Imperial College London

MSc Quantum Fields and Fundamental Forces

Author:

Paul Boës

Closure and Stability in Quantum Measure Theory

Supervisor:
Prof. Fay Dowker

Co-Supervisors:
Dr. Miguel Navascués
Dr. Joe Henson

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Science of Imperial College London

September 22, 2014

Abstract. This thesis has two aims: Firstly, it attempts to answer the following question: What are the implications of the requirement that sets of behaviors allowed by quantum measure theory should be closed under general physical operations? Secondly, it reviews the existing literature and results relevant to this question. Following a brief and general introduction of the background and motivation for such a project in Section 1, Sections 2- 5 will introduce the reader to the concepts and formalism necessary to understand the thesis question. Section 6 is an attempt to answer it in the light of the current state of knowledge. This section also contains most of the original material. In particular, it discusses a proof that $SPJQM/SPJQM_b$ are $\otimes$-maximal under PPI and possible strengthenings thereof. Furthermore the axioms underlying the above notions are discussed and related to the path integral interpretation behind quantum measure theory. Section 7 discusses these results and their relevance to answering the thesis question. Finally, Section 7.2 summarizes the findings of this work and suggests a number of directions for future research implied by them.

Contents

1. Introduction and Motivation

The interpretation of quantum theory remains one of the biggest topics in foundational research. While remarkable progress has been made in formally developing the quantum mechanics of Heisenberg, Dirac and Schrdinger in the 1920's and 1930's into the impressively corroborated standard model and beyond, no attempt to answer John Wheeler's question "Why the quantum?" could, as of yet, establish itself canonically. While answering this question may be of interest for its own sake, it will also almost certainly be necessary to develop a satisfactory theory of a unified quantum gravity. This is because the two basic building blocks of such a unification, quantum theory and general relativity, seem utterly incompatible in a number of ways.

Bell inequalities and non-locality. One of these concerns the propagation of effects over regions of spacetime: While the principle that no effect propagates faster than the speed of light is the resting stone of relativity theory, quantum mechanics allows for states in which a "spooky action at the distance", as Einstein has called it, is possible. These particular tensions between relativity and quantum theory and the extent to which they involve incompatible accounts of how effects are physically allowed to propagate between system in some spatiotemporal relationship can conveniently be studied via the concept of correlations: Consider two conditional probability distributions $P(a|x)$ and $P(b|y)$ for two systems A and B. It For the moment it does not matter whether one thinks of the sets $\{a\}, \{b\}$ and $\{x\}, \{y\}$ as possible states a, b of A and B upon the occurrence of interaction events x, y or, operationally, as being an experimenter's meter readings a, b upon setting a measurement device to x, y. We then define the correlator for a given x, y and finite sets $\{a\}, \{b\}$ as

$$C_{xy} := \sum_{a,b} a \cdot b \cdot P(a, b|x, y). \tag{1}$$

where $P(a, b|x, y)$ is the joint probability distribution for the two distributions. Correlators measure the strength of statistical dependence of the marginal distributions and if they are uncorrelated, i.e. the sum of correlators is zero, then $P(a, b|x, y) = P(a|x) \cdot P(b|y)$, i.e. the joint probability can be split into two entire independent distributions. For the case of the standard EPRB-experimental setup, in which two spacelike separated parties Alice and Bob share a bipartite system prepared in some state $|\psi\rangle$ which they can measure in two settings each to obtain possible outcomes $a, b = \pm 1$, John S. Bell famously showed that the correlators of every local, hidden variable theory have to satisfy certain simple inequalities, known as Bell inequalities, namely [1]

$$|C_{xy} + C_{x'y} + C_{xy'} - C_{x'y'}| \equiv S \leq 2, \tag{2}$$

where S is called the CHSH-value and three similar inequalities are obtained by circulating the LHS.

Apart from very weak ones like the assumption of Alice and Bob's free will to pick settings, these inequalities can be derived based only on the assumption that

$$P(a,b|x,y,\lambda) = P(a|x,\lambda)P(b|y,\lambda), \tag{3}$$

where λ is a hidden variable that represents a complete specification of the state $|\psi\rangle$ beyond the possibly incomplete quantum state description. (3), known as Bell locality, states that the joint probability distribution for the EPRB-setup is given by the product of the marginal distributions if one has full access to the complete state λ of the system being prepared. Different ways of breaking this condition into further assumptions are possible and have been discussed since Bell's first publication of it: These include parameter and outcome independence [2] or as a "screening off" condition resulting from the combination of a principle of Relativistic Causal Structure and a form of Reichenbach's Principle of Common Cause which states that correlations require causal explanation [3]. In this case λ is identified with a full description of the causal past shared by the margins, i.e. the intersection of their past light cones, see [4] for a discussion. What all discussions share in common is to identiy (3) as expressing commitment to a combination of locality and realism.

The paradigmatic example of how quantum mechanics allows for the violation of Bell locality in the EPRB-setup is then given by the measurement of the axial spin of two spin-$\frac{1}{2}$ systems and is quantum mechanically represented by the four projective sets formed by elements

$$\begin{aligned} E_a^x &= \frac{1}{2}(1 + a\alpha_x \cdot \sigma) \\ E_b^y &= \frac{1}{2}(1 + b\beta_y \cdot \rho), \end{aligned} \tag{4}$$

where $x, y \in \{0,1\}, a, b \in \{-1,1\}$, α_x, β_y are 3-vectors that give the direction of measurement for x, y and σ, ρ are the Pauli matrices. By preparing the systems in the singlet state $|\psi_-\rangle$ and choosing directions such that

$$\alpha_0 \cdot \alpha_1 = \beta_0 \cdot \beta_1 = 0$$

$$\alpha_0 \cdot \beta_0 = \alpha_0 \cdot \beta_1 = \alpha_1 \cdot \beta_0 = -\alpha_1 \cdot \beta_1 = \frac{1}{\sqrt{2}},$$

the resulting correlators for this measurement violate the Bell inequalities.

Bell's derivation articulates very clearly and generally the classical assumptions that are disobeyed by quantum mechanics, and that were discussed already by Einstein, Podolsky and Rosen in their 1935-paper [5]. Since experimental results, first by Aspect [6] but also more recently, e.g. [7], confirm the quantum mechanical predictions (disregarding the notorious loophole debate), this shows nature violates either realism or locality, or a peculiar mixture thereof, as they appear in (3).

Bell's result initiated research on the strength of correlations possible under the assumption of physical principles or other formal conditions for general setups in which several parties are space-like separated. This field of research is known as "non-locality"-research. A number of important results here were the generalization of the Bell inequalities for more general scenarios, with the ones discussed here corresponding to CHSH-inequalities [8] and Tsirelson's exact result for the maximal violations of quantum theory for CHSH-inequalities known as the Tsirelson's bound which is given by $S \equiv 2\sqrt{2}$ and produced also by the procedure above [9]. Finally,

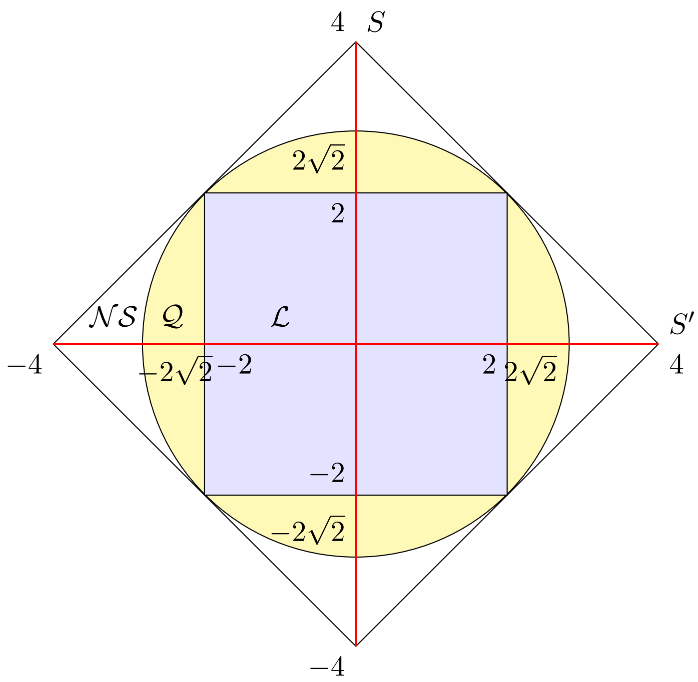


FIGURE 1. Schematic representation of the correlation bounds for the EPRB-setup: S, S' are the CHSH-values for two Bell inequalities, while the sets $\mathcal{L}, \mathcal{Q}, \mathcal{NS}$ contain distributions consistent with locality, quantum mechanics and the non-signalling requirements respectively (see Sec. 2.1).

Popescu and Rohrlich gave another boost to the field with their construction of probability distributions for the EPRB-setup which maximally violate the Bell inequalities with $S = 4$ while remaining consistent with relativity theory in the sense that they do not allow for superluminal signalling. While all of these concepts will be discussed in more detail in Sec. 2.1, we can already schematically represent these values for two CHSH inequalities in Fig. 1.
Contextuality. Bell's results were in fact later given a new twist by the Kochen-Specker theorem [10] in which it is proven that no hidden variable theory that assumes both λ to be independent of a given measurement context and to assign values to every observable that could be measured on the system at every moment can reproduce the predictions of quantum theory. In the standard quantum mechanical formalism the reason for this is the non-commutativity of the observables in terms of which [11] give an elegant and simple theorem reproducing the result.

Without going into any detail, it should be noted that since contextuality "attacks" hidden variable theories directly and does not require spacelike separation of margins it is a more general concept than non-locality. At the same time it is more difficult to study in that, for example, principles such as the one of Relativistic Causal Structure cannot be used to motivate the causal independence between margins as in the case of Bell locality (3). In this text we will be mostly concerned with non-locality and connections to contextuality will be pointed out where necessary.

Replacing by reconstructing. As we have seen, Bell arrived at his inequalities on the basis of assumptions about the physical world that seem reasonable from everyday, i.e. classical, experience (ignoring for the moment the further subtleties one encounters when delacing Bell locality further). With quantum mechanics things are the other way around. At least for the case of the

EPRB-setup, the Tsirelson bound tells us the exact limits of correlations but not the physical principle(s) that yield(s) these limits. To seek physical principles, or simple formal conditions, that produce the same correlation bounds as those of quantum theory then is a natural endeavor and is the subject of "reconstructions" research which produces axiomatic derivations of the quantum mechanical formalism.

Of course, despite its enormous degree of experimental corroboration, chances are that quantum theory will be superseded, although any successor theory would have to produce predictions and correlation bounds very close to the it. Nevertheless, this means that reconstructions that produce quasi-quantum theories can be interesting in their own right if they are more intelligible in terms of their physical content than quantum theory (conditional of course on their performance in experiment).

In recent years, especially with the advent of quantum information theory, many such principles have been suggested, such as non-trivial communication complexity (NTCC) [12], Macroscopic Locality (ML) [13] ($\equiv Q^1$), no advantage for nonlocal computation [14] and local orthogonality (LO) [15] or information causality [16], see [17] for a review. However, most of these are formulated operationally in terms of limits on the exchange of information between distant experimenter having access to measurement devices and their output readings. Regardless of the question whether these operational principles could, even in principle, can satisfyingly capture a theory's *physical* content, they are certainly bound to lose their meaning in the extreme scenarios for the description of which a unified quantum gravity is needed.

The Fine trio and conistent histories. One alternative reconstructive approach bases on a deep result in non-locality known as Fine trio: In his [18], Fine proved that, for the EPRB-scenario, the statements that

- $P(a,b|x,y)$ satisfies the Bell inequalities,
- $P(a,b|x,y)$ satisfies Bell locality (3) and
- there exists a probability measure $P(a_x, a_{\overline{x}}, b_y, b_{\overline{y}})$ over the whole space of possible outcomes for all measurements such that it reproduces the correct experimental probabilities, i.e.

$$P(a_x, b_y|x,y) = \sum_{a_{\overline{x}}, b_{\overline{y}}} P(a_x, a_{\overline{x}}, b_y, b_{\overline{y}}), \tag{5}$$

are equivalent. Here a_x denotes the ath outcome for setting x, $\overline{x}$ the alternative setting and similarly for $y, b_y, b_{\overline{y}}$. An analogue trio holds true for more general scenarios as well [19].

The third statement, the existence of a well-defined joint probability measure, is of natural interest to researchers in quantum gravity since, under the identification of elements of the space of possible outcomes above with "histories" or possible space-time "paths" of a system to take, it allows for the study of correlation limits as manifestations of the dynamics and kinematics of path integral-type theories, which, due to their inbuilt Lorentz covariance, are much better suited for quantum gravitational research than the above operational framework.

This is exactly the idea behind the histories approach to quantum theory which builds on Feynman's path integral formulation [20] and was first introduced by Griffiths, [21] and Hartle [22]. In the context of non-locality research the interesting question for this kind of approach then becomes how one can reconstruct a (quasi-)quantum theory that admits a histories interpretation. As will be shown in Sec. 3, this naturally leads to quantum measure theory, which, in a nutshell, is a histories based formulation of quantum theory in which the three members of the Fine trio are replaced by a quantum generalized analogue.

Closure for many copies of nonlocal distributions. Even though the operational principles stemming from information theoretic reconstructive research may be considered bad candidates for the development of a unified theory of quantum gravity, the information theoretic approach to the study of non-locality has certainly been fruitful, the PR-box being only one of many examples. Another one are a number of insights obtained by a resource theoretic approach to non-locality [23], in which the PR-box figures as a unit of non-locality [24] (and which should be distinguished from the treatment of entanglement as a resource [25]). In any resource theory the most natural question to ask is what states one can built from any number of identical copies of the basic unit state by acting on them with the free operations available in that resource theory (cf. LOCC) and by "wiring" them together in different ways.

Two important spin offs of asking this question about PR-boxes are, firstly, that quantum correlations require multipartite information principles [26], i.e. it is impossible to encompass all possible quantum correlations by reconstructive principles that are only bipartite, at least information theoretically; Secondly, that from the perspective of reconstructing quantum theory it seems clear that if some distribution is consistent with a fundamental principle about the universe, so should several copies of it. This gives rise to the notion of "closure" of sets of distributions [27], which will be discussed in detail in Ch. **??**.

For somebody working quantum measure theory to which this last requirement sounds reasonable, it is then of interest to study how it and notions related to it affect the sets of probability distributions deemed physical by quantum measure theory. To do so is the basic project of this thesis.

2. Non-signalling scenarios

In the study of non-locality in particular it is useful to consider the space of probability distributions for a given non-local scenario, i.e one in which parties are space-like separated. Denote a general scenario in which there are n parties being able to perform m measurements each with d possible outcomes on a jointly shared system as a (n, m, d)-scenario $\mathcal{S}_{n,m,d}$. Of course, more general scenarios exist but all the features studied here are covered in these scenarios which are much more convenient to study. One can then construct a space $\mathfrak{P}_{n,m,d}$ of $(dm)^n$ dimensions (reduced by a few normalization constraints) in which every point is a vector $P \equiv P(a_1, \dots, a_n|x_1, \dots, x_n), k \in \{1, \dots, n\}, x_k \in \{1, \dots, m\}, a_k \in \{1, \dots, d\}$, containing a complete specification of the probability for any possible combination of measurement choice and outcome for all parties. This *correlation space* space is bounded by the normalization constraints $\sum_{a_1,\dots,a_n} P = 1$ for every measurement setting. Also, every axial direction in this space corresponds to a sum of correlators of the form of (2), however generalized appropriately to the scenario. The study of correlations allowed under some physical principle then becomes the study of regions in $\mathcal{S}_{n,m,d}$ that satisfy this principle's formal representations.

2.1. **Local, quantum and non-signalling distributions: The NS-polytope.**

We begin by defining the set of distributions that satisfy Bell locality.

Definition 2.1 (Local Correlations $\mathcal{L}$)**.** Given a scenario $\mathcal{S}_{n,m,d}$, a probability distribution is *local* or *classical* iff [1]

$$P(a_1, \ldots, a_n|x_1, \ldots, x_n) = \sum_\lambda P(\lambda)P(a_1|x_1, \lambda), \ldots, P(a_n|x_n, \lambda), \tag{6}$$

where $\lambda \in \Lambda$ is an element of a hidden variable set satisfying $P(\lambda) \geq 0$ and $\sum_{\lambda\in\Lambda} P(\lambda) = 1$ and $P(a_i|x_i)$ are vectors in $\mathfrak{P}_{(1,m,d)}$.

(6) clearly corresponds to 3. Any distributions in $\mathcal{L}$ can be prepared in any classical theory, using only local operations and classical, i.e. (sub-)luminal, communication.

This set is convex, meaning that for any two elements of $\{P_i\} \subseteq \mathcal{L}$, their convex mixture $\sum_i \alpha_i P_i \in \mathcal{L}, \sum_i \alpha_i = 1$. Its extreme points are $P(a_1, \ldots, a_n|x_1, \ldots, x_n) = \delta_{a_1,f_1(x_1)} \ldots \delta_{a_n,f_n(x_n)}$, so-called deterministic boxes[2], where f_k are any function of the measurement settings. Since there exist maximally $d^{(nm)}$ of these, $\mathcal{L}$ is furthermore a convex polytope, i.e. the convex hull of finitely many extremal points [19]. There are two types of "facets", the surfaces connecting extremal points, for this polytope. One warrants positivity of the probabilities in (6), while the other is given by the Bell inequalities for this scenario [28].

Next consider quantum correlations.

Definition 2.2 (Quantum Correlations $\mathcal{Q}$)**.** For a (n, m, d)-scenario, a conditional probability distribution is *quantum* iff there exist a Hilbert spaces $\mathcal{H}$, a normalized state $|\psi\rangle \in \mathcal{H}$ and projector operators $\{E_k^{a,x}\} \subset B(\mathcal{H})$ with the properties:

(1) $\sum_a E_k^{a,x} = I_k$, $\forall x, k$
(2) $\left[E_k^{a,x}, E_{k'}^{a',x'}\right]$ $\forall k, k', x, x'$ if $k \neq k'$
(3) $P(a_1, \ldots, a_n|x_1, \ldots, x_n) = \langle\psi|E_k^{a_k,x_k}|\psi\rangle$

An alternative definition can be given in terms of individual Hilbert spaces to form the set $\mathcal{Q}'$:

Definition 2.3 (Quantum Correlations $\mathcal{Q}'$)**.** For a (n, m, d)-scenario, a conditional probability distribution is in $\mathcal{Q}'$ iff there exist Hilbert spaces $\{\mathcal{H}\}_{k=1}^n$, a normalized state $|\psi\rangle \in \bigotimes_{k=1}^n \mathcal{H}_k$ and projector operators $\{E_k^{a,x}\} \subset B(\mathcal{H}_k)$ with the properties:

(1) $\sum_a E_k^{a,x} = I_k$, $\forall x, k$
(2) $P(a_1, \ldots, a_n|x_1, \ldots, x_n) = \langle\psi| \bigotimes_{k=1}^n E_k^{a_k,x_k} |\psi\rangle$

These two definitions are known to be equivalent for finite dimensional systems, however it is an open question whether the same is true for the infinite-dimensional case [29]. In any case, since any measurements with a tensor product form $\mathcal{Q}' \subseteq \mathcal{Q}$, so here the main definition will be $\mathcal{Q}$.

$\mathcal{Q}$ is not a polytope. In fact, its characterization is not known to be a decidable problem and it is in general very difficult to decide membership of a given set to either $\mathcal{Q}$ or $\mathcal{Q}'$. This

[1] Throughout the whole text we consider only finite dimensional probability spaces. For continuous variables the same definitions can be made using integration instead of sums

[2] Distributions are in the information theoretic community often referred to as boxes, with the idea that any behavior can be abstractly thought of as a box with every party having access to a set of knobs to a measurement and outcome displays on the box such that statistics on sets of measurements and outcomes are given by marginals of the probability function. In this text "box" and "distribution" are used interchangeably.

is partly because correlation inequalities can only sample single points on its boundary, due to the absence of facets.

There is a third set that received much interest by researchers in non-locality.

Definition 2.4 (Non-signalling Correlations $\mathcal{NS}$)**.** For a (n, m, d)-scenario, a conditional probability distribution is *non-signalling* iff

$$(7) \quad \sum_{a_k} P(a_1, \dots, a_k, \dots, a_n | x_1, \dots, x_k, \dots, x_n) = \sum_{a_k} P(a_1, \dots, a_k, \dots, a_n | x_1, \dots, x'_k, \dots, x_n) \quad \forall k, k'.$$

(7) expresses the property of a behavior that the statistics of of the individual parties are independent from another's *measurement* settings. The name derives from the fact that, for spacelike separated parties, if a behavior did not obey the condition, the parties could signal superluminally by choosing different measurement settings on their end of the system. In this sense the condition is motivated from relativity theory. The NS-polytope has two kinds of extremal boxes. These are, firstly, the deterministic boxes from $\mathcal{L}$ and, secondly, a finite number of so-called PR-boxes. In the $(2, 2, 2)$-scenario there exist eight PR-boxes, which are all equivalent to

$$(8) \qquad P_{PR}(a, b|x, y) = \begin{cases} \frac{1}{2} & \text{if } a \oplus b = xy \\ 0 & \text{otherwise} \end{cases}$$

up to local operations. Overall, we have that, in any non-trivial scenario $\mathcal{L} \subseteq \mathcal{Q} \subseteq \mathcal{NS}$, with all containments being strict for the standard $(2, 2, 2)$-scenario. That $\mathcal{NS}$ contains $\mathcal{Q}$ is remarkable in at least two ways: On the one hand, it is not obvious from the algebraic structure of quantum mechanics that it should be sensitive to relativistic constraints. Secondly, the fact that this containment relation is strict for some scenarios implies that non-signalling itself is not enough to characterize the latter: The principles of relativity theory are not enough to explain physics on the quantum level. How far from achieving a full characterization of the physically possible correlations the NS-condition is, illustrates [30] nicely, where it is shown that PR-boxes trivialize communication complexity. Looking back at Fig. 1, some of the various features can now be read more clearly, for example the positivity facets indicated by the coincidence points of the three sets.

