Imperial College London

IMPERIAL COLLEGE LONDON

DEPARTMENT OF PHYSICS

Exploring the operational, device independent, and histories approaches to quantum foundations

---

Robin Simmons

Date: September 20, 2023

"Reality, this spiteful snake
     Shedding its smothering veil"

-Meshuggah

Abstract. Three approaches to quantum foundations are explored, operational (i.e General Probabilistic Theories) and device independent (e.g. Popescu-Rohrlich boxes), and known and novel results that advance the understanding of the relationships between them are presented. Finally, a tentative step towards a GPT for relativistic Quantum Field Theories is taken, and evidence that it violates the no-restriction hypothesis is provided. This result, combined with those from operational research, hints that abandoning theories due to violating the no-restriction hypothesis (e.g. the 'almost quantum' set), is premature and should be reevaluated.

Contents

Chapter 1

Introduction

Despite 100 years of stunning success, a truly deep understanding of quantum
theory seemingly remains as elusive as ever. The textbook mathematical
formalism for quantum mechanics (QM) and quantum field theory (QFT) were
all but finalised in the 1930s and 60s [1, 2, 3], respectively, yet both leave
significant interpretational and foundational issues unresolved. This is
highlighted by the EPR debate[4], Sorkin's impossible measurements[5], the
proliferation of interpretations of quantum theory[6, 7, 8], and quantum
gravity[3, 9, 10]. Together, they highlight an uncomfortable problem with trying
to take quantum theory beyond a phenomological model of tabletop or collider
experiments. Simultaneously, the difficulties of balancing rigour, utility, and
manifest Lorentz invariance have lead to at least three mainstream formalisms
for QFT[1]. The principles emphasised by each of these approaches have inspired
increasingly diverging solutions to the measurement problem and quantum
gravity.

For these and other reasons, quantum foundations has emerged as a small but
dedicated research effort. Some modern branches have been inspired by the
mathematical formalisms of quantum information (via the operational view of
quantum channels) and classical gravity (via causal structure), and building up
a number of impressive results with wide implications, yet often little-known by
those outside the field. These hint at an extremely rich structure underlying the
'space' of physical theories and the principles that define them. They have shone
much light both on quantum theory, and the feasibility of theories that could
replace it.

The aim of this report is to provide a short introduction to three approaches to
quantum foundations, and to present a number of novel and known results that
provide tentative links between them. The scope of this project is wide, and so
in many places we are not able to give existing results the attention they

---

[1]Canonical, algebraic, and path integral

deserve, and will often not be able to state them at all. In these places, we urge the reader to look at the suggested review articles to build up a better picture of the field.

In Chapter 2, the device independent, operational, and histories approaches to quantum foundations are introduced, highlighting the differences and known connections between them. In Chapter 3 we provide a proof that a principle known as post-processing implies a number of conditions used in the device independent and histories approaches, and then present work towards better understanding branching and the relationship between the NPA hierarchy, almost quantum set, and strongly positive decoherence functionals. Finally, Chapter 4 presents a tentative look at an operational model of QFT, and gives evidence that the textbook definition of QFT violates a commonly held principle in the operational framework. To do so, we introduce Sorkin's impossible measurements[5], which suggest that the standard QFT definition of measurement is flawed.

Chapter 2

Approaches to quantum foundations

While this work does not have the scope to touch on every approach, we hope that this section, which focuses on three frameworks, gives a suitably wide cross section. Throughout this thesis we will comment on the strengths and weaknesses of each approach, and possible connections between them.

2.1   Device independent

Based on the assumption that the internal specifics of an (ideal) experiment and its physical realisation are irrelevant to the outcomes, the device independent formalism is our first quantum foundations framework. Consider a Bell-type experiment[11], where projective measurements of spins at different relative angles are conducted by spacially separated observers. The results of (an ideal) experiment are independent of whether the particles are entangled electron-position pairs or photon pairs. In fact, any maximally entangled two party state where each party's local state is two level, will have the same (ideal) measurement statistics. This could correspond to qubits in a superconductor or vibrational modes of an oscillator. The most important property is that (assuming ideal from now) the measurement statistics are the same, and the correlations of certain measurements violate Bell's inequality. A device corresponds to a particular experimental setup to generate the measurement statistics, which is what we wish to abstract away.

A *scenario* is a device independent experimental setup, where a (from now, finite) number of (n) parties each have a lab. In and out of the labs flow classical information[1], in the form of measurement settings and measurement outcomes, and any allowed communication. In the quantum case, inside the labs we can imagine each party having access to pre-prepared entangled states

---

[1]While typically the labs are considered to input and output classical info, called a C-C box, it is possible to consider cases where one or more are replaced by other forms of information. See [12], and [13] for investigations of classical to quantum (C-Q) and Q-Q behaviours respectively, as well as their relationships to C-C behaviours.

distributed between them, and a full experimental setup where different settings, denoted by n-dim vectors $\vec{x}$ correspond to different measurements on those states. The results of those measurements are our measurement outcomes, given by the n-dim vectors $\vec{a}$. By abstracting, the inside of the lab is ignored, and the focus is solely on the statistics, given by the conditional probabilities

$$P(a,b|x,y) \text{ bi-partite, } P(\vec{a}|\vec{x}) \text{ } n\text{-partite.} \tag{2.1}$$

We refer to $(\vec{a}|\vec{x})$ as a finegrained event, and so P is a probability distribution on those events. A coarsegrained event is one in which some subset of the parties are margined out, specifically their outputs.

For a given scenario, which corresponds to a choice of number of parties, their relationship between each other (the form of allowed communication), a choice of number of measurement settings and outcomes for each party, P describes a device independent *behaviour*. If we demand causality, then spacelike seperation of the parties, requires no-signalling be imposed, by demanding that the behaviours themselves cannot be used to signal. By partitioning the parties in two sets R, S (Receiver and Sender), *no-signalling*[1] demands that the Receiver parties cannot learn anything about the measurement settings of the Sender party as follows:

$$P_R(\vec{a}_R|\vec{x}_R, \vec{x}_S) = P_R(\vec{a}_R|\vec{x}_R) \tag{2.2}$$

where $\vec{x}_R, \vec{x}_S$ is vector of measurement outcomes for the parties in R and S, and $\vec{a}_R$ are the outcomes for the parties in R. For two parties this reduces to

$$P(a|x,y) = P(a|x), \quad P(b|x,y) = P(b|y). \tag{2.3}$$

As we can see, when coarsegraining, non-signalling implies that coarsegraining is over the entire state of the other parties, and not just the outputs.

A *Bell scenario* is a non-signalling scenario where the number of settings m is the same for each of the n parties, and each measurement has p outcomes. We denote it as an $(n \ m \ p)$ scenario, and refer to any behaviours for that scenario as $(n \ m \ p)$ behaviours.

Definition 1. *An $(n \ m \ p)$ behaviour $P(\vec{a}|\vec{x})$ is quantum if $\exists \, |\psi\rangle$ in a Hilbert space $\bigotimes_{i=1}^{n} \mathcal{H}_i$, and for each party i and measurement setting x_i, there is a complete (sum to the identity and are orthogonal) set of projection operators*

---

[1] We highlight an interesting observation, credited to Fay Dowker: while causality, and thus Lorentz invariance, is taken as an unbreakable axiom in quantum foundations, it is not considered sacrosanct in quantum gravity. The assumption of non-signalling is so fundamental to foundations that there is almost no literature exploring its absence. Yet for quantum gravity, the presentation of a model that violates Lorentz invariance is most often only met with 'by how much?'

$\{E_{a_i}^{x_i}\}$ on $\mathcal{H}_i$, for each x_i, such that for any partition R, S

$$P(\vec{a}_R|\vec{x}_R) = \langle\psi| \bigotimes_{i\in R} E_{a_i}^{x_i} |\psi\rangle \tag{2.4}$$

In the finite dimensional case this is sufficient to define the quantum set, and the Tsirel'son bound. However, in the infinite dimensional case, another definition may provide a larger set of quantum behaviours which has lead to Tsirel'son's problem [14], recently solved by Slofstra [15].

A key observation in the history of quantum foundations, by Tsirel'son [16] and independently by Popescu and Rohrlich [17], is that the set of non-signalling behaviours is significantly greater than even the set of quantum behaviours. The bound between quantum and super-quantum is known as the Tsirel'son bound[1]. And in some senses, there are behaviours that are more non-local than those allowed by quantum theory[2]. The canonical example for the (2 2 2) scenario is the behaviour known as the PR (Popescu Rohrlich) box, given by

$$P_{PR}(a,b|x,y) = \frac{1}{2}\delta_{x\cdot y, a\oplus b}. \tag{2.5}$$

where $\oplus$ is addition modulo 2. The CHSH game [19] requires two parties to outcome binary numbers a, b respectively in response to receiving x, y respectively, such that $x \cdot y = a \oplus b$, without communicating.
Define

$$P_{succ} = \sum_{a,b,x,y} P(a\oplus b = x\cdot y) = \sum_{a,b,x,y} P(a\oplus b = x\cdot y|x,y)\frac{1}{4} \tag{2.6}$$

to be the probability of winning the game, where we have assumed that (x,y) are evenly distributed. It is easy to see that the PR box perfectly solves the CHSH game. Further, the marginal probabilities are $\frac{1}{2}$ for either party, and so it is also non-signalling. Since classical theories can only achieve a P_{succ} of 75% and quantum theory is bounded by approximately 85%, we can see an implied hierarchy of physical theories, distinguished by their non-locality. This leads to three sets of behaviours, L, Q and NS, the local, quantum, and non-signalling sets respectively. The local behaviours are those that have a local hidden variable interpretation, while the non-signalling are those that obey non-signalling only.

---

[1]The term Tsirel'son bound is sometimes used to refer specifically to the (2 2 2) bound, however we use the more general name for the bound of behaviours to any scenario

[2]We say in some senses, because it has been shown that for two particles in a CHSH like scenario, there are no possible correlation function that dominates the quantum one for all angles. This is in contrast to the fact that the optimal quantum correlation function dominates the classical one for all angles [18]

There is a strict inclusion

$$L \subset Q \subset NS \tag{2.7}$$

and for a fixed scenario, each forms a convex set. Further, for a given scenario, L, NS are polytopes, i.e. they have finitely many extremal points. However, the points of the Tsirel'son boundary, which are the extremal points of Q, are uncountable.

Popescu and Rohrlich raised the interesting question of what, if anything, singles out quantum theory against other sets of correlations, and if any other sets correspond to physical theories [17]. These questions can be considered a main focus of device independent research, outside of quantum information.

A number of 'device independent principles' have been proposed, many of which successfully single out the bound for the CHSH game (see Section 2.1). However, none have yet been proposed that are known to single out the Tsirel'son bound as a whole, for all games and all Bell scenarios. Even more curiously, computing Tsirel'son's bound is known to be undecidable [20]. This raises philosophical questions-as well technical questions about proving if a given distribution is quantum or not-should we expect the set of behaviours to be computable?

2.1.1   NPA hierarchy and almost quantum

Introduced by Navascues, Pironio, and Acin [21], the NPA hierarchy gives a sequence of nested subset of the non-signalling behaviours, which is convergent to the quantum set. Each set is strictly a subset of the one before, and it is thus sufficient to show that any behaviour is not an element of one of the NPA sets to show that it is not an element of the quantum set. We do not give the full machinery to define all of the NPA sets, however we will define the first set, as well as the almost quantum set. The following definition and notation are inspired by [22] and the definition of the almost quantum set in [23].

Definition 2. *[22] An n-partite behaviour $P(\vec{a}|\vec{x})$ for a Bell scenario (n m p) is in Q^1 if there exists a Hermitian positive semi-definite matrix Γ^1, with matrix elements labelled by the fine and coarse grained events $(\vec{a}|\vec{x})$ and the event Ω (i.e. something happens), with*

1. $\Gamma^1_{\Omega,\Omega} = 1$ normalised

2. $\Gamma^1_{\Omega,(\vec{a}|\vec{x})} = P(\vec{a}|\vec{x})$.

3. $\Gamma^1_{(\vec{a}|\vec{x}),(\vec{a}'|\vec{x}')} = P((\vec{a}|\vec{x}) \cap (\vec{a}'|\vec{x}'))$ where $\vec{x}, \vec{x}'$ do not contain the same parties

4. $\Gamma^1_{\Omega,(\vec{a}|\vec{x})} = \Gamma^1_{(\vec{a}|\vec{x}),(\vec{a}|\vec{x})}$

11

5. $\Gamma^1_{(\vec{a}|\vec{x}),(\vec{a}'|\vec{x})} = 0$ if $(\vec{a}|\vec{x}) \neq (\vec{a}'|\vec{x})$.

The NPA hierarchy gives a sequence of sets $Q \subset Q^n$, defined by Hermitian positive semi-definite matrices Γ^n, such that $Q^n \subset Q^{n-1}$ and

$$\lim_{n\to\infty} Q^n = Q. \tag{2.8}$$

Note that any Hermitian positive semi-definite matrix can be interpreted as an inner product matrix, called a Gram matrix, for vectors labelled by the rows/columns of the matrix, so that there exists a Hilbert space and set of vectors $|i\rangle$ such that

$$\Gamma_{ij} = \langle i|j\rangle. \tag{2.9}$$

Another device independent set of significant interest is the "almost quantum set" $\tilde{Q} \supset Q$ [23]. This set has been shown to be closed under classical post-processing (see Section 3.1.1) and satisfies 5 of the 6 major device independent principles (see Section 2.1.2), with evidence that it satisfies the remaining one (information causality). As such, within the device independent framework, there is little reason to prefer the quantum set over the almost quantum set, and it has been proposed that it represents a physically meaningful theory or even a model of our own universe. The operational and histories frameworks provide differing perspectives on these conjectures, which we will explore.

For our next definition, the notation $\perp$ must be introduced to describe the relationship between two (possibly coarsegrained) events $(\vec{a}|\vec{x}), (\vec{a}'|\vec{x}')$:

$$(\vec{a}|\vec{x}) \perp (\vec{a}'|\vec{x}') \text{ if } \exists \text{ at least one party } i \text{ with } x_i = x_i' \text{ and } a_i \neq a_i'. \tag{2.10}$$

Definition 3. *[23] An n-partite behaviour $P(\vec{a}|\vec{x})$ for a Bell scenario $(n\ m\ p)$ is in $\tilde{Q}$ if there exists a Hermitian positive semi-definite matrix Γ^{AQ}, with matrix elements labelled by the fine and coarse grained outcomes $(\vec{a}|\vec{x})$ and the null event Ω, with*

1. $\Gamma^{\tilde{Q}}_{\Omega,\Omega} = 1$ normalised

2. $\Gamma^{\tilde{Q}}_{\Omega,(\vec{a}|\vec{x})} = P(\vec{a}|\vec{x})$

3. $\Gamma^{\tilde{Q}}_{(\vec{a}|\vec{x}),(\vec{a}'|\vec{x}')} = 0$ if $(\vec{a}|\vec{x}) \perp (\vec{a}'|\vec{x}')$.

4. $\Gamma^{\tilde{Q}}_{(\vec{a}|\vec{x})\cap(\vec{a}'|\vec{x}'),(\vec{a}|\vec{x})\cap(\vec{a}''|\vec{x}'')} = \Gamma^{\tilde{Q}}_{(\vec{a}'|\vec{x}'),(\vec{a}|\vec{x})\cap(\vec{a}''|\vec{x}'')} = \Gamma^{\tilde{Q}}_{(\vec{a}|\vec{x})\cap(\vec{a}'|\vec{x}'),(\vec{a}''|\vec{x}'')}$ if $\vec{x}$ shares no parties with $\vec{x}', \vec{x}''$.

We can quickly see that $\tilde{Q} \subseteq Q^1$, as $3, 4$ of $\tilde{Q}$ imply $3, 4, 5$ of Q^1. In fact, this inclusion is strict and related to local orthogonality, as we shall see below.

2.1.2 Device independent principles

While it is out of the scope of this work to introduce all of the device independent principles in detail, two will come up in our discussions later, and so it is helpful to review them briefly.

Macroscopic locality

Macroscopic locality imagines a situation where some large number N of copies of a behaviour exist, and can detect the plurality of each measurement outcome with sensitivity of order $\mathcal{O}(\sqrt{N})$. For example, Alice and Bob share N copies of a device described by a $(2\,2\,2)$ behaviour, and Alice's detector measures the numbers of $a = 0, a = 1$, with a sensitivity of $\mathcal{O}(\sqrt{N})$, and likewise for Bob. A behaviour is macroscopically local if the new $(2\,2\,N)$ effective behaviour exhibited by the macroscopic detector satisfies Bell's inequality, i.e. it can be described by a local hidden variable model, in the limit of large N. It has been shown that the set of macroscopically local behaviours is exactly Q^1[24]. Hence, as claimed above, almost quantum is macroscopically local, as it is a strict subset of Q^1.

We point out that a principle derived *after* almost quantum was proposed, called macroscopic noncontextuality, picks out $\tilde{Q}$, and so can be seen as stronger than macroscopic locality [25]. This is not unexpected, as contextuality is a broader concept than nonlocality, and thus it might be expected that noncontextuality is a tighter constraint than locality. Certainly, we expect any post-quantum theory with any hope of being correct to contain a noncontextual (classical) limit. As with macroscopic locality, it is of course using our current theories to demand properties of new theories, a sort of bottom up approach.

