{X^*}$, for the remaining outcome $X^*$ of $M_i$, by

$$\sum_{X \in M_i} E^X = \mathbf{1}, \tag{29}$$

where $\mathbf{1}$ is the identity operator on the Hilbert space $\mathcal{H}$. Since the other $E^X$'s are orthogonal projectors, so is $E^{X^*}$. $E^{X^*}$ projects onto the orthogonal complement, $\mathcal{H}^{X^*}$, of the space spanned by all the subspaces $H^X$ with $X \in M_i$ such that $X \neq X^*$. We have, therefore,

$$E^X E^Y = \delta_{XY} E^X, \quad \forall\, X \in M_i. \tag{30}$$

Thus, the set of operators $\left\{ E^X \right\}_{X \in M_i}$, indexed by the outcomes of $M_i$, forms a projective measurement. Note this does not prove that there is an ordinary quantum model, because these operators may not satisfy (9).

Using these projectors, we define the vector for an element $\gamma \in \Xi$ as

$$\left| \{\gamma\} \right\rangle := E^{X_1} E^{X_2} \dots E^{X_p} \left| \Xi \right\rangle, \tag{31}$$

where $X_i$ is the outcome for $M_i$ such that $\{\gamma\} = X_1 \cap X_2 \cap \dots X_p$.

We now show that the original $Q^{1+AB}$ vectors are sums of the vectors $| \{\gamma\} \rangle$. Let $M \in \tilde{\mathcal{M}}$ be a joint measurement of the set of $m$ basic measurements $\left\{ M_{i_1}, M_{i_2}, \dots M_{i_m} \right\}$. Let $X \in M$ be a fine-grained outcome of $M$, $X = X_{i_1} \cap X_{i_2} \cap \dots X_{i_m}$ where $X_{i_k} \in M_{i_k}$. By using (29) for all basic measurements *not* in $M$, we obtain,

$$\sum_{\gamma \in X} \left| \{\gamma\} \right\rangle = E^{X_{i_1}} E^{X_{i_2}} \dots E^{X_{i_m}} |\Xi\rangle \tag{32}$$

$$= E^{X_{i_1}} E^{X_{i_2}} \dots E^{X_{i_m}} \sum_{X' \in M} |X'\rangle, \tag{33}$$

where the second line is an application of (16). Note that $X$ is one of the fine-grained outcomes of $M$ being summed over here. Each $E^{X_{i_k}}$ either leaves $|X'\rangle$ invariant or annihilates it depending on whether or not $X' \subset X_{i_k}$. Of all the outcomes of $M$ only $X$ itself is left invariant by all the projectors and so

$$\sum_{\gamma \in X} \left| \gamma \right\rangle = \left| X \right\rangle, \quad \forall\, X \in M. \tag{34}$$

From (16) we can see that this holds for all coarse-grained outcomes,

$$\sum_{\gamma \in X} \left| \gamma \right\rangle = \left| X \right\rangle, \quad \forall\, X \in \mathfrak{A}_M, \tag{35}$$

and, since we chose an arbitrary measurement $M$ to consider, this holds for all outcomes $X \in \tilde{C}$. Therefore we have

$$\left| X \right\rangle = \sum_{\gamma \in X} \left| \{\gamma\} \right\rangle, \quad \forall\, X \in 2^\Xi, \tag{36}$$

where this serves as the *definition* of $|X\rangle$ for each $X \in 2^\Xi \backslash \tilde{C}$.

This set of vectors $\left\{ |X\rangle \right\}_{X \in 2^\Xi}$ satisfies (23) by construction. Because it restricts to the original $Q^{1+AB}$ vectors $\left\{ |X\rangle \right\}_{X \in \tilde{C}}$, condition (17) in the definition of $Q^{1+AB}$ ensures that it also obeys (24). Lemma 10 states that if there exists a set of vectors $\left\{ |X\rangle \right\}_{X \in 2^\Xi}$ satisfying these two conditions, the behaviour is in *SPJQM*. □

Note that the order in which the projectors are applied to the state $|\Xi\rangle$ in (31) was chosen in the proof to be the reverse of the order of the labelling of the basic measurements, $1 \dots p$, but any

other order would have worked also. Thus there is no unique set of vectors $\left\{\left|\{\gamma\}\right\rangle\right\}$ satisfying the conditions.