2.2. **Complexity of characterizing sets in $\mathfrak{P}$: The NPA-hierarchy.** Often it is of interest to decide whether a given distribution is an element of some set or not. For the case of sets which are convex polytopes, this problem can be formulated as a linear program (LP). Linear programs are used in the field of convex optimization. A problem is (equivalent to) a **linear program** if it can be stated in the form:

$$(9) \qquad \begin{array}{ll} \text{maximize} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & A\mathbf{x} \leq \mathbf{b} \\ \text{and} & \mathbf{x} \geq 0 \end{array}$$

where $\mathbf{x}$ is the vector to be determined and $A, \mathbf{b}, \mathbf{c}$ are known.

The connection to the polytope sets in $\mathcal{S}$ becomes evident if we identify the first line with an area, the second with a facet and the third with a positivity constraint on the polytope.

As has already been noted, for $\mathcal{Q}$ membership cannot be decided in this way. Indeed, no algorithm is known to decide membership of it for an arbitrary behavior with certainty. To aid this situation, [31, 29] introduced a hierarchy of increasingly strong tests of membership of $\mathcal{Q}$ for any bipartite scenario $(2, m, n)$.[3] The idea is to define sets of correlations $Q^n, n \in \mathbb{N}$, containing $\mathcal{Q}$ with $\underset{n\to\infty}{lim} Q^n \to \mathcal{Q}$ such that membership of any of them is decidable. In this way, failure of membership in Q^n for any n immediately implies that the distribution is not quantum.

This goal is achieved in the following way: Take any quantum distribution satisfying Def. 2.2 for bipartite case $n = 2$. By convexity we can form a mixed quantum state $\rho \in \mathcal{H}_A \otimes \mathcal{H}_B$ and sets of projectors $E_a = \hat{E}_a \otimes I, E_b = I \otimes \hat{E}_b$ describing a single measurement per party. We also have $[E_a, E_b] = 0$ and

$$P_{ab} = tr[E_a E_b \rho]. \tag{10}$$

Now consider $mathttS = S_1, \ldots, S_n$ as a set of n operators obtained by taking products and linear combinations of E_μ with $\mu \in a, b$. Define the hermitian $n \times n$-matrix

$$\Gamma_{ij} = tr(S_i^\dagger S_j \rho) \tag{11}$$

Γ is positive-semidefinite, since for any $\mathbf{v} \in \mathbb{C}^n$

$$\begin{array}{rcc} \mathbf{v}^\dagger \Gamma \mathbf{v} & = & \sum_{i,j} v_i^* tr(S_i^\dagger S_j \rho) v_i \\ & = & tr\left[(\sum_i v_i S_i)^\dagger \left(\sum_j v_j S_j \right) \rho \right] \\ & \geq & 0, \end{array} \tag{12}$$

since ρ is positive.

Furthermore, by the linearity of the trace,

$$\sum_{i,j} c_{ij} \Gamma_{ij} = 0 \text{ if } \sum_{i,j} c_{ij} S_i^\dagger S_j = 0 \tag{13}$$

and, if $\sum_{i,j} c_{ij} S_i^\dagger S_j = \sum_{a,b} d_{ab} E_a E_b$, then

$$\begin{array}{rcl} \sum_{i,j} c_{ij} \Gamma_{ij} & = & \sum_{i,j} c_{ij} tr(S_i^\dagger S_j \rho) \\ & = & tr(\sum_{i,j} c_{ij} S_i^\dagger S_j \rho) \\ & = & \sum_{a,b} d_{ab} tr[E_a E_b \rho] \\ & = & \sum_{a,b} d_{ab} P_{ab} \end{array} \tag{14}$$

Therefore, the existence of a matrix $\Gamma$with the above properties is a necessary feature of any behavior in $\mathcal{Q}$. The possibility of constructing this question given a behavior is an instance of a semidefinite program (SDP).

A **semidefinite program** is a problem of the form [33]:

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & F(x) \geq 0 \end{array} \tag{15}$$

where

$$F(x) := F_0 + \sum_{i=1}^{m} x_i F_i, \tag{16}$$

[3] The NPA hierarchy is generalized to any contextuality scenario in [32], however considering the bipartite case will be sufficient in this context.

$\mathbf{x} \in \mathbb{R}^m$ is the vector to be determined, $\mathbf{c} \in \mathbb{R}^m$ and the $m+1$ symmetric and positive semi-definite matrices $F_j \in \mathbb{R}^{n\times n}$ are known. Like linear programs SDPs are convex optimization programs for the solution to which powerful methods exist, even though they are more complex than the former. They have also been previously employed in the context of non-locality, specifically to prove extensions of the Tsirelson inequalities to multipartite scenarios [34] and, implicitly, in older work on correlation functions [35].

Since clearly there are no a priori limitations on the number of projectors that could be combined into $\mathtt{S}$, for any quantum behavior the matrix Γ, also called a certificate for $\mathtt{S}$, exists for all $|\mathtt{S}| = n \geq 1$. Conversely, a non-quantum behavior may yield a certificate for some $n \leq k_P$, where is some critical value specific to P. Call then Q^n the set of behaviors for which a certificate (denoted Γ^n) exists given that $|\mathtt{S}| = n$. Clearly, if $P \in Q^n$ then $P \in Q^k, 1 \leq k \leq n$ and $Q^{n+1} \subseteq Q^n$ for all $n \geq 1$.

This hierarchy is very useful given the difficulty to investigate membership to the quantum set directly, at least for bipartite scenarios. Moreover, [29] show that it is complete in the sense that a behavior is quantum if and only if there exists a certificate Γ^n for all $n \geq 1$. In other words

$$\lim_{n\to\infty} Q^n = \mathcal{Q} \tag{17}$$

At the same time, the complexity of testing the existence of a certificate of order n increases exponentially with n. This is because the size of Γ^n for a scenario with m measurements and d outcomes is $(md)^n$ and SDP algorithms used to construct the certificate have a running time polynomial in the latter's size.

2.2.1. *Almost quantum correlations.* While the NPA hierarchy from Sec.2.2 was initially devised to compuationally aid the characterization of the quantum set, two of the sets in the hierarchy have found interest in a reconstructive context: Q^1 which defines the principle of macroscopic locality [13] and the set Q^{1+AB} which, in its multipartite generalization defines the *almost quantum* distributions as

Definition 2.5 (Almost quantum set $\tilde{Q}$)**.** Given any (n, m, d)-scenario, a conditional probability distribution is *almost quantum* iff there exists a Hilbert space $\mathcal{H}$, a normalized state $|\psi\rangle \in \mathcal{H}$ and projector operators $E_k^{a,x} \subset B(\mathcal{H})$ with the properties:

(1) $\sum_a E_k^{a,x} = I_k, \forall x, k$
(2) $E_1^{a_1,x_1} \dots E_n^{a_n,x_n}|\psi\rangle = E_{\pi(1)}^{a_{\pi(1)},x_{\pi(1)}} \dots E_{\pi(n)}^{a_{\pi(n)},x_{\pi(n)}}|\psi\rangle$ for any permutation $\pi \in S_n$
(3) $P(a_1, \dots, a_n|x_1, \dots, x_n) = \langle\psi| \prod_{k=1}^n E_k^{a_k,x_k}|\psi\rangle$

Clearly, $\mathcal{Q} \subseteq \tilde{Q}$, since for any quantum distribution according to Def. 2.3 the projective set formed of projectors $\tilde{E}_k^{a,x} = \mathbb{I}^{\otimes k-1} \otimes E_k^{a,x} \otimes \mathbb{I}^{\otimes n-k}$ satisfies the requirements of Def. 2.5. In [36] this result is proven to be strict for the $(2, 2, 2)$-scenario.

Almost quantum correlations approximate the quantum set very closely: [36] prove that $\tilde{Q}$ implies[4] all of the operational principles that were listed in the introduction, i.e. non-trivial communication complexity (NTCC) [12], Macroscopic Locality (ML) [13] ($\equiv Q^1$), no advantage for nonlocal computation [14] and local orthogonality (LO) [15] (Sec.**??** and all of which are

[4]Here "A implies P" means that the set A is contained by the set of correlations satisfying a given principle P in any scenario.

proposed information theoretic principles to reconstruct the quantum set. The authors also present strong numerical evidence that $\tilde{Q}$ implies information causality [16].

Lemma 2.1. *$\tilde{Q}$ recovers the Tsirelson bound*

Proof. This result is obtained in [29] who show that for $(2, 2, d)$-scenarios the maximal violations of the generalized Bell inequalities (the CGLMP inequalities [37]) by $\tilde{Q}$ coincide with those of $\mathcal{Q}$ at least for $2 \leq d \leq 8$. For the case $d = 2$ these are the Tsirelson's bound. $\square$

3. Quantum Measure Theory

3.1. **The Quantum Fine Trio.** Recall that the Fine trio presented in Sec. 1 proves the equivalence of (i) the (original) Bell inequalities (2), (ii) the satisfaction of Bell locality (3) and (iii) the existence of a probability function reproducing the right statistics,

$$P(a_x, b_y|x, y) = \sum_{a_{\overline{x}}, b_{\overline{y}}} P(a_x, a_{\overline{x}}, b_y, b_{\overline{y}}). \tag{18}$$

Note that the two functions on the LHS and RHS are completely different objects: While the conditional probability on the LHS is a vector in $\mathfrak{P}$, the domain of the equation on the RHS is a set, or corresponding vector space, whose elements specify the outcomes of of all possible measurements in the scenario, including alternative measurements for the same party. They are denoted by the same symbol because they both relate to probabilities and can be clearly distinguished by their arguments. The existence of a probability function over this space expresses both a kind of realism, asserting that measurement statistics are the product of having only partial access to this well-defined statistics on this other space, and a non-contextuality assumption in that the probability function on the right hand-side is independent of what is being measured (i.e. the outcomes over which one traces). This intuition becomes clearer if one takes the elements of this space, atoms $\gamma_i \equiv \{a_1^{x_1}, \ldots, a_1^{x_m}, a_2^{x_1}, \ldots, a_n^{x_m}\}$, which we call the "non-contextuality (NC-) space" Ξ to be a full specification of the state of a system over all time, its "history" or "path" (through spacetime or phase space).

Apart from articulating this expectation, Fine's trio also connects locality as a physical principle of the world with a clear formal statement about the ability to define a function over a certain space. Since this is easier to study and at the same time admits a neat interpretation in terms of histories or paths, being the most natural kinematic objects of path integral-type theories, it is a reasonable hope to seek a similar trio on the quantum level, a quantum Finte trio, even though in the light of the past section it is easy to see that such a generalization might be more difficult than for the classical case.

The obvious starting point to seek such a this is to investigate possible generalizations of the original trio. Of course, there is no unique way of carrying out such a generalization. It is suggestive to first generalize Fine's result about joint probability spaces to the quantum level and then look for the corresponding physical statement, by analogy or otherwise. But what does it mean to "generalize probabilities to the quantum level"? One possible way of carrying out such a generalization results in Quantum Measure Theory (QMT) and is the subject of this section.

3.2. **The quantum measure.**

3.2.1. *Measures.* [5] To carry out this generalization some terminology is required. Adopting the above history-type approach, given scenario $\mathcal{S}_{n,m,d}$ call the set formed by "atoms" $\gamma_i \equiv \{a_1^{x_1}, \ldots, a_1^{x_m}, a_2^{x_1}, \ldots, a_n^{x_m}\}$ with $i \in \{1, \ldots, d^{(mn)}\}$ the "non-contextuality (NC-) space" Ξ.

It is easy to see from a generalized form of (18) that every single outcome event e$\equiv (a_1, \ldots, a_n | x_1, \ldots, x_n)$ corresponds to a subset A of Ξ, namely the set of atoms over which one sums to obtain the corresponding probabilities. For a given measurement, set of all subsets for the different outcomes in fact forms a partition $M \equiv M(x_1, \ldots, x_n)$ of Ξ which we can identify with that measurement. Denote $A(e)$ and $A(I)$ the bijective maps from events and sets I of atom labels to the power set 2^{Ξ} respectively (yielding a bijective map $e(I)$ as well) and also call $\mathcal{M}$ the set all measurements M. Next define the notion of a sigma algebra [39]:

Definition 3.1 (Sigma algebra $\mathfrak{A}$)**.** A class of subsets $\mathfrak{A}$ of some set X forms a sigma algebra of X if it contains X itself and is closed under the formation of complements and countable unions:

(1) $X \in \mathfrak{A}$
(2) if $A \in \mathfrak{A}$ then $\overline{A} \in \mathfrak{A}$
(3) if $A_1, A_2, \dots \in \mathfrak{A}$ then $A_1 \cup A_2 \cup \dots \in \mathfrak{A}$

where $\overline{A}$ denotes the complement of A. The largest possible sigma algebra over X is the power set of X, 2^X and the smallest is the union of X with the empty set $\emptyset$. The set $\mathfrak{A}_M$ of all "coarse-grainings", i.e. all possible unions of outcomes for a given measurement M, together with the empty set $\emptyset$ forms such an algebra. For notational convenience can further form $\mathcal{C} := \bigcup_{M \in \mathcal{M}} \mathfrak{A}_M$. That outcomes of measurements form a (sigma) algebra is important because it is necessary to define a measure for every M.

Definition 3.2 (Measure μ)**.** Let X be a set and $\mathfrak{A}$ an algebra over X. Then the function $\mu : \mathfrak{A} \to \overline{\mathbb{R}}$ is a measure if it satisfies:

(1) $\mu(A) \geq 0 \quad \forall A \in \mathfrak{A}$
(2) $\mu(\emptyset) = 0$
(3) if $\bigsqcup_{i=1}^{\infty} A_i \in X$ then

$$\mu\left(\bigsqcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i) \tag{19}$$

Here $\overline{\mathbb{R}}$ denotes the extended real line, i.e. $\mathbb{R}$ together with $\pm\infty$. Also $\sqcup$ represents pairwise disjoint union, i.e. it acts like the binary $\cup$ operator, however only on sets that are disjoint. Its use in the above definition (and henceforth) then implies that we require all members of the collections on which it acts to consist of disjoint members. The condition in the third property is called *countable additiviy*.[6] Here, really we are interested in defining a probability measure as a special kind of measure.

[5]This section introduces some of the concepts from the joint measurement scenarios (JMS) [38] necessary to discuss quantum measure theory. Since it is not necessary here to distinguish between basic and joint measurements the set of all measurements is denoted $\mathcal{M}$, i.e. the tilde is dropped, following the example of [?]

[6]It is required since $\mathfrak{A}$ does not need to be a sigma algebra. Of course, for any finite collection of n elements of $\mathfrak{A}$ we are warranted that

$$\mu\left(\bigsqcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i), \tag{20}$$

Definition 3.3 (Probability Measure P)**.** Let Ξ be a NC- space. Then a function $P : \mathfrak{A} \to \overline{\mathbb{R}}$ is a probability measure over Ξ if it is a measure over Ξ such that

(1) $P(A) \geq 0 \quad \forall A \in \Xi$
(2) $P(\Xi) = 1$

For every probability measure one has that $0 \leq P(A) \leq 1$ for every member of the algebra. The upper limit follows since $1 = P(\Xi) = P(A) + P(\overline{A})$ implies that $P(A) \leq 1$.

The reason for these considerations is that we *know* from the Fine trio and the violation of the Bell inequalities that already for the simple $(2, 2, 2)$-scenario one cannot define a probability *measure* over the whole of 2^{Ξ} such that it produces the correct conditional probabilities on $\mathfrak{P}_{(2,2,2)}$ for every possible state. Nevertheless, consistency with the observation of classicality of any macroscopic measurement devices imposes the strict requirement on any probability *function* on 2^{Ξ} that at least the conditional probabilities it produces are probability measures. This is the case if we know that for any $M \equiv M(x_1, \dots, x_n), P(A)$ behaves like a probability measure for all $A \in \mathfrak{A}_M$.

A probability measure is defined over an algebra on the NC-space and not the NC-space itself or its power set because probabilities are assigned to alternative outcomes of possible measurements (where these measurements can sometimes be thought as representing physical processes that happen outside any laboratory or operational context). The definition of an algebra, in particular its closure under set-theoretic operations, warrants the minimal requirement for collections of subsets of Ξ to fulfill exactly this representative function as sets of measurement outcomes, with sigma algebras allowing for the treatment of measurements with infinitely many (but countable) outcomes.

Subsets of the NC-space, or its power set, in general correspond to events involving incompatible measurements or outcomes, for which it makes no sense to attribute probabilities to them under any interpretation of classical probability. To take the standard example from quantum mechanics this would be the case for the description of the spin measurement of an electron. That the spin along some canonical reference axis, call it the x-axis, and an axis orthogonal to the x-axis, call it the y-axis, should be both "up" at time t or both "down" at the time t are not alternative outcomes of a possible measurement but both events are members of the power set of Ξ. In this example it wouldn't make sense to impose the conditions of a classical probability measure on any collection of sets including these events as members.

If one is interested in studying the possibility of making statements about sets of events that are *not* alternative measurement outcomes - the basic premise of non-contextuality research - then the above however is by no means the end of the game. Indeed, if one reformulates the subject of non-contextuality research as the study of functions that are defined over 2^{Xi}, then the above only imposes the restriction on the these functions to reduce to probability measures over any algebra corresponding to possible measurements with classical outcomes. In this sense, the recovery of probability measures at the level of classical events in probability theory is exactly analogous to the requirement on any physical theory to reproduce classical behavior in the corresponding limit - a requirement of *empirical* (as opposed to *logical*) consistency.

called *finite additivity*, by setting $A_{n+1}, A_{n+2}, \dots = \emptyset$. The requirement of countable additivity for measures on general algebras also nicely illustrates the independence between finite/countable additivity and closure under finite/countable unions.

3.2.2. *The Kolmogorov sum rule and higher-order interference.* This empirical motivation of the definition of a probability measure[7] may invoke the question at which part of the definition of a probability measure classical assumptions enter. The answer lies in the additivity-property. Countable additivity is a consequence of the axioms of axiomatic probability theory, in particular of the "Kolmogorov Sum rule" that for any pair of disjoint elements A, B of some algebra X

$$P(A) + P(B) = P(A \sqcup B) \tag{21}$$

In general of course, this needn't be true for functions on algebras. For a general linear function $f : X \to \overline{\mathbb{R}}$ on a set X we have for any $A, B \in X$ that

$$f(A) + f(B) = f(A \cup B) - f(A \cap B) \tag{22}$$

It is obvious that (21) postulates for $f = P$ that $P(A \cap B) = P(A \cup B) - P(A) - P(B) = 0$, i.e. that the interference term vanishes. In this sense the Kolmogorov sum rule is a physical or logical decoherence requirement in the sense that, if A and B correspond to alternative measurement outcomes or logical propositions and $\cap$ represents conjunct occurrence or logical binary-AND respectively, the rule postulates that the probability of the latters' occurrence or being true is zero. And this postulate, for the physical case at least, is motivated by the fact that we never see alternative measurement outcomes occur both at once (see [40, 4] for investigations on relaxing logic in quantum measure theory or [41] for more general consistent histories).

If one considers situations in which the alternatives in question are histories of the universe, that is, specifications of the state of the universe over some space-time volume, then this decoherence postulate becomes the requirement that there exists no interference between pairs of alternative histories for the universe in question or the two universes in which either history is realized depending on one's underlying ontology.

In [42], Sorkin first discussed the possibility of generalizing this prohibition of interference of two histories to higher-orders, that is to n histories. Define the interference term for a function $\mu : X \to \overline{\mathbb{R}}$

$$\begin{aligned} I_n(A_1, A_2, \ldots, A_n) &:= \sum_{i=0}^{n-1} \sum_{j=1}^{N_i} (-1)^i \mu\left(\bigsqcup_{k\in K} A_k\right) \\ &= \mu(A_1 \sqcup A_2 \sqcup \ldots A_n) - \sum \mu(A_1 \sqcup A_2 \sqcup \ldots A_{n-1}) \\ &\qquad - \mu(A_2 \sqcup A_3 \ldots A_n) + \cdots + (-1)^{n-1} \sum_{k=1}^{n} \mu(A_k), \end{aligned} \tag{23}$$

where $K(i, j)$ is the subset of $N = 1, \ldots, n$ with $|K(i, j)| = n - i$ members corresponding to the choice of A_k for the pair (i, j) and $N_i = \binom{n}{i}$.