It is, however, not clear that *every* (or any) correlation must have a local or noncontextual interpretation the statistical limit discussed above for a well defined classical limit to exist. In fact, as emphasised by decoherence, the classical limit is far more closely related to coarsegraining over systems (parties) and not over many independent correlations (behaviours). We can see that it is generically true that the outcomes of one party are given exactly by local hidden variable models for any behaviour, this is imposed directly by non-signalling. It would seem that demanding macroscopic locality or noncontextuality would suggest that: 1-strong correlations arise generically and are not suppressed in realistic situations by effects analogous to monogamy of entanglement (which has analogues in all non-local theories), 2-coarsegraining of multiple correlations plays a role in, or is closely related to, the existence of a classical limit. The first is hard to rule out without a more complete model. The second is the main

claim behind both principles, and should be backed up if we are to make the assumption that all theories with a classical limit must be subsets of Q^1 or $\tilde{Q}$.

Local orthogonality

Consider a set of events $S = \{E^i\} = \{(\vec{a}^i|\vec{x}^i)\}$ for a fixed scenario, where the events are *pairwise* orthogonal,

$$E^i \perp E^j \forall i \neq j. \tag{2.11}$$

Then we say S is an orthogonal set. The Local Orthogonality principle (LO) is the requirement that

$$\sum_{E \in S} P(E) \leq 1, \tag{2.12}$$

for every orthogonal set in a scenario [26].
For two events, this simple. Consider the events $E^1 \perp E^2$, e.g. Alice has setting x and measures either a or a'. Then

$$P(E^1) + P(E^2) = P(a, b, ...|x, y, ...) + P(a', b', ...|x, y'....) \leq 1 \tag{2.13}$$

by normalisation, as we can consider this to a *branching* (see Sections 2.3.5,3.2) measurement, where Alice distributes her result to the other parties, who then measure $y, z...$) or $y', z'...$ depending on if Alice measures a or a'. Since the only orthogonal sets for a (2 2 2) scenario are at most 2 events, all (2 2 2) behaviours are LO. However, for effectively any more complex scenario, this is not the case, and the LO principle excludes extremal behaviours (those at the boundary of the NS set) for (3 2 2) and (2 m p). Moreover, if two parties have more than one copy of a PR-box, then LO is violated[26].

Other principles

Principles that have not yet been mentioned include information causality, non-trivial communication complexity, consistent exclusivity and no advantage for non-local computation [27, 28, 29]. For the latter three, proof has been given that almost quantum does not violate those principles, and there is evidence that it does not violate information causality [23][25]. This has lead to the realisation that quantum correlations are not bounded by these principles, and opened the possibility of a super-quantum theory, such as almost quantum, that satisfies many 'reasonable' principles.
Again, it is not entirely clear that there are physical reasons to expect non-trivial communication complexity or no advantage for non-local computation, beyond discomfort at the prospect by some researchers. In a

discussion by Wim van Dam, one of the first to point out that PR boxes allow trivial communication complexity (TCC), he asks why we should not want TCC: "*It is not clear if there is a convincing answer to this question, as it does not seem to conflict with any physical intuition...Such hierarchies are at the core of theoretical computer science, and their 'collapse'...goes against the intuition of most researchers in the field of complexity theory.*"[30]. The existence of TCC has been compared the computer scientists' version of a violation of causality [27]. However, there seems to be no *physical* reason to exclude it, unlike violations of causality.

2.2 Operational foundations

Unlike the previous section, we will focus now on the general probabilistic theories (GPT) formalism that attempts to study properties of abstract theories, based on their state and effect (read:measurement) spaces, as well as maps between them [31].
There are a number of approaches that share a number of similarities with GPTs, which are not discussed here. However, for the sake of completion we mention Hardy's and others derivations of quantum theory from axioms [32]. Like GPTs, these are motivated by Ludwig's theorem [33] (see Chapter 4).

2.2.1 GPTs

GPTs are the prototypical operational tool for studying general physical theories. They are motivated by Ludwig's embedding theorem, the requirements (Ludwig's axioms) of which are so general to be thought to apply to any theory with an operational formalism. Ludwig's embedding theorem is so fundamental to GPTs that it is often completely omitted from discussions, and its consequences, namely the formulation of physical theories as convex sets of vector spaces, are taken as axioms themselves. More time is devoted to a technical discussion of Ludwig's embedding theorem in Chapter 4.
Here, we simply state that for a set of states S, effects E, and a bi-linear form $P : S \times E \to \mathbb{R}$ with $0 \leq P \leq 1$, obeying some very reasonable requirements, there exists a GPT. A GPT is a convex subset $\mathcal{S}$ of a real vector space V, almost always taken to be finite dimensional, and the effect space $\mathcal{E}$,

$$\mathcal{E} \subseteq \mathcal{E}' \subset V^*, \tag{2.14}$$

is some subset of all effects on $\mathcal{S}$ such that the probability is normalised, $0 \leq p = \omega(s) \leq 1$ for $\omega \in \mathcal{E}', s \in \mathcal{S}$ [31]. We shall give these statements rigorous

definitions in Chapter 4.

This gives an extremely general framework, effectively any theory with states and measurements is equivalent to a GPT, where many properties of the theory are encoded in the convex structure of the GPT. One surprising result is that many of the main features of quantum theory that are not present in classical theory, ones which we often take as distinguishing them, are generic to any no classical theory. This includes entanglement, non-locality, and contextually, and so allow behaviours that violate Bell's and similar inequalities. This goes some way to supporting Popescu and Rohrlich's conjecture of a set of super-quantum theories.

While the GPT framework has demonstrated a number of interesting and surprising results, its generality is also one of its weaknesses. In order to constrain the space of theories, one has to make additional assumptions. These assumptions are often motivated by what we know from quantum and classical theory, and so are biased by our physical intuition, something we may wish to avoid when trying to imagine a theory for which we have no physical evidence. The question arises: if we had discovered GPTs before quantum theory, would we have relaxed the correct assumptions to open the plausible space of theories to include quantum mechanics *before* we were forced to by the discoveries of the 20th century?

As with the device independent formalism, GPTs give little information about dynamics or physical interpretation-how to translate a physical experiment into a state and effect and *vice versa*. These issues are more manageable in a histories based approach, with sacrifices in other areas, as we shall see.

2.2.2   No-restriction, Specker's principle and GTTs

Within GPTs, axioms often must be imposed in order to restrict the space of theories. While behaviours are not the central objects in the GPT framework, several axioms have been shown to constrain the sets of behaviours associated to the theory. However, there are also reasons to doubt that these axioms are well motivated.

No-restriction hypothesis

The no-restriction hypothesis [34] states that the inclusion in Eq 2.14 is an equality, i.e. there are no restrictions on the physical effects. All well defined effects are physically realisable. This has been a common assumption in many studies of GPTs, as it is satisfied by classical and quantum mechanics. At the

level of QM it is given by the duality between density operators and projection operators.

It has been shown that any theory that recreates the almost quantum set cannot satisfy the no-restriction hypothesis [35]. This no-go theorem is non constructive, as there is no guarantee that a set of behaviours corresponds to a GPT at all, and finding one is no trivial task.

We discuss the no-restriction hypothesis in greater detail, and in the context of QFT where its satisfaction is far less obvious, in Chapter 4.

Specker's principle

Specker's principle can be stated as follows:'If in a set of measurements every pair is compatible, then all the measurements are compatible.' It has been shown that while quantum theory satisfies this, almost quantum does not [36]. Certainly, this follows by the existence of a mutual eigenbasis for any set of compatible operators, however it appears to be a bottom up principle rather than a well motivated axiom. It has been shown that Specker's principle implies consistent exclusivity, while the converse does not hold, as exemplified by almost quantum satisfying the latter and not the former.

Gleason type theorem

Despite often being taken as an important postulate in GPTs, there is growing evidence that meaningful physical theories can still be constructed even when the no-restriction hypothesis is violated [37][38], as long as they have a "Gleason type theorem". The above works show that no-restriction is neither sufficient nor necessary for 'indeterminate determinism'.

Inteterminate determinism implies the existence of a set of properties for each state that can be determined by measurements, and are sufficient to determine the probabilities of any measurement. This is guaranteed for QM by Gleason's theorem, hence the name. It appears to be enough for a well behaved operational theory, and implies that no-restriction is too strong a postulate, which has been discussed before [39].

Some work has shown that a form of GPTs that can violate no-restriction, known as self-dual, can reproduce the Tsirel'son bound for maximally entangled bi-partite scenarios [40]. For more general Bell scenarios, the equivalent result is unknown, however there is no evidence that a GPT without the no-restriction cannot recreate the entire Tsirel'son bound.

2.3   Histories foundations

The histories approach to quantum foundations is motivated by the histories approach to quantum mechanics, which aims to be a coherent interpretation for quantum theory that is in the spirit of relativity and path integrals[41, 42]. The branch of research we shall discuss is based on Sorkin's axiomatisation of path integral-like theories into quantum measure theories (QMT), for which there is a very similar branch of research instigated by Hartle called generalised quantum mechanics (GQM) [43]. Since almost all results are the same, we stick only to QMT.

Unlike in the previous approaches, we need to introduce a very different perspective for quantum theory in order to appreciate this approach. The necessary concepts are presented as an extremely short, self-contained, introduction to the histories approach to quantum foundations. For more general reviews and discussions of the histories approach to interpreting quantum theory or to quantum gravity, see the following references [41, 42, 44, 45, 46, 47].

2.3.1   Histories

The histories approach is an interpretation of quantum theory, initiated by Hartle [48] (and motivated by Griffith's consistent histories[49]), where states and measurements are replaced by histories and events as the central concepts. Being path integral inspired, this places manifest Lorentz invariance front and centre, however it of course suffers from the same issues of mathematical rigour that plague almost all studies of path integrals. The path integral, however, is useful despite this lack of rigour, and its ubiquity in high energy physics justifies this further investigation.

Path integrals are often introduced simply as a way to calculate the main objects of interest in QFT, expectation values of time ordered operators acting on a Hilbert space (or functionals thereof). This is flipped on its head by the successful formulation of quantum theory based entirely on the path integral, with no reference to the Hilbert spaces or operators[46]. For QM a derivation of the Hilbert space and Schrödinger equation from that formalism has been demonstrated. The equivalent derivation is less rigorous for QFT for the same reasons that the path integral evades proper definition.

We focus on single particle mechanics (quantum, classical, or otherwise), but generalisations to field theories are the same as those from QM to QFT path integrals. For a more in depth introduction, see [50].

A 'history' γ is a spacetime trajectory, and an 'event' corresponds to something

18

happening. A particle being at $(t, \vec{x})$ is an event E, and we define it as the set

$$\{\gamma | \gamma(t) = \vec{x}\} \tag{2.15}$$

so that $\gamma \in E$.

The set of all histories is denoted Ω, and the set of all events (sets of histories) is the powerset 2^Ω.

2.3.2 Quantum measure theory

The word 'event' is borrowed from probability theory, and one may be tempted to appeal to the Kolmogorov axioms by assuming some measure $\mu : 2^\Omega \to \mathbb{R}$ on the powerset of events satisfies

1. $0 \le \mu \le 1$ positive

2. $\mu(\Omega) = 1$ normalised

3. $\mu(A \sqcup B) - \mu(A) - \mu(B) = 0$ classical sum rule,

where $\sqcup$ is disjoint union. However, we can quickly see that this cannot be the case. Consider the double slit experiment, and label γ_L (γ_R) to be histories where the particle goes left (right) slit and ends up at a particular destructive fringe[1]. Then, let the events L (R) to be the events that the particle goes through the left (right) slit, so that $\gamma_L(\gamma_R) \in L$ (R). Finally, let D be the event that the particle is at the destructive fringe, see Fig 2.1.

Clearly, $\mu(D) = 0$, as the particle is never observed there. This is consistent with $L \cap R = \emptyset$. Each history can only go through one slit. However, $\mu(L) = \mu(R) = 0.5$. Hence $\mu(L \sqcup R) = \mu(D) = 0 \ne \mu(L) + \mu(R) = 1$. Clearly, we must relax the Kolmogorov axioms if we wish to describe quantum theory in this way. For classical mechanics, where there is no sum over histories, there is no interference and so the Kolmogorov axioms hold, hence we will call normal probability theory 'classical'. The relaxed quantum measure theory is

1. $0 \le \mu$ positive

2. $\mu(\Omega) = 1$ normalised

3. $\mu(A \sqcup B \sqcup C) - \mu(A \sqcup B) - \mu(A \sqcup C) - \mu(B \sqcup C) + \mu(A) + \mu(B) + \mu(C) = 0.$ quantum sum rule

---

[1]One might argue that it is begging the question to consider histories that arrive at a destructive fringe, as there is obviously zero probability of that happening. However, if we cover up the left slit, the histories going through the right slit should not be effected, by locality, and experiment tells us that the previously dark fringe will now be illuminated due to the lack of interference. So we must include all the histories, including those that contribute to impossible events.
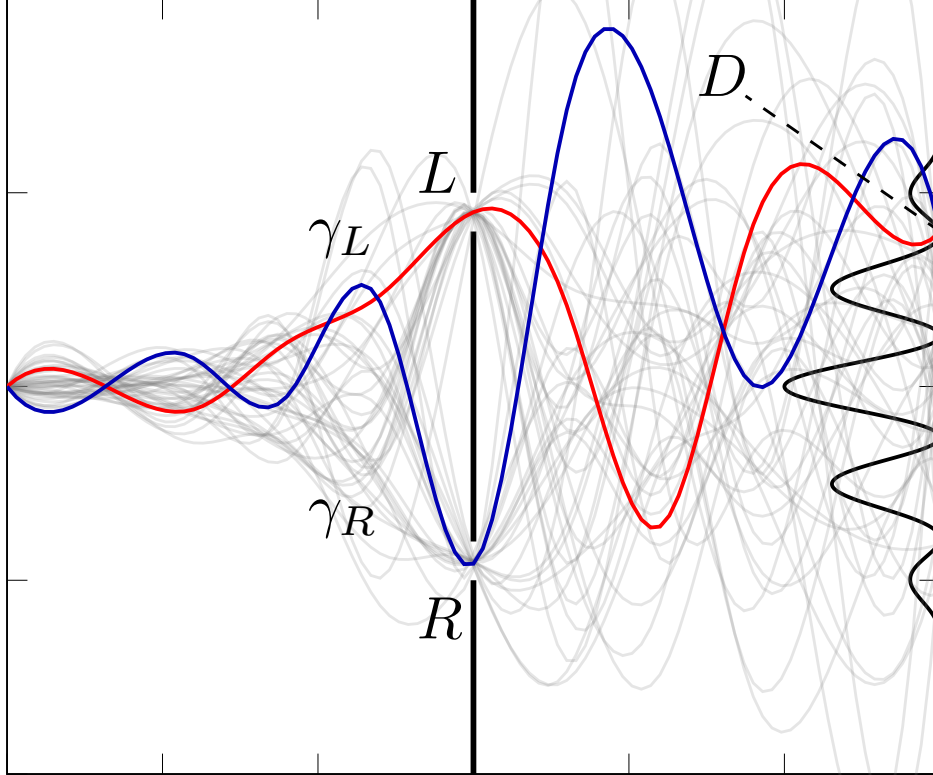
Figure 2.1: A histories perspective on the double slit experiment. A sample of histories are shown in grey, while a history that passes through the left (right) slit and ends at a destructive fringe is shown in red (blue). The events $L(R), D$ represent the set of all histories passing through the left (right) slit and arriving at the destructive fringe respectively. From left to right are the source, wall with slits, and screen. The intensity is shown on the right.

where the final axiom, the quantum sum rule, expresses the lack of so called 3rd order interference [41]. In this language, axiom (3) of the Kolmogorov axioms expresses the lack of 2nd order interference, e.g. classical particles do not interfere in the double slit experiment. To see how the double slit might satisfy this, we introduce one final event N, where the particle goes through no slits. We have assumed that the setup is such that the particle always goes through some slit, and so N contains no histories and is disjoint with L, R, D. Hence

$$\mu(L \sqcup R \sqcup N) - \mu(L \sqcup R) - \mu(L \sqcup N) - \mu(R \sqcup N) + \mu(L) + \mu(R) + \mu(N) \quad (2.16)$$

$$= \mu(D) - \mu(D) - \mu(L) - \mu(R) + \mu(L) + \mu(R) = 0 \quad (2.17)$$

as required. We have come across the major difference between quantum and classical (measure theory), where quantum events, i.e. the particle making it to the destructive fringe, not occurring does not imply that a subevent, i.e. a particle travelling through the left slit to the destructive fringe, from happening. To make contact with the path integral, we note that for nonrelativistic

quantum mechanics

$$\mu(A) = \int_{\gamma \in A^T} d\nu(\gamma) \int_{\bar{\gamma} \in A^T} d\nu(\bar{\gamma}) e^{iS(\gamma) - iS(\bar{\gamma})} \delta(\gamma(T), \bar{\gamma}(T)) \qquad (2.18)$$

where A^T is the set of histories truncated at some time T after the point of measurement, and it is expected that some boundary conditions must be specified. We note that this is a double path integral, unlike the one usually derived from the Hilbert space formalism. The double integral form

$$\int_{\gamma \in A^T} d\nu(\gamma) \int_{\bar{\gamma} \in A^T} d\nu(\bar{\gamma}) \qquad (2.19)$$

is implied by the quantum sum rule. In this sense, we have already found a constraint on the dynamics of 2nd order theories, given by integrals over 2 path interference of histories.