### 5.1. A complication: branching measurements

In the formalism above, we have not allowed measurement choices to depend on other measurement outcomes, though one can imagine setting up such a situation in a lab. In the CHSH scenario, an example would be using the outcome of whatever measurement is done in the A wing to determine the measurement setting in the B wing, if the former is done in the causal past of the latter. Such 'branching' measurements play an important role in the alternative FLS formalism for contextuality [42], where they are necessary for the definition of Bell scenarios, and so a comparison with results in this formalism requires us to consider them.

We will investigate the inclusion of these branching measurements and the altering of the non-contextuality condition to reflect the assumption that if measurements $M$ and $M'$ are jointly performable then they are jointly performable in any causal order—either may causally precede the other, thus allowing branching, or they may be performed at spacelike separation—without altering an underlying measure on the NC space. This is to allow the classical communication between wings necessary for the branching choices. Note that this assumption is incompatible with an assumption that the basic measurements have fixed locations in spacetime and jointly performable basic measurements are spacelike separated from each other. We argued above that such fixed spacetime locations for the basic measurements justifies our joint measurement scenario and behaviour framework. Including branching measurements means that the non-contextuality assumptions inherent in the joint measurement scenario and behaviour framework must be made without the argument from locality.

We will show below that the altered non-contextuality condition in the classical case is no stronger than JPM. However in the quantum case the altered non-contextuality condition makes a difference to our conclusions. To illustrate the reason for this, consider two outcomes in the CHSH scenario, $X = \left(a = +1, b = +1; x = 0, y = 0\right)$ and $Y = \left(a = -1, b = +1; x = 0, y = 1\right)$. These are disjoint in the sense that $X \cap Y = \varnothing$ but there is no measurement partition in $\tilde{\mathcal{M}}$ that contains them both. Thus, a joint decoherence functional, $D_J$, for a behaviour for this scenario, as defined in definition 8 or 9, need not satisfy $D_J\left(X, Y\right) = 0$: these outcomes need not decohere with respect to each other because they are not alternative outcomes of a possible measurement. If, however, we add to the collection of possible measurements a branching measurement with both $X$ and $Y$ as alternative outcomes—one in which the outcome, $a$, of the measurement $x = 0$ in the A wing is used to determine the $y$ setting in the B wing—then the corresponding quantum non-contextuality condition requires that the joint decoherence functional restricts to a probability measure on the sigma algebra of outcomes for this branching measurement. This implies $D_J\left(X, Y\right) = 0$, an additional condition on the joint quantum measure. All the extra conditions of this sort from branching measurements will allow us to prove a stronger result.

In general, given a joint measurement scenario $\left(\varXi, \tilde{\mathcal{M}}\right)$, the possible branching measurements form an extended set of measurement partitions. We will need to consider only branching measurements in which one decision is made; we call these single branching measurements. A single branching measurement consists of an initial (not necessarily basic)

measurement $M^1 \in \tilde{M}$ and, depending on its outcome $X^1$, a second measurement $M^2(X^1)$ from the set of all those measurements in $\tilde{M}$ that include $M^1$. This defines a new measurement partition of $\Xi$, $M_b(M^1, M^2(X^1))$. Explicitly, $M_b$ contains all $X^2 \in M^2(X^1)$ such that $X^2 \subset X^1$, for all $X^1 \in M^1$. We denote the set of all such single branching measurements by $\tilde{M}_b$. Note that $\tilde{M}_b$ is uniquely determined given $\tilde{M}$. Also $\tilde{M} \subseteq \tilde{M}_b$ since each measurement, $M \in \tilde{M}$, is (trivially) a single branching measurement where $M^1 = M$ and $M^2(X^1) = M$ for all $X^1$.[8]

For each $M_b \in \tilde{M}_b$, the sigma algebra $\mathfrak{A}_{M_b}$ is that generated by the fine-grained outcomes in the partition $M_b$. The set of all outcomes for all measurements in $\tilde{M}_b$ will be denoted $\tilde{C}_b := \bigcup_{M \in \tilde{M}_b} \mathfrak{A}_M$.

Only one consequence of these new definitions, expressed in the following lemma, will be salient to the discussion below.