Clearly, For the case of $n = 2$ this reduces to the Kolmogorov sum rule (21) if we set I_2 to zero. For $n = 3$ we have

$$I_3(A, B, C) = \mu(A \sqcup B \sqcup C) - \mu(A \sqcup B) - \mu(B \sqcup C) - \mu(A \sqcup C) + \mu(A) + \mu(B) + \mu(C) \tag{24}$$

[7]The following considerations are, to some extent, also valid for general measures but we will restrict attention to the relevant case of probability measures only.

For $I_3 = 0$ this equation, called the **Sorkin sum rule**, allows for the possibility of pairwise interference but not third-order interference. In fact, for any n, we have that, if $I_n = 0$ then $I_k = 0$ for all $k \geq n$. This follows by the following lemma

Lemma 3.1. *For any $n \geq 1$,*

$$(25)\quad I_{n+1}(A_1, A_2, \ldots, A_{n+1}) = I_n(A_1, \ldots, A_n \sqcup A_{n+1}) - I_n(A_1, A_{n-1}, A_{n+1} - I_n(A_1, A_2, \ldots, A_n)$$

Proof. This is proven in [43]. □

One can classify stochastic theories by the order of interference which they allow. As we saw above, classical probability theory forbids any non-trivial interferences. Quantum theory allows for pairwise interference but forbids interferences for $n \geq 3$.

To see this, consider a three-split experiment as a simple generalization of the double-slit experiment in which installing blinds at each slit gives $2^3 = 8$ settings of the slits leading to different interference patterns. One can represent possible paths of a quantum system in this setup by defining and 7-dimensional projective set $\{E_J\}$ where $E_J = |E_J\rangle\langle E_J|$ and each J is an element of the power set of $\{1, 2, 3\}$ ignoring the empty set which corresponds to all slits being closed. Then, the quantum mechanical probability that any prepared state $|\psi\rangle = \sum_k c_k |k\rangle$ follows some path J (in the sense of interfering histories) is

$$(26)\qquad P(J||\psi\rangle) = tr[E_J|\psi\rangle\langle\psi|] = \sum_k \langle E_J|c_k|k\rangle\langle k|c^*|E_J\rangle = \sum_k a_k^{(J)*}$$

and we have $a_k^{(ij)} = a_k^{(i)} + a_k^{(j)}$ etc.. Furthermore, for a set of POVMs $\{D_i\}$ given by $D_i = |D_i\rangle\langle D_i|$ which correspond to detectors d_i on a screen, the quantum mechanical probability that a system will first follow a certain path through the slits and then be measured by detector d_i is

$$(27)\qquad P(d_i, J||\psi\rangle) = P(d_i|J, |\psi\rangle) \times P(J||\psi\rangle) = tr\left[D_i|E_J\psi\rangle\langle\psi|E_J\right]$$

Superimposing the probabilities for all J and a given d_i such that probabilities with $|J|$ even or odd destructively interfere then gives

$$\begin{aligned}
(28)\quad & tr\left[D_i E_{123}\psi\rangle\langle\psi|E_{123}\right] - \sum_{1=m\leq n}^{3} tr\left[D_i E_{mn}|\psi\rangle\langle\psi|E_{mn}\right] + \sum_{m=1}^{3} tr\left[D_i E_m\psi\rangle\langle\psi|E_m\right] \\
&= \sum_k \langle D_i| \left[a_k^{(123)*} a_k^{(123)} - \sum_{1=m\leq n}^{3} a_k^{(mn)*} a_k^{(mn)} + \sum_{m=1}^{3} a_k^{(m)*} a_k^{(m)} \right] |D_i\rangle \\
&= \sum_k \langle D_i| \left[\left(a_k^{(1)} + a_k^{(2)} + a_k^{(3)} \right)^* (a_k^{(1)} + a_k^{(2)} + a_k^{(3)}) \right. \\
&\quad - \sum_{1=m\leq n}^{3} \left(a_k^{(m)} + a_k^{(n)} \right)^* \left(a_k^{(m)} + a_k^{(n)} \right)^* + \sum_{m=1}^{3} a_k^{(m)*} a_k^{(m)} |D_i\rangle \\
&= 0,
\end{aligned}$$

where we used the cyclicity and linearity of the trace and the fact that the projectors sum up to identity. (28) is clearly true for any D_i and state preparation, it can also easily be generalized to mixed states, since interferences will cancel for every state in an ensemble. The generality of this result and recognizing its form as that of $I_3 = 0$ in the hierarchy lets one classify quantum mechanics as a stochastic theory forbidding third and higher order interference. It is also in this way that we can understand the coincidence with the classical probabilities P_J for the three-slit experiment which yields

$$P_{123} - P_{12} - P_{23} - P_{13} + P_1 + P_2 + P_3 = \sum_{j=1}^{3} 2(P_j - P_j) = 0. \tag{29}$$

See [44] for a discussion of the three-slit experimental in an operational framework. Motivated by this interesting property of quantum statistics which automatically produces the marginal classical behavior we are seeking for our non-contextual probability function $P(A)$, we can formulate the following principle to study the correlations it allows.

Principle of pairwise interference (PPI)
The only kind of interference that is possible between
histories of systems is pair-wise interference.

3.2.3. *The quantum measure.* To formalize PPI the natural step is to weaken the Kolmogorov sum rule to allow for second-order interference:

Definition 3.4 (Quantum measure μ)**.** Let Ξ be a NC-space for some scenario $\mathcal{S}_{(n,m,d)}$. Then a function $\mu : 2^{\Xi} \to \overline{\mathbb{R}}$ is a quantum measure over Ξ if

(1) $\mu(A) \geq 0 \quad \forall A \in 2^{\Xi}$
(2)
$$\begin{gathered}\mu(A) + \mu(B) + \mu(C)\\ -\mu(A \sqcup B) - \mu(A \sqcup C) - \mu(B \sqcup C)\\ +\mu(A \sqcup B \sqcup C) = 0 \quad \forall A, B, C \in 2^{\Xi}\end{gathered}$$
(3) $\mu(A(e)) = P(e) \equiv P(a_1, \ldots, a_n | x_1, \ldots, x_n) \quad \forall A \in \mathbb{C}$

We recognize 2 as the Sorkin sum rule (24) and find the third condition 3 to express exactly the empirical consistency condition discussed above. Also μ is a proper measure in that it is defined over an algebra on Ξ.

In the literature one finds ambiguous usage of the terms "quantum measure" and "quantal measure". Here we adopt the convention that any generalized measure satisfying the Sorkin sum rule will be called a "quantal measure", while a quantal measure that also reduces to the classical probabilities for a given behavior is called a "quantum measure".

The definition for the interference term in Eq. (23) and the proof for Lemma 3.1 still work out for the function in question being a quantum measure. The Sorkin sum rule then implies that the quantum measure is bi-additive in its argument in the sense that

$$I_2(A \sqcup B, C) = I_2(A, C) + I_2(B, C), \tag{30}$$

which follows immediately from Lemma 3.1 and can also be generalized to n-additivity for $I_{n+1} = 0$.

3.3. **The decoherence functional.** The bi-additiviy of I_2 allows for the construction of a convenient equivalent formal representation of a quantal measure. Consider the following definition

Definition 3.5 (Decoherence functional D)**.** A map $D : 2^{\Xi} \times 2^{\Xi} \to \overline{\mathbb{R}}$ is called decoherence functional if, for all sets $A, B, C \in 2^{\Xi}$

(1) $D(A, B) = D(B, A)^*$
(2) $D(A \sqcup B, C) = D(A, B) + D(A, C)$
(3) $D(A, A) \geq 0$
(4) $D(\Xi, \Xi) = 1$,

where we remind the reader that the $\sqcup$-operation is defined for disjoint sets only. Clearly, a decoherence functional corresponds to a *quantum* measure if it also satisfies

$$D(A(e), A(e)) = P(e) \equiv P(a_1, \ldots, a_n | x_1, \ldots, x_n) \quad \forall A \in \mathbb{C} \tag{31}$$

This equivalence is particularly useful in that, by defining the matrix $D_{AB} := D(A, B)$, the quantum measure can be studied using simple algebraic tools.

3.3.1. *Atomic decoherence functional.* As D_{AB} grows exponentially with the size of the NC-space, it is often even more convenient to define an "atomic" decoherence functional over the NC-space.

Definition 3.6 (Atomic decoherence functional $\hat{D}$)**.** A map $D : \Xi \times \Xi \to \overline{\mathbb{R}}$ is an atomic decoherence functional for a quantum measure if, for all atoms $\gamma_i \in \Xi$ with $I \subset \{1, 2, \ldots, n\}$

(1) $\hat{D}(\gamma_i, \gamma_j) = \hat{D}(\gamma_j, \gamma_i)^*$
(2) $\sum_{i,j \in I} \hat{D}(\gamma_i, \gamma_j) \geq 0$
(3) $\sum_{i,j}^{n} \hat{D}(\gamma_i, \gamma_j) = 1$
(4) $\hat{D}(\gamma_i \sqcup \gamma_j, \gamma_k) = \hat{D}(\gamma_i, \gamma_k) + \hat{D}(\gamma_j, \gamma_k)$
(5) $\sum_{i,j \in I} \hat{D}(\gamma_i, \gamma_j) = \mu(A(I)) = P(e(I)) \quad \forall A \in \mathbb{C}$

The existence of an atomic decoherence functional follows by the simple relation, which we already used, that

$$A(I) = \bigsqcup_{i \in I} \gamma_i \quad I \subset \{1, \ldots, d^{(mn)}\}. \tag{32}$$

Again, we can represent $\hat{D}$ with the hermitian matrix $\hat{D}_{ij} := \hat{D}(\gamma_i, \gamma_j)$.

3.3.2. *Strong positivity.* The properties of the decoherence functional are very weak, which makes their properties difficult to study in all generality. Most of the existing literature on quantum measure theory has therefore focussed on a family of decoherence functionals, characterized by their being *strongly positive*.

Definition 3.7 (Strong positivity)**.** A decoherence functional D, or its corresponding atomic decoherence functional $\hat{D}$, is strongly positive if $\hat{D}_{ij} \equiv \hat{D}(\gamma_i, \gamma_j)$ is a positive semi-definite matrix.

There is no obvious a priori reason for requiring strong positivity of the decoherence functional. Nevertheless most of the research on quantum measures has focussed on this family of functional. Besides the fact that by requiring strong positivity the full array of tools in the linear algebra of positive matrices becomes available, this is because strong positivity is also sufficient for the construction of a Hilbert space in the following way.

3.3.3. *Hilbert space construction from strongly positive decoherence functional.* Define a Hilbert space as a complex vector space with non-degenerate Hermitian inner product which is complete with respect to the induced norm. The construction proceeds in the fashion of "GNS" (Gel'fand-Naimark-Segal) who use C^*-algebras and is employed similarly in [45] and [46]. Given a behavior $(\Xi, \mathcal{M}, P)$, define the vector space $\mathfrak{H}_1$ that is spanned by basis vectors $\{|[\gamma_i]\rangle\}$ where γ_i are the atoms of the partition scenario induced by the behavior. If $P \in SPJQM$ then we can define a (possibly degenerate) Hermitian inner product on $\mathfrak{H}_1$ as

$$\langle[\gamma_i]|[\gamma_j]\rangle := \hat{D}_{ij}, \tag{33}$$

where $\hat{D}$ is the positive-semidefinite matrix given by the atomic decoherence functional for this behavior as defined in Def. 3.6.

The double $(\mathfrak{H}_1, \langle\cdot|\cdot\rangle)$ forms an inner product space only, since it in general contains vectors with zero norm (see [47] for an example). Since a true Hilbert space does not contain vector with zero norm we define $\mathcal{H} = \mathfrak{H}_1/\mathfrak{H}_0$ where $\mathfrak{H}_0$ is the vector subspace of states with zero norm. Then the vectors $|\gamma_i\rangle \in \mathcal{H}$ corresponding to $|[\gamma_i]\rangle \in \mathfrak{H}_1$ form a complete basis for $\mathcal{H}$, making it a proper Hilbert space as defined above [48]. In the case of infinite-dimensional systems this construction is more subtle but still possible as is shown in [47] for path integrals, i.e. for a non-relativistic particle moving in d-dimensions and over possibly infinite time. Since here we consider finite-dimensional systems only, the above construction is sufficient.

3.4. **Correlations in Quantum Measure Theory.** Following the idea presented in the introduction to this section our main concern lies in investigating the possibility of formulating a quantum analogue of the Fine trio based on the quantum measure. To do so the first step is to study the extent to which the Sorkin sum rule characterizes quantum theory, how could is PPI as a reconstructive principle? In terms of non-locality research this question asks about the relationship between the sets whose member distributions allow for a quantum measure and the quantum set. First define the following natural quantum measure theoretic sets, where we use the power set 2^Ξ since we are only considering finite-dimensional noncontextuality-spaces.

Definition 3.8 (JQM)**.** For a given (n, m, d)-scenario a conditional probability distribution $P(e)$ is in JQM if there exists a quantum measure μ on 2^Ξ for this scenario such that

$$\mu(A(e)) = P(e) \quad \forall A \in \mathcal{C} \tag{34}$$

We can equivalently define the set in terms of the decoherence functional as above.

Definition 3.9 (JQM)**.** For a given (n, m, d)-scenario a conditional probability distribution $P(e)$ is in JQM iff there exists a decoherence functional $D(A, B)$ on 2^Ξ such that

(1) $D(A(e), A(e)) = P(e) \quad \forall A \in \mathcal{C}$
(2) $D(A(e), B(e')) = 0 \quad \forall A, B \in \mathfrak{A}_M, A \cap B = \emptyset, \forall M \in \mathcal{M}$

Here,$B(e') \equiv A(e')$, i.e. it is given by the same mapping, however denoted differently for convenience. Note that for decoherence functionals whose entries are real, property 2 is redundant, however for the complex case this adds the non-trivial requirement that $Im[D(A,B)] = 0$ for the above sets. Another set of interest further involves strong positivity.

Definition 3.10 ($SPJQM$)**.** For a given (n,m,d)-scenario a conditional probability distribution $P(e)$ is in $SPJQM$ if there exists a *strongly positive* decoherence functional D on 2^{Ξ} such that

- $D(A(e),A(e)) = P(e) \quad \forall A \in \mathcal{C}$
- $D(A(e),B(e')) = 0 \quad \forall A,B \in \mathfrak{A}_M, A \cap B = \emptyset, \forall M \in \mathcal{M}$

Since the possibility to construct a Hilbert space is not only a necessary but also a sufficient condition for the existence of a Hilbert space, one can define $SPJQM$ alternatively as

Definition 3.11 ($SPJQM$)**.** For a given (n,m,d)-scenario a conditional probability distribution $P(e)$ is in $SPJQM$ iff there exists a Hilbert space $\mathcal{H}$ spanned by a set of vectors indexed by elements A of 2^{Ξ}, $\{|A\rangle\}$ such that, for any $I \subseteq \{1,2,\ldots,d^{(mn)}\}$

(1) $|A(I)\rangle = \sum_{i\in I}|\gamma_i\rangle$
(2) $\langle A(e)|A(e)\rangle = P(e) \quad \forall A \in \mathcal{C}$
(3) $\langle A(e)|B(e')\rangle = 0 \quad \forall A,B \in \mathfrak{A}_M, A \cap B = \emptyset, \forall M \in \mathcal{M}$,

where the $|\gamma_i\rangle$ are the vectors constructed in 3.3.3.

These sets are defined in [38], however for bipartite scenarios only. Even though this makes sense in that these scenarios are the best studied, our investigation of scenarios with multiple copies makes it necessary to define them for any scenario. Note however, that this does not mean that Fine's trio applies for any scenario, since it does not.

3.4.1. *Relations between sets.* We have the following basic relationships for quantum measure theoretic sets.

Lemma 3.2. *$SPJQM \subseteq JQM$*

The containment relation is clear. In the case of the $(2,2,2)$-scenario this containment is strict, which follows by Theorem **??** and Lemma 3.7 to be discussed in the upcoming subsection. It is not clear whether there exist scenarios in which $JQM \equiv SPJQM$.

Lemma 3.3. *$\mathcal{Q} \subseteq SPJQM$*

Proof. Take any behavior $P \in \mathcal{Q}$. Then there exists a normalized state $|\psi\rangle \in \mathcal{H}$ together with projection operators $E_k^{a_k,x_k} \in B(\mathcal{H})$ with the properties and statistics of Def. 2.2. Now define vectors $|\gamma_i\rangle$ as

$$|\gamma_i\rangle := \prod_k \prod_{l=1}^{m} E_k^{a_k(i),l_k}|\psi\rangle \tag{35}$$

for every atom γ_i where $a_k(i)$ specifies the measurement outcomes for every measurement setting and party corresponding to the ith element of the NC-space. We can then define vectors for coarser events as

$$|A(I)\rangle := \sum_{i\in I}|\gamma_i\rangle \tag{36}$$

These vectors allow for the construction of a Hilbert space with the properties given in Def.3. $\square$

Lemma 3.4. $\mathcal{Q} \subseteq JQM$

Proof. This is implied by the two lemmata above. □

These results are very interesting: They show that for any scenario the weakening of the Kolmogorov Sum Rule according to the PPI, that is weakening it by only one order, already allows for the recovery of all quantum correlations and distributions together with the existence of a well defined, albeit quantized, measure on the non-contextuality space. Since this corresponds to a precise principle that can be interpreted in terms of a physical path integral framework it is a promising first step towards a quantum Fine trio. Of course, one first needs to investigate how closely these sets approximate the quantum set. Here a number of results have been obtained for the $(2,2,2)$-scenario.
Correlation bounds for the CHSH scenario. If one considers only single copies of boxes the following results exist for the above sets:

Lemma 3.5. *For the $(2,2,2)$-scenario, the correlators of behaviors in $SPJQM$ are given by the inner product of unit vectors given by*

$$|x\rangle := \sum_{a_x, a_{\overline{x}}, b_y, b_{\overline{y}}} a_x |a_x, a_{\overline{x}}, b_y, b_{\overline{y}}\rangle \in \mathcal{H} \tag{37}$$

and similarly for the other outcomes.

Proof. This is simply proven using the Hilbert space construction ang done explicitly in [48]. □

From this we can prove the following:

Theorem 3.6. *$SPJQM$ recovers the Tsirelson bound: The correlators* (1) *of any distribution in $SPJQM$ in the $(2,2,2)$-scenario obey*

$$|C_{x=0y=0} + C_{10} + C_{01} - C_{11}| \equiv S \leq 2\sqrt{2} \tag{38}$$

and similarly for the three other CHSH inequalities.

Proof. The full proof is again given in [48]. It suffices to proof (38) without considering the absolute value as can be seen by flipping the y-settings.

Using (33) we can express the correlators defined in (1) as

$$C_{xy} = \langle x|y\rangle. \tag{39}$$

Denoting $|x=0\rangle \equiv |0_x\rangle$ etc., this gives

$$\begin{aligned} S &= \langle 0_x|0_y\rangle + \langle 1_x|0_y\rangle + \langle 0_x|1_y\rangle - \langle 1_x|1_y\rangle \\ &= (\langle 0_x| + \langle 1_x|)|0_y\rangle + (\langle 0_x| - \langle 1_x|)|1_y\rangle \\ &\leq ||0_y\rangle + |1_x\rangle|^2 + ||0_y\rangle - |1_x\rangle|^2, \end{aligned} \tag{40}$$

where we used the fact that the vectors are normalized and that S is maximal when $|0_y\rangle$ and $(|0_y\rangle + |1_x\rangle)$ as well as $|1_y\rangle$ and $(|0_y\rangle - |1_x\rangle)$ are parallel. Now for any normalized vectors u, v such that $S = |u+v|^2 + |u-v|^2 \equiv \|u+v\| + \|u-v\|$ then, writing $\xi \equiv Re[\langle u|v\rangle]$,

$$\begin{aligned} S^2 &= \|u+v\| + \|u-v\| + 2\|u+v\|\|u-v\| \\ &= (2+2\xi) + (2-2\xi) + 2\sqrt{(s+2\xi)(2-2\xi)} \\ &= 4 + 2\sqrt{4-4\xi^2} \leq 4 + 2\sqrt{4} = 8 \end{aligned} \tag{41}$$

Substituting this into (40) then gives $S \leq 2\sqrt{2}$. $\square$

[48] also provide a decoherence functional which satisfies this bound by constructing the decoherence functional for the standard EPRB-setup with quantum model given by the singlet and the projectors and settings given in the introduction. The functional to this behavior is constructed as in the proof to Lemma 3.3 as

$$D(a, a', b, b'|c, c', d, d') = \langle \psi_-|E_c^{x=0} E_{c'}^{x=1} E_d^{y=0} E_{d'}^{y=1} E_{b'}^{y=1} E_b^{y=0} E_{a'}^{x=1} E_a^{x=0}|\psi_-\rangle. \tag{42}$$

As expected by continuity this decoherence functional has 12 null directions corresponding to an effectively smaller 4-dimensional Hilbert space for the measured system than the 16 dimensions of the NC- space, thus making the atomic decoherence functional only positive *semi*-definite.