2.3.3 Decoherence functionals

From the above discussion it is possible to re-derive the usual rules of quantum mechanics. However, it still requires us to have specified from Eq 2.19 to Eq 2.18, or a relativistic, or non scalar equivalent. Questions still stand about what theories allowed by the quantum sum rule/double path integral, and how they relate to regular quantum theory. To investigate this, it is useful to introduce the 'decoherence functional'.
Let $D : 2^\Omega \times 2^\Omega \to \mathbb{C}$ obey the following axioms

1. $D(\Omega, \Omega) = 1$ normalised

2. $D(A, B) = D(B, A)$, i.e. Hermitian

3. $D(A, B \sqcup C) = D(A, B) + D(A, C)$ bi-additive

4. $D(A, A) \geq 0$ weakly positive

then D is equivalent to a quantum measure via $D(A, A) = \mu(A)$, so that any decoherence functional defines a unique quantum measure, and a family of decoherence functionals exist that are picked out by any quantum measure [22]. It is clear that

$$D(A, A') = \int_{\gamma \in A^T} d\nu(\gamma) \int_{\bar{\gamma} \in A'^T} d\nu(\bar{\gamma}) e^{iS(\gamma) - iS(\bar{\gamma})} \delta(\gamma(T), \bar{\gamma}(T)) \qquad (2.20)$$

is a decoherence functional associated with μ from Eq 2.18.

We can also impose that when restricted to a finite subset of the event space, $D \geq 0$, i.e. D is positive semidefinite as a (restricted) matrix. This is referred to as strongly positive or SP. As we shall see, this is necessary and sufficient for the existence of a Hilbert space, and is satisfied by QM and QFT. It also allows us to interpret a finite dimensional decoherence functional as a Gram matrix, allowing cross-talk between histories and the NPA hierarchy. The physical motivation for this step is not entirely clear [51], yet the resulting decoherence functional is significantly more physically reasonable, as we shall see. 'Decoherence functional' implies a relation to the concept of decoherence introduced by Zeh and others [52]. In the usual decoherence, a system coupled to an environment evolves such that they are entangled, and so that 'pointer' states diagonalise the local density operator of the system. These pointer states are robust against the interaction with the environment, and are seen as the set of allowed classical sates, those that are 'selected' for during measurement, or rather interaction. This process of 'einselection' is the key observation of decoherence, which proposes a solution to the question of the preferred basis: why do we only observe quantum states in an example of a particular basis? It does not provide an answer to why we see only one example and not some mixture. Decoherence provides a mechanism by which quantum states evolve towards states that appear, with coarsegraining, as mixtures of semi-classically meaningful states.

On the other hand, in the histories approach, non-classicality is related to the quantum sum rule. Decoherence here referrers to situations involving coarsegraining such that 2nd order interference becomes negligible and μ can be interpreted as a classical probability distribution due to an approximate satisfaction of the Kolmogorov sum rule. We can see that this is similar to the usual definition of decoherence, as a theory that satisfies Kolmogorov's axioms is a theory of classical mixtures. Further, since histories attempts to be a complete interpretation, it claims to solve the entire measurement problem, unlike decoherence.

2.3.4 n-hopper model and strong positivity

To build intuition for the history and event spaces, and the decoherence functional, we can consider the n-hopper model[53], where a single particle is confined to points of the finite set $\mathbb{Z}_n$ and evolving in discrete time steps, so that for every time τ, its position is given by $\gamma(\tau) \in \mathbb{Z}_n$. The full trajectory γ gives a

history, and so the history space Ω is given as

$$\Omega := \underbrace{\mathbb{Z}_n \times \mathbb{Z}_n \times \cdots \times \mathbb{Z}_n}_{T-\text{times}} \qquad (2.21)$$

where T is the number of time steps considered.
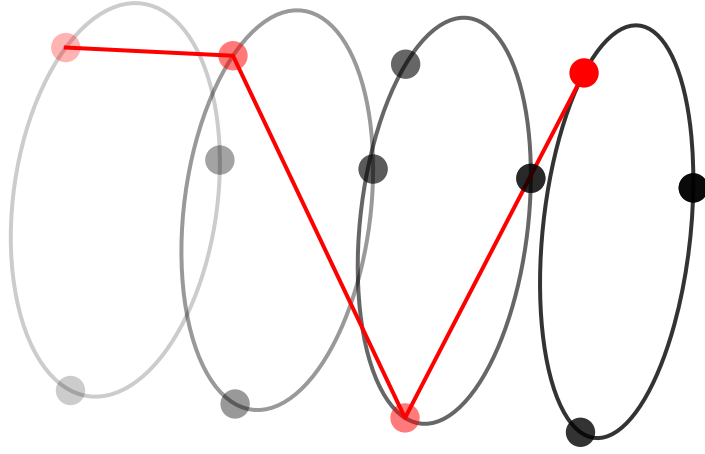
Figure 2.2: A diagram showing a 4 timestep history for a 3-hopper.

$$D_t(A, B) = \sum_{\gamma_1 \in A} \sum_{\gamma_2 \in B} a_t \left[\gamma_1\right] a_t \left[\gamma_2\right]^* \delta_{\gamma_1(t)\gamma_2(t)} \qquad (2.22)$$

where the amplitudes a are given by

$$a_t[\gamma] = \psi_{\gamma(0)} \prod_{\tau=0}^{t-1} U_{\gamma(\tau)\gamma(\tau+1)}$$

$$U_{ab} = \frac{1}{\sqrt{n}} \exp\left(\frac{i\pi(a-b)^2}{\alpha n}\right) \quad (2.23)$$

and $\alpha = 1$ for odd n and 2 for even. Note $\{\psi_{\gamma(0)}\}$ acts as the initial conditions discussed earlier and the form analogous to the non-relativistic propagator. The

time t corresponds the truncation time, some late time chosen to be after the events in consideration. Clearly, D is Hermitian and weakly positive Consider $n = T = 3$, so that there are only 3 position considered at 3 time steps, and that we are looking only at events at $\tau = 0, 1$, i.e. that the truncation time is set as $t = 2$, the final timestep. In order to calculate D explicitly, consider initial conditions fixing $\gamma(0) = 0$, so that $\psi_{\gamma(0)} = \delta_{\gamma(0),0}$. We now calculate $D_2(\gamma, \bar{\gamma})$ as $D_2(E, E')$ can then be found by bi-additivity. Each $\gamma = (z_0, z_1, z_2)$ and the initial conditions imply that only $\gamma = (0, z_1, z_2)$ do not decohere. $D_2(\gamma, \bar{\gamma})$ decoheres if $z_2 \neq \bar{z}_2$, by the delta function. Hence, D_2 takes the form

$$D_2(\gamma, \bar{\gamma}) = a_t \left[\gamma_1 \right] a_t \left[\gamma_2 \right]^* \delta_{\gamma_1(t)\gamma_2(t)} \tag{2.24}$$

There are 27 histories, however D_2 vanishes on any with $z_0 = 1, 2$, so only D_2 need only be evaluated on 9, giving a 9×9 block. We can consider the matrix rows and columns each labelled by (z_1, z_2) (i.e. $((0,0), (1,0), (2,0), (0,1), ...)$, and find that the 9×9 block has the form

$$D_2((z_1, z_2), (\bar{z}_1, \bar{z}_2)) = \frac{1}{9} \begin{pmatrix} \begin{array}{ccc} 1 & a & b \\ a^* & 1 & c \\ b^* & c^* & 1 \end{array} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \begin{array}{ccc} 1 & \alpha & \beta \\ \alpha^* & 1 & \delta \\ \beta^* & \delta^* & 1 \end{array} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \begin{array}{ccc} 1 & x & y \\ x^* & 1 & z \\ y^* & z^* & 1 \end{array} \end{pmatrix} \tag{2.25}$$

where we have used that D_2 is Hermitian, the decoherence of histories with differing z_2 and that $D_2(\gamma, \bar{\gamma}) = \frac{1}{n^2} = \frac{1}{9}$ for any $\gamma = \bar{\gamma}$ by unitarity of the propogator, to reduce the elements to explicitly calculate from 81 to 9. Due to only having 3 sites, at each timestep the particle can take either a step either of length 0 or 1 (in either clockwise or anticlockwise directions). They contribute each a phase of

$$0 \text{ and } \pm i\frac{\pi}{3}, \tag{2.26}$$

respectively. We can now fill in the remaining off diagonals.

$$D_2((z_1, z_2), (\bar{z}_1, \bar{z}_2)) = \frac{1}{9} \begin{pmatrix} \begin{matrix} 1 & \omega^2 & \omega^2 \\ \omega & 1 & 1 \\ \omega & 1 & 1 \end{matrix} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \begin{matrix} 1 & 1 & \omega \\ 1 & 1 & \omega \\ \omega^2 & \omega^2 & 1 \end{matrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \begin{matrix} 1 & \omega & 1 \\ \omega^2 & 1 & \omega^2 \\ 1 & \omega & 1 \end{matrix} \end{pmatrix} \tag{2.27}$$

where

$$\omega = e^{i\frac{\pi}{3}}. \tag{2.28}$$

Calculating the eigenvalues, we find them to be $0, 1, 2, 3$, and so D_t is positive semi-definite as a matrix, and thus strongly positive as a decoherence functional. While it is not immediately obvious in this form of the matrix, the remaining symmetry left over from the global $\mathbb{Z}_3$ broken by our choice of initial condition, swapping site 1 and 2, is a symmetry of D_2 (both the form in Eq 2.27 and the 27×27 full form). For example $D_2((1,1),(2,1))) = \omega = D_2((2,2),(1,2))$. Since there was nothing special about the choice $z_0 = 0$, we can see that the full D_2, defined on entire histories (z_0, z_1, z_2) will be the same form, upto re-ordering, regardless of which point the particle starts on. The general form of D_2 can then be found by complex linear combinations of D_2 for $z_0 = 0, 1, 2$, and is thus also strongly positive.

Finally, note that D_2 is normalised, as the total sum of all its entries is 1 (the off-diagonals in each block form the roots of unity and so cancel).

2.3.5 Histories perspective on non-signalling scenarios

As with the double slit experiment, we can investigate non-signalling scenarios by constructing the histories and from them, the history and event spaces, as well as important subspaces. To do this we use a modified version of the non-contextuality (NC) space used by Dowker *et al* [22], with some notational and structural modifications.

For an n party non-signalling scenario, consider a history γ to be an array of numbers γ_{ij}, which denote the outcome of the ith experiment conducted by the jth party. For a Bell $(n\ m\ p)$ scenario, these take the form of an $n \times m$ matrix over $\{0, 1, 2, ..., p\}$, however in the most general case we cannot make this association as the number of measurements and outcomes can vary between parties. For example, the $(2\ 2\ 2)$ scenario has histories that are represented by

2×2 binary matrices e.g.

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \tag{2.29}$$

represents an outcome of 0 for $x = 0$, 1 for $x = 1$, and 1 for both $y = 0, 1$. We again denote Ω to be the set of all histories, and 2^Ω is its powerset.

Of course, a behaviour is not defined on histories, but on events corresponding to single outcomes of a single experiments per party. As before, consider such an event to be the set of all histories in Ω that correspond to a certain outcome to a certain experiment for each party. Borrowing notation from the device independent formalism, we say the event[1] $(\vec{a}|\vec{x})$ is the set

$$(\vec{a}|\vec{x}) = \{\gamma | \gamma_{x_i i} = a_i, \forall i \in \{0, 1, ..., n\}\}. \tag{2.30}$$

In the $(2\ 2\ 2)$ case again, we can see the example for $(0, 0|1, 0)$ is

$$(0, 0|1, 0) = \left\{ \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix} \right\}, \tag{2.31}$$

where the highlighted elements correspond to the fixed outcomes. Clearly, this is a subset of 2^Ω. A coarsegraining (by partitioning into R, S) of such an event is the union of all

$$(\vec{a}_R|\vec{x}_R) = \{\gamma | \gamma_{x_i i} = a_i, \forall i \in R\} = \bigcup_{(\vec{x}'_R|\vec{a}'_R) = (\vec{a}_R|\vec{x}_R)} (\vec{a}'|\vec{x}'). \tag{2.32}$$

For $(2\ 2\ 2)$, an example is given where we take the union of $(0, 0|1, 0)$ and $(0, 1|1, 0)$,

$$(a = 0|x = 1) = \{ \begin{pmatrix} 0 & 0 \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ \mathbf{0} & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \mathbf{0} & 1 \end{pmatrix}, \tag{2.33}$$

$$\begin{pmatrix} 0 & 1 \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ \mathbf{0} & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ \mathbf{0} & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ \mathbf{0} & 1 \end{pmatrix} \}, \tag{2.34}$$

and note that by taking the subsets of the $1, 3, 5, 7$th elements and $2, 4, 6, 8$th elements we get the events $(0, 0|1, 1), (0, 1|1, 1)$ respectively, as expected by non-signalling. For a fixed $\vec{x}$, the set of all fine and coarsegrained $(\vec{a}|\vec{x})$ is a subset of 2^Ω. Denote this as $\mathfrak{u}_{\vec{x}}$, and the union as

$$\tilde{\mathcal{O}} = \bigcup_{\vec{x}} \mathfrak{u}_{\vec{x}} \tag{2.35}$$

over all $\vec{x}$ is the set of all fine and coarsegrained outcome events of the form $(\vec{a}|\vec{x})$. Finally, we discuss one more subset of 2^Ω. The *set* $\vec{x}$ is defined as the set

---

[1] Now our use of the word event earlier has paid off.

of all $(\vec{a}|\vec{x})$ for all $\vec{a}$. We have overloaded $\vec{x}$, using it both as a set and as a vector. However, since there is a natural bijection between them, there is no ambiguity, and $\vec{x}$ will be treated as both depending on the context. The *union* of all $\vec{x}$'s, fine and coarsegrained, is $\tilde{\mathcal{S}}$ (contains all sets $(\vec{a}|\vec{x})$ of histories), and represents the set of all measurement setting events, while $\tilde{\mathcal{S}}$ denotes the set of all events of measurement outcomes conditional on measurement settings. We note the following important set relationships

$$(\vec{a}|\vec{x}) \in \vec{x}, \quad (\vec{a}|\vec{x}) \in \mathfrak{u}_{\vec{x}} \subset \tilde{\mathcal{O}}. \tag{2.36}$$

$$(\vec{a}|\vec{x}) \subseteq (\vec{a}'|\vec{x}'), \quad \vec{x} \subseteq \vec{x}' \tag{2.37}$$

where $\vec{x}', (\vec{a}'|\vec{x}')$ are coarsegrainings of $\vec{x}, (\vec{a}|\vec{x})$.

Some histories sets

We state the definitions of $\tilde{Q}$, JQM and SPJQM in the histories formalism. While we deliberately made sure the notation lined up with the device independent one, we will at times use $E = (\vec{a}|\vec{x})$ etc. for convenience, compactness, or to refer to elements of the event algebra that do not have a joint measurement interpretation, i.e. are not of the form $(\vec{a}|\vec{x})$. In the following definitions, the translation to the usual notation should be immediate.

Definition 4. *Given a behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$, *it is in the set JQM (joint quantum measure) if there is a quantum measure* μ *on* 2^{Ω} *such that*

$$\mu(E) = P(E) \tag{2.38}$$

for all $\in \tilde{\mathcal{O}}$,*(recall that* $\tilde{\mathcal{O}}$ *is the set of all fine and coarsegrained outcomes, which is a subset of* 2^{Ω}, *and* $\tilde{\mathcal{S}}$ *is the set of all fine and coarsegrained settings).*

Equivalently,

Definition 5. *Given a behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$, *it is in the set JQM (joint quantum measure) if there is a decoherence functional on* $2^{\Omega} \times 2^{\Omega}$ *such that*

1. $D(E, E) = P(E), \quad \forall E \in \tilde{\mathcal{O}}$

2. $D(E, E') = 0, \quad \forall E, E' \in \mathfrak{u}_{\vec{x}}, E \cap E' = \emptyset, \forall \vec{x} \in \tilde{\mathcal{S}}$

Note that the PR box, and the entire $(2\ 2\ 2)\ NS$ polytope, is included in JQM [54], and so we have reason to believe, as laid out by the device independent principles, that we need to consider a subset of JPM.

Definition 6. *Given a behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$, *it is in the set SPJQM (strongly positive joint quantum measure) if there is a decoherence functional on* $2^\Omega \times 2^\Omega$ *such that*

    1. $D(E, E) = P(E), \quad \forall E \in \tilde{\mathcal{O}}$

    2. $D(E, E') = P(E \cap E'), \quad \forall E, E' \in \mathfrak{u}_{\vec{x}}$

    3. D is strongly positive

It has been shown that for the CHSH game, SPJQM coincides with the Tsirel'son bound. However, even for general (2 2 2) scenarios, SPQJM contains behaviours that are not in Q [54].