**Lemma 17.** *If $X, Y \in \tilde{C}$ and there exists a basic measurement $M_i$ and outcomes $X', Y' \in M_i$ such that $X' \cap Y' = \varnothing$, $X \subset X'$ and $Y \subset Y'$ then there exists a single branching measurement $M^* \in \tilde{M}_b$ such that $X, Y \in \mathfrak{A}_{M^*}$.*

**Proof.** If the premise holds, there exist two measurements $M_X, M_Y \in \tilde{M}$ of which $X$ and $Y$ are, respectively, outcomes. Both $M_X$ and $M_Y$ must include basic measurement $M_i$. Consider the branching measurement $M^*$ in which $M^1 = M_i$ and $M^2(X') = M_X$, $M^2(Y') = M_Y$ and $M^2(Z) = M_i$ for $Z \in M_i$, $Z \neq X'$ and $Z \neq Y'$. Then $X, Y \in \mathfrak{A}_{M^*}$. □

We have extended the joint measurement scenario $(\Xi, \tilde{M})$ to $(\Xi, \tilde{M}_b)$. Given a behaviour $(\Xi, \tilde{M}, P)$ we also extend the probability function $P$ to cover the larger set of outcomes $\tilde{C}_b$. Note here that each *fine-grained* outcome for a branching measurement is a fine-grained outcome for some joint measurement in $\tilde{M}$, so the only elements in $\tilde{C}_b$ and not in $\tilde{C}$ are coarse-grained outcomes of the new measurements. This means that the probability function, $P$, on $\tilde{C}$ can be uniquely extended to a probability function on $\tilde{C}_b$, by applying the Kolmogorov sum rule. We call the behaviour thus extended, $(\Xi, \tilde{M}_b, P)$, the branching extension of $(\Xi, \tilde{M}, P)$.

Does including these types of measurements make any difference to the strength of the conditions discussed above? That is, if in the definitions of our sets *JPM*, *Q*, *JQM*, *etc.*, we were to replace $\tilde{M}$ with $\tilde{M}_b$ and $\tilde{C}$ with $\tilde{C}_b$, would the sets remain the same? For the set of classically non-contextual behaviours *JPM* given by definition 1, it is not hard to see that nothing changes: the alternative definition would be the same except with equation (7) replaced with

---

[8] One could also consider multiple branching measurements, the conceptually obvious (but notationally cumbersome) generalization of single branching measurements. However, this would have no effect of the argument below. With such an altered definition of $\tilde{M}_b$ lemma 17 would still be true, and so the altered definition of *SPJQMb* would still be equivalent to $Q^{1+AB}$. That is, once single-branching measurements have been added to set of possible measurements, adding multiple branching measurements does not alter the quantum non-contextuality condition.

$$P(X) = P_J(X), \quad \forall \, X \in \tilde{C}_b \,, \tag{37}$$

whereas before $P_J$ is a probability measure over the sigma algebra $2^\Xi$. Obviously this implies equation (7) because $\tilde{C} \subset \tilde{C}_b$. It is also implied by equation (7), because as just discussed, for sets $X \in \tilde{C}_b$ which are not in $\tilde{C}$, the value of $P(X)$ is determined by the probabilities of sets in $\tilde{C}$ by the Kolmogorov rule, as is the value of $P_J(X)$.

It is a straightforward exercise to show that the same is true for the sets $Q$ and $Q^1$. Here we spell out how this works for $Q^{1+AB}$.

**Lemma 18.** *A behaviour* $\left(\Xi, \tilde{\mathcal{M}}, P\right)$*, whose branching extension is* $\left(\Xi, \tilde{\mathcal{M}}_b, P\right)$*, is in* $Q^{1+AB}$ *if and only if there exists a Hilbert space* $\mathcal{H}$ *and a set of vectors* $\left\{|X\rangle\right\}_{X \in \tilde{C}_b}$ *indexed by all the coarse-grained outcomes of all branching measurements, that span* $\mathcal{H}$ *and satisfy the following conditions:*

*(i) if* $M_b \in \tilde{\mathcal{M}}_b$ *and* $X, Y \in \mathfrak{A}_{M_b}$ *are disjoint outcomes,* $X \cap Y = \varnothing$*, then*
$$\left|X \cup Y\right\rangle = \left|X\right\rangle + \left|Y\right\rangle; \tag{38}$$

*(ii) for each branching measurement* $M_b \in \tilde{\mathcal{M}}_b$ *and all outcomes* $X, Y \in \mathfrak{A}_{M_b}$
$$\langle X|Y \rangle = P\left(X \cap Y\right). \tag{39}$$

**Proof.** Suppose a behaviour is in $Q^{1+AB}$. Construct vectors $\left\{|X\rangle\right\}_{X \in 2^\Xi}$ as in the proof of theorem 16. These vectors satisfy (38) when $X \cap Y = \varnothing$. If $X$ and $Y$ are fine-grained outcomes of $M_b$ then either they disagree on the outcome of the first measurement $M^1$ in the branching measurement in which case their corresponding vectors are orthogonal, or they agree on that first outcome in which case they are outcomes of the same measurement $M^2 \in \tilde{\mathcal{M}}$. Therefore (39) holds for fine-grained outcomes of $M_b$. The vector for a coarse grained outcome is the sum of the vectors for the fine-grained outcomes of which it is the union and and so (39) also holds for coarse-grained outcomes of $M_b$.