Concerning JQM, the continuity of (38) together with the fact that 42 does not lie on the boundary of $WPJQM$ (i.e. for no binary 16-vector $\mathbf{x}$ is $\mathbf{x}^T D\mathbf{x} = 0$) are already sufficient to show that JQM does not, for single copies of systems, recover Tsirelson's bond. In fact, every PR-box admits a valid decoherence functional with examples provided in [48, 49], e.g.

$$\hat{D}(0000|0000) = \hat{D}(1010|1010) = \hat{D}(1001|1001) \tag{43}$$

$$= \hat{D}(0101|0101) = \frac{1}{2} \tag{44}$$

$$\hat{D}(0101|0100) = \hat{D}(0001|1100) = \hat{D}(1001|1010) = \hat{D}(0101|1001) = \frac{1}{4} \tag{45}$$

$$\hat{D}(0001|0000) = \hat{D}(1010|1000) = \hat{D}(1001|1100) \tag{46}$$

$$= \hat{D}(0101|1100) = \hat{D}(1001|0001) = \hat{D}(0101|0001) \tag{47}$$

$$= -\frac{1}{2} \tag{48}$$

This immediately implies the following [49].

Lemma 3.7. *For single copies of boxes in the* $(2, 2, 2)$*-scenario,* $\mathcal{NS} \equiv JQM$

Proof. First show that for this scenario $\mathcal{NS} \subseteq JQM$. Recall first of all that NS is a convex polytope with 24 extremal boxes, 8 of which are PR-boxes and the other 16 local. This means that any box in $\mathcal{NS}$ can be expressed as a convex mixture of these extremal boxes with some 24-vector p of weights. By the above we know that every PR-box admits a decoherence functional and from $\mathcal{L} \subset \mathcal{Q} \subseteq SPJQM$ we also know that every local box admits a (diagonal) decoherence functional. It is furthermore easy to see that any convex mixture of the decoherence functionals corresponding to these extremal boxes defines a new decoherence functional satisfying Def. 4 and producing the right statistics for the box that is produced by the same weights p. Therefore every non-signalling distribution in this scenario admits a decoherence functional.

Finally, $JQM \backslash NS = \emptyset$ for any scenario since the additivity of the decoherence functional prevents the construction of decoherence functionals for any non-signalling distribution. This implies $\mathcal{NS} \equiv JQM$. $\square$

This is not true for more than one copy as will be discussed in Sec.5.

Finally, [49] also prove the following strengthening of Theorem 3.4.1.

Theorem 3.8. *For any distribution with a strongly positive decoherence functional in the* $(2,2,2)$*-scenario the correlators obey*

$$|\arcsin C_{x=0y=0} + \arcsin C_{10} + \arcsin C_{01} - \arcsin C_{11}| \leq \pi, \tag{49}$$

where each angle C_{xy} *lies between* $\frac{-\pi}{2}$ *and* $\frac{\pi}{2}$*.*

Proof. As in the proof to Theorem we need to prove the above expression without the absolute signs only. By defining $\theta_{xy} := \frac{\pi}{2} - \arcsin C(x,y)$ the expression becomes

$$\theta_{11} - \theta_{00} - \theta_{01} - \theta_{10} \leq 0 \tag{50}$$

By the strong positivity and Lemma 3.5 we further have $\cos\theta_{xy} = \langle x|y\rangle$ where $|x\rangle, |y\rangle$ are unit vectors. Now consider the two cases

- $\Theta := \theta_{00} + \theta_{01} + \theta_{10} \leq \pi$: If all vectors are coplanar, then $\theta_{11} = \Theta$, otherwise $\theta_{11} \leq \Theta$
- $\Theta > \pi$: Since $\theta_{11} \leq \pi$, $\Theta - \theta_{11} \geq 0$

In either case, (50) follows. □

Tsirelson proved that the satisfaction of (49) is necessary and sufficient for the existence of an "ordinary quantum model for the coorelators" (OQMC) of a given behavior, while it is only a necessary condition for the existence of an "ordinary quantum model for the probabilities" (OQMP) [50, 51, 9] which is just $\mathcal{Q}$ for this scenario. Theorem 3.8 then shows that for any behavior in $SPJQM$ for the $(2,2,2)$-scenario there exist two commuting pairs of hermitian operators $\{S^x\}$ and $\{S^y\}$ in a model with a normalized state ψ in a Hilbert space $\mathcal{H}$ such that, for any (x,y),

$$C_{xy} = \langle\psi|S^x S^y|\psi\rangle. \tag{51}$$

The fact that $SPJQM$ satisfies the Tsirelson bound means that, at least in certain directions of $\mathfrak{P}_{(2,2,2)}$, it exactly coincides with the quantum set, further motivating the suitability of PPI to produce a quantum analogue of Bell locality. However, by itself it surely cannot be sufficient, as $JQM \equiv \mathcal{NS}$ for the CHSH scenario also impressively illustrates. At least in this scenario, then, the "magic ingredient" to tighten the bound seems to be the strong positivity property, for the origin of which still no intuition could be provided, the constructability of the Hilbert space being a motivation but not an explanation for the prominence of this set. The next sections develop one possible approach to such an explanation.

4. Multiple boxes: Wirings and branching

One possibility that has so far been disregarded is that of combining several distributions, whether they be elements of the same correlation space or not. The simplest combinatory action is composition of independent systems. In order to discuss even only composition it is first necessary to consider the formation of joint scenarios.

4.0.2. *joint scenarios.* We have seen that the correlations space $\mathfrak{P}_\mathcal{S}$, the pair$\mathcal{S} = (\Xi, \mathcal{M})$ and a (n,m,d)-triple can all be constructed from another. It is convenient to look at the composition of scenarios in terms of the NC-space.

For a set of scenarios $\mathcal{T} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n\}$ we construct $\mathcal{S}(\mathcal{T}) = \{\Xi_\mathcal{T}, \mathcal{M}_\mathcal{T}\}$ by defining $\Xi_\mathcal{T} = \bigtimes_{i=1}^{n} \Xi_{\mathcal{S}_i}$ and $\mathcal{M}_\mathcal{T} = \{M_\mathcal{T}\}$ where $\times$ is the Cartesian product and $M_\mathcal{T}$ are the sets given by

$$A_{\mathcal{T}} = A \times \left(\bigtimes_{i \neq k}^{n} \Xi_{\mathcal{S}_i} \right) \tag{52}$$

for every outcome A of every measurement M in every scenario S_k. These $M_{\mathcal{T}}$ again form partitions of the new, larger NC-space of the joint scenario. $S(\mathcal{T})$ then defines a new correlations space. The CHSH scenario is then an example of a joint scenario composed of two marginal $(1, 2, 2)$-scenarios whose spacelike separation motivates their treatment as being independent.
tensor product and composition. We here adopt the tensor product as the standard representation of composition of distributions, see [52] for a general derivation of this rule for distributions under very weak assumptions.

Definition 4.1 (Tensor product rule)**.** For any two scenarios $\mathcal{S}_1$, $\mathcal{S}_2$, the joint distribution of any two conditional probability distributions $P \in \mathfrak{P}_{\mathcal{S}_1}, P' \in \mathfrak{P}_{\mathcal{S}_2}$ is given by [8]

$$P_{\mathcal{T}} := P \otimes P' \in \mathfrak{P}_{\mathcal{S}(\mathcal{T})}, \tag{53}$$

where $\mathcal{T} = \{\mathcal{S}_1, \mathcal{S}_2\}$.

4.1. **Wirings.** Less trivial ways of combining systems to produce a smaller number of effective boxes are described by means of "wirings". Introduced in [27], they have been studied in several contexts. In general one can distribute r copies of an n-partite box among up to rn parties which can then form $s \leq r$ groups, which, for group inputs $y_1, \ldots, y_s \in \{1, \ldots, p\}$ produce group outputs $b_1, \ldots, b_s \in \{1, \ldots, q\}$ to give the effective wired distribution. For a group k of l members its output b_k is determined by a strategy of subsequent measurements of each member. Upon reception of y_k, a first member measures $x_1 = f_1(y_1)$ to obtain outcome a_1 which is then passed o to a second party (and so on) who determine their measurement settings as $x_{j+1} = f_{j+1}(y_1, a_1, \ldots, a_j)$. Finally, the group output is determined as $b_k = g_k(y_1, a_1, \ldots, a_l)$ [55]. Taking for example the group $k = 1$, the wired box is then given as

$$\begin{aligned} &P^r(b_1, \ldots, b_s | y_1, \ldots, y_s) \\ &= \sum_{\substack{a_1,\ldots,a_l \\ \text{s.t.}\, g(a_1,\ldots,a_l)=b_1}} P^{\otimes r}(a_1, \ldots, a_l, b_2, \ldots, b_s | f_1(y_1), \ldots, f_l(y_1, a_1, \ldots, a_{l-1}), y_2, \ldots, y_s) \in \mathfrak{P}_{(s,p,q)}, \end{aligned} \tag{54}$$

This process, together with the example of a grouping is given in Fig. 2. In fact, these wirings, which so far are *deterministic* and in which the order of measurement within a group is taken to be fixed beforehand (meaning they are *static*), can be generalized to *stochastic, dynamic* wirings in which the group members share randomness and in which measurement outcomes determined the ordering of parties. Furthermore the wiring of boxes that are not identical copies, or originate from different scenarios, is described in the same way. The general form of all of these generalizations is in general notationally very cumbersome but perfectly possible. Here we will write $P_i^r, s \equiv \mathcal{W}_{\otimes s}(\otimes_i^s P_i)$ to denote any wiring that can be obtained by the wiring

[8]For a more general discussion about how joint distributions should be formed from marginals it is informative to consider the writings of the "minimalist" school around Fine that questions the necessity to produce the joint probability function by the arithmetic product, in particular surrounding the discussion of the interpretation of the Bell inequality violations. See [53, 54] for this debate.

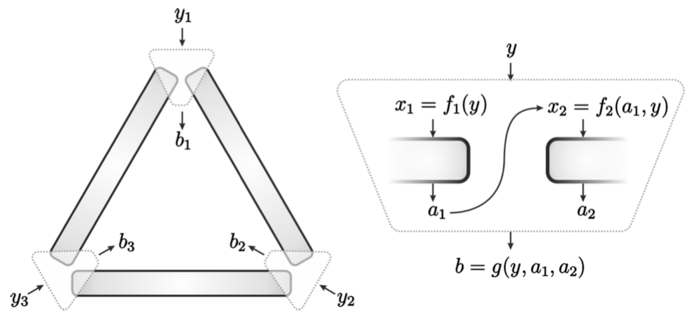


FIGURE 2. **Left:** Three bipartite boxes distributed among three groups of two parties each. **Right:** Determining the group output by passing on outcomes to subsequent paties. Taken from [55]
.

of r boxes $\{P_i\}$ by s parties. The tensor product index on the wiring map $\mathcal{W}$ is meant to indicate that these wirings are *local* operations.

Every wiring can be represented as a combination of grouping, composition and post-selection on outcomes, which together exhaust the set of classical operations one can perform on sets of distributions [36].

4.2. **Branching.** The introduction of multiple boxes and the possibility of wiring it brings along adds complications to defining the joint scenario that define the correlations spaces or wired boxes. This possibility of conditionalizing on measurement outcome of previous measurements enlarges the number of measurements that parties can perform. In terms of the partitions $\mathcal{M}$ of the NC-space Ξ, this possibility of "branching" is taken into account by extending $\mathcal{M}$, and the various sets defined from it. This is illustrated here for the case of a single decision but can easily be generalized for any number of decisions: Given two elements, $M, M' \in \tilde{\mathcal{M}}$ where M is included by M', meaning that one can produce the partition M by taking unions of the partition elements of M', and $M' \equiv M'(A)$, i.e. M' is determined by the outcome of M. Together these define a new partition $M_b(M, M')$ which contains all outcomes A' of $M'(A)$ such that $A' \subset A$ for all $A \in M$. Then, $\tilde{\mathcal{M}_b} = \{M_b\}$ is the set of all these branching partitions which, for trivial branching, clearly includes $\mathcal{M}$. we then define $\mathfrak{A}_b, \tilde{\mathbb{C}}_b$ as before. Since every fine-grained outcome for a branching measurement is also a fine-grained outcome for the element of $\tilde{\mathcal{M}}$, we can otherwise otherwise construct the branching joint scenario $\mathcal{S}_b(\mathcal{T})$ for $T = \{\mathcal{S}, \mathcal{S}\}$ as before, with $\mathfrak{P}_b$ also being uniquely determined by $\mathfrak{P}_{\mathcal{S}}$.

4.2.1. *The QMT branching sets.* The possibility of branching also imposes further decoherence constraints on the decoherence functionals corresponding to branching scenarios, in which which events appear as alternatives of a single measurement when before they did not (in other words branching reduces the number of orthogonal events). We can therefore introduce the following strengthenings of JQM and $SPJQM$.

Definition 4.2 (JQM_b)**.** For a given (n, m, d)-scenario $\mathcal{S}$ a conditional probability distribution $P(e)$ is in JQM_b iff there exists a decoherence functional $D(A, B)$ on 2^Ξ such that

(1) $D(A(e), A(e)) = P(e) \quad \forall A \in \mathcal{C}_b$
(2) $D(A(e), B(e')) = 0 \quad \forall A, B \in \mathfrak{A}_M, A \cap B = \emptyset, \forall M \in \mathcal{M}_b$,

where $\mathcal{C}_b$ and $\mathcal{M}_b$ are given by the joint branching scenario $\mathcal{S}_b$ which is determined by $\mathcal{S}$.

and similarly for $SPJQM_b$.

Definition 4.3 ($SPJQM_b$)**.** For a given (n, m, d)-scenario $\mathcal{S}$ a conditional probability distribution $P(e)$ is in $SPJQM_b$ if there exists a *strongly positive* decoherence functional D on 2^{Ξ} such that

- $D(A(e), A(e)) = P(e) \quad \forall A \in \mathcal{C}_b$
- $D(A(e), B(e')) = 0 \quad \forall A, B \in \mathfrak{A}_M, A \cap B = \emptyset, \forall M \in \mathcal{M}_b$,

where $\mathcal{C}_b$ and $\mathcal{M}_b$ are given by the joint branching scenario $\mathcal{S}_b$ which is determined by $\mathcal{S}$.

Of course one can also define branching extended versions for any other set considered so far but that they remain invariant under this extension is a simple corollary of Lemma 5.1 proven in the next section. We also get the following central result for the branching set $SPJQM_b$.

Theorem 4.1. *In any bipartite scenario* $SPJQM_b \equiv \tilde{Q}$.

Proof. This is proven in [38]. □

This is interesting because these sets were derived in very different contexts, convex optimization using semidefinite programming and quantum gravity research. This not only motivates, at least to some extent, that these sets capture an interesting structure that goes beyond either approach, and also, more pragmatically, allows for the translation of a number of results from information theoretic research to quantum measure theory (and vice versa) that are central to the thesis question, especially in the next section.

5. Closure and Stability

Given that wirings are taken to be (exhaustively) representing the possible actions local experimenters can perform on their part of a shared state, it is a natural requirement that sets corresponding to physical theories should be closed under them.

First define some basic notions of closure

Definition 5.1 (Closure under composition)**.** A set of conditional probability distributions $S(\mathcal{S}) \subseteq \mathfrak{P}_{\mathcal{S}}$ that is defined for some set of scenarios $\mathscr{S} = \{\mathcal{S}\}$ is *closed under composition* iff, for any $P \in S(\mathcal{S}), P' \in S(\mathcal{S}')$,

$$P \otimes P' \in S(\mathcal{S}_J), \tag{55}$$

where $\mathcal{S}_J, \mathcal{S}, \mathcal{S}' \in \mathscr{S}$ and $\mathcal{S}_J$ is the joint scenario constructed from $\{\mathcal{S}, \mathcal{S}'\}$.

Definition 5.2 (Closure under post-selection)**.** A set of conditional probability distributions $S(\mathcal{S}) \subseteq \mathfrak{P}_{\mathcal{S}}$ that is defined for some set of scenarios $\mathscr{S} = \{\mathcal{S}\}$ is *closed under post-selection* iff, for any $P(a_1, \ldots, a_n | x_1, \ldots, x_n) \in S(\mathcal{S})$,

$$P(a_2, \ldots, a_n | x_1, \ldots, x_n, a_1) \in S(\mathcal{S}') \tag{56}$$

where $\mathcal{S}, \mathcal{S}' \in \mathscr{S}$.

One immediate and useful result for closure under post-selection is

Lemma 5.1. *A set S such that $S \subseteq JQM$ is closed under post-selection if and only of $S \equiv S_b$, i.e. it is equivalent to its branching extension*

Proof. $\Rightarrow$: Assume S is closed under post-selection. Then for any decoherence functional D corresponding to some $P \in S$ there exists a valid decoherence functional D_b which reproduces the right statistics for any post-selection $P_c \in S$ of P. But this is equivalent to $S \equiv S_b$. $\Leftarrow$: Assume $S \equiv S_b$. Then for any decoherence functional D corresponding to some $P \in S$ there exists a valid decoherence functional D_b which reproduces the right statistics for any post-selection P_c of P with $P_c \in S_b \equiv S$. Therefore S is closed under post-selection. □

Since we already know that $\mathcal{L}, \mathcal{Q}$ and $\mathcal{NS}$ are closed under wirings, it follows that $\mathcal{L} \equiv \mathcal{L}_b$ etc. in any scenario. Whether $SPJQM \equiv SPJQM_b$ and $JQM \equiv JQM_b$ is however unclear. It is also easy to see that the construction of the branching joint scenario $\mathcal{S}_b$ defines a new kind of composition for scenarios that is different from the one in Sec. 4.0.2. This will be investigated further in Sec. 6.4.

5.1. **Physical Closure.** With the same motivation but by also including convexity (possible states should also be possible to prepare probabilistically) [56] define "physical closure" as a necessary property of any set of distributions that is meant to correspond to a physical theory.

Definition 5.3 (Physical Closure)**.** A set of n-partite boxes S defined for any scenario $\mathcal{S}$ is "physically closed"if

(1) S is convex
(2) S is closed under wiring, i.e. for any collection of r boxes $\{P_i\} \subseteq$ that are equally distributed among by s parties, $\mathcal{W}_{\otimes s}(\otimes_i^s P_i) \in \mathfrak{P}'_{\mathcal{S}}$ is an element of S as it is defined for $\mathcal{S}'$.

In this text we will refer to "physically closed" simply as "closed" unless stated otherwise and we will denote any general physical operation $\mathcal{W}_p(\cdot)$. The exact sense in which these operations are physical, however, is possibly unclear to someone who seeks the notion of "physicality" outside an operational framework, as is the case in quantum measure theory. In Sec. 7 and, in particular, in Sec. **??** this will be investigated further.

The requirement on sets to be closed under physical operations can then be formulated in the following axioms.

Axiom of Physical Closure (APC)

The set of behaviors derived from any physically
feasible theory is physically closed.

Even though this seems to be a very weak requirement on reconstructive principles and the sets they produce, APC rules out a surprising number of sets in the NS-polytope, as the next section shows.

5.2. **Some results about closed sets.** $\mathcal{L}, \mathcal{Q}$ and $\mathcal{NS}$ are all instances of closed sets. Indeed, $\mathcal{L}$ is the smallest closed set because it is the convex hull of the deterministic extremal boxes and every deterministic box can be generated by using a trivial wiring.

Given APC, it is a natural question to ask for the general properties of closed sets inside the non-signalling polytope. Given their simple structure and the closure of $\mathcal{L}$ and $\mathcal{NS}$, general polytopes seem a good starting point for such an investigation. [27] discuss two families of non-signalling polytopes for the CHSH scneario that violate APC and are illustrated in Fig. 3.

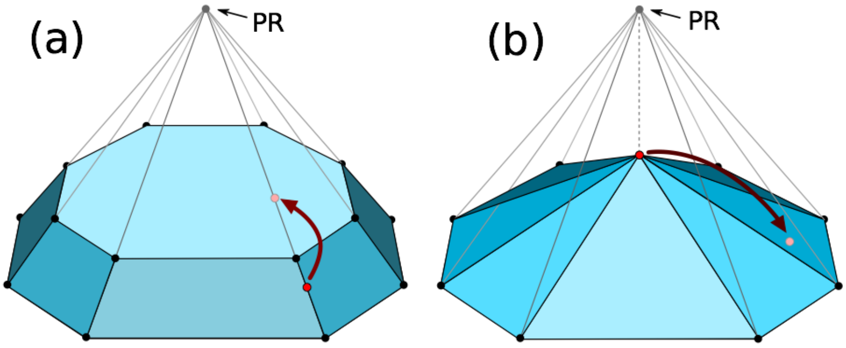


FIGURE 3. Two simple polytope families that are ruled out by APC, in particular the requirement of closure under wiring: (a) the polytope with a CHSH-cutoff, (b) the squeezed polytope given by isotropic noisy PR-boxes. Taken from [27]

.