As previously mentioned, a strongly positive decoherence functional is also a Gram matrix. This allows the following lemma:

Lemma 7. *[22] A behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$ *is in SPJQM if and only if there is a Hilbert space* $\mathcal{H}$ *with a set of vectors* $|\gamma\rangle$ *labelled by the atoms (singleton sets of histories) of* 2^Ω, *with*

    1. $E \in 2^\Omega$

$$|E\rangle = \sum_{\gamma \in E} |\{\gamma\}\rangle, \quad \forall \tag{2.39}$$

    2. $\forall \vec{x} \in \tilde{\mathcal{S}}$

$$\langle E | E' \rangle = P(E \cap E'), \quad \forall E, E' \in \mathfrak{u}_{\vec{x}} \tag{2.40}$$

We note that for a fixed $\vec{x}$, different outcomes $\vec{a}, \vec{a}'$ decohere for behaviours in JQM and SPJQM, i.e. $D((\vec{a}|\vec{x}), (\vec{a}'|\vec{x})) = 0$. However, decoherence for locally orthogonal events has not been imposed. Consider $\vec{x}, \vec{x}'$ such that they agree for at least one party, i.e. there is at least one party where in both cases they perform the same measurement. We denote one of these parties as i, so that $x_i = x_i'$. If they get different outcomes, $a_i \neq a_i'$, then we say these are locally orthogonal events, $(\vec{a}|\vec{x}) \perp (\vec{a}'|\vec{x}')$. Since $\tilde{Q}$ is consistent with local orthogonality, we know that these events do decohere in these sets of behaviours, as shown by, however it has not been imposed that on JQM or SPQJM.

We are now able to define $Q^1, \tilde{Q}, \tilde{Q}'$ in the Hilbert space formalism, where $\tilde{Q}'$ is the set investigated in [22] and denoted Q^{1+AB}. The reason we provide new notation is that generally Q^{1+AB} has been defined only for bi-partite cases, and this has lead to some confusion in the literature.

Lemma 8. *[22] A behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$ *is in* Q^1 *if there is a Hilbert space* $\mathcal{H}$ *with a set of vectors* $|A\rangle$ *labelled by the fine grained outcomes* $\{|E\rangle\}_{E \in \tilde{\mathcal{O}}}$, *with for every* $\vec{x}, \vec{x}' \in \tilde{\mathcal{S}}$

1. $|E \cup E'\rangle = |E\rangle + |E'\rangle$ if $E, E' \in \mathfrak{u}_{\vec{x}}$ are disjoint

2. $\langle E|E'\rangle = P(E \cap E')$ for all $E, E' \in \mathfrak{u}_{\vec{x}}$

The alternate definition of the almost quantum set is given in [23].

Lemma 9. *A behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$ *is in* $\tilde{Q}$ *if there is a Hilbert space* $\mathcal{H}$ *with a set of vectors* $|A\rangle$ *labelled by the fine grained outcomes* $\{|E\rangle\}_{E \in \tilde{\mathcal{O}}}$, *with for every* $\vec{x}, \vec{x}' \in \tilde{\mathcal{S}}$

1. $\langle \Omega|E\rangle = P(E)$ for all $E \in \mathfrak{u}_{\vec{x}}$

2. $\langle E|E'\rangle = 0$ for all $E \in \mathfrak{u}_{\vec{x}}, E' \in \mathfrak{u}_{\vec{x}'}$ with $E \perp E'$ *(local orthogonality)*

3. $\langle E \cap E'|E \cap E''\rangle = \langle E'|E \cap E''\rangle = \langle E \cap E'|E''\rangle$ *where any parties in* E *are not in* E' *or* E''.

Finally, we state the definition of almost quantum given by Dowker *el al.* in [22], in order to prove that the two sets are indeed equivalent.

Definition 10. *A behaviour* $(\Omega, \tilde{\mathcal{S}}, P)$ *is in* $\tilde{Q}'$ *if there is a Hilbert space* $\mathcal{H}$ *with a set of vectors* $|A\rangle$ *labelled by the fine grained outcomes* $\{|E\rangle\}_{E \in \tilde{\mathcal{O}}}$, *with for every* $\vec{x}, \vec{x}' \in \tilde{\mathcal{S}}$

1. $|E \cup E'\rangle = |E\rangle + |E'\rangle$ if $E, E' \in \mathfrak{u}_{\vec{x}}$ are disjoint

2. $\langle E|E'\rangle = P(E \cap E')$ for all $E, E' \in \mathfrak{u}_{\vec{x}}$

3. $\langle E|E'\rangle = 0$ for all $E \in \mathfrak{u}_{\vec{x}}, E' \in \mathfrak{u}_{\vec{x}'}$ with $E \perp E'$ *(local orthogonality)*

It is then clear that $\tilde{Q}$ is a strengthening of Q^1 to have local orthogonality. This implies that macroscopic locality, which is satisfied if and only if if a behaviour is in Q^1, does not imply local orthogonality. However, we cannot use this result to infer that local orthogonality is a strictly stronger constraint than macroscopic locality as it is known that there are correlations outside of Q^1 that are locally orthogonal [55].

In order to fix the confusion referenced earlier, we prove the following result.

Lemma 11. *For a behaviour in SPJQM, conditions* (1) $\implies$ (2)

1. $\langle X|Y\rangle = 0$ if $X \perp Y$,

2. $\langle X \cap Y|X \cap Z\rangle = \langle Y|X \cap Z\rangle = \langle X \cap Y|Z\rangle$ *where any parties in* X *are not in* Y *or* Z.

Proof. For SPJQM we have the following conditions

1. (a)$\langle X|Y\rangle = 0$ if $X \cap Y = \emptyset$

2. (b) if $X \cap Y = \emptyset$, then $|X\rangle + |Y\rangle = |X \cup Y\rangle$.

To see this, first assume that the behaviour is in SPJQM.

$$\langle \vec{a}', \vec{x}' | \vec{a}, \vec{x}, \vec{a}'', \vec{x}'' \rangle = \sum_{a'''} \langle \vec{a}''', \vec{x}, \vec{a}', \vec{x}' | \vec{a}, \vec{x}, \vec{a}'', \vec{x}'' \rangle \tag{2.41}$$

$$= \sum_{a''' \neq a} \langle \vec{a}''', \vec{x}, \vec{a}', \vec{x}' | \vec{a}, \vec{x}, \vec{a}'', \vec{x}'' \rangle + \langle \vec{a}, \vec{x}, \vec{a}', \vec{x}', | \vec{a}, \vec{x}, \vec{a}'', \vec{x}'' \rangle \tag{2.42}$$

$$= \langle \vec{a}, \vec{x}, \vec{a}', \vec{x}' | \vec{a}, \vec{x}, \vec{a}'', \vec{x}'' \rangle \tag{2.43}$$

where we have used (b) to break up the first line into a sum over all possible measurement outcomes $\vec{a}'''$ given the measurements $\vec{x}$, and then assumed that $(\vec{a}|\vec{x})$ shares no parties with $(\vec{a}'|\vec{x}')$ or $(\vec{a}''|\vec{x}'')$ and used (1) between Eq. 2.42 and Eq. 2.43 to see that each term in the restricted summation vanishes. This demonstrates that (1) $\implies$ (2). □

Theorem 12. *[22] SPJQM $\supseteq \tilde{Q}'$*

Proof. Let $(\Omega, \tilde{\mathcal{S}}, P)$ be in $\tilde{Q}$. It is easy to see that 2 of SPQJM and $\tilde{Q}$ are equivalent.
What we are left with is showing that we can find a set of vectors labelled by the atoms of 2^Ω such that 1 of SPQJM is satisfied. This is left to Appendix 6.2. □

Lemma 13. $\tilde{Q}' = \tilde{Q}$.

Proof. Consider a behaviour in $\tilde{Q}'$. Then use the fact that for any $\vec{x}$, the complete set of finegrained events, $\{\vec{x}\}$ associated with that measurement setting are disjoint, and union to make Ω. Hence,

$$\langle \Omega | E \rangle = \sum_{E \in \{\vec{x}\}} \langle E' | E \rangle = \sum_{E \in \{\vec{x}\}} P(E' \cap E) = P(E) \tag{2.44}$$

where we have used that E', E are disjoint to get $P(E' \cap E) = 0$ if $E \neq E'$, and that 1 of $\tilde{Q}'$ to get the first equality. Hence, we have demonstrated 1 of $\tilde{Q}$, and 2 is obviously equivalent to 3 of $\tilde{Q}'$. Since SPJQM $\supseteq \tilde{Q}'$ by Theorem 12, we can use Lemma 11, to see that 3 of $\tilde{Q}'$ implies 3 of $\tilde{Q}$. Hence, $\tilde{Q}' \subseteq \tilde{Q}$.
Consider a behaviour in $\tilde{Q}$. 2 of $\tilde{Q}$ implies 3 of $\tilde{Q}'$. Using the fact that $\tilde{Q} \subset Q^1$ [23], and lemma 23 of [22] (which we have stated as definition), if a behaviour is in $\tilde{Q}$, then there exists

1. $|E \cup E'\rangle = |E\rangle + |E'\rangle$ if $E, E' \in \mathfrak{u}_{\vec{x}}$ are disjoint

2. $\langle E|E'\rangle = P(E \cap E')$ for all $E, E' \in \mathfrak{u}_{\vec{x}}$

and that they coincide with the vectors given in the definition of $\tilde{Q}$. Hence $1, 2$ of $\tilde{Q}'$ is implied. Thus, $\tilde{Q} \subseteq \tilde{Q}'$, and $\tilde{Q} = \tilde{Q}'$. $\square$

We will thus refer only to $\tilde{Q}$ from now on, and can see that Theorem 12 implies that $\tilde{Q} \subseteq \text{SPJQM}$. We note that historically there appears to have been some confusion in the literature between Q^{1+AB} and $\tilde{Q}$. We want to stress that the results of [22] refer to $\tilde{Q}$ in general, and the use of Q^{1+AB}, which before was only defined for bi-partite scenarios [21], and extended by Dowker *el al*, is historical only, as the original almost quantum paper, and thus the notation $\tilde{Q}$, had not yet been published. Unfortunately, this seems to have been misinterpreted, and Theorem 16 has been quoted as showing that the two sets in question coincide only for the two party case [23], which is not true. We also stress that while the proof of 13 and the supporting lemmas are original, the result was stated by Dowker *el al.* [22] as a simple extension of a similar result for Q^1, and that we have explicitly proven it only to clarify issues in the literature.

Introducing branching

While we have defined the NC space by considering parties that are spacelike, it is possible to use this formalism to consider more general spacetime setups. This allows us to investigate branching, where the outcomes of one party are used to pick the measurement settings for another. This of course requires that the labs are no longer spacelike in general. However, a nonsignalling behaviour will remain nonsignalling.

The main issue with branching is that we end up with experimental setups (choices of measurement settings) where there should be decoherence, that has not previously been imposed. To see this, imagine $f(a, y) = a$. Then, for $\vec{x} = (0,0)$ and $\vec{x}' = (0,1)$, $\vec{a} = (a, b) = (0,0), \vec{a} = (1,0)$, where we stress that $\vec{x}, \vec{x}'$ denote the chosen measurement settings and not the value that is given from the far causal past (these coincided in the non-branching case). While in the non-branching case we have no reason to expect decoherence as the measurement settings are not the same, we must have it in the branching cases, as $(\vec{x}|\vec{a}), (\vec{a}'|\vec{x}')$ are both events that are possible given the same initial event (the measurement setting x being given from the far causal past).

We can now consider an expanded set of measurements, which correspond to all allowed branching setups. The new set $\tilde{\mathcal{S}}_b$, called a branching extension, consists of the usual measurements, as well as those where the measurement settings depend on those before them. Denote the measurement setting of the ith party

depending on the jth as $a_i \prec x_j$. In the branching scenario we discussed above, we can write

$$x \prec (a|x) \prec (a,b|x,y) = (a,b|x,a), \text{and } x \prec (a'|x) \prec (a',b'|x,y') = (a',b'|x,a') \tag{2.45}$$

Note that the use of the causal order symbol is more than just suggestive, the selection of a measurement strictly causally precedes the outcome, and in the branching scenario, the outcome of one party's measurement strictly causally precedes the choice of measurement for another party. In the non-branching setup, all we can say is that

$$x \text{ or } y \text{ or } (x,y) \prec (a|y), (a,b|x,y) \text{ etc.} \tag{2.46}$$

For two parties, the set $\tilde{\mathcal{S}}$ consists of all (x,y), and each (x,y) consists of all $(a,b,|x',y')$ that strictly precede it which is trivially the set where $x = x', y = y'$. For the branching extension, this triviality is lost, and we must consider $(x,y), x, y$ and the events that they precede. The main new condition is that we now have additional probabilities that must be given by the decoherence functional. We will consider only single branches, see [22] for a discussion of why this sufficient.

We note that because branching introduces no new finegrained probabilities, $\tilde{\mathcal{O}}_b$ differs only from its old non-branching version by coarsegrained events. Hence, all branching probabilities follow by the usual coarsegraining/marginalling via the classical sum rule [22].

JQM is already compatible with branching. This is because the decoherence functional $D(E, E')$ can be extended unambiguously to cover the new coarsegrained events in $\tilde{\mathcal{O}}_b$ via the bi-additive property, and decoherence is only required for events E, E' that have the same measurement setting.

Similarly, Q^1 and $\tilde{Q}$ are compatible with branching [22]. The argument for Q^1 is very similar to JQM, while for $\tilde{Q}$, it rest on the fact we have local orthogonality, so all the additional decoherence conditions are automatically met. As discussed by Dowker *el al.* these arguments are not so obviously applied to SPJQM. This is because we do not have LO in general.

Definition 14. *[22] Given a behaviour $(\Omega, \tilde{\mathcal{S}}, P)$ with branching extension $(\Omega, \tilde{\mathcal{S}}_b, P)$, it is in SPJQM$_b$ if there is a decoherence functional on $2^\Omega \times 2^\Omega$ such that*

1. *$D(E, E) = P(E) \forall E \in \tilde{\mathcal{O}}$*

2. *$D(E, E') = P(E \cap E') \forall E, E' \in \mathfrak{u}_{\vec{x}}, E \cap E' = \emptyset, \forall \vec{x} \in \tilde{\mathcal{S}}_b$*

3. *D is strongly positive*

Lemma 15. *Given a behaviour $(\Omega, \tilde{\mathcal{S}}, P)$ with branching extension $(\Omega, \tilde{\mathcal{S}}_b, P)$, it is in $SPJQM_b$ if and only if there is a Hilbert space $\mathcal{H}$ with a set of vectors $|\gamma\rangle$ labelled by the atoms of 2^Ω, with*

1. $\forall E \in 2^\Omega$

$$|E\rangle = \sum_{\gamma \in E} |\{\gamma\}\rangle\,,, \tag{2.47}$$

2. $\forall \vec{x} \in \tilde{\mathcal{S}}_B$

$$\langle E|E'\rangle = P(E \cap E'), \quad \forall E, E' \in \mathfrak{u}_{\vec{x}} \tag{2.48}$$

The following was first shown in [22], however we provide a slightly different version.

Theorem 16. $SPJQM_b = \tilde{Q}$.

Proof. Let $(\Omega, \tilde{\mathcal{S}}, P)$ be a behaviour in $SPJQM_b$. We note that with the new measurements in the branching extension, that $SPJQM_b$ is locally orthogonal, and so 3 of $\tilde{Q}$ is satisfied, 2 of $\tilde{Q}$ follows immediately from non-branching measurements in 2 of $SPQJM_b$. Finally, note that for two disjoint $E, E' \in \mathfrak{u}_{\vec{x}}$, that 1 of $SPQJM_b$ implies
$|E \cup E'\rangle = \sum_{A \in E \cup E'} = \sum_{\gamma \in E} |\{\gamma\}\rangle + \sum_{\gamma' \in E'} |\{\gamma'\}\rangle = |E\rangle + |E'\rangle$ and so we have 1 of $\tilde{Q}$. Hence $SPJQM_b \subseteq \tilde{Q}$.
To see the other direction, we follow Theorem 12 and note that $2, 3$ of $\tilde{Q}$ implies 2 of $SPQJM_b$, as we consider only single branching measurements. In order to get 1 of $SPQJM_b$ we again use Appendix 6.2 to show the existence of atomic vectors. $\qquad\square$

While it has been shown that $\tilde{Q}$ and $SPJQM_b$ coincide, and that $SPJQM_b \subseteq SPJQM$, if this inclusion is strict or not has been left an open question. While consistency with branching is a natural requirement, showing that it follows directly from SP would be a much stronger result.

Chapter 3

Connecting the formalisms

3.1 SPJQM and the device independent sets

The following section explores a number of "closure" conditions, some explored in a sequence of operational papers [56, 57, 24, 55, 23], some in histories [58, 22], and the rest being novel definitions. The aim is connect several concepts together; as shown in the spiderweb diagram Fig 3.1, to give sharp definitions of each; something often lacking in the literature, and to show that one of the more recent concepts, post-processing, is indeed the most general.