Conversely, if vectors $\left\{|X\rangle\right\}_{X \in \tilde{C}_b}$ exist satisfying (38) and (39) then those for $X \in \tilde{C}$ satisfy conditions *(i)* and *(ii)* in definition 5. It remains to show that these vectors satisfy *(iii)* in definition 5. Suppose $X, Y \in \tilde{C}$ and there exists a basic measurement $M_i$ and outcomes $X', Y' \in M_i$ such that $X' \cap Y' = \varnothing$, $X \subset X'$ and $Y \subset Y'$. Then lemma 17 implies there is a branching measurement $M^* \in \tilde{\mathcal{M}}_b$ such that $X, Y \in \mathfrak{A}_{M^*}$. $X$ and $Y$ are disjoint and so by (39) we have
$$\langle X|Y \rangle = 0. \tag{40}$$
$\square$

Thus, the conditions *JPM*, $Q$, $Q^1$ and $Q^{1+AB}$, although originally expressed in terms of the joint measurements $\tilde{\mathcal{M}}$, also contain the information necessary to incorporate branching measurements in a non-contextual manner: once you have a *JPM*—or $Q$, $Q^1$ or $Q^{1+AB}$—model

for the non-branching behaviour there is a unique way to extend it to a model of the behaviour extended by adding all branching measurements.

In contrast, however, the corresponding alteration of the set *SPJQM* is not obviously the same.

**Definition 19.** (SPJQM$_b$) A behaviour $\left( \Xi, \tilde{M}, P \right)$, whose branching extension is $\left( \Xi, \tilde{M}_b, P \right)$, is in the set *SPJQM$_b$* if there exists a joint decoherence functional, $D_J$ over $2^\Xi$ such that

(i) $D_J(X, Y) = P(X \cap Y), \quad \forall \ X, Y \in \mathfrak{A}_{M_b}, \quad \forall \ M_b \in \tilde{M}_b;$

(ii) $D_J$ is strongly positive (SP).

*SPJQM$_b$* $\subseteq$ *SPJQM*, but as mentioned above there is reason to think that the two are not equivalent. There is a modified version of lemma 10.

**Lemma 20.** *A behaviour* $\left( \Xi, \tilde{M}, P \right)$*, whose branching extension is* $\left( \Xi, \tilde{M}_b, P \right)$*, is in SPJQM$_b$ if and only if there exists a Hilbert space* $\mathcal{H}$ *spanned by a set of vectors indexed by elements of the sigma algebra* $2^\Xi$*,* $\left\{ |X\rangle \right\}_{X \in 2^\Xi}$*, such that their inner products satisfy the following conditions:*

*(i) For any* $X \in 2^\Xi$*,*

$$|X\rangle = \sum_{\gamma \in X} |\{\gamma\}\rangle; \tag{41}$$

*(ii) for each measurement* $M_b \in \tilde{M}_b$*,*

$$\langle X|Y\rangle = P(X \cap Y) \quad \forall \ X, Y \in \mathfrak{A}_{M_b}. \tag{42}$$

*In particular, if two outcomes* $X$ *and* $Y$ *of a measurement are disjoint then* $\langle X|Y\rangle = 0$*: disjoint outcomes decohere.*

**Theorem 21.** *SPJQM* $_b = Q^{1+AB}$.

**Proof.** Lemmas 20 and 18 imply that $SPJQM_b \subseteq Q^{1+AB}$ and $Q^{1+AB} \subseteq SPJQM_b$ can be proved following the proof of theorem 16. $\qquad \square$

## 6. Discussion

We have found close relations between the NPA hierarchy of sets of experimental behaviours, and sets of behaviours of interest in quantum measure theory. In particular a quantum non-contextuality condition, *SPJQM*, was defined, subsuming the previous definition for the CHSH scenario [35]. This condition was found to be intermediate in strength between $Q^1$ and $Q^{1+AB}$. Thus, if $Q^{1+AB}$ is strictly larger than $Q$, as indicated by the computational evidence, then non-local correlations beyond those achievable in ordinary QM are achievable within strongly positive quantum measure theory. However, it is possible that additional constraints may arise within this framework as it is developed.