CHSH cutoff. One way of producing polytopes is by considering the polytopes $\mathcal{R}^S_{CHSH}$ inside $\mathcal{NS}$ of boxes whose correlators satisfy a certain CHSH value S where $2 \leq S \leq 4$. The structure of these polytopes looks just like $\mathcal{L} = \mathcal{R}^2_{CHSH}$ with "postivity" facets connecting deterministic extremal boxes and a "CHSH"-facet for every CHSH-inequality $CHSH \leq S$ with the LHS being a function of the correlators. Specifically, every non-deterministic vertex of $\mathcal{R}^S_{CHSH}$ can be specified as a nested convex mixture of PR-boxes and deterministic, local boxes. One can produce wired boxes lying outside of this polytope (for the wirings considered in [27], the functions g and inputs y are Boolean and all boxes grouped into two parties so that the resulting box will again lie in $\mathfrak{P}_{\mathcal{S}_{(2,2,2)}}$) by "distilling" boxes lying on the edges of the polytope $\mathcal{NS}\backslash\mathcal{L}$, which overlaps with $\mathcal{R}^S$ for any $2 \leq S \leq 4$, i.e. wiring them such that the resulting box lies on the same edge but with increased CHSH value.

Squeezed polytopes. The second class of general non-signalling polytopes that are found not to be closed are isotropic noisy PR-boxes $\mathcal{R}^S_I$. These are convex mixtures of a PR-box with isotropic (white) noise, i.e. for the $(2, 2, 2)$-scenario,

$$P_{PR}(\epsilon) = \epsilon P_{PR} + (1-\epsilon)\frac{1}{4}\mathbb{I}. \tag{57}$$

Importantly, every bipartite box can be turned into the form (57) by local operations without changing its CHSH value S [28], with $S = 4\epsilon$. Therefore $\mathcal{R}^S_I$ geometrically gives a "squeezed" version of the NS-polytope in which all local boxes remain invariant (see Fig.**??**). A single AND wiring is sufficient to produce boxes outside $\mathcal{R}^S_I$ for any $\frac{8}{3} < S < 4$ For smaller values of S, no proof exists of their not being closed but it has been conjectured that they're not closed as well.

Other relevant sets that have been shown to not be closed are the Uffink [57] and Pitowsky set [58], both in [27].

That the above two classes of polytopes are not closed motivates the question whether APC allows for only very small number of closed sets including the local, quantum and non-signalling sets, or indeed only those three. This question has to be answered in the negative as the following results show.

5.2.1. *There exists an infinity of closed super-quantum sets in the CHSH scenario.* While the last section gives an example of a super-quantum set that is closed under wiring, there exist other examples that conclusively prove the existence of *physically* closed sets.

Lemma 5.2. *Q^n is closed for all $n \geq 1$.*

Proof. This is proven in [56]. □

We also have the following Lemma

Lemma 5.3. *Assume that $P \neq NP$ and that Kirchberg's conjecture holds. Then, for any $k \in \mathbb{N}$, there exist k numbers $N1, 1 \ldots, N_k$ and a bi-partite scenario with binary outputs where $\left\{Q^{N_i}\right\}_{i=1}^{k}$ are all different.*

Proof. This is proven in [56]. □

Together these imply that, under the weak assumptions of the lemma, there exists an infinity of super-quantum closed sets in the CHSH scenario.

As is proven in [56], there also exists an infinity of bipartite closed super-quantum polytopes, as there at least two families of polytopes, called "Ghost world" and "Twin world" which are closed and of are infinite in size. Furthermore, [56] provide an example of a set that is tripartite, super-quantum and closed. In fact, [36] prove that the more general n-partite $\tilde{Q}$ is closed under wiring. Together with its convexity in any scenario this implies that

Corollary 1. *$\tilde{Q}$ is physically closed.*

meaning that there exists at least one super-quantum closed set (that isn't $\mathcal{NS}$) in every scenario in which $\tilde{Q} \neq \mathcal{Q}$.

In the light of the above results one might hope that the quantum set then is characterizable as the smallest closed set exhibiting non-locality, at least in the CHSH-scenario? The answer is no.

5.2.2. *There exist an infinite number of closed non-local subquantum sets.* The question about closed sets is closely related to the research program of non-locality distillation. Recall first the isotropic noisy PR-box $P_{NL}(\epsilon)$ from (57) which was is equivalent to all boxes of with the same CHSH value $S = 4\epsilon$ up to local operations and where $\epsilon = \frac{1}{2}, \frac{1}{\sqrt{2}}, 1$ give extremal CHSH boxes of the $\mathcal{L}, \mathcal{Q}, \mathcal{NS}$ sets respectively.

The problem of non-locality distillation is whether identical copies of isotropic boxes with $\epsilon > \frac{1}{2}$ can be wired to give a box with $\epsilon' > \epsilon$. All attempts to solve this problem so far have failed. One result from [59] of interest to us is that there exists an infinity of values $\epsilon \in (\frac{1}{2}, \frac{1}{\sqrt{2}})$, forming the set $\mathcal{E}$, for which distillation is impossible. This implies that each of the infinite number of sets $Q_\epsilon, \epsilon \in \mathcal{E}$ which we define to be the sets obtainable by locally acting on $P_{NL}(\epsilon)$ are closed and for each two elements $\epsilon, \epsilon' \in \mathcal{E}, \epsilon < \epsilon'$ we have the strict nesting $\mathcal{L} \subset Q_\epsilon \subset Q_{\epsilon'} \subset \mathcal{Q}$. It is an open question whether $mathcalE$, or its extension $\overline{\mathcal{E}}$ to the all values of $\frac{1}{2} < \epsilon < 1$ for which distillation is impossible, is discrete or continuous. This is an interesting question in that one of either of turned out to be discrete (countable and complete in the corresponding regime, then this would allow to label every physical theory (in the sense of CPS) by some $\epsilon \in \overline{\mathcal{E}}$.

Finally, from Corollary 1 and Lemma 2.1 we also have that $\mathcal{Q}$ is not even the only set recovering the Tsirelson bound, another property of quantum correlations one might have hoped for APC to single out. All together, it is clear that even though closure is certainly far from being a trivial requirement on sets of behaviors it is also highly insufficient to single out out the quantum set by itself.

5.3. **Stability under composition - Maximality.** So far we have neglected an important subtlety: A reconstructive principle may allow for several closed sets that however intersect non-trivially. To discuss this first introduce the *closure* of a set R as the smallest closed set C such that $R \subseteq C$. Define then for any two closed sets S_1, S_2 the set $S_1 + S_2$ as the closure of $S_1 \cup S_2$. We also have that $S_1 \cap S_2$ is also a closed set. Characterization of $S_1 + S_2$ and $S_1 \cap S_2$ are in general problems of very different complexity, the latter being much easier to decide than the former [56]. In this terminology the above subtlety is easily expressed as the fact that for a number of closed sets S_i that are consistent with with some principle Z, i.e. $S_i \subseteq \mathcal{Z}$, $\forall i$, where $\mathcal{Z}$ is the set of all behaviors consistent with Z, it is in general not the case that $\bigcup_i S_i$ is compatible with the principle, i.e. $\sum_i S_i \nsubseteq \mathcal{Z}$.

The additional property required of the *principle* to ensure that this is not the case is that it is stable under composition [56].

Definition 5.4 (Stability under composition)**.** A device-independent principle Z is stable under composition iff, for any scenario $\mathcal{S}$ and any pair of physically closed sets S_1, S_2 compatible with Z, the set $S_1 + S_2$ is also compatible with Z.

For convenience, refer to principles that are stable under composition as stable principles. One can translate this, as in the case of closure, into an axiom:

Axiom of Stability under composition (ASC)
Any device-independent principle from which a feasible theory
can be constructed is stable.[9]

To clarify the structure that ASC imposes on sets of behaviors further consider the following definition.

Definition 5.5 ($\times$-Maximality)**.** For some scenario $\mathcal{S}$, a set $S \subset \mathfrak{P}_{\mathcal{S}}$ that is closed under some action $\times$ is $\times$-maximal under a principle Z with corresponding set $\mathcal{Z}$ if $S \subset \mathcal{Z}(\mathcal{S})$ and $\exists P \in S$ such that

$$P \otimes P' \notin \mathcal{Z}(\mathcal{S}'(\mathcal{T})) \quad \forall P' \in Z \backslash S, \tag{58}$$

where $\mathcal{T} = \{S, S\}$.

In simple terms, a set is $\times$-maximal under a principle if it cannot be enlarged without yielding violations of that principle under the $\times$-action. Note that the existence of a maximal set trivially implies that $\mathcal{Z}$ is not closed under $\times$.

The connection between maximality and stability is made clear by the following Lemma.

Lemma 5.4. *A principle Z is stable iff it has no non-trivially intersecting $\mathcal{W}_p$-maximal sets.*

Proof. By definition, for any set R that non-trivially intersects with a $\mathcal{W}_p$-maximal set S, $\mathcal{Z} \subset R + S$. So if R is closed and satisfies Z, then Z is unstable under composition. Conversely,

[9] Note that one could have defined the notion of stability under composition for sets that are not closed, however, it is implicit in calling a set R "compatible with Z" only if $R \subseteq \mathcal{Z}$. Therefore we can w.l.o.g. define stability for these closures.

suppose that all closed sets in $\mathcal{Z}$ are nested. Then for any closed sets $S_1, S_2 \subset \mathcal{Z}$, $S_1 + S_2$ is either S_1 or S_2 and therefore Z is stable under composition. $\square$

Note that we required $\mathcal{W}_p$-maximality of the closed sets only to prove the right implication. A theorem equivalent to the above for the CHSH scenario, namely that there exists a bipartite Bell-type linear inequality unstable under composition iff there exist two bipartite closed sets with non-trivial intersection, is proven in [56]. Also it is clear that for stable principles there exists only one $\mathcal{W}_p$-maximal set by $\mathcal{W}_p$-closure of the maximal set. Call this set then the *wiring domain* of Z in the sense that any physical wiring in the sense of Def. 5.3 that is guaranteed to produce behaviors that are consistent with Z can have this set as its largest domain. Call the corresponding $\otimes$-maximal set for principles stable under $\otimes$-composition its *compositional domain*.

We an therefore see that APC and ASC impose, despite their weakness, impose highly non-trivial requirements on the candidate principles for a physical theory and single out a natural set, its wiring domain, for physical distributions in a universe governed by these principles.

5.4. **Example: $\tilde{Q}$ as the wiring domain of local orthogonality.** One highly relevant example of W_p-maximality that has recently been found serves to illustrates these concepts.

5.4.1. *Local Orthogonality.* Recently is was shown proven that any principle that is to single out quantum correlations for an arbitrary number of parties cannot be based on bipartite information concepts [26]. In reaction to this [15] introduce an intrinsically multipartite principle into reconstructive research with similar ideas having been discussed previously in [60]. The basic notion for this principle is that of exclusiveness between possible events: For any (n, m, d)-scenario, two events $e \equiv (a_1, \dots, a_n | x_1, \dots, x_n)$ and $e' \equiv (a'_1, \dots, a'_n | x'_1, \dots, x'_n)$ are "(locally) orthogonal" if, for at least one party k, $x_k = x'_k$ while $a_k \neq a'_k$. Sets of events $\{e_i\}$ are then called "orthogonal" or "exclusive" if all their members are (pairwise) orthogonal, denoted $e_i \perp e_j$. With this terminology [15] then define

Definition 5.6 (Local Orthogonality (LO))**.** For any exclusive set, every element of the set of physically allowed distributions satisfy

$$\sum_i P^{\otimes k}(e_i) \leq 1 \tag{59}$$

for any number of copies k.[10]

Call $\mathfrak{S} \equiv \{S_i\}$ the set of all exclusive sets S_i for a given scenario. By definition, $\mathcal{M} \subset \mathfrak{S}$, with every $M \in \mathcal{M}$ yielding equality in (59). The same is true for any coarse-graining of elements in $\mathcal{M}$.

The origin and interpretation of LO, in particular its relation to logic, have been investigated in several studies, see [61, 62, 63]. In terms of sets of distributions, one can define

Definition 5.7 ($\mathcal{LO}^k$)**.** For a given (n, m, d)-scenario a conditional probability distribution $P(e)$ is in the set $\mathcal{LO}^k$ if $P^{\otimes k}$ satisfies Local Orthogonality in the (kn, m, d)-scenario, where

[10]Note that considering composition (as opposed to any general $\mathcal{W}_p(\cdot)$ here is enough because the number of events doesn't change and the question which events are alternative outcomes of some branching measurement can easily be inferred from the original scenario, especially the orthogonality graph, see below.

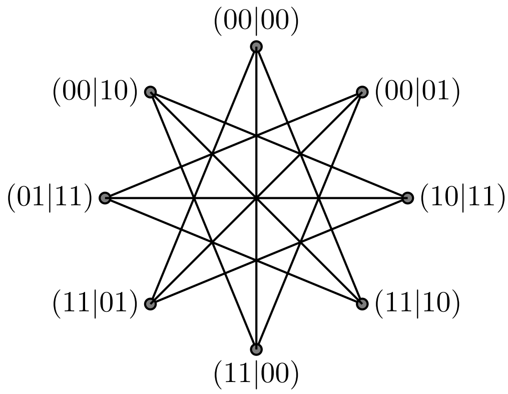


FIGURE 4. Orthogonality graph of all non-zero weighted events for the PR-box in the CHSH-scenario. Labels denote events

$k \in \mathtt{N}$. In particular,

$$\mathcal{LO}^\infty := \bigcap_{k \in \mathtt{N}} \mathcal{LO}^k. \tag{60}$$

Local Orthogonality can be conveniently studied in terms of graphs to represent the relations between different events [15, 64].

Definition 5.8 (Orthogonality Graph)**.** Given a (n, m, d)-scenario, its "orthogonality graph" $O(\mathcal{S})$ is an undirected graph whose vertex set is the set of all events for this scenario with adjacency relation

$$e_i \sim e_j \iff e_i \perp e_j. \tag{61}$$

One can study distributions over the corresponding scenario by constructing non-negative-weighted subgraphs (G, p_i) of $O(\mathcal{S})$. The assumption of non-contextuality in this setting is the assumption that the weights (interpreted as probabilities of the events' occurrence) are independent of G. Local orthogonality can be re-stated in graph theoretic terms as the requirement that the sum of weights of every clique in $O(\mathcal{S})$, that is, every subset of vertices inducing a complete graph (a graph in which every vertex is connected to every other vertex) is less or equal to one. By the assumption of non-contextuality a sufficient condition for the satisfaction of local orthogonality by any clique is then its satisfaction by all maximal cliques, that is, cliques from which no new clique can be generated by adding vertices.

A first result one can prove easily here is that $\mathcal{LO}^1 \equiv \mathcal{NS}$. This is proven in [15] and, differently, in [60]. However, the intrinsically multipartite character of LO becomes apparent for $k \geq 2$. For instance, the PR-box is ruled out by LO already for two copies [15]. To see this, first note that, by the definition of the PR-box in (8), the only non-zero probabilities (weights) of events for k identical copies of the PR-box are 2^{-k} (for any event such that $a_i \oplus b_i = x_i y_i, i \in \{1, \ldots, k\}$). This means that in the exclusivity graph consisting only of vertices with non-zero weight finding a clique with 2^k members is sufficient to prove the above. [15] find cliques with 5 members for $k = 2$ (that is, in the graph given by the strong product of two copies of the graph given in Fig.4). Hence, LO excludes the PR-box.

Really, of course, one is interested in characterizing $\mathcal{LO}^\infty$ as this is the only "physical" set. [55] prove the following:

Theorem 5.5. *$\mathcal{LO}^\infty$ is closed under wirings*

Proof. This follows by Lemma 5.6 below for $k = \infty$ together with the fact that if $P^r \in \mathcal{LO}^1$ can be produced by wiring, then so can any number of copies $(P^r)^{\otimes r'}$, meaning that $P^r \in \mathcal{LO}^\infty$. $\square$

Lemma 5.6. *For any $P \in \mathcal{LO}^k, P^r \in \mathcal{LO}^1$*

Proof. This is proven in [55]. $\square$

Note that this does not prove that $\mathcal{LO}^\infty$ is *physically* closed. Indeed, it is not clear whether $\mathcal{LO}^\infty$ is convex in every scenario.[11] An interesting corollary of Lemma 5.6 is that wirings are "useless" to produce violations of local orthogonality, i.e. if a wiring of r boxes P violates LO^1 then P violates LO^k, meaning that $P^{\otimes k}$ already violates LO^1. In [32] it is proven that $\mathcal{Q} \subset \mathcal{LO}^\infty$ for the CHSH-scenario, meaning that LO cannot characterize the quantum set.

5.4.2. *Maximality under LO.* The most important result and the one illustrating maximality, however, is that $\tilde{Q}$ is the wiring domain of an extended LO under the assumption that any physical set should include $\mathcal{Q}$. First prove $\otimes$-maximality.

Theorem 5.7. *$\tilde{Q}$ is $\otimes$-maximal under*

Proof. Given any collection of events $\{e_i\}$ for some scenario $\mathcal{S}$ consider another collection of events $\{f_i\}$, generally of another scenario $\mathcal{S}'$, such that $e_i \not\perp e_j \Rightarrow f_i \perp f_j$. Then, by definition $O(\mathcal{S}) = \overline{O(\mathcal{S}')}$, i.e. the orthogonality graphs of the two scenarios are complementary to another. By virtue of the independence between events e_i and f_i implied by this relationship, the joint probabilities of the events $e_i f_i$ for *any* distributions over the joint scenario $\mathcal{T} = \{\mathcal{S}, \mathcal{S}'\}$ are $P_i P_j'$, following the tensor product rule for distributions.

Since LO applies to this joint behavior as well we require

$$\sum_i P_i P_i' \leq 1 \tag{62}$$

This inequality can be rephrased as follows: By Local Orthogonality a distribution $\{P_i\}$ has to satisfy (62) for *any* other P_i' allowed by $\overline{O(\mathcal{S})}$. Given a normalized vector $|\psi\rangle$ and orthonormal basis $\{|f_i\rangle\}$ this bounds the P_i as

$$\left(\sum_i P_i\right) \min |\langle\psi|f_i\rangle|^2 \leq \sum_i P_i |\langle\psi|f_i\rangle|^2 \leq 1 \tag{63}$$

$$\leftrightarrow \quad \sum_i P_i \leq \frac{1}{\max|\langle\psi|f_i\rangle|^2} \tag{64}$$

Since this has to hold for any $|\psi\rangle$ and $\{|f_i\rangle\}$ on $\mathcal{S}'$, we further minimize over all models such that

$$\sum_i P_i \leq \frac{1}{\min \max|\langle\psi|f_i\rangle|^2}. \tag{65}$$

The RHS of (65) is the Lovsz function on $O(\mathcal{S})$ [65, 66] which in [64] is proven to be the maximum CHSH-value allowed by almost quantum correlations. $\square$

[11] In fact, in [32] it is proven that there exist some contextuality scenarios for wich $\mathcal{CE}^\infty$ is not convex.

We have skimmed over one detail: That the LO principle should apply to the product scenario of two different scenarios is not actually part of the original formulation of LO, where copies of the same distribution are considered. Given their independence by construction it nevertheless seems like a reasonable requirement. This shows that $\tilde{Q}$ is $\otimes$-maximal but not that it is the compositional domain of LO. Noting that the models on the second scenario used in the proof to Theorem 5.7 to obtain the bound were any quantum model, assuming that any physically feasible theory should contain $\mathcal{Q}$ then rules out any other maximal set that non-trivially intersects $\tilde{Q}$ and, by Lemma 5.4, makes $\tilde{Q}$ the compositional domain of LO.[12] Finally, from the above results that wirings are useless to produce violations of LO ([55]) together with the physical closure of $\tilde{Q}$ it follows that the almost quantum is also the wiring domain of LO!

This argument provides an illustration of how local orthogonality (extended to apply to combinations of different scenarios), APC, ASC together with the assumptions of the tensor product rule for independent distributions and containment of the quantum set in any future theory single out the almost quantum set. We now see that the recovery of the Tsirelson bound by almost quantum correlations (Lemma 2.1) can be explained as a special case of the maximal correlations that are allowed by these various principles for the CHSH scenario if the further assumptions hold true.

What about the quantum measure theoretic sets? In the next section we show that, under the assumption of a composition rule for decoherence functionals, a similar result can be obtained for $SPJQM_b$.

6. Closure and maximality in quantum measure theory

This section brings together the different concepts and results discussed in previous chapters by asking what one learns about quantum set theory by studying in the light of APC and ASC

6.1. **Closure.** Are JQM and $SPJQM$ closed? Since both these sets are defined by the existence of a decoherence functional, it is necessary to discuss closure and maximality in terms of the latter. Firstly, it is clear that both sets are convex by convexity of the decoherence functional. Regarding composition, by definition the joint behavior $P_J = P \otimes P' \in \mathfrak{P}_{\mathcal{T}=\{\mathcal{S},\mathcal{S}'\}}$ of two independent behaviors $P \in \mathfrak{P}_{\mathcal{S}}, P' \in \mathfrak{P}_{\mathcal{S}'}$ that admit decoherence functionals $D \in B(\mathfrak{H}_{\mathcal{S}}), D' \in B(\mathfrak{H}_{\mathcal{S}'})$ defined over the two vector spaces $\mathfrak{H}_{\mathcal{S}}, \mathfrak{H}_{\mathcal{S}'}$ respectively, will be an element of JQM if there exists a valid decoherence functional $D_J \in B(\mathfrak{H}_{\mathcal{T}})$ for P_J.