3.1.1 Closure conditions

Each closure condition corresponds to demanding that 'classical' actions groups of parties can perform to the data to and from the labs should not create new behaviours that are outside of a set S of behaviours if S is associated with a real theory. Deciding which behaviour from a physical set to generate based on the outcome of a coin toss, distributed to each party, should not generate a behaviour that is not in the physical set. This is the requirement that the behaviours form a convex set, or closure under convex combinations.
We note that while we have previously been imprecise about sets of behaviours—either referring to a set S as the set of behaviours obeying some conditions, regardless of the scenario; and other times implicitly considering subsets of S that correspond to the same fixed scenario—we no longer have that privilege. As we shall see, wiring and other post-processing allows us to take behaviours and produce new behaviours for different scenarios, and so from now on we are considering the set of behaviours S to contain behaviours for all scenarios, unless stated otherwise. In general these will not be Bell scenarios, and in fact there may be cases that even for a fixed measurement setting the number of measurement outcomes of the output vector could vary, depending on wirings. Our formalism is sufficiently flexible to account for this, as we can

consider the outcome of each measurement to be defined over the largest possible range, e.g. if some wiring allows output of $a_1 = \{0, 1, 2\}$ or $\{0, 1, 2, 3, 4\}$ depending on some internal random process, we can simply consider outcomes to be defined over the larger of the two sets.

Finally, we stress that the new behaviours generated by classical processing are *effective*. While we have taken great care to ensure that the outcomes are conditional on the actual measurement settings, this is no longer something we can have, and often the measurement settings of the individual behaviours will be non-deterministic. Instead, the behaviours will be defined on *effective* settings and outcomes, which in general will only have an interpretation as real settings and outcomes if we imagine that all the classical processing is abstracted away, and view the processed behaviours as device independent behaviours themselves, with complete ignorance of the internals and processing.
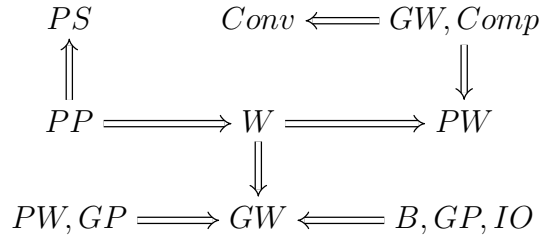
$$
\begin{array}{ccc}
PS & Conv \Longleftarrow GW, Comp \\
\Uparrow & \Downarrow \\
PP \Longrightarrow W \Longrightarrow PW \\
& \Downarrow \\
PW, GP \Longrightarrow GW \Longleftarrow B, GP, IO
\end{array}
$$

Figure 3.1: A spiderweb diagram showing the implications of the closure conditions and some of their combinations. The $\Longrightarrow$ and $\Longleftrightarrow$ arrows represent the usual "implies" and "if and only if" respectively. This is *not* a commutative diagram. The acroynms are defined below from Def to

To begin with, we define a number of closure conditions for sets $S \supseteq L$. This is because we are focusing on classical operations. The first is post-selection, where we condition the behaviour on the input and output of some subset of the parties. Requiring closure under this is natural, we would not expect a GHZ experiment to generate super-quantum 2 party behaviours if we conditioned the 3 party behaviour on $c = 0, z = 0$ etc.

Definition 17. *A set of behaviours S is closed under post-selection (PS) if for all $P(\vec{a}|\vec{x}) \in S$, $P(\tilde{\vec{a}}|\vec{x}, a_i) \in S$, where $\tilde{\vec{a}}$ is $\vec{a}$ with the ith party removed.*

Composition is also quite natural, for example, two independent entanglement experiments should not (and do not) give an effective behaviour that is not explainable by quantum theory.

Definition 18. *A set of behaviours S is closed under composition (Comp) if for all $P(\vec{a}|\vec{x}), P'(\vec{a}'|\vec{x}'), PP'(\vec{A}|\vec{X}) \in S$, where P, P' are n, n'-partite and PP' is $n + n'$-partite, $\vec{A}, \vec{X}$ are concatenations of $\vec{a}, \vec{a}'$ etc, and juxtaposition is normal multiplication.*

This allows us to consider independent composition of multiple behaviours, with no grouping or sharing of the behaviours between parties. For instance, two (2 2 2) behaviours shared by Alice and Bob, and Charlie and Danny can be composed to give a (4 2 2) behaviour shared by Alice, Bob, Charlie and Danny, where there is no 4 party correlation due to the independence of the two original behaviours. Clearly, this is a very weak condition, as no novel correlations can be generated by composition alone. It is, however, a useful stepping stone for defining and proving closure under more complex conditions.
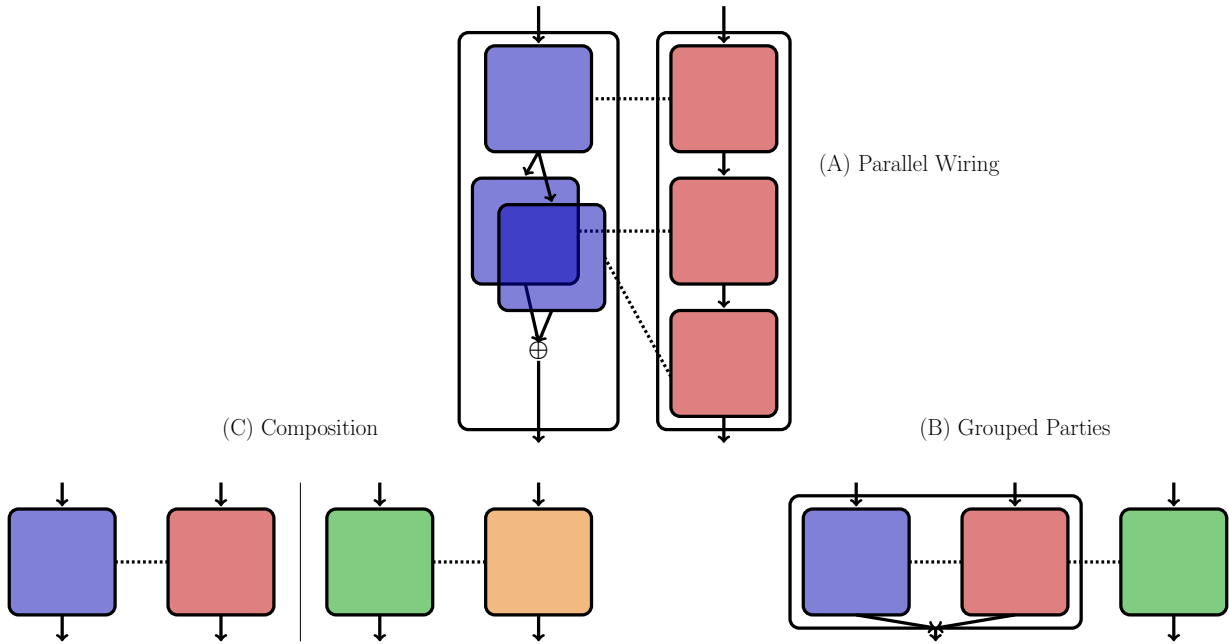

Figure 3.2: A representation of three classical processing scenarios. A box represents a party (or effective party if multiple boxes are contained within it), arrows in (out) represent, possibly effective, measurement settings (outcomes), and dashed lines show how each party in a behaviour is related(A) Parallel wiring of three bi-partite behaviours to create one effective bi-partite behaviour. Note that the outcomes of each of the original behaviours is not outputted, and so there is some identification of outputs. (B) grouping of a tri-partite behaviour to create an effective bi-partite behaviour. (C) Composition of two bi-partite behaviours to create an effective quad-partite behaviour. The solid line emphasises that composition does not introduce any additional correlations, the behaviours are composed in an independent way.

Definition 19. *A set of behaviours S is closed under identification of outcomes (IO), if a new effective behaviour defined by coarsegraining over the measurement outcomes is also in S. I.e. one or more parties can decide to reduce the number of effective outcomes by associating multiple together, e.g. Alice could decide that her outputs, defined over $\{0, 1, 2, ..., p-1\}$ are too numerous, and that $a_1 = p-2$ and $a_1 = p-1$ should be associated together, and output an effective output $\tilde{a}_1$ defined only over $\{0, 1, 2, ...p-2\}$*

Mathematically, if $P(\vec{a}|\vec{x}) \in S$,

$$P(\tilde{\vec{a}}|\vec{x}) = \sum_{\{\vec{a}'\}} P(\vec{a}'|\vec{x}) \tag{3.1}$$

must be in S, where we are summing over all the outcomes that are associated together.

The following definition is split into 3 parts in order to address some complications and overloading of terminology in the literature.

Definition 20. *A set of behaviours S is closed under*

1. *parallel wiring (PW), if N parties each have access to their part of n N-partite boxes corresponding to behaviours $P(\vec{a}_i|\vec{x}_i)_i \in S$ such that they can locally wire their part of the boxes together in any way, and the new effective behaviour $P(\tilde{\vec{a}}|\tilde{\vec{x}})$ is in S.*

2. *grouped wiring (GW), if $M < N$ parties partition themselves into C subsets and group together into those C groups, and wire their parts of the box together, creating $N - M + C$ effective parties, so that we have coarsegrained over the parties within the groups.*

3. *wiring (W), if it closed under 1 and 2.*

Wiring involves any (possibly non deterministic) local strategy of taking the inputs and outputs of boxes and using them (perhaps after feeding them into some function) to determine others. To see an example, imagine Alice and Bob both share three behaviours. Alice takes the inputs to all 3, adds them modulo 3, and then uses that to determine which of the three behaviours to output. Meanwhile, Bob uses the outputs of the first behaviour modulo 2 to determine if he should output the results of behaviour 2 or 3, which have had their proper unaltered inputs.

Definition 20.1 is the original definition of wiring, as introduced in [57] and explored further in [59], 20.2 is used in [24],[23] while the broader 20.3 is implied in [23]. As we shall see, under some safe assumptions, $2 \iff 1$ and thus 1 or $2 \implies 3$.

Clearly, there are a huge number of ways to wire even a modest number of (2 2 2) behaviours. To make this manageable, we can consider only complete (with no identification of outputs) deterministic wirings, and use closure under convex combinations and identification of outputs to cover the remaining cases. The example given above is deterministic but non-complete, as Alice and Bob

both identify outputs together, both margining out two of the behaviours. To make it complete, Alice and Bob must provide some inputs to all 3 behaviours and output the results for all 3.

Definition 21. *[23] A set S of behaviours is closed under post-processing (PP) if it closed under post-selection, composition, and wiring.*

Finally, we introduce an almost trivial closure concept and corollary, which will allow us to connect our discussion to branching.

Definition 22. *A set of behaviours is closed under grouped parties (GP) if $M < N$ parties partition themselves into C subsets and group together into those C groups, creating $N - M + C$ effective parties, so that we have coarsegrained over the parties within the groups.*

Note there is no wiring involved here, the parties in a group simply group their outcomes and measurement settings to give an effective setting and outcome for the group, thus acting as one party.

Corollary 23. *Grouped wiring implies grouped parties.*

We now present the non-trivial results expressed in Fig 3.1, with the aim of fully justifying some common claims in the literature.

Lemma 24. *Grouped wiring and composition implies parallel wiring.*

Proof. Suppose we have some parallel wiring scenario created by wiring n lots of N party boxes together. We can recreate this by composing the n boxes together, and then grouped wiring them in the obvious way. Since we then have closure under grouped and parallel wiring, we have closure under wiring. □

Lemma 25. *Parallel wiring and grouped parties implies grouped wiring.*

Proof. By grouping parties together, we can use parallel wiring to recreate any grouped wiring. □

Since grouped parties and composition are both relatively weak conditions, and wiring is satisfied if both parallel and grouped are satisfied, we can now understand the somewhat confusing terminology and definitions used in the literature. Navascués *el al* were able to prove wiring for the almost quantum set by proving grouped wiring after composition in [23]. While their proof is technically of a different closure condition to the original wiring proposal [57], we have shown that they are equivalent under the right conditions.

Lemma 26. *Convex follows from composition and grouped wiring.*

Proof. Suppose we want to show that

$$\sum_{i=1}^{n} p_i P_i(\vec{a}|\vec{x}) \tag{3.2}$$

is in S for any set of probabilities p_i and set of N party behaviours with equal sized inputs/outputs $P_i \in S$.

A set S that is closed under grouped wiring is closed under wiring 24, and so satisfies $\mathcal{L} \subseteq S$ 27. Let $P_R(\vec{a}|\vec{x}) \in \mathcal{L}$ be an N-party source of shared randomness, so that

$$P_R(\vec{a}|\vec{x}) = p_{\vec{a}} \tag{3.3}$$

such that the probability is non zero only if $\vec{a}_i = i$ for all parties i and $p_{\vec{a}} = p_i$. Such a box is clearly local and so is in S. Take the following composition

$$P_R \Pi_{i=1}^{n} P_i \tag{3.4}$$

and group wire the boxes so that each party gets the result i of the shared randomness, uses it to pick P_i and then inputs $\vec{x}_i$ and returns $\vec{a}_i$. Clearly, this is equivalent to Eq.3.2. $\qquad \square$

These results underpin much of the discussion of post-processing, yet, to the author's knowledge, have never been stated. They also lead us to the following interesting result.

Lemma 27. *The smallest set closed under grouped wiring and composition (or post-processing) is L.*

Proof. The trivial behaviour is $P(\vec{a}|\vec{x}) = \delta_{\vec{a}\vec{x}}$. By Lemmas2624 and 26, grouped wiring and composition implies wiring and convexity. By [57], parallel wiring of the trivial behaviour is sufficient to generate all extremal points of L, and L is convex, so we can produce the full set L from it. Since L is closed under all post-processing, it is the smallest set closed under grouped wiring and composition (or post-processing). $\qquad \square$

This reassures us that post-processing can exactly pick out the local behaviours. This leads us to conclude that post-processing is indeed the most general of the above closure conditions. While we are still left with branching, which we can view as a classical process on behaviours, it takes a somewhat different flavour, and was not strictly introduced as a closure condition.

We can define the branching completion of a set S to be S^b. This is the set of behaviours that are generated by all branching extensions of all behaviours in S.

It is clear that $S^b \supseteq S$, as the branching extension includes the non-branching scenario. We also know that for $Q, Q^1, \tilde{Q}$, and JQM, $S = S^b$.

The remaining question is if SPJQM=SPJQM^b, or equivalently if SPJQM$_b$=SPQJM. We know that (SPJQM$_b$)^b=SPJQM$_b$ by definition. Hence, we have the following relation

$$\text{SPJQM}_b \subseteq \text{SPJQM} \subseteq \text{SPQJM}^b. \tag{3.5}$$

Corollary 28. *For a subset $S \subseteq$ SPJQM, branching and grouped parties implies wiring.*

The following results are applications of the closure conditions to some of the device independent sets we have been discussing. The proofs are not dissimilar to our own proofs for SPJQM in Subsection 3.1.2, and so we do not include them.

Theorem 29. *[55] Q^1 is closed under wiring and convex combinations.*

Theorem 30. *[23] $\tilde{Q}$ is closed under post-processing.*

It was proven by Lang *el al* [55] that there is a countably infinite nested hierarchy of sets closed under wiring and convex, the NPA hierarchy sets. The work also raised two questions regarding the structure of correlations 1) Are the wiring and convex closed sets a continuum and 2) are all wiring and convex closed sets are in the nested hierarchy [1]? The latter was shown to be false by Lang *el al*, while the former was, indirectly, swiftly answered in the positive by Beigi *el al* [60]. It is not known if these results extend to the more general post-processing.

One thing we should note about post-processing, is that we are considering only classical processes. Unlike in a full theory, we are not able to use the device independent formalism to consider the most general transforms under which a set of behaviours should be closed. For instance, in QM, we expect that for 1 $E_{a_i}^{x_i} \mapsto U^{x_i} E_{a_i}^{x_i} U^{x_i \dagger}$ changes the behaviour $P(\vec{a}|\vec{x})$, yet is still in Q. Since every operation can be considered a unitary (up-to conditioning), this is sufficient to show full physical closure of the theory with respect to sets of behaviours. To show a similar result for other sets, we need a definition of them in terms of states, measurements and maps between states (kets, POVMs, unitaries). This is the realm of GPTs, yet unfortunately, it is not simple, or guaranteed to be possible, to construct GPTs from a set of behaviours.

---

[1]These appear as questions 3 and [negation of] 5 in [55]

3.1.2 Applying device independent closure to histories sets

Since we expect device independent closure conditions to hold for a physical theory, we should investigate sets that have appeared in the histories framework. Firstly, we can compared JQM and SPJQM, and see that JQM fails one of the most basic closure conditions.

Theorem 31. *[54] JQM is not closed under composition*

Theorem 32. *[58][22] SPJQM is closed under composition.*

Proof. Suppose P_1, P_2 are behaviours in SPJQM, then they are equivalent to strongly positive decoherence functionals D_1, D_2, which are naturally isomorphic with positive semi-definite matrices $\hat{D}_1, \hat{D}_2$ on their respective spaces of atoms. The composition is given by $\hat{D}_1 \otimes \hat{D}_2$ which is itself a positive semi-definite on the tensor product of the two atomic spaces. Thus there is a behaviour P_J that is equivalent to a strongly positive decoherence functional D_J, and so P_J is in SPJQM.