We also saw that an interesting modification of *SPJQM*, which we called *SPJQM$_b$*, is equal to the NPA set $Q^{1+AB}$. The necessary change is to require that for any joint measurement behaviour allowed by the physics, its branching extension is also physical. Including branching

measurements, where the outcome of one measurement determines which future measurement is performed, imposes more conditions because the quantum measure must restrict to a probability measure, and therefore decohere, on a larger collection of measurement subalgebras. The set $SPJQM_b$ is therefore 'closed under wiring' meaning that any behaviour created by applying 'classical operations'—coarse-graining measurement choices and/or outputs, post-selection, composition or 'branching' as described above—to a collection of behaviours in the set is also in the set [38]. It is not yet known whether $SPJQM_b$ is strictly contained in $SPJQM$ or whether they are equal. We will assume the former is the case, for the sake of the following discussion.

Returning to our main motivation, the use of quantum analogues of the Fine trio of conditions on a behaviour to discover quantum Bell causality, we see now that we have a choice for the 'joint measure' condition in the trio: $SPJQM$ or $SPJQM_b$. In contrast, as we saw, in the original Fine trio there is no choice to be made because the non-contextuality conditions are equivalent when the joint measure is a probability measure: $JPM \equiv JPM_b$. For the purpose of identifying a quantum Bell causality condition, the relevant scenarios are Bell scenarios where the jointly performable measurements are spacelike separated which therefore rules out the possibility of branching measurements and it would seem therefore that $SPJQM$ is the appropriate condition[9]. It would be prudent to bear $SPJQM_b$ in mind, however. We hope that the quantum non-contextuality condition will be a guide to discovering quantum Bell causality because the latter should imply the former. It is possible that if we find a scientifically fruitful condition of quantum Bell causality—perhaps by another route—it will turn out to imply the stronger condition and then we will have discovered that $SPJQM_b$ was the appropriate quantum non-contextuality condition after all.

Many other questions remain. The exact strength of the $SPJQM$ condition has not been determined: it is not known, for example, whether or not it is equivalent to $Q^1$. An exploration of the relation of the weaker, but simpler $JQM$ condition to other principles such as local orthogonality and a proof that if $JQM$ is strengthened to $JQM_b$ by including branching measurements it implies local orthogonality will appear elsewhere [55]. The assumption of strong positivity of the quantum measure was motivated by arguing that it guarantees composability of quantum measures on two independent systems. If it can be proved that strong positivity is necessary for composability then the assumption of strong positivity would have very strong physical motivation. The connection to local orthogonality/consistent exclusivity might lead to new insights here. What of the higher levels of the NPA hierarchy? As mentioned above, these conditions require the existence of a matrix with the same properties as the inner product matrix of vectors corresponding to strings of projection operators acting on the initial state in an ordinary quantum model. Strings of projection operators are known as *class operators* in the literature on decoherence functionals and are basic entities from which decoherence functionals are constructed for sequences of events ordered in time. For Bell scenarios therefore, it is possible that there is a connection between the higher level NPA

---

[9] This is the main reason we used a formalism based on the ideas of LSW [39] rather than FLS [42]. In the latter's framework, the natural definition of the existence of a strongly positive joint quantum measure would have automatically given $SPJQM_b$ for Bell scenarios, because branching measurements are used by FLS in their 'Randall–Foulis product'. There would have been no way to define $SPJQM$ in the FLS formalism.

conditions and decoherence functional based conditions arising from consideration of sequences of possible measurements.

## Acknowledgments

## Appendix A. Definitions of the NPA heirarchy

Here we demonstrate the equivalence between the original definition of the sets of behaviours $Q^1$ and $Q^{1+AB}$ given by NPA and the definitions in the main text.

Firstly we restate the original NPA form of the $Q^1$ condition adapting it to our joint measurement scenarios. For Bell scenarios, the first three conditions below are equivalent to equation (19) in [8] and the remaining two conditions to (20) in [8] for $n = 1$. NPA's 'null string' 1 is more naturally referred to as $\Xi$ in our terminology.

**Definition 22.** A behaviour $\left( \Xi, \tilde{\mathcal{M}}, P \right)$ is in $Q^1_{\mathrm{NPA}}$ if there exists a positive semi-definite matrix $\Gamma^1$ indexed by the set of all fine-grained outcomes $X \in M_i$ of all basic measurements $M_i \in \mathcal{M}$, and by $\Xi$, such that:

- $\Gamma^1_{\Xi,\Xi} = 1;$ (A.1)

- if $M_i$ is a basic measurement and $X \in M_i$ is a fine-grained outcome,
$$\Gamma^1_{\Xi,X} = P(X);$$ (A.2)