For functionals in $SPJQM$ this construction is easy as implied by this simple theorem.

Lemma 6.1. *$SPJQM$ is closed under composition*

Proof. For any two behaviors $P, P' \in SPJQM$ with atomic decoherence functionals $\hat{D}, \hat{D}'$ we can produce a valid decoherence functional for $P_J = P \otimes P'$ as $\hat{D}_J = \hat{D} \otimes \hat{D}'$. It is easy to check the validity of this functional: The tensor product of two positive semi-definite matrices is again a positive semi-definite matrix and also hermiteanity, additivity and the production of the right joint behavioral statistics are warranted by this construction. □

[12][32] implicitly make this argument in defining an extended consistent exclusivity/ local orthogonality principle involving closure under composition with any quantum model.

For a general element of JQM this is not as simple in that the positivity property of decoherence functionals is not in general conserved. In analogy to the tensor product rule for distributions we therefore make the following assumption:

Existence of product decoherence functional (EPD)
For any two independent distributions $P \in \mathfrak{P}_{\mathcal{S}}, P' \in \mathfrak{P}_{\mathcal{S}'} \in JQM$,
if $P_J = P \otimes P' \in JQM$, then there exists a joint atomic decoherence functional of the form
$\hat{D}_J = \hat{D} \otimes \hat{D}' \in B(\mathfrak{H}_J)$ where
$\hat{D} \in B(\mathfrak{H}), \hat{D}' \in B(\mathfrak{H}')$ are two valid atomic decoherence functionals for the two distributions.
In particular, if either one of the distributions is an element of $\mathcal{Q}$, then the correct
marginal decoherence functional is given by the construction
used in the proof of Lemma 3.3.

Some remarks EPD: Recall first that these vector spaces are not Hilbert spaces in that they do not have a well defined positive inner product. Secondly, the axiom does not state that for *any* valid decoherence functionals D_A, D_B, if their tensor product is not a valid decoherence functional, then the joint behavior is not in JQM. This would be an unreasonable and overly strong requirement. Because of this, the axiom is weaker than a general composition rule in that it does not provide a recipe for producing a joint decoherence functional from given marginal functionals (except for the case of two quantum models). At the same time, it prescribes the tensor product algebra for the cases in which a joint decoherence functional exists. In this sense it does establish a “chopping” rule.

Finally, the specification for the form of the functionals for quantum models is first of all necessary for the upcoming proof of maximality but can also be justified otherwise by considering a convex mixture of the decoherence functionals for the PR-box given in (48) and local noise for $(2, 2, 2)$-scenario,

$$\hat{D}(\epsilon) = \epsilon \hat{D}_{PR} + (1 - \epsilon)\frac{1}{16}\mathbb{I}, \tag{66}$$

which gives a valid decoherence functional for the isotropic PR-box from (57) for any $0 \leq \epsilon \leq 1$. Of course, at least for $\epsilon \leq \frac{1}{\sqrt{2}}$ we know that there exists a strongly decoherence functionals for this distribution, while (66) is not a positive matrix for any ϵ. Without adding the additional clause in EPD the tensor product $\hat{D}^{\otimes 2}(\epsilon)$ for two independent noisy PR-boxes would be assumed to be a valid decoherence functional for this range of ϵ because we know that $\mathcal{Q}$ is closed. However, due to the sparseness of (48) it is easy to check that this tensor product does not produce valid decoherence functionals for $P_{PR}^{\otimes 2}(\epsilon)$. Indeed simple testing in Matlab gave this result for values of ϵ as low as 0.275, illustrating the importance of keeping in mind non-uniqueness of the decoherence functional and also motivating the above as a sufficiently weak but adequate assumption.

What about the physical closure of these sets? In the following section it is proven that JQM (and by implication JQM_b) is not closed under composition and therefore also not closed physically. For $SPJQM$ it is still unclear whether it is physically closed but we can easily prove physical closure of $SPJQM_b$ by recalling the connection between branching and closure under post-selection from Lemma 5.1.

Lemma 6.2. *$SPJQM_b$ is physically closed.*

Proof. For the bipartite case we know this already from Theorem 4.1 and Corollary 1. More generally, since $SPJQM_b$ is by definition equivalent to its branching extension, it is closed under post-selection by Lemma 5.1. Since $SPJQM_b \subseteq SPJQM$, closure under wiring is warranted by the additional decoherence requirements imposed on a decoherence functional in the definition of $SPJQM_b$ together with its positive-semidefiniteness, the combination of which ensures that any joint decoherence functional that is both strongly positive and closed under post-selection. Finally, $SPJQM_b$ is convex because any convex mixture of positive-semidefinite matrices is again a positive-semidefinite matrix and every convex mixture of behaviors that are closed under post-selection is again closed under post-selection. □

6.2. **Maximality.** Consider next maximality, where the natural principle to study is PPI. [67] proves the following theorem which links the example developed in the last subsection to quantum measure theory.

Theorem 6.3. *$JQM \subseteq LO^\infty$ for any non-locality scenario.*

Proof. See [67] for a proof. □

This is not only interesting in its own right, opening up for the possibility of the two principles actually being equivalent, at least for those families of scenarios for which they are both defined. It also suggests the possibility of simply translating the result of the last section onto $SPJQM_b$ and PPI since we know that for bipartite scenarios the equivalence result between $\tilde{Q}$ and $SPJQM_b$ from Theorem 4.1 and furthermore the simple property that, if a set S is $\times$-maximal under some set $\mathcal{Z}$, then it is also $\times$-maximal under any set $\mathcal{Z}'$ for which $S \subseteq \mathcal{Z}' \subseteq \mathcal{Z}$. However, things are not that easy because LO is a general contextuality principle and the maximality proof from the last section works for any contextuality scenario, meaning that it proves W_p-maximality for sets of contextuality scenarios for which the quantum measure theoretic sets are not defined. This is more rigorously discussed in Appendix A, since it requires some more graph theoretic considerations that otherwise are not relevant for the main text.

6.2.1. *$SPJQM_b$ is the wiring domain of PPI..* Despite the unavailability of the above maximality result to quantum measure theoretic sets, it is nevertheless suggestive and very likely to hold similarly here. Indeed one can prove the following number of results.

First of all consider the following simple matrix lemma.

Lemma 6.4. *Let $\mathcal{N}$ be the set of all positive-semidefinite hermitian matrices of any dimension $n \times n$ and let $\mathcal{M}$ be the set of all hermitian matrices M_{ij} of any dimension $m \times m$ such that, for any vector $\mathbf{w} \in \{0,1\}^m$, $w_i M_{ij} w_j \geq 0$ where we use Einstein summation. If for some $M \in \mathcal{M}$, $M \otimes N \in \mathcal{M}$, $\forall N \in \mathcal{N}$, then $M \in \mathcal{N}$.*

Proof. Assume that $M \notin \mathcal{N}$. Then there exists a vector $\mathbf{v} \in \mathbb{C}^m$ such that $v_i^* M_{ij} v_j < 0$. Now define the $m \times m$-dimensional hermitian matrix $N_{kl} = v_k^* v_l \equiv v_{kl}$ together with the vector $\mathbf{w} \in \{0,1\}^{m^2}$ as $w_{ik} = \delta_{ik}$. Then

$$w_{ik} M_{ij} N_{kl} w_{jl} = \delta_{ik} M_{ij} v_{kl} \delta_{jl} = v_i^* M_{ij} v_j < 0. \tag{67}$$

But this contradicts the assumption since $u_i^* v_k^* v_l u_j = |\mathbf{uv}^*|^2 \geq 0$ so that $N \in \mathcal{N}$ □

The analogy to the definition of $\otimes$-maximality is clear, however in the case of decoherence functionals one needs to be more careful. This is because $N = |v^*\rangle\langle v^*|$ is not in general a valid atomic decoherence functional for any model since it does not warrant decoherence of any alternative outcomes. Still, only a slightly more analogous proof does the job.

Theorem 6.5. *$SPJQM$ is $\otimes$-maximal under PPI if EDP is true.*

Proof. For $SPJQM$ to be $\otimes$-maximal under PPI means that, if for any two scenarios $\mathcal{S}, \mathcal{S}'$, $P_J = P_{\mathcal{S}} \otimes P_{\mathcal{S}'} \in JQM\ \forall P_{\mathcal{S}'} \in SPJQM$, then $P_{\mathcal{S}} \in SPJQM$. Now let $\hat{D}$ be any valid $m \times m$-dimensional atomic decoherence functional for a behavior $P_{\mathcal{S}} \in JQM$ in scenario $\mathcal{S}$. Assume that $P_{\mathcal{S}} \notin SPJQM$. Then there exists a vector $|v\rangle \in \mathbb{C}^m$ which we can take to be normalized w.l.o.g., such that $\langle v|\hat{D}|v\rangle < 0$.

Now consider a quantum model consisting of a normalized vector $|v^*\rangle \in \mathbb{C}^m$ and two projective sets $\{E_a\}, \{F_0, F_1\}$ defined by $E_a = |a\rangle\langle a|$, for $a = 1, ..., m$, i.e 1-rank projectors onto the computational basis of $\mathbb{C}^m$, and $F_0 = \frac{1}{m}\sum_{j,k} |j\rangle\langle k|$, $F_1 = \mathbb{I}_m - F_0$.

Construct the decoherence functional for this model following the proof to Lemma 3.3 by defining $\hat{D}' \in B(\mathbb{C}^m \otimes \mathbb{C}^2)$ as

$$\hat{D}'(a, b|a', b') \equiv \langle v^*|E_{a'}F_{b'}F_bE_a|v^*\rangle. \tag{68}$$

Denoting the basis of $\mathbb{C}^2$ as $\{|0\rangle, |1\rangle\}$ we then have

$$\hat{D}' = \frac{1}{m}|v^*\rangle\langle v^*| \otimes |0\rangle\langle 0| + \left(\sum_{a=1}^{m} |\langle v^*|a\rangle|^2 |a\rangle\langle a| - \frac{1}{m}|v^*\rangle\langle v^*|\right) \otimes |1\rangle\langle 1|. \tag{69}$$

Since $\mathcal{Q} \subset SPJQM$, this decoherence functional is strongly positive. By assumption of the existence of a product joint decoherence functional (EPD) one has that *if* $P_J \in JQM$ (otherwise the proof would be finished), then $D \otimes D'$ gives a valid decoherence functional for the joint distribution P_J and so with $|w\rangle \equiv \sum_{a=0}^{m-1} |a\rangle \otimes |a, 0\rangle$,

$$\langle w|\hat{D} \otimes \hat{D}'|w\rangle = \sum_{a,b=0}^{m} \langle a|\hat{D}|b\rangle\langle a, 0|\hat{D}'|b, 0\rangle = \frac{1}{m}\langle v|\hat{D}'|v\rangle < 0. \tag{70}$$

Since this construction is always, this means that there is always at least one scenario $\mathcal{S}'$ such that $P_J \notin JQM$ for $\mathcal{T} = \{\mathcal{S}, \mathcal{S}'\}$. ☐

One immediate result of this is that

Corollary 2. *JQM is not closed under composition*

It is particularly interesting that only a small subset of $SPJQM$, namely the set $\mathcal{Q}_S$ of quantum models describing single-site experiments with two (incompatible) measurements, was needed to prove maximality. Thus, analogously to the proof for $\tilde{Q}$, this can be used to strengthen the result.

Corollary 3. *$SPJQM$ is the compositional domain of PPI if any physically feasible theory contains the set contains $\mathcal{Q}_S$, if EPS is true.*

Proof. That $SPJQM$ contains $\mathcal{Q}_S$ follows simply by $\mathcal{Q}_S \subset \mathcal{Q} \subseteq SPJQM$ and any set that non-trivially intersects $SPJQM$, whether or not it contains $\mathcal{Q}_S$, is ruled out by Theorem 6.5. ☐

As in the case of $\tilde{Q}$, we would like to strengthen this result to $\mathcal{W}_p$-maximality. Given that the exact relationship between $SPJQM$ and $SPJQM_b$ is not yet known this is involves more assumptions. Following the strategy of the maximality theorem above we can again consider a reductio argument. For any behavior in JQM that lies outside $SPJQM_b$, does there exist an element of $SPJQM_b$ such that $\mathfrak{P}_{\mathcal{S}_J} \notin JQM$? For any $P \in JQM \backslash SPJQM$ we can apply Theorem 6.5 together with $Q_2 \subset SPJQM_b$ since composition falls under physical maps. If $P \in SPJQM \backslash SPJQM_b$ (and $SPJQM \not\equiv SPJQM_b$) and $SPJQM$ is not physically closed then $SPJQM_b$ would very likely be $\mathcal{W}_p$-maximal under SPJQM. Otherwise, if $SPJQM$ turned out to be physically closed (and nevertheless different from $SPJQM_b$), then it is itself the wiring domain of PPI.

Proving Cor. 3 without single-site experiments. While the assumption that every physical theory should include $\mathcal{Q}$ is certainly not very strong, ideally one could prove $\otimes$-maximality of $SPJQM$ without it.

Under the assumption of EDP, to prove this means to show that for any m-dimensional $\hat{D} \in JQM \backslash SPJQM$, there exists a $|w\rangle \in 0, 1^{n\times m}$ such that for finite n,

$$\langle w|\hat{D}^{\otimes n}|w\rangle < 0. \tag{71}$$

This is more difficult to show than the above since we don't have the freedom of constructing a suitable D'. The problem can be made more explicit in the following way:

Denote the elements of D as $d_{i_1 j_1}$ where $i_1, j_1 \in 1, 2, \ldots, m$. Then the elements of of $D^{\otimes n}$ are given by $D_{\imath\jmath}$, where $\imath = i_1 i_2 \ldots i_n$, $\jmath = j_1 j_2 \ldots j_n$. In exact analogy to the reasoning leading to (67) one has that, using Einstein summation convention, one requires a binary vector $|w\rangle$ such that

$$w_{a\imath} D_{\imath\jmath} w_{a'\jmath} = v_{aa'} \ \forall a, a' \in 1, 2, \ldots, m, \tag{72}$$

the proposition was proven true.

This is essentially an eigenvalue problem, as is easily seen by transforming into the eigenbasis.Since D is Hermitian, there exists a unitary similarity transformation S such that

$$0 > \langle v|D|v\rangle = \langle v|S^\dagger S D S^\dagger S|v\rangle = \langle \tilde{v}|\tilde{D}|\tilde{v}\rangle = \sum_a |\tilde{v}_a|^2 \lambda_a, \tag{73}$$

since,in terms of an orthonormal eigenbasis $\{|\lambda_a\rangle\}$, we have $\tilde{D} = \sum_a \lambda_a |\lambda_a\rangle\langle\lambda_a|$ and $|\tilde{v}\rangle = \sum_a \tilde{v}_a |\lambda_a\rangle$. In this form it becomes clear that one solution to this problem is to find a procedure that approximates to arbitrary precision and in finite time a vector in $\Lambda_-^{(S)} = span\{S|\lambda_a\rangle\}, \lambda_a < 0$ in the basis of the original decoherence functional. The matrices corresponding to general atomic decoherence functionals have, however, not received much attention by mathematicians, leaving their eigenspectrum unexplored. Once the latter was to be found, then developing an algorithm should be straightforward since, as is easy to see, as n grows large, a general element v_a of the sought after vector can be constructed from an almost entirely general polynomial of products of the form $v_a = \sum_{\imath,\jmath \in I(a)} D_{\imath\jmath}$ for some $I(a) \subset \{1, 2, \ldots, m^n\}$ with only a couple of constraints on the number of diagonal elements. This is, in a nutshell, because the number of different elements (multisets) of $D_{\imath\jmath}$ converges as n grows, i.e. for the ratio of elements per successive steps one has

$$(74)\qquad \binom{d^2+n}{n+1}\Big/\binom{d^2+n-1}{n} = \frac{d^2+n}{n+1} \overset{n\to\infty}{\rightarrow} 1,$$

while the number of elements continues growing exponentially.

Since an analytic result to the eigenspectrum might in general not be obtainable, it is also worth considering an approximation using inverse iteration, for which only an upper limit on the positive eigenvalues would be required.

6.3. **Probing CHSH inequality violations of JQM in the light of APC..** As discussed in earlier sections, the characterization of sets of behaviors in terms of their violation of CHSH inequalities, in particular the maximum CHSH-value they obey, is an important tool in non-locality research. Lemma 3.7 proved that for single copies in the CHSH scenario $JQM \equiv \mathcal{NS}$. However we know from Theorem 6.3 and Sec. 5.4.1 that, when requiring closure under composition, this relationship for the same scenario becomes $JQM \subseteq \mathcal{LO}^\infty \subset \mathcal{NS}$. One immediate question then is whether JQM or $\mathcal{LO}^\infty$ recover the Tsirelson bound. In [15], and in more detail in [55], this is studied for two copies of isotropic PR-boxes. Recall that these are defined as convex mixtures of the PR-box and local noise as

$$(75)\qquad P_{PR}(\epsilon) = \epsilon P_{PR} + (1-\epsilon)\frac{1}{4}\mathbb{I}.$$

and are distinguished by the fact that every bipartite box can be turned into the form (75) by local operations without changing its CHSH value S [28]. This useful property means that one can probe whether a principle, LO or PPI for example, recover the Tsirelson bound by maximizing the ϵ for which a noisy PR-box (or copies thereof) satisfies that principle. Similarly to the proof that the PR-box violates LO presented in Sec. **??**, one seeks maximal cliques on the orthogonality graph of $P_{PR}(\epsilon)^{\otimes 2}$ such that the sum of its weights violates the corresponding LO inequality. Unlike the case of $P_{PR}^{\otimes 2}$, for noisy PR-boxes all vertices carry non-zero weight which means that none of the 5-cliques used to rule out the PR-box is maximal anymore. However, by extending the cliques to new, maximal ones [15, 55] find cliques that are for $\epsilon_{LO} > (\sqrt{10}-1)/3 \approx 0.721$. This is close to Tsirelson's bound at $\epsilon_T = 1/\sqrt{2} \approx 0.707$ and lead the authors to conjecture that as $k \to \infty, \epsilon_{LO} \to \epsilon_T$.

To prove this conjecture is a very difficult problem even for equiprobable distributions. [55] illustrate this difficulty by relating the search for maximal cliques to the calculation of a graph invariant called the "Shannon capacity of a graph" [68], whose computability is in fact unknown (that is, membership of this problem to any computational complexity class is unclear), see [32].

The existing results on the maximal CHSH-value by these various sets is summarised in Fig. 5. Given that PPI implies LO, it is an interesting project to study whether JQM gives a tighter bound for the same setup. This is, firstly, because it could help to decide the question posed in [67] whether $JQM \equiv \mathcal{LO}$. Secondly, it could throw interesting light on the relation between CHSH-values and the requirement of closure under composition or physical wirings. In Sec. 5.2 we saw that there exist many sub- and superquantum sets that are closed. Since it is already known that $SPJQM$, being a $\otimes$-maximal set, recovers the Tsirelson bound, showing that JQM already, or $\mathcal{LO}^\infty$, do so as well could throw new light on the general relationship between the maximal CHSH-value of sets and their ability to be wiring or compositional domains of physical principles.

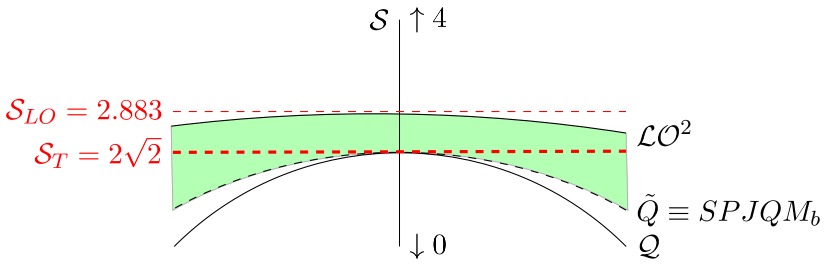


FIGURE 5. Maximal Bell inequality violation by the quantum set $\mathcal{Q}$, the almost quantum relations $\tilde{Q} \equiv SPJQM_b$ and the local orthogonality set $\mathcal{LO}^2$ in the $(2,2,2)$-scenario. The green area indicates the possible region for JQM.

First set up some suitable formalism. The NC-space (Sec. 3.1) is given by

$$\Xi = \{a_0^0, a_1^0, b_0^0, b_1^0, a_0^1, a_1^1, b_0^1, b_1^1\} \equiv \{\gamma_i\}, \tag{76}$$

where the upper index denotes the PR-box, the lower the measurement setting and the invertible function $i : \Xi \to \mathbb{N}$ translates the binary strings $\overline{\gamma_i} \equiv a_0^0 a_1^0 b_0^0 b_1^0 a_0^1 a_1^1 b_0^1 b_1^1$ formed by elements of Ξ into an integer $i \in \{0, 1, \ldots, 255\}$.