$\square$

At the level of the Hilbert space construction, this follows from the fact that $\mathcal{H}_1 \otimes \mathcal{H}_2$ inherits an inner product $\langle \cdot, \cdot \rangle_J$ that is the product of the inner products $\langle \cdot, \cdot \rangle_i$ on each $\mathcal{H}_i$, giving the probabilities as $P_1(\cdot)P_2(\cdot)$ as required for independent scenarios.

Having seen that weak positivity is far too weak of a constraint, we can focus on SPJQM.

Theorem 33. *SPJQM is closed under post selection.*

Proof. Suppose $(\Omega, \tilde{\mathcal{S}}, P)$ is in SPJQM. Then there exists a Hilbert space $\mathcal{H}$ with vectors obeying the following conditions,

1. $\forall X \in 2^\Omega$,

$$\sum_{\gamma \in E} |\{\gamma\}\rangle \tag{3.6}$$

2. $\forall \vec{x} \in \tilde{\mathcal{S}}, \forall E, E' \in \mathfrak{u}_{\vec{x}}$

$$\langle X | Y \rangle = P(X \cap Y). \tag{3.7}$$

Consider a partition of the parties into R, S, where we post-selecting for $(\vec{a}_R | \vec{x}_R)$. Then the new set of histories, Ω', is the set of all histories with $\gamma_{\vec{x}_R 1} = \vec{a}_R$. The event space

$$2^{\Omega'} \tag{3.8}$$

is a subset of the non-post-selected event space, so that every event $E' \in 2^{\Omega'}$ is in 2^{Ω}. All the usual subsets of the event space, e.g. $\tilde{\mathcal{O}}', \tilde{\mathcal{S}}'$, are defined similarly, and distinguished from the non-post-selected versions by a inverted comma (we note that all post-selected subsets are subsets of the non-post-selected subsets). Consider that the probability distribution P' is defined on $\tilde{\mathcal{O}}'$ as

$$P'(E') = \frac{P(E')}{P(\vec{a}_R | \vec{x}_R)}, \quad \forall E' \in \tilde{\mathcal{O}}'. \tag{3.9}$$

Hence, there is a behaviour $(\Omega', \tilde{\mathcal{S}}', P')$. For every $E' \in 2^{\Omega'}$, we define

$$|E'\rangle_{ps} = \frac{1}{P(\vec{a}_R | \vec{x}_R)^{\frac{1}{2}}} |E'\rangle \tag{3.10}$$

and so for every history $\gamma' \in \Omega'$

$$|\{\gamma\}\rangle_{ps} = \frac{1}{P(\vec{a}_R | \vec{x}_R)^{\frac{1}{2}}} |\{\gamma\}\rangle . \tag{3.11}$$

The span of all $|E'\rangle_{ps}$, $\mathcal{H}_{ps}$, is a subspace of $\mathcal{H}$. Hence condition (1) is satisfied. Consider any two physical events $E', F' \in \mathfrak{u}_{\vec{x}'}$ for any $\vec{x}' \in \tilde{\mathcal{S}}'$ (since these are the post-selected subsets, these are automatically post-selected events). Then to satisfy (2), we need

$$P'(E' \cap F') = {}_{ps}\langle E' | F'\rangle_{ps} . \tag{3.12}$$

which follows directly from Eqs 3.9 and 3.10. Hence, SPJQM is closed under post-selection. $\qquad \square$

This result does not appear surprising, however there are nontrivial sets that are not closed under post-selection [57].
The following result is rather trivial, yet notationally cumbersome.

Lemma 34. *SPJQM is closed under identification of outputs.*

Proof. Suppose $(\Omega, \tilde{\mathcal{S}}, P)$ is in SPJQM. Then we have the usual Hilbert space construction. Suppose, without loss of generality, that the first party decides to identify two of its outputs for measurement 1, labelled $\vec{a}_1 = 1$ and 2, together to give a new effective set outputs. The 'vector' of outputs $\vec{a}$ has

$$\vec{a}_1 \sim \vec{a}_1' \text{ if } \vec{a}_1, \vec{a}_1' = 1, 2, \text{ and } \vec{x}_1 = 1, \tag{3.13}$$

where $\sim$ denotes the identification choice. Then we have a new behaviour described by $\bar{P}(\vec{\alpha} | \vec{x}) = \bar{P}(\bar{E})$ where $\vec{\alpha}$ is the new effective output vector due to identification such that $\vec{\alpha}_1 = 1$ denotes the new effective output due to the

identification of the 1 and 2 outputs to measurement 1, and $\vec{\alpha}_1 = 2, ..., m_1 - 1$ denotes $\vec{a}_1 = 3, ..., m_1$. Hence, if $\vec{\alpha}_1 = 1 = \vec{x}_1$

$$\bar{P}(\vec{\alpha}|\vec{x}) = P(\vec{a}|\vec{x}) + P(\vec{a}'|\vec{x}) \tag{3.14}$$

and if $\vec{\alpha}_1 \neq 1$ or $\vec{x}_1 \neq 1$ then we identify $\bar{P}(\alpha|\vec{x}) = P(\vec{a}|\vec{x})$. or more generally

$$\bar{P}(\bar{E}) = \sum_{E' \in T(\bar{E})} P(E') \tag{3.15}$$

where $T(\bar{E})$ is the set of all events E that are identified together to give $\bar{E}$. We define the new atomic vectors as

$$|\{\bar{\gamma}\}\rangle = \sum_{\gamma' \in S(\bar{\gamma})} |\{\gamma'\}\rangle \tag{3.16}$$

where $S(\bar{\gamma})$ is the set of all atoms of the original behaviour that are identified with $\bar{\gamma}$. Then for $\bar{E} \in \vec{x} \in \bar{\tilde{\mathcal{S}}}$

$$|\bar{E}\rangle = \sum_{\bar{\gamma} \in \bar{E}} \sum_{\gamma' \in S(\bar{\gamma})} |\{\gamma'\}\rangle = \sum_{E' \in T(\bar{E})} \sum_{\gamma' \in E'} |\{\gamma'\}\rangle = \sum_{E' \in T(\bar{E})} |E'\rangle \tag{3.17}$$

where the second equality follows by Appendix 6.1 and the final equality follows from the original Hilbert space. Finally, given $\bar{E}, \bar{F} \in \tilde{\mathcal{O}}$

$$\bar{P}(\bar{E} \cap \bar{F}) = \sum_{E' \in T(\bar{E})} \sum_{F' \in T(\bar{F})} \langle E'|F'\rangle = \sum_{E' \in T(\bar{E})} \sum_{F' \in T(\bar{F})} P(E' \cap F')$$

$$= \sum_{G' \in T(\overline{E \cap F})} P(G') = \sum_{G' \in T(\bar{E} \cap \bar{F})} P(G') \tag{3.18}$$

where it is easy to see that the sum in the final expression is over all events in the original behaviour that have been identified together. Hence SPJQM is closed under one party identifying two outputs. By induction we have our final result. $\qquad \square$

We require the following trivial corollary, immediate from the definition of SPJQM.

Corollary 35. *SPJQM is closed under grouping of parties.*

Lemma 36. *Any parallel wiring of SPQJM behaviours with effective settings and outcomes $(\tilde{\vec{a}}|\tilde{\vec{x}})$ has a representation $|\tilde{\vec{a}}, \tilde{\vec{x}}\rangle$ that obeys condition 2 of SPJQM, i.e. that for any measurement outcomes (fine or coarsegrained), E, E'*

$$P_w(E \cap E') = \langle E|E'\rangle \tag{3.19}$$

Proof. We will consider only deterministic complete wirings, where n devices are shared between N parties. We can create the pre-wiring behaviour by sharing n sets of N-party behaviours. This can be represented by composing all n behaviours to produce an nN-party behaviour. Then group the parties as follows. The kth effective party consists of the $k, K + N, k + 2N...$th parties, giving N parties each made of n grouped parties. By Theorem 32 and Lemma 35 this new behaviour is an element of SPJQM. Finally, we must show that any deterministic wiring gives another SPJQM behaviour. Non deterministic wirings and non complete wirings follow from using a random source as one of the shared devices, and Lemma 34 respectively.

We clearly have the correct atomic Hilbert space structure, following the tensor product used in composition, and so 1 of SPJQM is satisfied. Hence, we need only show that $P(E \cap E') = \langle E | E' \rangle$ for any E, E' corresponding to the same inputs (since we are allowing arbitrary functions to be applied to the inputs, and using outputs to pick other inputs, we distinguish inputs from measurement settings). Hence, consider a fixed input, denoted $\vec{I}$. This contains the *potential* settings for all nN parties, however in general they will be modified before becoming *actual* measurement settings. We can consider the wirings as a chain of deterministic functions, inputs and outputs etc.

Consider an effective event with effective settings $\tilde{\vec{x}}$ and effective outcomes $\tilde{\vec{a}}$. For each party, the settings of each behaviour are either decided directly from the effective settings, or by a deterministic function of the effective settings and the outcomes of other behaviours. Since it is a complete wiring, the effective outcomes are simply a concatenation of the individual outcomes of the behaviours, so the values of any functions of the effective settings and outcomes of previous behaviours are fully determined, and hence the actual settings for all behaviours are full determined by $(\tilde{\vec{a}}|\tilde{\vec{x}})$. Hence, we define the finegrained vector associated to $(\tilde{\vec{a}}|\tilde{\vec{x}})$ to be

$$|\tilde{\vec{a}}, \tilde{\vec{x}}\rangle = \bigotimes_{i=1}^{n} |\vec{a}^i, \vec{x}^i\rangle \tag{3.20}$$

where i represents each behaviour in the wiring, x_j^i in general depends on a $a_j^{i'}$ and $\tilde{x}_j^{i'}$ for some i' representing preceding behaviours in the wiring chain, and the concatenation of $\vec{a}^i$ gives $\tilde{\vec{a}}$. Clearly, the inner product between any two fine grained outcomes will either vanish if the $\tilde{\vec{a}}$ do not match, or give $P_w(\tilde{\vec{a}}|\tilde{\vec{x}})$. As before, to coarsegrain out any other parties, we sum all finegrained vectors that agree on the parties we are not coarsegraining. The correct probabilities follow.

In order to satisfy condition 1 of Def 7, invert

$$|\tilde{\vec{a}}, \tilde{\vec{x}}\rangle = \sum_{\gamma \in (\tilde{a}|\tilde{x})} |\{\gamma\}\rangle, \tag{3.21}$$

to get the atomic vectors. From there, define

$$|E\rangle = \sum_{\gamma \in E} |\{\gamma\}\rangle, \tag{3.22}$$

and we are done. $\qquad\square$

Corollary 37. *SPJQM is closed under post-processing.*

3.2 Branching

We can now return our focus to branching, now using the lens of closure conditions.

Lemma 38. *A behaviour in SPJQM is in SPJQM$_b$ if and only if a vector space construction can be found that satisfy the non-branching probabilities and the additional decoherence (i.e. local orthogonality) conditions.*

Proof. The backwards direction is immediate via contradiction. To see the forward direction, take a behaviour in SPJQM and construct its unique branching extension. All additional probabilities are coarse grained [22], and given by the classical sum rule, and derivable from the non-branching vectors by the vector sum rule, or they vanish due to LO, giving decoherence conditions. After fixing the decoherence conditions, the atomic vectors can be found via inverting the relations

$$|E\rangle = \sum_{\gamma \in E} |\{\gamma\}\rangle. \tag{3.23}$$

This is possible as the number of atoms and physical events is equal. $\qquad\square$

Hence, a proof that all behaviours in SPJQM are in SPQJM$_b$ could involve finding a set of unitary matrices that rotate a normal SPJQM set of event vectors, so that the new LO conditions are met, e.g.

$$|\vec{a}, \vec{x}\rangle \mapsto U_{\vec{x}} |\vec{a}, \vec{x}\rangle \tag{3.24}$$

We then see that

$$\langle \vec{a}', \vec{x}'| U_{\vec{x}'}^\dagger U_{\vec{x}} |\vec{a}, \vec{x}\rangle \tag{3.25}$$

recovers the original probabilities when $x = x'$ by unitarity, and we aim to pick the operators such that the inner products vanish if $\vec{x}, \vec{x}'$ agree on at least one

party, and $\vec{a}, \vec{a}'$ disagree on the same party, i.e. local orthogonality. Finding such a set of operators is rather trivial,

$$U_{\vec{x}} = \sum_{\vec{a}''} |\vec{a}''\rangle \langle \vec{a}'', \vec{x}|_N \tag{3.26}$$

can be seen to satisfy exactly that, where $|\vec{a}, \vec{x}\rangle_N$ is a normalised and complete basis constructed from the associated event vectors (which are orthogonal, but not normal or complete), and $|\vec{a}\rangle$ is any orthonormal basis.

$$\langle \vec{a}', \vec{x}' | U_{\vec{x}'}^\dagger U_{\vec{x}} | \vec{a}, \vec{x}\rangle = \sum_{\vec{a}''} \langle \vec{a}', \vec{x}' | \vec{a}'', \vec{x}' \rangle_N \,_N\langle \vec{a}'', \vec{x} | \vec{a}, \vec{x}\rangle = \sum_{\vec{a}''} \delta_{\vec{a}, \vec{a}''} \delta_{\vec{a}', \vec{a}''} = \delta_{\vec{a}, \vec{a}'} \tag{3.27}$$

as required. However, in order to meet the conditions we must also have the correct coarsegrained vectors and inner products. Consider the (2 2 2) scenario with finegrained vectors $|a, b, x, y\rangle$. To get all coarsegrained probabilities it is sufficient to ensure that

$$U_{xy} |a, x\rangle = |a, x\rangle \tag{3.28}$$

etc, where we have absorbed the obvious global freedom $U_{xy} \mapsto V U_{xy}$. This imposes 4 additional conditions on each unitary, in the form of eigenvector/values, and cannot be met in general, or even for all classical behaviours, by Eq 3.26:

$$U_{\vec{x}} = \sum_{\vec{a}''} |\vec{a}''\rangle \langle \vec{a}'', \vec{x}|_N |a, x\rangle \tag{3.29}$$

$$= P(a, b|x, y) |a, b, x, y\rangle_N + P(a, \bar{b}|x, y) |a, \bar{b}, x, y\rangle_N \tag{3.30}$$

$$= P(a, b|x, y)^{\frac{1}{2}} |a, b, x, y\rangle + P(a, \bar{b}|x, y)^{\frac{1}{2}} |a, \bar{b}, x, y\rangle \tag{3.31}$$

$$\neq |a, b, x, y\rangle + |a, \bar{b}, x, y\rangle \tag{3.32}$$

unless either $P(a, b|x, y)$ or $P(a, \bar{b}|x, y)$ vanish. Note that for fixed x, $|a, x\rangle$ form an orthogonal set, and so must span an eigenspace, and likewise for $|b, y\rangle$. However, the two eigenspaces are, in general, not equal or orthogonal, and so typically the joint space they span is dimension 3. Let the space $\mathcal{H}_{x,y} \subset \mathcal{H}$ be the eigenspace

$$\mathcal{H}_{x,y} = \operatorname{span}_{a,b}\{|a, x\rangle, |b, y\rangle\}, \tag{3.33}$$

and $|i, x, y\rangle$ be an orthonormal basis spanning it. We can expand this basis to the whole of $\mathcal{H}$, and define the new unitaries

$$U_{xy} = \sum_i |i\rangle \langle i, x, y| \tag{3.34}$$

where $|i\rangle$ is any fixed orthonormal basis. We then have

$$\sum_i \langle a, b, x, y | i, x, y \rangle \langle i, x', y' | a', b', x', y' \rangle = \langle a, b, x, y | V_{xyx'y'} | a', b', x', y' \rangle. \quad (3.35)$$

Clearly, when $x = x', y = y'$, V reduces to the identity and all the probabilities are as expected.

Assume that $x = x'$, and $a \neq a' = \bar{a}$. Then

$$\sum_i \langle a, b, x, y | i, x, y \rangle \langle i, x, y' | \bar{a}, b', x, y' \rangle = \langle a, b, x, y | V_{xyxy'} | \bar{a}, b', x, y' \rangle. \quad (3.36)$$

must vanish.

It is important to note that non-signalling implies that for the (2 2 2) scenario,

$$| \langle a, b, x, y | \bar{a}, b', x, \bar{y} \rangle | = | \langle a, b, x, y | a', \bar{b}, \bar{x}, y \rangle | \quad (3.37)$$

where $\bar{y}$ is the opposite of y, is constant with respect to a, b, b', x, y, e.g.

$$\langle a, b, x, y | \bar{a}, b', x, \bar{y} \rangle + \langle a, \bar{b}, x, y | \bar{a}, b', x, \bar{y} \rangle = \langle a, x | \bar{a}, b', x, \bar{y} \rangle = 0. \quad (3.38)$$

This implies that if one LO inner product is zero, and all condition 2 of Def 7 is satisfied, the all LO inner products are zero. Since any unitary U_{xy} leaves condition 2 alone, we simply need a family of U_{xy} that send any of the inner products of the form Eq 3.37 to zero.