- if basic measurements $M_i, M_j \in \mathcal{M}$ are jointly performable and $M_i \neq M_j$, and $X \in M_i$ and $Y \in M_j$ are fine-grained outcomes, then
$$\Gamma^1_{X,Y} = P(X \cap Y);$$ (A.3)

- if $M_i$ is a basic measurement and $X \in M_i$ is a fine-grained outcome,
$$\Gamma^1_{\Xi,X} = \Gamma^1_{X,X};$$ (A.4)

- if $M_i$ is a basic measurement and $X, Y \in M_i$ are fine-grained outcomes,

$$\Gamma^1_{X,Y} = 0 \text{ if } X \neq Y. \tag{A.5}$$

The translation to the terminology of NPA is fairly direct. For example, for two fine-grained outcomes $X$ and $Y$ taken from jointly performable basic measurements, $P(X \cap Y)$ corresponds to $P(a, b)$, because the only jointly performable pairs of measurements in bipartite Bell scenarios are those taken from different wings. The following reformulation brings us closer to that used in the main text:

**Lemma 23.** *A behaviour* $\left(\Xi, \tilde{\mathcal{M}}, P\right)$ *is in* $Q^1_{\mathrm{NPA}}$ *if and only if there exists a positive semi-definite matrix* $\Gamma^1$ *indexed by the set of all fine-grained outcomes* $X \in M_i$ *of all basic measurements* $M_i \in \mathcal{M}$, *and by* $\Xi$, *such that:*

- *For all basic measurements* $M_i \in \mathcal{M}$ *and for all* $X \in M_i \cup \{\Xi\}$,

$$\Gamma^1_{X,\Xi} = \sum_{Y \in M_i} \Gamma^1_{X,Y}; \tag{A.6}$$

- *if basic measurements* $M_i, M_j \in \mathcal{M}$ *are jointly performable and* $X \in M_i$ *and* $Y \in M_j$ *are fine-grained outcomes, then*

$$\Gamma^1_{X,Y} = P(X \cap Y). \tag{A.7}$$

**Proof.** First we show that (A.1)–(A.5) imply (A.6, A.7). Equation (A.7) is the same as (A.3) in the case that $M_i \neq M_j$. When $M_i = M_j$ and $X \neq Y$ (A.7) follows from (A.5) (in this case $X \cap Y = \varnothing$), and for $X = Y$ (A.7) follows from (A.4) and (A.2). To obtain (A.7) when $X = \Xi$, we apply (A.1) and then (A.2):

$$\Gamma^1_{\Xi,\Xi} = 1 = \sum_{Y \in M} P(Y) = \sum_{Y \in M} \Gamma^1_{\Xi,Y} \tag{A.8}$$

for all basic measurements $M \in \mathcal{M}$. Otherwise $X$ is some outcome of the basic measurement $M_i$, and we apply (A.2) and then (A.7):

$$\Gamma^1_{X,\Xi} = P(X) = \sum_{Y \in M_i} P(X \cap Y) = \sum_{Y \in M_i} \Gamma^1_{X,Y}. \tag{A.9}$$

Now we show the converse, that (A.6, A.7) imply (A.1)–(A.5). Equation (A.7) implies (A.3) and (A.5). Then (A.5) and (A.6) together imply (A.4). Then (A.4) and (A.7) together imply (A.2). Finally (A.2) and (A.6) together imply (A.1). $\square$

It is an elementary result that a matrix is positive semidefinite (has no negative eigenvalues) if and only if it is the inner product matrix (or 'Gram matrix') of some set of vectors in a Hilbert space, and we rely on this to prove the following result.

**Lemma 24.** $Q^1 = Q^1_{\mathrm{NPA}}$.

**Proof.** Consider a behaviour satisfying the conditions of definition 4 and define $\Gamma^1_{X,Y} := \langle X|Y \rangle$ for $X$, $Y$ fine-grained outcomes of basic measurements or $\Xi$. Then $\Gamma^1$ is positive semidefinite, and equations (A.6, A.7) are implied by (14, 15) respectively.

Conversely, consider a behaviour in $Q^1_{\mathrm{NPA}}$. There exists a set of vectors indexed by the set of fine-grained basic measurement outcomes and $\Xi$, spanning a Hilbert space $\mathcal{H}$, such that $\Gamma^1_{X,Y} = \langle X|Y \rangle$. We can extend this set of vectors to one indexed by $\mathcal{C}$—that is, all coarse-grained outcomes of basic measurements—by summation according to (14). This is consistent with the property (A.6) of the original vectors. Condition (15) follows from (A.7), noting that if this equation holds for the fine-grained basic measurement outcomes then it will also hold for unions of them, the coarse-grained outcomes. $\square$

Very similar reasoning can be applied to the case of $Q^{1+AB}$. Similar reasoning can be applied to the case of $Q^{1+AB}$. There, the equation analogous to (A.6) is exactly the normalization condition of FLS [58] specialized to the first level of the hierarchy, while (A.7) implies their orthogonality condition (and also gives the connection to the experimental probabilities).