To study this problem then amounts to the following problem: Maximize ϵ such that there exists a atomic decoherence functional $\hat{D}(\gamma_i|\gamma_j)$ over Ξ, such that,

$$\begin{aligned} \sum_{i,j \in I} \hat{D}(\gamma_i|\gamma_j) &\equiv \langle I|\hat{D}|I\rangle \\ &= P(a_{x^0}^0, b_{y^0}^0, a_{x^1}^1, b_{y^1}^1 | x^0, y^0, x^1, y^1), \end{aligned} \tag{77}$$

where, denoting $a_{x^0}^0 \equiv a_x^0$ for convenience, $I = \{i\}$ for all $\gamma_i \in \Xi$ for which $\{a_x^0, b_y^0, a_x^1, b_y^1\} \subset \gamma_i$ and, in an alternative representation, $|I\rangle \in \{0,1\}^{\otimes 256}$ is the binary vector with ones on the ith elements.

Since the aim was to produce distributions for which such a decoherence functional does *not* exist we can make things easier by considering real decoherence functionals only. Otherwise we would have to impose explicitly the conditions that for alternative outcomes given by sets $I, J, I \cap J = \emptyset$, $\sum_{i \in I, j \in J} \hat{D}_{ij} = 0$.

Using the symmetries to reduce the complexity of the positivity constraint. As with most membership problems for quantum measure theoretic sets the difficulty lies in the positivity of the measure: Already for this very simply scenario we have $|2^\Xi| = 2^{256}$ positivity constraints on the decoherence functional. One way to tackle this problem is by using the high degree of symmetry of the two PR-box scenario, in which many different marginal settings produce the same statistics.

A convenient way of studying these symmetries is by considering the action of the corresponding group elements $g \in G$ represented as map on the strings $\overline{\gamma_i}$ or, alternatively, on the corresponding i-integer $g(\gamma_i) = \gamma_{\overline{g}(i)} = \gamma_{i'}$. The generating maps of the group, together with examples, are given in Table 1. We find that there are two commuting generators for each PR-box g_1, g_2, g_3, g_4, all of which commute, and one that swaps the two PR-boxes, g_5, which doesn't commute with any of the others, with the commutation relation $[g_i, g_5] = [g_5, g_{i+2 \pmod 2}], i \in \{1,2,3,4\}$. Moreover, g_1, g_3, g_5 have period 2 (i.e. they are their own inverse) while g_2/g_4 have period 4. We therefore have $|G| = 2^7 = 128$

Generator	Action	$g(a_0a_1b_0b_1)$	E.g. $g(1000)$	$\overline{g}(i=8)$
g_1/g_3	swap $A \leftrightarrow B$	$b_0b_1a_0a_1$	0010	1
g_2/g_4	flip conditional on xy	$\overline{a_1}a_0b_1\overline{b_0}$	1101	13
g_5	swap PR-boxes	$g(a_0^0a_1^0b_0^0b_1^0a_0^1a_1^1b_0^1b_1^1) = a_0^1a_1^1b_0^1b_1^1a_0^0a_1^0b_0^0b_1^0$		

TABLE 1. Symmetries of the two PR-box scenario. All group actions are given for the original string $\overline{\gamma_i} \equiv a_0^0a_1^0b_0^0b_1^0a_0^1a_1^1b_0^1b_1^1$.

Recalling that the statistics for the single PR-box are

$$P_{PR}(a,b|x,y) = \begin{cases} \frac{1}{2} & \text{if } a \oplus b = xy \\ 0 & \text{otherwise} \end{cases} \tag{78}$$

it is easy to see that

$$\begin{aligned} \langle I|\hat{D}_g|I\rangle &\equiv \langle I|\hat{D}(g(\gamma_i)|g(\gamma_j))|I\rangle \\ &= \langle I'|\hat{D}(\gamma_{i'})|\gamma_{j'})|I'\rangle \\ &= \langle I|\hat{D}(\gamma_i)|\gamma_i)|I\rangle \\ &= \langle I|\hat{D}|I\rangle \quad \forall g \in G, \end{aligned} \tag{79}$$

where in the penultimate step we used the fact that I and I' produce the same statistics. Therefore every $\hat{D}_g$ is a valid decoherence functional for this scenario.

Now consider the matrix formed by taking the uniform convex sum of the decoherence functionals obtainable by the group action

$$\tilde{D} = \sum_{g \in G} \frac{1}{|G|} \hat{D}_g. \tag{80}$$

Since each $\hat{D}_g$ is a valid decoherence functional, so is $\tilde{D}$. The key property of $\tilde{D}$ is that it is invariant under group action since

$$g \cdot \tilde{D} = \sum_{g' \in G} \frac{1}{|G|} g \cdot (\hat{D}_{g'}) = \sum_{g' \in G} \frac{1}{|G|} (\hat{D}_{g \cdot g' = g''} = \sum_{g'' \in G} \frac{1}{|G|} (\hat{D}_{g''}) = \tilde{D}, \quad \forall g \in G \tag{81}$$

by the Rearrangement Theorem. This means that $\tilde{D}$ is symmetrized and has a set $\mathcal{O} = \{o_k\}$ of k invariant subspaces, or orbits, o_k, under the group action, for which we write $\parallel o_k \parallel$ to denote the value that all elements of $\tilde{D}$ in this orbit have in common. For any collection of atoms $\{\gamma_i\}$ we can then translate the sum $\sum_{i \in I, j \in J} \tilde{D}_{ij}$ into a linear equation

$$X_{\mathbf{n}(IJ)} = \sum_{o_k \in \mathcal{O}} n_k(I \sqcup J) \parallel o_k \parallel, \tag{82}$$

where $n_k(I \sqcup J)$ are the numbers of elements in the original sum that are members of o_k and $\mathbf{n}(IJ) = (n_0, n_1, \ldots, n_{255})$ is the vector constructed from them. In this formulation the problem becomes to maximize ϵ such that there exists a function $\parallel \cdot \parallel : \emptyset \to \bar{R}$ such that $X_{\mathbf{n}(\Xi)} \equiv \sum_{o_k \in \emptyset} |o_k| \parallel o_k \parallel = 1$, $X_{\mathbf{n}(II)} \geq 0$, $\forall I \subseteq \{0, \ldots, 255\}$ and

$$X_{II} = P(a_x^0, b_y^0, a_x^1, b_y^1 | x^0, y^0, x^1, y^1), \tag{83}$$

if $I = \{i\}$ for all $\gamma_i \in \Xi$ such that $\{a_x^0, b_y^0, a_x^1, b_y^1\} \subset \gamma_i$, $|o_k|$ being the cardinality of o_k.

For the group G we find these orbits by studying the eigenspaces of the individual generators. The orbits of the whole group are given by the intersection of the eigenspaces of the individual generators. For example,

Generator	Eigenspace decomposition
g_1/g_3 (on single PR-box)	$\underline{1}_4 \oplus \underline{2}_6$
$g_1 \otimes g_2$	$\underline{4}_{96} \oplus \underline{2}_{144} \oplus \underline{1}_{16}$

and similarly for the other generators. Unfortunately some eigenspaces are contained in others, meaning that they cannot be used to "bridge" between subspaces, so that all together we find $|\mathcal{O}| = 1062$ with $|o_k| \in \{32, 64\}$.

All of this served to make use of the fact that many I produce the same $\mathbf{n}(II)$ so conversely positivity of a single $X_{\mathbf{n}(II)}$ ensures positivity for all members of the equivalence class $[I]$ of $\mathbf{n}(II)$. Further, since for any decoherence functional one can construct a $\tilde{D}$, showing that for some $\epsilon\tilde{D}$ does not exists is sufficient to show that no valid decoherence functional exists.

Further remarks. Since the aim was to achieve a tighter bound than the existing result for local orthogonality, it seemed reasonable to further include the possibility of branching by requiring decoherence, i.e. $X_{IJ} = 0$, whenever there exists a measurement in the branching extension of the scenario such that I, J correspond to alternative outcomes. This could easily be done by noting that two events e, e' are alternative outcomes of *some* possible branching measurement if and only if $e \perp e'$. This is simply because any two events that involve some measurement by some party with different outcomes (which, recall, is the definition of exclusivity) allow for these results to be the alternative outcomes of a strategy involving a conditioning on these outcomes. The decoherence requirement imposed by branching is therefore easily implemented via the orthogonality graph of the the scenario.

Finally, one can note that in any of the above guises the problem is a semi-definite program as introduced in Sec. 2.2 since $P_{PR}^{\otimes 2}(\epsilon)$ is a second order polynomial in the variable to be optimized, ϵ. In order to solve this problem as a linear program instead, we simply form another convex mixture

$$P(\zeta) = \zeta P_{PR}^{\otimes 2}(\epsilon) + (1-\zeta)\frac{1}{16}\mathbb{I} \tag{84}$$

and maximize $0 \leq \zeta \leq 1$ for a given value ϵ and with the same constraints otherwise. We use the Matlab-package YALMIP [69] to do this and treat an optimization for a given ϵ as successful if $\zeta_{max} - 1 \leq \varepsilon$ where ζ_{max} and ε are the solution and output uncertainty given by YALMIP. Furthermore, since, despite the symmetrization, we cannot cover the whole range of constraints, we follow a randomization strategy in which, following an initial set of randomly generated $\mathbf{n}$ (and, of course, the vectors imposing the branching statistics of the behavior), new $\mathbf{n}$ are constructed after run such that $n_k \neq 0$ for the most negative $\| o_k \|$.

Despite these various strategies, no strenghtening of the existing bound of $\epsilon \approx 0.721$ was achieved, leaving it open whether JQM recovers the Tsirelson bound, whether $LO \equiv JQM$ and how APC and ASC influence correlation limits.

6.4. **Closure under temporal composition.** So far the implications of requiring physical closure and stability under composition for quantum measure theoretic sets have been studied. To complete the translation of the ideas connected with closure into quantum measure theory, however, one also needs to make sure that it is clear what these axioms mean in an interpretative context for quantum measure theory. In particular, it is not yet clear what some of the aspects of physical closure correspond to in the histories setting suitable for the latter. In the following an attempt to clarify this is made. It is proven that closure under pre- and post-selection corresponds to the ability to describe any kind of future or past directed timelike joint histories that have been composed from shorter (atomic or timelike) ones, yielding a clearer idea of how to interpret these properties in the context of quantum measure theories.

6.4.1. *Generalizing joint measurement scenarios for timelike scenarios.* To study pre- and post-selection in the JMS-framework it is necessary to generalize the latter to incorporate time. It is convenient to first discuss this generalization operationally: Consider a single party Alice that has access to n identical copies of some box with statistics $P(a|x)$ in an $(1, m, d)$-scenario $S_k, k \in \{1, 2, \ldots, n\}$. To begin with, assume that Alice has access to only one button such that, if she presses it N random generators output measurement settings x_k, which is used to measure box k yielding output a_k. Further assume that all the boxes and random generators are situated in a way that the only timelike related events are the kth random generator and the kth box (and of course, the pressing of the button to all others).

Given the spacelike relation between all remaining events the joint behavior for this scenario will be given by $P_{\mathcal{T}} = \otimes_k P(a_k|x_k)$, defined on the joint scenario $\mathcal{T} = \{\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_n\}$ [38]. If $P(a|x) \in JQM$, then by EPD we have that there exist marginal decoherence functionals $\{D_{P_k}\}$ such that $D_{P_{\mathcal{T}}} = \bigotimes_k D_{P_k}$. In fact the joint decoherence would be obtained in the exact same way if one took away the random generators and allowed Alice to choose the settings according to any strategy, i.e. $x_k \equiv x_k(x'_{k'})$ for any k, k'.

Now consider the scenario in which the Alice instead has only one copy of the box $P(a|x)$, however instead of measuring the box once she gets to measure it n successive times without resetting it (this would artificially impose independence). Since every prior measurement will in general influence the statistics of the prior measurement, the distributions at the different times now form a set $\{P(a_t|x_t) \equiv P_t\} \subset \mathfrak{P}_{\mathcal{S}_t}$ for the $(1, m, d)$-scenario $\mathcal{S}_t$. If we assume that no more than one measurement can be conducted on the box at a single moment then $t \in \{1, 2, \ldots, n\}$.

The analogy between the first setup with various boxes at a single time and this one becomes clear when one exchanges $t \leftrightarrow k$.[13] For instance, since at every t Alice has the choice of $|\mathcal{M}| = m$ measurements the NC- space for this process grows as $|\Xi(t)| = d^{(tm)}$ with NC-space elements given by collections $\{a^t_1, a^t_2, \ldots, a^t_m\} \quad \forall t$, which is exactly analogous to the growth and format of the joint NC- space for $\mathcal{T}$. At the same time it is clear that the timelike relation between successive measurement events in the second setup introduces complications that weren't present in the first case. These complications are exactly the possibility of branching described earlier.

6.4.2. *Closure under temporal composition is equivalent to closure under pre/post-selection.* It is not (and cannot be) the purpose of this section to fully discuss the differences between the two setups, instead the focus here will lie only on implications of the above complications on closure of sets. In particular, in the second setup we can ask what additional requirements we have to impose on sets that are closed under composition to ensure that the joint behavior over times $0 \leq t \leq T$ is still a member of this set. Formalize this closure as follows:

[13]The change of index position is to allow for their convenient notational combination in general setups later on.

Definition 6.1 (Closure under timelike composition)**.** For a one party $(1, m, d)$-scenario $\mathcal{S}$, a set $S(\mathcal{S}) \subset \mathfrak{P}_{\mathcal{S}}$ that is defined for any non-local scenario is *closed under timelike composition* iff, for any $P(a^t|x^t) \in S(\mathcal{S}_t), P(a'^{t'}|x'^{t'}) \in S(\mathcal{S}_{t'}$ and successive times t, t',

$$P(a^t, a'^{t'}|x'^{t'}, x^t) \equiv P(a'^{t'}|x'^{t'}, x^t, a^t) \cdot P(a^t|x^t) \in S(\mathcal{S}_{t\times t'}), \tag{85}$$

where $\mathcal{S}_{t\times t'}$ is the joint scenario constructed from $\{\mathcal{S}_t, \mathcal{S}_{t'}\}$.

One can further distinguish between *future-directed* and *past-directed* timelike composition for the cases in which $t < t'$ and $t' < t$ respectively. Closure under composition clearly is insufficient to warrant closure under timelike composition. However, closure under wiring is, in fact one can prove that or closed sets closure under future-directed timelike composition is equivalent to closure under post-selection, as is shown in the following.

Lemma 6.6. *If a set $S(\mathcal{S}) \subset \mathfrak{P}_{\mathcal{S}}$ is closed under composition and for the tensor product rule, any (non-trivial) joint probability distribution in $S(\mathcal{S}_1)$ can be decomposed into two elements of S, i.e. for any $P_J \in S(\mathcal{S}_J)$, $\exists P \in S(\mathcal{S}), P' \in S(S')$, such that $P \otimes P' = P_J$*

Proof. This is quite obvious, since we exclude single setting probabilities as trivial. □

We can then prove the following theorem:

Theorem 6.7. *A set $S \subseteq JQM$ that is closed under composition is closed under future directed temporal composition iff it is closed under post-selection*

Proof. $\Rightarrow$: Assume S is closed under future directed temporal composition. Then by Lemma 6.6 we know that if $P(a^t, a'^{t'}|x'^{t'}, x^t) \in S(\mathcal{S}_{t\times t'}$, then also $P(a'^{t'}|x'^{t'}, x^t, a^t) \in S(\mathcal{S}')$ for $t \leq t'$, where $\mathcal{S}'$ is the scenario formed by the branching construction from Sec. 4.2.1 But this is the only possible post-selection on the former, so we can immediately conclude that S is closed under post-selection.

$\Leftarrow$: Assume S is closed under post-selection. Then by Lemma 5.1 we know that $S \equiv S_b$ so for any D corresponding to some $P(a'^{t'}, x'^{t'}) \in S(\mathcal{S}_t)$ there exists a D_b which is valid for $P(a'^{t'}|x'^{t+1}, x^t, a^t)$, implying that the latter is in $S(\mathcal{S}')$. Then by closure under composition and (85) we have that S is closed under future directed temporal composition. □

Interestingly the exact same proof goes through for "pre-selection" and past-directed timelike composition. Indeed, this similarity becomes apparent in JMS in that future and past directed branching produce the same branching extension. It should furthermore be possible to generalize this proof to the composition of $\{t_i\}, i \in \{0, \ldots, T\}$ times, i.e. a set $S \subset JQM$ that is closed under composition is closed under future/past-directed timelike composition of $P(a^{t_i}, x'^{t_i}) \in S$ into $P(a^{t_0}, a^{t_1}, \ldots, a^{t_T}|x^{t_0}, x^{t_1}, \ldots, x^{t_T})$ iff S is closed under post/pre-selection, although one expects that further assumptions about "markovianity" of the process are necessary.

That the requirement of closure is not itself sufficient to introduce a temporal or causal arrow is not very surprising, in that it is an "epistemic" requirement on a theory's ability to jointly describe systems as opposed to a physical one, see the next section. It is tempting to relate the question of closure under timelike composition to the dynamics of a theory, for instance by connecting the requirement of closure under both future and past directed temporal composition to the reversibility of a theory but this might be a dangerous conflation of these physical and epistemic aspects of theory making (in the above sense). A more fruitful investigation of is to relate the causality condition on decoherence functional presented in [48] and the growth dynamics of decoherence functionals governed by the transfer matrix [70].

7. Discussion

Both the study of the properties of closure and maximality in the context of quantum measure theory has shown that bringing them together allows one to obtain sets of distributions, that approximate the quantum set very closely for different non-locality scenarios. The last subsection has also further attempted to turn these properties into statements that are easily interpretable in the path integral kinematic framework underlying quantum measure theory. In order to assess whether the singling out of the set $SPJQM_b$ under the various axioms presented in this text can be interpreted as a promising approach to reconstructing quantum theory (or a possible successor thereof) depends mainly on the feasibility of the axioms, which will be discussed in this section.

7.1. **Feasibility of Axioms.**

Axiom of Physical Closure (APS). In March 1948, Einstein returned a manuscript sent to him by Born with a long comment which contains the following passage (cited in [71, p.191], transl. by Howard)

> "However, if one renounces the assumption that what is present in different parts of space has an independent, real existence, then I do not at all see what physics is supposed to describe. For what is thought to be a system is, after all, just conventional, and I do not see how one is supposed to divide up the world objectively so that one can make statements about the parts."

As interpreted by Howard, Einstein here suggests that the possibility to regard different parts of the universe as independent is an *epistemic* necessity in that it is required to make theories that describe the world, in other words, "to do science". One can reformulate this requirement by saying that a theory, at least a unified one and in principle, should be able to discriminate between parts of the universe. It should be able to allow for any scope and graining of description of the physical world. The possibility of producing joint descriptions from pairs of given smaller descriptions, i.e. the notion of closure, is an expression of this requirement on any theory.

However, this is only one aspect of APC: To require *physical* closure is stronger in that closure is required only for joint descriptions that correspond to physically possible actions. In the case of APS these actions are convexity and wiring, which trivially subsumes composition, and *post*-selection, of which we have seen that it corresponds to *future*-directed timelike composition. So while one could, in principle require closure under both pre- *and* post-selection, corresponding to both future and past-directed timelike composition, we restrict our requirement to closure under post-selection because only future-directed timelike composition describes a *physical* process. These two aspects, the requirement of being able to produce joint descriptions that are consistent with the possible physical processes then motivate APC, although it is theoretically possible that the notion of physical closure in this text does not capture all physical processes that are possible.

It is interesting to note in this context that [32] employ the "Foulis-Randall" product as the proper construction of joint scenarios. This product corresponds to the branching construction of joint scenarios discussed in Sec. 4.2 and produces a smaller set of physically possible distributions than the general joint scenario construction of 4.0.2. This is noteworthy because it is known that this constructions automatically produces the $\mathcal{NS}$ polytope for the joint scenario, linking these axioms in some sense even the locality appearing in Bell locality.

Finally, the apparent tension between the motivation of closure as an *epistemic* condition on theories which is restricted to kinds of composition that are *physical* is a fine illustration of the subtle role of that spacetime plays in physics, as both a physical entity and an "organizing principle" that characterizes what description of the world by physical theories are possible, and which is apparent in Einstein's comment.

Axiom of stability under composition (ASC). That principles should be stable under composition is equivalent to the requirement that they should have compositional domain, it should not allow for two physical theories in which one is not a special case of the other. While this may seem like an unjustified requirement, it can be motivated from the point of view of the set of behaviors corresponding to the real world which determines what behaviors are "physical". Then, ASC states that no principle under which this set is the compositional domain should also admit *un*physical behaviors. This is nevertheless a very strong axiom in that, if a number of principles was to jointly produce the physical world, none of them individually would in general satisfy it but only the whole lot.

Assumption of existence of a product joint decoherence functional (EPD). The purpose of EPDin this reconstruction is different from the two others: It was introduced as a necessary condition to prove theorem 6.5. It could be dropped if $\mathcal{W}_p$-maximality could be proven without it, while the other two are motivated on the above grounds and not meant to be dropped. Nevertheless, relating finally these results to the sought after quantum fine trio, it is noteworthy that EPD seems to have a relationship to the quantum screening off condition developed in [48], in which the conditioning of decoherence functionals for marginal systems on their shared causal past is taken to decorrelate them, that is similar to the assumption of separability in the case of classical screening off.