This gives us the following conditions on the family U_{xy} where we take a, b, b' fixed

1. $\langle a, b, x, y | U_{xy}^\dagger U_{x\bar{y}} | \bar{a}, b', x, \bar{y} \rangle = 0 = \langle a, b, x, y | U_{xy}^\dagger U_{\bar{x}y} | a', \bar{b}, \bar{x}, y \rangle = 0$

2. $U_{xy} | a, x \rangle = | a, x \rangle$ and $U_{xy} | b, y \rangle = | b, y \rangle$

or

1. $\langle a, b, x, y | U_{xy}^\dagger U_{x\bar{y}} | \bar{a}, b', x, \bar{y} \rangle = 0$

2. $U_{xy}^\dagger U_{x\bar{y}} | a, x \rangle = | a, x \rangle$ and $U_{xy}^\dagger U_{\bar{x}y} | b, y \rangle = | b, y \rangle$

While we do not complete a proof of existence for these unitaries, these results seem to imply that at least for the (2 2 2) scenarios, SPJQM is compatible with branching. This does not go against our intuition, as we have already seen that (2 2 2) is a special scenario with respect to local orthogonality (see Section 2.1.2) and as pointed out earlier there is a relationship between local orthogonality and branching that has not yet been fully investigated. While scenarios with more parties, measurements, or settings do not have the same automatic local orthogonality as (2 2 2), this may not be an issue, as branching seems more closely related to orthogonal sets of size 2, which always satisfy the normalisation property. This suggests that branching may always be possible for SPJQM, however we have not shown this.

Chapter 4

Operational QFT

4.1 GPTs and Ludwig's theorem

4.1.1 The no-restriction hypothesis

As well as the reasons provided in section 2.2.2, another reason to believe that the no-restriction hypothesis is not a good choice is that its analogue is far less obvious in the case of QFT. As shown by Sorkin, a naive extension of the projective measurement postulate allows for signalling measurements[5]. Further research has shown that this result holds even when considering (smeared) local operators [61]. This is in stark contrast to QM, where no-signalling is an almost trivial theorem. GPTs do appear not to have been extended in the literature in such a way to cover QFTs, often assuming state spaces that are isomorphic to $\mathbb{R}^d$ or at the very least separable. The space of states in typical, or rather; realistic, QFTs are neither. Since QFTs have been known to be a better model of reality than QM for over 70 years, it is imperative that QFT is folded into the mix. It is often implied that quantum foundations research is consistent with special relativity, mainly due to the ironclad loyalty to non-signalling, however non-signalling is only one of many consequences of relativity[12], and many papers explicitly accept that including QFT is a significant complication due to the many difficulties posed by the theory[63]. It is not clear how many authors are thinking of measurements, and how many are thinking of the complications of uncountable dimensional Hilbert spaces, such as divergent physical observables, unitarily inequivalent representations aka the choice problem, or any other host of issues. Of course, the latter *are* significant technical issues, however a general consensus has been reached by field theorists, due to the success of Wilson's and others' work on effective field theory, that such issues stemming from the Hilbert

---

[1] For instance, non-relativistic QM obeys non-signalling in the device independent framework, yet allows arbitrarily high velocities and thus violates causality.

[2] See [62] for a discussion on the relationship between non-signalling and causality

space can almost always be absorbed into counter-terms, and are not pathological [3]. Since these counter terms are fixed by a choice of finite cut-off, with a finite number of measurements the theory is fixed, becomes fully predictive, and can be studied in an operational sense. With measurement however, it does not seem clear how many of the operational approaches carry over in any sense to a theory that has yet to have an operational definition of measurement formulated. Non-signalling is imposed immediately by QM, yet must be put in by hand in QFT. This 'embarrassment' has been avoided both by field theorists and foundations researchers, who have considered only single measurements on compact or boundary spacelike surfaces (e.g. detector geometries) and finite dimensional QM respectively.

4.1.2 QFT as a GPT

To the author's knowledge, there are no published explorations of QFT in the GPT framework, as the vast majority of works assume a finite dimensional state space. There are, however, studies based on infinite dimensional spaces, and Ludwig's theorem applies equally to infinite dimensional vector spaces as it does to finite dimensional. Here, we look at the minimal requirements a QFT[1] must meet as a GPT. Finally, we provide evidence towards a conjecture, based on [61], that no QFT can satisfy the no-restriction hypothesis.

To begin with, we give an argument that any QFT satisfies Ludwig's theorem, which gives sufficient conditions for a theory to be a GPT. The GPTs considered are based on Banach spaces [64, 65], rather than the more typical finite dimensional real spaces [31].

We will work in the algebraic QFT framework (AQFT), and consider only free real scalar fields. Observables and states in AQFT are given by self adjoint operators and linear functionals of them. More concretely, consider the usual unital $*$-algebra of all observables $\mathcal{A}$, generated by (smeared) field operators $\phi(f)$ and the identity, and the set of positive linear functionals (we do not take these to be normalised) of $\mathcal{A}$, denoted $\mathcal{S}(\mathcal{A})$.

To begin, we need the following three axioms, which are taken, in inexact form, from. More precise formulations can be found in the source.

1. States that are not distinguishable by any effect are, to all intents, the same.

2. There are effects that 'always accept' and 'never accept', and for any effect, its negation is also an effect.

---

[1]From here onwards, 'QFT' refers only to relativistic QFTs.

3. One can perform probabilistic mixtures of states (effects) and obtain a valid state (effect), and purposes the same state, and vice versa.

We will assume that a representation exists, so that every state can be associated to a vector on the Hilbert space. Firstly, two states $|\psi\rangle, |\psi'\rangle$ cannot be distinguished by an effect are in the same ray, and so we can consider the state space to be the set of rays.

Secondly, there is always an identity operator $\mathbb{I}$, and a null operator $\mathbb{O}$. The actions $\frac{1}{\mathrm{tr}(\rho)} \mathrm{tr}(\mathbb{I}\rho) = 1$ and $\frac{1}{\mathrm{tr}(\rho)} \mathrm{tr}(\mathbb{O}\rho) = 0$ give us our always yes and no operators. Suppose ρ, ρ' are non trivial states, satisfying

$$\frac{\mathrm{tr}(\mathbb{P}\rho')}{\mathrm{tr}(\rho)} = 0, \quad \frac{\mathrm{tr}(\mathbb{P}\rho')}{\mathrm{tr}(\rho')} = 1 \tag{4.1}$$

then, by linearity

$$\frac{\mathrm{tr}((1 - \mathbb{P})\rho')}{\mathrm{tr}(\rho')} = 1, \quad \frac{\mathrm{tr}((1 - \mathbb{P})\rho')}{\mathrm{tr}(\rho')} = 0 \tag{4.2}$$

and so there exists a negation.

Finally, the state space automatically includes density operators, and so we are only left with showing that mixtures of effects are effects. However, effects are clearly also closed under convex operations as they are an algebra.

To investigate GPTs in more detail we must introduce several concepts from convex analysis.

Definition 39. *An ordered vector space is a real vector space V with an ordering $\leq$ such that obeys translation*

$$x \leq y \implies x + z \leq y + z, \tag{4.3}$$

and positivity

$$x \leq y \implies \lambda x \leq \lambda y \tag{4.4}$$

for all $\lambda > 0$.

An important subset of the real vector space can be generated by a convex set.

Definition 40. *A subset C of a real vector space V is a cone if it is closed under addition, multiplication by positive scalars, and contains no non-trivial vector subspaces, i.e. $C \cap -C = \{0\}$*
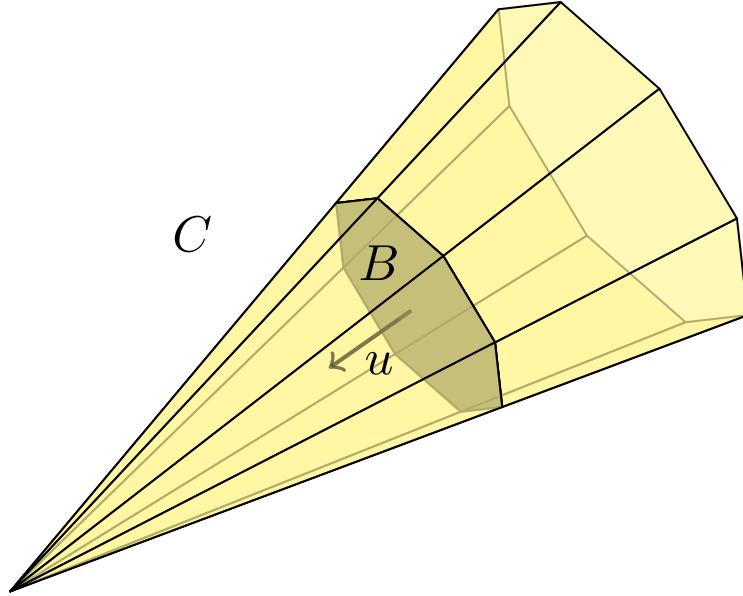
We have the following trivial result.

Figure 4.1: A convex cone C, with base B, generated by u.

Lemma 41. *The subset C of an ordered vector space V*

$$C = \{x \geq 0 : x \in V\} \tag{4.5}$$

is a cone.

Proof. Let $x, y \in C$, $\lambda > 0$. Then $x + y \geq x \geq 0$ and $\lambda x \geq 0$, so it is closed under addition and multiplication by positive scalars. Finally, $C \cap -C = \{0 \geq x \geq 0\} = \{0\}$. Hence C is a cone. $\qquad\square$

We call such a cone a positive cone, and denote it as V_+, as it is uniquely picked out by the choice of ordering of V.

The following definition and lemma support our assertion that we can generate cones from convex sets.

Definition 42. *For an ordered vector space V the set K is a base of the positive cone if for all $x \in V_+$, there is a unique $t \geq 0$*

$$xt \in K. \tag{4.6}$$

Lemma 43. *For a real vector space V with dual V^*, bilinear form $\langle \cdot, \cdot \rangle$, and positive cone V_+, every base K, where defined, is*

$$K = \{x \geq 0 : \langle u, x \rangle\}. \tag{4.7}$$

See Fig 4.1.

Given a cone in a vector space, we want to be able to discuss its dual, the set of positive functionals on the cone, as this will allow us to define our effect set.

Definition 44. *For a real vector space V with dual V^*, bilinear form $\langle \cdot, \cdot \rangle$, and positive cone V_+, the dual cone V_+^* is a subset of V^* such that*

$$\forall s \in V, e \in V_+^*, \ \langle e, s \rangle \geq 0. \tag{4.8}$$

Finally, we define the form of normed vector space that Ludwig's theorem embeds the state and effect spaces into.

Definition 45. *A Banach space is a real vector space with a norm such that it is complete, i.e. every Cauchy sequence converges to a limit point in the space.*

Example 4.1.1. *To tie these concepts together, we can take the example of a single qubit. Pure states live the Hilbert space $\mathcal{H} \cong \mathbb{C}^2$, while the mixed states ρ act on it. Any pure qubit can be described by 2 numbers, giving us the Bloch sphere, while mixed states live in its interior. Hence, any normalised mixed state is described by an element of $\mathbb{R}^3$ with norm $|\vec{x}| \leq 1$,*

$$\rho = \frac{1}{2} \vec{v} \cdot \vec{\sigma}. \tag{4.9}$$

This Bloch ball is an example of the convex set of states.
To account for non normalised states, we include a 4th coefficient, so that the state space is the set of all $\vec{w} \in \mathbb{R}^4$ with $\vec{w}_0 \in [0, 1]$ and $\vec{w}_i \in$ the Bloch ball, so that

$$\rho = \frac{1}{2} \vec{w} \cdot \vec{\sigma}^+ \tag{4.10}$$

where $\vec{\sigma}^+ = (I, \vec{\sigma})$. The variation of $\vec{w}_0$ corresponds to constructing a cone from the convex set, where we note that the measurement $\mathbb{I}$ is (trivially) isomorphic to I, and acts as the effect u.
See Fig 4.2.

Theorem 46. *Given a set of states and effects satisfying Ludwig's axioms, there exists a GPT, where*

1. *The (reliable) states form a convex set Ω, with the boundary given by the pure states and the interior by the mixed.*

2. *The effects are elements of $V_+^* \cap (u - V_+^*)$ where u is the unit effect.*

3. *Closure(Ω) and u form a base for the cone V_+, and the unreliable states are in the cone.*

4. *The canonical form between the two vector spaces, evaluated on the states and effects, coincides with the probability P.*
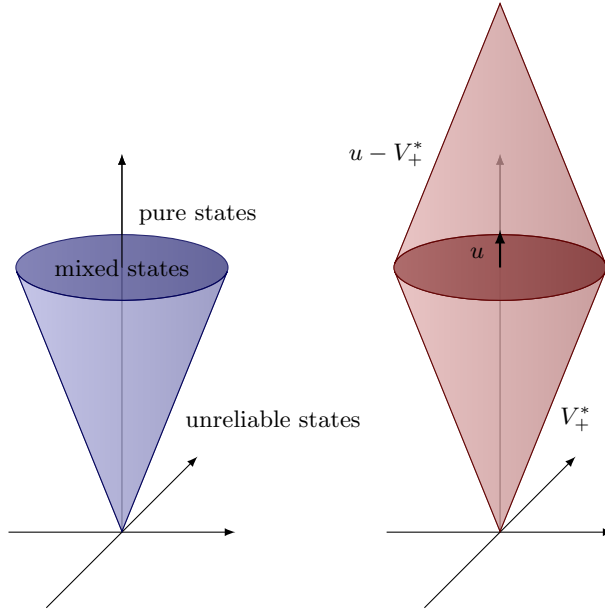
Figure 4.2: The state (left) and effect (right) spaces of a GPT.

By Ludwig's theorem, a QFT can thus be expressed as a GPT where $\mathbb{S} \subseteq \mathbb{V}, \mathbb{E} \subseteq \mathbb{V}^*$ are the images of $\mathcal{S}(\mathcal{A})$ and $\mathcal{A}$ under the embeddings, respectively, and there is a closed positive cone $V \subseteq \mathbb{S}$ of which Closure($\mathbb{S}$) is the base. The canonical bilinear form $\langle \cdot, \cdot \rangle : \mathbb{V}^* \times \mathbb{V} \longrightarrow \mathbb{R}$, restricted to $\mathbb{E} \times \mathbb{S}$, is the 'pushforward' of μ by the embeddings. The element u is mapped into $\mathbb{E}$ such that it defines the base.

Definition 47. *A GPT satisfies the no-restriction hypothesis if every $e \in [0, u]$ is a physically valid effect.*

We state the following theorem, proven in the case of finite dimensional real vector spaces. We do not attempt to prove it in the case of general Banach spaces, however it stands to reason that it should follow.

Theorem 48. *[39] For a GPT with state cone V which satisfies the NRH, the effect space $\mathcal{E}$*

$$V_+^* \cap (u - V_+^*). \tag{4.11}$$

See Fig 4.2.

4.2   Ideal measurements in QFT

Unlike in the case of QM, ideal measurements in QFT are a somewhat thorny subject. Few QFT textbooks cover them, perhaps accidentally implying that the
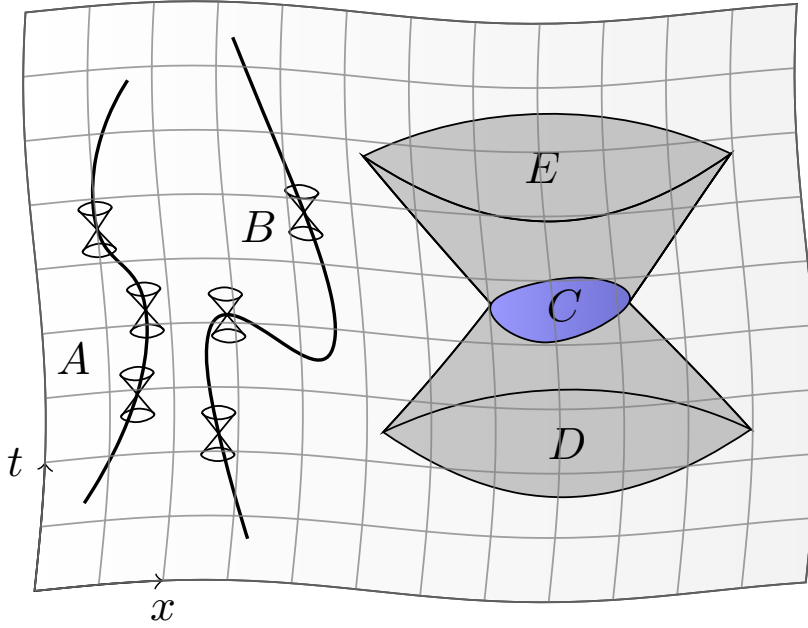
Figure 4.3: Patch of spacetime showing a causal curve (A), acausal curve (B), spacetime region C, and its causal past and future $(D), (E)$

measurement formalism from QM is sufficient. As we have discussed, this is not the case. Even without Sorkin's impossible measurements, the standard model is a local QFT, and so we need a concept of local measurement. In QM, this is often handed by having separated labs be in a tensor product. However, the Hilbert space associated with a QFT is (always) non-separable, and so we must be construct a measurement paradigm with a different, Lorentzian, notion of locality in mind.