## References

[1] Bacciagaluppi G and Valentini A 2009 *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference* (Cambridge: Cambridge University Press)

[2] Bell J 1990 La nouvelle cuisine *Between Science and Technology* (Amsterdam: Elsevier) pp 216–33

[3] Cirel'son B 1980 Quantum generalisations of Bell's inequality *Lett. Math. Phys.* **4** 93–100

[4] Tsirel'son B 1987 Quantum analogues of the Bell inequalities: the case of two spatially separated domains *J. Sov. Math.* **36** 557–70 (Translated from a source in Russian of 1985)

[5] Tsirelson B 1993 Some results and problems on quantum Bell-type inequalities *Hadronic J. Suppl.* **8** 329–45

[6] Popescu S and Rohrlich D 1994 Nonlocality as an axiom *Found. Phys.* **24** 379

[7] Navascués M, Pironio S and Acín A 2007 Bounding the set of quantum correlations *Phys. Rev. Lett.* **98** 010401

[8] Navascués M, Pironio S and Acín A 2008 A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations *New J. Phys.* **10** 073013

[9] van Dam W 2013 Nonlocality and communication complexity *Nat. Comput.* **12** 9

[10] Linden N, Popescu S, Short A J and Winter A 2007 Quantum nonlocality and beyond: limits from nonlocal computation *Phys. Rev. Lett.* **99** 180502

[11] Pawłowski M, Paterek T, Kaszlikowski D, Scarani V, Winter A and Żukowski M 2009 Information causality as a physical principle *Nature* **461** 1101–4

[12] Navascues M and Wunderlich H 2009 A glance beyond the quantum model *Proc. R. Soc.* A **466** 881–90

[13] Foulis D J and R C H 1981 *Interpretations and Foundations of Quantum Theory* ed H Neumann (Mannheim: Wissenschaftsverlag Bibliographisches Institut) pp 216–33 (Empirical logic and tensor products)

[14] Fritz T, Sainz A B, Augusiak R, Brask J B, Chaves R, Leverrier A and Acín A 2013 Local orthogonality: a multipartite principle for correlations *Nat. Commun.* **4** 2263

[15] Cabello A 2013 Simple explanation of the quantum violation of a fundamental inequality *Phys. Rev. Lett.* **110** 060402

[16] Leifer M S and Spekkens R W 2013 Formulating quantum theory as a causally neutral theory of Bayesian inference *Phys. Rev.* A **88** 052130

[17] Hawking S W 1978 Quantum gravity and path integrals *Phys. Rev.* D **18** 1747–53

[18] Hartle J B 1991 *Quantum Cosmology and Baby Universes: Proceedings of the 1989 Jerusalem Winter School for Theoretical Physics* ed S Coleman, J B Hartle, T Piran and S Weinberg (Singapore: World Scientific) (The quantum mechanics of cosmology)

[19] Sorkin R D 1997 Forks in the road, on the way to quantum gravity *Int. J. Theor. Phys.* **36** 2759–81

[20] Isham C 1993 Canonical quantum gravity and the problem of time *Integrable Systems, Quantum Groups and Quantum Field Theories* ed L A Ibort and M A Rodriguez (Dordrecht: Kluwer) pp 157–288

[21] Kuchař K V 2011 Time and interpretations of quantum gravity *Int. J. Mod. Phys.* D **20** 3–86

[22] Sorkin R 2008 Solved and unsolved problems of time in quantum gravity *Talk given at the Clock and the Quantum* www.youtube.com/watch?v=Kp27zFc4Sz0

[23] Dirac P A M 1933 The Lagrangian in quantum mechanics *Physikalische Z. Sowjetunion* **3** 64–72

[24] Feynman R P 1948 Space-time approach to nonrelativistic quantum mechanics *Rev. Mod. Phys.* **20** 367

[25] Caves C M 1986 Quantum mechanics of measurements distributed in time. a path-integral formulation *Phys. Rev.* D **33** 1643–65

[26] Gell-Mann M and Hartle J B 1990 Quantum mechanics in the light of quantum cosmology *Complexity Entropy and the Physics of Information (SFI Studies in the Sciences of Complexity* vol VIII) ed W Zurek (Reading, MA: Addison-Wesley) pp 150–73