7.2. **Conclusion.** In this thesis the recently introduced notions of physical closure and stability under composition of principles were applied to quantum measure theoretic sets of behaviors. The text also consists of a review of the current literature surrounding this subject. The main result is that that the principle of pair-wise interference which underlies quantum measure theory together with the axiom of physical closure singles out the at least one of the sets $SPJQM_b \stackrel{?}{\equiv} SPJQM$ as its wiring domain under two additional explicit assumptions, namely the existence of quantum single-site experiments in any feasible physical theory and the assumption that for independent behaviors in JQM the decoherence functional corresponding to their joint behavior is given by the tensor product on the the marginal decoherence functional.

Since both the axiom of physical closure as well as the assumption concerning the product form of joint functionals can both be related to very basic aspects of physical theories, this result motivates the principle of pair-wise interference as a candidate for an "explanation" of the limits of physical correlations analogous to the way in which the assumptions underlying Bell locality are meant to explain the Bell inequalities. This is further supported by the principle's satisfaction of another axiom, the axiom of stability under composition.

The special properties of this set, for example the fact that it is the largest set that warrants the possibility of defining an inner product on it which is necessary for the definition on a Hilbert space.

7.2.1. *Future work.* Of the many possible directions of research that can take these early insights as a starting point, the following seem particularly promising:

(1) The argument that infers the physical maximality of $SPJQM_b$ under JQM from the $\otimes$-maximality of $SPJQM$ under the latter requires formal underpinning. A key step here will be whether $SPJQM \equiv SPJQM_b$, which itself requires a further investigation of the branching extension and its consequences for the allowed correlations
(2) Given that $SPJQM_b \equiv \tilde{Q}$ is closer to $\mathcal{Q}$ than all other reconstructive sets known to the author, and that is was arrived at from independent approaches, it is worth considering the option to develop a physical theory producing this set in order to possibly derive a feasible critical test of quantum theory. Here quantum measure theory, being closely connected to the path integral formalism, may be more suitable for such a job than for example a general probability theory (GPT) framework. Indeed, there exist attempts to develop such a theory for quantum measure theory for example in [70, 72] and furthermore experimental investigations of the third-order interferences are already being carried out [73, 74]. The (possible) strengthening from $SPJQM$ to $SPJQM_b$ may here yield new insights.
(3) [75] have recently introduced a characterization of quantum theory in terms of four postulates that seem closely related to the ones discussed here. Their postulates are (i) Classical decomposability ($\equiv$ EP?), (ii) no higher-order interference ($\equiv$ PPI), (iii) strong symmetry and (iv) observability of energy ($\approx$ APC/ASC ?). It would be interesting to study the relationship between the various postulates and axioms in more detail.
(4) Finally, an interesting general idea, that arises only implicitly in the context of closure and stability, is to investigate the possibility of the allowing for n-order interference, with higher order interference being enormously more difficult to observe (as illustrated already in the step from classical to quantum theory for example). However, depending on their accuracy the above experiments could already rule out such a possibility.

Acknowledgements. I would like to thank Fay Dowker for her patient supervision and Joe Henson for his many helpful remarks. I also especially thank Miguel Navascués and the whole Grup d'Informació Quàntica at Universitat Autònoma de Barcelona for hosting me during a whole month of this research. I hope that the exchange of ideas induced by my little exchange makes up for any inconveniences caused by it.

.

References

[1] J. S. Bell, "On the einstein-podolsky-rosen paradox," in *Speakable and Unspeakable in Quantum Mechanics*, Cambridge: Cambridge University Press, 1964.

[2] J. P. Jarrett, "On the physical significance of the locality conditions in the bell arguments," *Nous*, vol. 18, no. 4, pp. 569–589, 1984.

[3] J. Butterfield, “Stochastic einstein locality revisited,” *ArXiv e-prints*, Aug. 2007.

[4] J. Henson, “Causality, Bell’s theorem, and Ontic Definiteness,” *arXiv:1102.2855*, Feb. 2011.

[5] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?,” *Physical Review*, vol. 47, p. 777, 1935.

[6] A. Aspect, P. Grangier, and R. Gérard, “Experimental tests of realistic local theories via bell’s theorem,” *Phys. Rev. Lett.*, vol. 47, p. 460, 1981.

[7] M. Giustina, A. Mech, S. Ramelow, B. Wittmann, J. Kofler, J. Beyer, A. Lita, B. Calkins, T. Gerrits, S. W. Nam, R. Ursin, and A. Zeilinger, “Bell violation using entangled photons without the fair-sampling assumption,” *Nature*, vol. 497, pp. 227–230, 05 2013.

[8] J. Clauser and M. Horne, “Experimental consequences of objective local theories,” *Physical Review D*, vol. 10, no. 2, pp. 525–35, 1974.

[9] B. Tsirelson, “Quantum generalizations of bell’s inequality,” *Letters in Mathematical Physics*, vol. 4, no. 2, p. 93, 1980.

[10] S. Kochen and E. P. Specker, “The problem of hidden variables in quantum mechanics,” *Journal of Mathematics and Mechanics*, vol. 17, pp. 59–87, 1967.

[11] A. Cabello, J. Estebaranz, and G. García-Alcaine, “Bell-kochen-specker theorem: A proof with 18 vectors,” *Physics Letters A*, vol. 212, no. 4, pp. 183 – 187, 1996.

[12] 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.*, vol. 96, p. 250401, Jun 2006.

[13] M. Navascues and H. Wunderlich, “A glance beyond the quantum model,” *Proc.R. Soc. A*, vol. 466, p. 881, 2010.

[14] N. Linden, S. Popescu, A. J. Short, and A. Winter, “Quantum nonlocality and beyond: Limits from nonlocal computation,” *Phys. Rev. Lett.*, vol. 99, p. 180502, 2007.

[15] T. Fritz, A. B. Sainz, R. Augusiak, J. Bohr Brask, R. Chaves, A. Leverrier, and A. Acin, “Local orthogonality as a multipartite principle for quantum correlations,” *Nature Communications*, vol. 4, no. 2263, 2012.

[16] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, “Information causality as a physical principle,” *Nature*, vol. 461, p. 1101, 2009.

[17] S. Popescu, “Nonlocality beyond quantum mechanics,” *Nature Physics*, vol. 10, pp. 264–70, 2014.

[18] A. Fine, “Hidden variables, joint probability, and the bell inequalities,” *Phys. Rev. Lett.*, vol. 48, p. 291, 1982.

[19] R. Werner and M. Wolf, “Bell inequalities and entanglement,” *Quantum Inf. Comput.*, no. 1-25, 1.

[20] R. Feynman, “Space-time approach to non-relativitstic quantum mechanics,” *Rev. Mod. Phys.*, vol. 20, p. 367, 1948.

[21] R. Griffiths, “Consistent histories and the interpretation of quantum mechanics,” *Journal of Statistical Physics*, vol. 36, no. 1-2, pp. 219–272, 1984.

[22] J. B. Hartle, “The spacetime approach to quantum mechanics,” *Vistas in Astronomy*, vol. 37, no. 0, pp. 569 – 583, 1993.

[23] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, “Nonlocal correlations as an information-theoretic resource,” *Phys. Rev. A*, vol. 71, p. 022101, 2005.

[24] J. Barrett and S. Pironio, “Popescu-rohrlich correlations as a unit of nonlocality,” *Phys. Rev. Lett.*, vol. 95, p. 140401, Sep 2005.

[25] N. Brunner, N. Gisin, and V. Scarani, “Entanglement and non-locality are different resources,” *New J. Phys.*, vol. 7, p. 88, 2005.

[26] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, “Quantum correlations require multipartite information principles,” *Phys. Rev. Lett.*, vol. 107, p. 210403, Nov 2011.

[27] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. Vertesi, “Closed sets of non-local correlations,” *Phys. Rev. A*, vol. 80, p. 062107, 2009.

[28] L. Masanes, A. Acin, and N. Gisin, “General properties of nonsignaling theories,” *Phys. Rev. A*, vol. 73, p. 012112, Jan 2006.

[29] M. Navascues, S. Pironio, and A. Acin, “A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations,” *New J. Phys.*, vol. 10, p. 073013, 2008.

[30] W. van Dam, “Implausible consequences of superstrong nonlocality,” *arXiv:quant-ph/51159*, 2005.
[31] M. Navascués, Miguel, S. Pironio, and A. Acín, “Bounding the set of quantum correlations,” *Phys. Rev. Lett.*, vol. 98, p. 010401, Jan 2007.
[32] A. Acin, T. Fritz, A. Leverrier, and A. B. Sainz, “A combinatorial approach to nonlocality and contextuality,” *arXiv:1212.4084*, 2014.
[33] L. Vandenberghe and S. Boyd, “Semidefinite programming,” *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1996.
[34] S. Wehner, “Tsirelson bounds for generalized clauser-horne-shimony-holt inequalities,” *Phys. Rev. A*, vol. 73, p. 022110, Feb 2006.
[35] L. Landau, “Emprical two-point correlation functions,” *Found.Phys.*, vol. 18, no. 4, pp. 449–460, 1988.
[36] M. Navascues, M., Y. Guryanova, M. J. Hoban, and A. Acin, “Almost quantum correlations,” *arXiv:1403.4621*, 2014.
[37] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, “Bell inequalities for arbitrarily high-dimensional systems,” *Phys. Rev. Lett.*, vol. 88, p. 040404, Jan 2002.
[38] F. Dowker, J. Henson, and P. Wallden, “A histories perspective on characterizing quantum non-locality,” *New J. Phys.*, vol. 16, p. 033033, 2014.
[39] P. Billingsey, *Probability and measure.* New York: John Wiley and Sons, 1995.
[40] R. D. Sorkin, “An exercise in ”anhomomorphic logic”,” *Journal of Physics: Conference Series*, vol. 67, no. 1, p. 012018, 2007.
[41] C. J. Isham, “Quantum logic and the histories approach to quantum theory,” *J.Math.Phys.*, vol. 35, pp. 2157–2185, 1994.
[42] R. D. Sorkin, “Quantum mechanics as quantum measure theory,” *arXiv:gr-qc/9401003*, 1994.
[43] R. Salgado, “Some identities for the quantum measure and its generalizations,” *arXiv:gr-qc/9903015*, 2002.
[44] C. Ududec, H. Barnum, and J. Emerson, “Three slit experiments and the structure of quantum theory,” *arXiv:0909.4787*, 2009.
[45] V. Belavkin, “Reconstruction theorem for quantum stochastic processes,” *Teoreticheskaya i Matematicheskaya Fizika*, vol. 62, p. 409, 1985.
[46] D. Dutkay, “Positive definite maps, representations and frames,” *Rev. Math. Pys*, vol. 16, p. 451, 2004.
[47] F. Dowker, S. Johnston, and R. D. Sorkin, “Hilbert spaces from path integrals,” *Journal of Physics A*, vol. 43, no. 27, p. 275302, 2010.
[48] D. Craig, F. Dowker, J. Henson, S. Major, D. Rideout, and R. D. Sorkin, “A bell inequality analog in quantum measure theory,” *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 3, p. 501, 2007.
[49] M. Barnett, F. Dowker, and D. Rideout, “Popescu–rohrlich boxes in quantum measure theory,” *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 26, p. 7255, 2007.
[50] B. Tsirelson, “Quantum analogues of the bell inequalities. the case of two spatially separated domains,” *J. Soviet Math*, vol. 36, p. 557, 1987.
[51] B. Tsirelson, “Some results and problems on quantum bell-type inequalities,” *Hadronic Journal Supplement*, vol. 8, p. 329, 1993.
[52] J. Barrett, “Information processing in generalized probabilistic theories,” *eprint arXiv:quant-ph/0508211*, Aug. 2005.
[53] A. Fine and E. Winsberg, “Quantum life: Interaction, entanglement, and separation,” *Journal of Philosophy*, vol. 100, no. 2, p. 80, 2003.
[54] T. Mueller and T. Pacek, “Against a minimalist reading of bell’s theorem,” *Synthese*, vol. 128, p. 343, 2001.
[55] A. B. Sainz, T. Fritz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, “Exploring the local orthogonality principle,” *Phys. Rev. A*, vol. 89, p. 032117, Mar 2014.
[56] B. Lang, T. Vertesi, and M. Navascuez, “Closed sets of correlations: answers from the zoo,” *arXiv:1402.2850*, 2014.
[57] J. Uffink, “Quadratic bell inequalities as tests for multipartite entanglement,” *Phys. Rev. Lett.*, vol. 88, p. 230406, May 2002.
[58] I. Pitowsky, “Geometry of quantum correlations,” *Phys. Rev. A*, vol. 77, p. 062109, Jun 2008.
[59] D. Dukaric and S. Wolf, “A limit on non-locality distillation,” *arXiv:0808.3317*, 2008.

[60] A. Cabello, S. Severini, and A. Winter, "(non-)contextuality of physical theories as an axiom," *ArXiv e-prints*, Oct. 2010.

[61] J. Henson, "Quantum contextuality from a simple principle?," *arXiv:1219.5978*, 2012.

[62] E. Specker, "Die logik nicht gleichzeitig entscheidbarer aussagen," *Dialectica*, vol. 14, p. 239, 1960.

[63] Y.-C. Liang, R. W. Spekkens, and H. M. Wiseman, "Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity," *Physics Reports*, vol. 506, no. 1–2, pp. 1 – 39, 2011.

[64] A. Cabello, S. Severini, and A. Winter, "Graph-theoretic approach to quantum correlations," *Phys. Rev. Lett.*, vol. 112, p. 040401, Jan 2014.

[65] L. Lovász, "On the shannon capacity of a graph," *IEEE Tran*, vol. 25, no. 1, p. 1, 1979.

[66] M. Grötschel, L. Lovász, and A. Schrijver, "Relaxations of vertex packing," *Journal of Combinatorial Theory*, vol. 40, no. 3, p. 330, 1986.

[67] J. Henson, "Bounding quantum contextuality with lack of third-order intereference," *ArXiv e-prints*, June 2014.

[68] C. Shannon, "The zero error capacity of a noisy channel," *IRE Transactions on Information Theory*, vol. 2, no. 3, p. 8, 1956.

[69] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, (Taipei, Taiwan), 2004.

[70] X. Martin, D. O'Connor, and R. D. Sorkin, "Random walk in generalized quantum theory," *Phys. Rev. D*, vol. 71, p. 024029, Jan 2005.

[71] D. Howard, "Einstein on locality and separability," *Stud. Hist. Phil. Sci.*, vol. 16, no. 3, pp. 171–201, 1985.

[72] R. D. Sorkin, "Quantum dynamics without the wavefunction," *Journal of Physics A: Mathematical and Theoretical*, vol. 40, no. 12, p. 3207, 2007.

[73] U. Sinha, C. Couteau, Z. Medendorp, I. Söllner, R. Laflamme, R. Sorkin, and G. Weihs, "Testing Born's Rule in Quantum Mechanics with a Triple Slit Experiment," in *American Institute of Physics Conference Series* (L. Accardi, G. Adenier, C. Fuchs, G. Jaeger, A. Y. Khrennikov, J.-rA. Larsson, and S. Stenholm, eds.), vol. 1101 of *American Institute of Physics Conference Series*, pp. 200–207, Mar. 2009.

[74] U. Sinha, C. Couteau, T. Jennewein, R. Laflamme, and G. Weihs, "Ruling out multi-order interference in quantum mechanics," *Science*, vol. 329, no. 5990, pp. 418–421, 2010.

[75] H. Barnum, M. P. Mueller, and C. Ududec, "Higher-order interference and single-system postulates characterizing quantum theory," 2014.

Appendix A. Theorem 5.7 does not apply to $SPJQM$

Theorem 5.7, proving that $\tilde{Q}$ is $\otimes$-maximal under LO, is a very general result in that it applies to all contextuality scenarios. It is therefore not clear whether the same result follows for only a subclass of scenarios, namely the non-locality scenarios for which quantum measure theoretic sets are defined. This manifests itself implicitly in the use orthogonality graphs and their complements. Since it is unclear whether scenarios corresponding to these complementary graphs are always non-locality scenarios themselves. Indeed, the sole fact that $SPJQM$ and JQM are not define for some scenarios is not sufficient to rule out this possibility that the proof nevertheless applies. If it could be shown that from the complementary graph to every graph corresponding to a non-locality scenario one can construct another valid non-locality scenario, then the theorem would apply. In other words, if the set of graphs corresponding to all non-locality scenarios was closed under the complement action, the proof goes through. That this is, even under "generous" translation of non-locality into graphs, not the case we show in the following paragraphs.

Joint measurement scenarios to graphs. To every (non-trivial) non-locality scenario $\mathcal{S}$ corresponds a (non-trivial) orthogonality graph. In particular we can re-formulate Definition **??** in terms of the partition scenarios $(\Xi, \mathcal{M})$ that were introduced in Sec. 3.1.

Definition A.1 ($O_{PS}(\mathcal{S})$)**.** A graph with vertex set V is an orthogonality graph of a partition scenario $\mathcal{S} = \{\Xi, \mathcal{M}\}$ with coarse-graining set $\mathbb{C}$ and collection of exclusive sets $\mathfrak{S}$ if

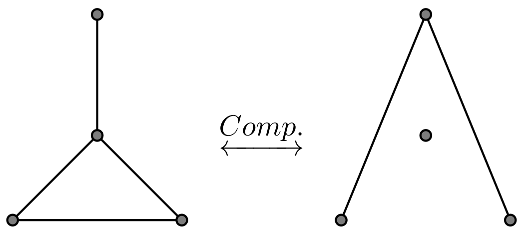
- the graph has adjacency relation
$$v(A) \sim v(B) \iff \exists M \in \mathcal{M} \text{ such that } A, B \in M$$
- there exists a bijective map $v : \mathbb{C} \to 2^V$ such that
  - $v(A) \in V \quad$ for all $A \in M \quad \forall M \in \mathcal{M}$
  - v is additive in the sense that, for any exclusive set of fine-grained outcomes $S \in \mathfrak{S}$ with elements A_i,[14]
$$v\left(\bigsqcup_i A_i\right) = \bigcup_i v(A_i)$$

The proof of equivalence is straightforward. Furthermore, this representation can be extended to behaviors in $\mathfrak{P}_{\mathcal{S}}$ by defining the weights $p(v(A)) \equiv P(A)$.

Illustrating non-applicability. The question of the applicability of the above maximality result in the JMS-framework then becomes whether for any scenario $\mathcal{S}$ there exists another scenario $\mathcal{S}'$ such that $\overline{O_{JMS}(\mathcal{S})} \equiv O_{JMS}(\mathcal{S}')$.

To investigate this consider first a simple P_3 path.

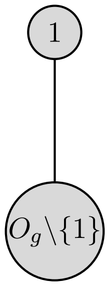
This graph cannot be an orthogonality graph of any scenario with non-empty $\mathcal{M}$. This is because we require by the additivity of v that every (maximal) clique corresponding to a measurement gives $v(\Xi)$. But this implies that the two outer vertices of P_3 map into the same outcome, which is forbidden by bijectivity. Therefore, no valid orthogonality graph may contain P_3 as induced path (which is sufficient to rule out induced paths P_k for any $k \geq 3$). Now consider the following pair of complementary graphs:



The left graph is a valid orthogonality graph according to Def.A.1 for a scenario with two measurements, while the right one clearly has P_3 as induced subgraph. This simple example is already sufficient to disprove applicability if one excepts Def.A.1 as a valid. However, looking closer at the left graph it becomes clear that, again by the additivity property, its top vertex corresponds to the union of the two bottom vertices. In other words, this scenario describes a setting in which one measurement is a (partial) coarse-graining over the other. One might therefore exclude graphs like this one by introducing a notion of "genuine fine-grained outcome" which form the set $\mathfrak{M} \subset \mathcal{M}$ of all fine-grained outcomes of all measurements in $\mathcal{M}$ except for those which can be obtained by the union of other members. Then define a "genuine orthogonality graph $O_g(\mathcal{S})$" similar to Def.A.1 with the only difference that $v(A) \in V$ only for all $A \in \mathfrak{M}$.

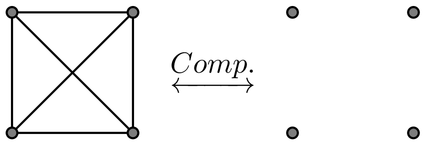
[14] $\sqcup$ is here slightly redundant in that orthogonal outcomes are by definition disjoint

One can then check that this additional requirement rules out all graphs with “appendices” of the form



and therefore exclude the potential pair above as counterexamples. Seeing that we excluded induced P_2 paths it is worth noting that there exists a family of graphs, called co-graphs that are closed under complementation and can be fully characterized by the non-existence of P_4 induced subpaths.

Nevertheless, this additional requirement is not sufficient: While it is possible to disprove applicability in general even for the case of genuine orthogonality graphs more rigorously, suffice here to consider an even simpler issue with the example of single measurement scenarios whose complementary graphs produce edgeless graphs that violate bijectivity, e.g.



This result is not surprising in that in that many orthogonality graphs can be constructed that correspond to scenarios that do not admit a single experimental probability distribution, see [32] for examples in terms of their more general hypergraph framework.