[27] Hartle J B 1995 Space-time quantum mechanics and the quantum mechanics of space-time *Gravitation and Quantizations: Proc. Les Houches Summer School 1992* ed B Julia and J Zinn-Justin (Amsterdam: North-Holland)

[28] Hartle J B 1993 The space time approach to quantum mechanics *Vistas Astron.* **37** 569

[29] Sorkin R D 1994 Quantum mechanics as quantum measure theory *Mod. Phys. Lett.* A **9** 3119–28

[30] Henson J 2005 Comparing causality principles *Stud. Hist. Philos. Mod. Phys.* **36** 519–43

[31] Rideout D P and Sorkin R D 2000 A classical sequential growth dynamics for causal sets *Phys. Rev.* D **61** 024002

[32] Sorkin R D 2000 Indications of causal set cosmology *Int. J. Theor. Phys.* **39** 1731–6

[33] Clauser J, Horne M, Shimony A and Holt R 1969 Proposed experiment to test local hidden-variable theories *Phys. Rev. Lett.* **26** 880

[34] Fine A 1982 Hidden variables, joint probability and the Bell inequalities *Phys. Rev. Lett.* **48** 291–5

[35] Craig D A, Dowker H F, Henson J, S M, Rideout D and Sorkin R D 2007 A Bell inequality analog in quantum measure theory *J. Phys.* A **40** 501–23

[36] Werner R F and Wolf M M 2001 Bell inequalities and entanglement *Quantum Inf. Comput.* **1** 1–25

[37] Barnett M, Dowker F and Rideout D 2007 Popescu Rohrlich boxes in quantum measure theory *J. Phys. A: Math. Gen.* **40** 7255–64

[38] Navascués M, Guryanova Y, Hoban M and Acín A 2014 Almost quantum in preparation

[39] Liang Y, Spekkens R W and Wiseman H M 2011 Specker's parable of the overprotective seer: a road to contextuality *Phys. Rep.* **506** 1–39

[40] Henson J 2013 Non-separability does not relieve the problem of Bell's theorem *Found. Phys.* **43** 1008–38

[41] Foulis D and Randall C 1974 Empirical logic and quantum mechanics *Synthese* **29** 81–111

[42] Fritz T, Leverrier A and Belén Sainz A 2012 A combinatorial approach to nonlocality and contextuality arXiv:1212.4084

[43] Cabello A, Severini S and Winter A contextuality of physical theories as an axiom arXiv:1010.2163

[44] Sorkin R D 1997 Quantum measure theory and its interpretation *Quantum Classical Correspondence Proc. 4th Drexel Symp. on Quantum Nonintegrability (8–11 September 1994)* ed D Feng and B-L Hu (Cambridge, MA: International Press) pp 229–51

[45] Dowker F, Johnston S and Sorkin R D 2010 Hilbert spaces from path integrals *J. Phys.* A **A43** 275302

[46] Sorkin R D Notes on a histories formulation of quantum mechanics www.perimeterinstitute.ca/personal/rsorkin/lecture.notes/path.integral.ps

[47] Sorkin R D 2012 Toward a fundamental theorem of quantal measure theorm *Math. Struct. Comput. Sci.* **22** 816

[48] Griffiths R B 1984 Consistent histories and the interpretation of quantum mechanics *J. Stat. Phys.* **36** 219–72

[49] Omnès R 1988 Logical reformulation of quantum mechanics. 1. Foundations *J. Stat. Phys.* **53** 893–932

[50] Dowker F and Kent A 1996 On the consistent histories approach to quantum mechanics *J. Stat. Phys.* **82** 1575–646

[51] Dowker F and Kent A 1995 Properties of consistent histories *Phys. Rev. Lett.* **75** 3038–41

[52] Sorkin R D 2007 Quantum dynamics without the wave function *J. Phys.* **A40** 3207–22

[53] Sorkin R D 2007 An exercise in 'anhomomorphic logic' *J. Phys.: Conf. Ser.* **67** 012018

[54] Martin X, O'Connor D and Sorkin R D 2005 The random walk in generalized quantum theory *Phys. Rev.* **D71** 024029

[55] Henson J in preparation

[56] Landau L 1988 Empirical two-point correlation functions *Found. Phys.* **18** 449–60

[57] Masanes L 2003 Necessary and sufficient condition for quantum-generated correlations arXiv:quant-ph/0309137

[58] Fritz T, Leverrier A and Belén Sainz A 2013 Probabilistic models on contextuality scenarios arXiv:1307.0145