To begin with, we recall some concepts from the study of Lorentzian manifolds.

Definition 49. *A curve $\Gamma : [a, b] \longrightarrow M$ in a Lorentzian manifold (M, g) is causal if its tangent vector is time or lightlike along its entire length, i.e.*

$$g\left(\frac{d}{dt}\Gamma(t), \frac{d}{dt}\Gamma(t)\right) \leq 0 \ \ \forall t \in [a, b], \tag{4.12}$$

where we are using a mostly positive signature. A curve is future (past) directed if its tangent is oriented into the future (past)[1].

See Fig4.3 curves A and B.

Definition 50. *The causal future (past) of a subset S of a Lorentzian manifold (M, g) is the set of all points M that can be reached by a future (past) directed*

---

[1]This requires the manifold be time-orientable. We assume this property

causal curve. Two subsets are spacelike, or causally disconnected, if they are not in each others causal past or future, i.e. all points in one subset are spacelike with all points in the other.

See Fig4.3 regions C and D.
Having defined some concepts of causal structure, we now wish to find a concept of local operators on the manifold.

Definition 51. *A subset S of a Lorentzian manifold (M, g) is causally convex if every causal curve $\Gamma : [a, b] \longrightarrow M$ with endpoints in S is entirely in S, i.e.*

$$\Gamma(a), \ \Gamma(b) \in S \implies \Gamma[a, b] \subseteq S. \tag{4.13}$$

A region is an open set that is a causally convex.

Example 4.2.1. *Consider a $1 + 1$ spacetime and a diamond set D with vertices at $(0, 0), (1, -1), (1, 1), (2, 0)$, see Fig 4.4. Pick any point P in D, and draw its future lightcone C. Then any causal curve with start point P must be entirely contained in that lightcone, including the end point Q. By picking endpoint Q in D and drawing its past lightcone, we see that any causal curve connecting them must be contained in the past lightcone C'. The intersection $C \cap C'$ is a subset of D, and so D is causally convex.*

It is with respect to regions of spacetime that the locality of a QFT algebra is defined. Given the algebra $\mathcal{A}(M)$ defined on the manifold M, the algebra $\mathcal{A}(R)$ defined on the region R is a subset of it, and is generated by the smeared local field operators $\phi(f)$, where f is a test function with compact support $\text{supp } f \subseteq R$. The relationship between algebras defined on disconnected regions is how locality is imposed. Hence, we need the following definitions.

Definition 52. *The causal compliment of a subset $S \subset M$ is the set of all points that are spacelike to all points in S.*

Definition 53. *Einstein causality states that for any two operators that are elements of algebras defined on spacelike regions, they commute, i.e.*

$$X \in \mathcal{A}(R), Y \in \mathcal{A}(R') \implies [X, Y] = 0, \tag{4.14}$$

if R, R' are spacelike. This is sometimes written as

$$[\mathcal{A}(R), \mathcal{A}(R')] = 0. \tag{4.15}$$

This is the textbook 'locality' condition in AQFT, and can be seen as a natural generalisation of the commutivity of local operators on a tensor product space.
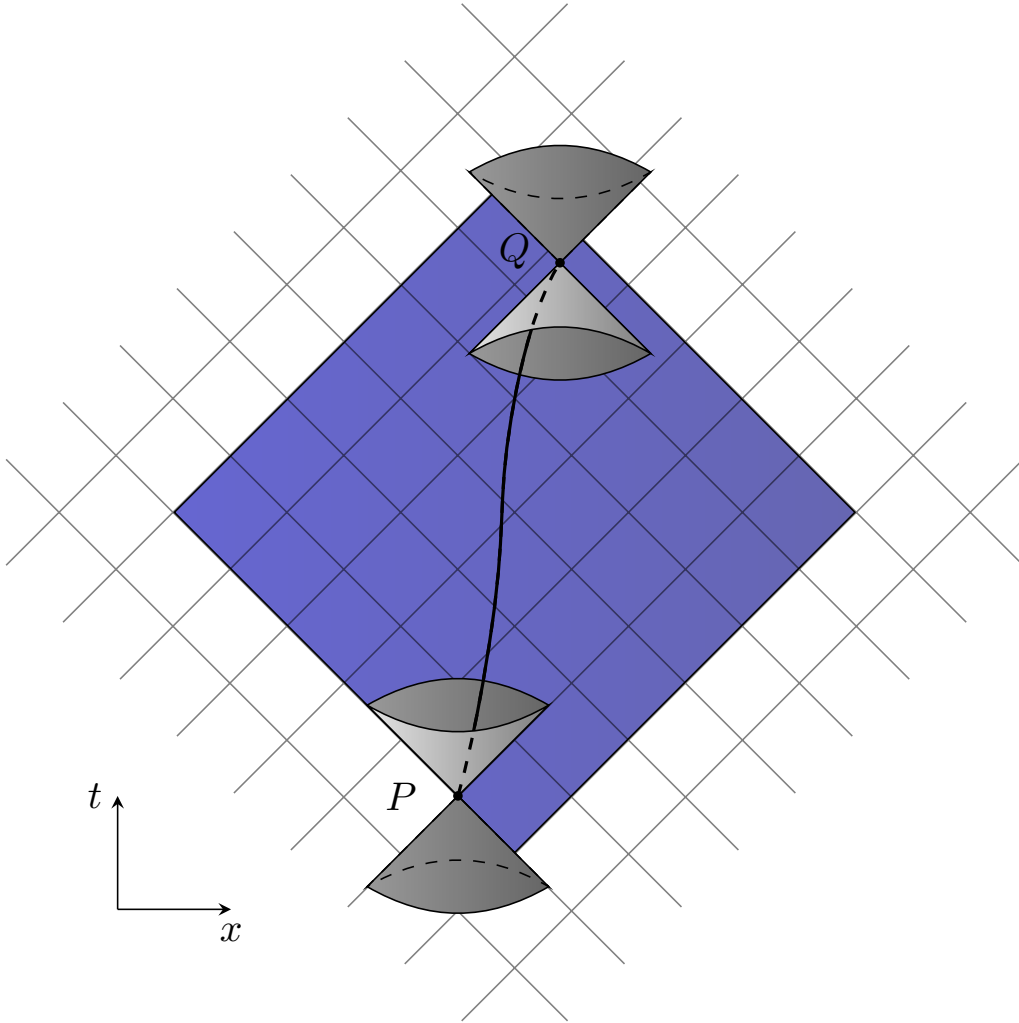
Figure 4.4: A spacetime diamond (blue) is causally convex, as the lightcone at any point of a curve in the diamond forbids the curve from leaving the diamond and returning.

It should be clear that such a property is necessary, as without it any two measurement operators acting on spacelike regions of the manifold could be applied in either order, and this ambiguity matters if and only if they do not commute.

Finally, we apply these concepts to measurement. To do so, we must define an ideal measurement of a (possibly unbounded) operator. Following [66] Borsten *et al*,

Definition 54. *A resolution $\mathcal{R}$ is a countable set (indexed by I) of mutually disjoint Borel sets of $\mathbb{R}$ or intervals of $\mathbb{R}$, which are sets that can be formed by countable union, countable intersection, and relative complement of open sets, and are complete, i.e.*

$$B_n \cap B_m = \emptyset, B_n, B_m \in \mathcal{R}, n \neq m \tag{4.16}$$

and

$$\bigcup_{n \in I} B_n = \mathcal{R}. \tag{4.17}$$

For an operator A with spectral decomposition

$$A = \int_\sigma \lambda dP^A(\lambda), \tag{4.18}$$

$P(B_n) = E_n$ is a projection operator, and

$$E_n E_m = E_n \delta_{n,m}, \sum_{n \in I} E_n = I. \tag{4.19}$$

As in regular QM, we consider ideal measurements, obeying the projection postulate, with the additional condition of locality. Consider a compact subset K of a spacetime M. Locality demands that we can only measure local operators defined on regions of K. We do not condition over each outcome, instead taking a mixture of each updated state, i.e.

$$\rho \xrightarrow{\text{measure A}} \sum_n E_n \rho E_n, \tag{4.20}$$

where E_n are the projections associated with a resolution $\mathcal{R}$ and $A \in \mathcal{A}(R)$ for some region $R \subset K$. This is the same as the usual projection postulate, without conditioning on a particular outcome n, so called non-selective ideal measurement. By the usual machinery of trace preserving completely positive maps (CPTP) we know that this is a map on the set of density operators, and we can represent it as an operator on the space of states,

$$\mathcal{E}(\rho) = \sum_{n \in I} E_n \rho E_n, \tag{4.21}$$

or equivalently on as an update map on the set of (bounded) operators

$$\mathcal{E}(X) = \sum_n E_n X E_n, \tag{4.22}$$

where the equivalence of these pictures follows from the cyclic property of the trace. The update map is defined by the choice of operator to measure and the choice of resolution, so we can label it as $\mathcal{E}_{A,\mathcal{E}}$.

Definition 55. *Given a compact subset K, an update map $\mathcal{E}$ is local to K if it is trivial when spacelike to K, i.e.*

$$\mathcal{E}(X) = X, \quad \forall X \in \mathcal{B}(K^\perp). \tag{4.23}$$

Lemma 56. *An ideal measurement of a smeared local field operator $\phi(f)$ is local to any compact subset K that contains* $\operatorname{supp} f$.

Proof. Let f be a test function with support $\operatorname{supp} f$, and K be a compact subset with $\operatorname{supp} f \subseteq K$. Suppose we act on some bounded operator X local to some region of $K^\perp$, then
From Einstein causality, $[\phi(f), X] = 0$ as the operators are local to regions spacelike to each other. The spectral theorem then implies $[E_n, X] = 0$. Hence,

$$\mathcal{E}(X) = \sum_{n \in I} E_n X E_n = \sum_{n \in I} E_n^2 X = \sum_{n \in I} E_n X = X. \tag{4.24}$$

$\square$

4.2.1 Sorkin scenarios

A Sorkin scenario consists of three compact, disjoint subsets of M, A, B, C, such that C is partly in the causal future of A (the overlap of the causal future of A with C is non-empty) and B is partly in the causal future of C, but B is not in the causal future of A, see Fig 4.5. Alice, Bob, and Charlie are free to perform local updates on A, B, C respectively. Sorkin pointed out that local projective update maps can allow signalling between Alice and Bob. These are referred to as Sorkin's impossible measurements, and demonstrate that locality does not ensure causality. Suppose Alice and Charlie update the state locally using O_A, O_C respectively. The expectation value that Bob measures is

$$\operatorname{tr}(\rho \mathcal{E}_{O_C}(\mathcal{E}_{O_A}(X))) \tag{4.25}$$

while we should expect by causality that

$$\operatorname{tr}(\rho \mathcal{E}_{O_C}(X)). \tag{4.26}$$

Since A, C are not causally disconnected, we cannot rely on locality arguments to ignore the action of $\mathcal{E}_{O_A}$. As shown by Sorkin, there are choices A, C, B and O_A, O_C such that Eq 4.25 and 4.26 are not equal, and so Bob can learn if Alice interacted locally or not, despite A, B being causally disconnected.
Despite being pointed out 30 years ago, there are still unanwsered questions regarding Sorkin's impossible measurements, namely: what measurements are allowed? A common line research is to construct a von-Neumann type setup, with the field coupled to another field or to some finite dimensional ancilla [67, 68]. As with the normal measurement problem, this leads to a recursive 'solution'. What measurements are allowed on the ancilla, and does it need its own ancilla etc? Another path involves constructing models of physical
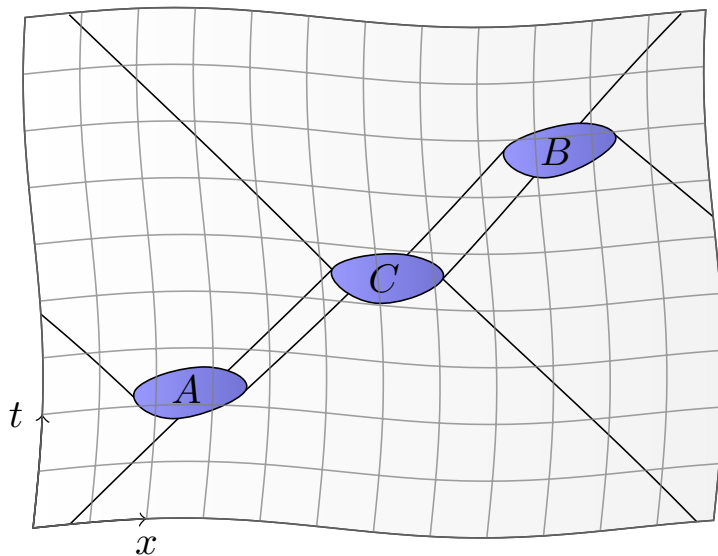
Figure 4.5: Spacetime diagram of a Sorkin scenario, showing the three regions A, B, C, where C is partly in the causal future of A, and B is partly in the casusal future of C, yet A and B are spacelike.

measurement apparatus, à la Unruh-Dewitt [69]. However, an operational definition of allowed measurements or maps is still missing [66], unlike in non relativistic QM, where any CPTP map is an allowed map, and can be imagined as a unitary acting on the system coupled to an ancilla via the Stinespring dilation theorem. An equivalent formulation for QFT is an open problem, and counter-intuitive results are still being found. It has been proposed numerous times that smeared local field operators are causal [5, 66, 68], yet recent research suggests otherwise [61, 70]. This provides some of the first evidence that *local* operators can be acausal, and leads us to the following:

Conjecture 57. *QFTs violate the no-restriction hypothesis.*

In aid of this conjecture we investigate smeared local field operators in the GPT framework.

Lemma 58. *Take a QFT described by a GPT $(\mathcal{S}, \mathcal{E})$, where $\mathcal{S}$ is the closure of the embedding of $S \subseteq \mathcal{S}(\mathcal{A})$ with S obeying the Luwdig axioms, and $\mathcal{E}$ is the full effect space in 48, as a subset of (embedded) $\mathcal{A}$.*
Then smeared local field operators are in $\mathcal{E}$.

Proof. Every state $\omega \in S$ is associated in a representation (GNS) with $|\Omega\rangle$. Hence, under the embedding given by Ludwig's theorem, $\omega \in \mathcal{S}$ for any valid QFT. Clearly, $0 \leq \text{tr}(|\Omega\rangle \langle\Omega| \phi(f)) \leq 1$. Hence, $\phi(f) \in \mathcal{E}$. $\qquad \square$

Lemma 58 is robust, as it holds even if we consider a restricted subset of states. As proposed in [70] and demonstrated for real scalar fields in [61], measurements of smeared local field operators are not causal. Hence, QFTs, which take Poincaré invariance as an axiom, can be expected not to satisfy the NRH. The acausality of local field operators has only been demonstrated for real scalars, and requires a number of, albeit reasonable, assumptions that have only been derived from first principles for $1 + 1$ theories. While 58 is robust with respect to choice of state subsets, we have not demonstrated its robustness with respect to effect subsets. One way that the conjecture could be sidestepped is that if, when it has been found, the set of causal measurements is embedded into the dual Banach space in such a way that the no-restriction hypothesis is satisfied. However, this requires further investigation, and there is no guarantee that the causal subset would be embed in such a way as to satisfy no-restriction.

4.2.2   Sharpening the conjecture

One possible criticism of the above argument is that unbounded operators do not correspond to physical measurements, as any real measurement device has bounded readout, e.g. CCDs have finite sensitivity and become saturated at high luminance, and physical dials have finite travel. Often, the algebra of interest in AQFT is the bounded algebra $\mathcal{B}$, of which $\phi(f)$ is not an element. Hence, to strengthen the conjecture we can introduce the limited local field operator,

$$\phi_a^b(f) = \int_{\sigma \cap [-a,b]} \lambda dP^{\phi(f)}(\lambda), \qquad (4.27)$$

where we have used the spectral theorem to expand the self-adjoint $\phi(f)$ into its spectrum σ and introduced a cut-off for the operator so that it has no spectrum outside of $[-a, b]$. This operator is bounded,

$$||\phi_a^b(f)|| = \sup\nolimits_{\lambda \in \sigma(\phi_a^b(f))}(|\lambda|) = \sup\nolimits_{\lambda \in \sigma(\phi(f)) \cap [-a,b]}(|\lambda|) \leq \max(a, b), \qquad (4.28)$$

where we have used that the spectral radius

$$\rho(A) = \sup\nolimits_{\lambda \in \sigma(A)}(|\lambda|) \qquad (4.29)$$

of a self adjoint operator is equal to its operator norm [71]. Thus, it is an element of the algebra $\mathcal{B}(R) \subset \mathcal{A}(R)$ where supp $f \subseteq R$. It also converges to $\phi(f)$ as $a, b \to \infty$ under the strong operator topology. To see this, let D be the dense subset of $\mathcal{H}$ for which $\phi(f)$ is defined. Then, strong convergence is equivalent to

$$||\phi_a^b(f)x - \phi(f)x|| \to 0, \ \forall x \in D. \qquad (4.30)$$

For any $x \in D$, we have $\phi(f)x \in \mathcal{H}$, and so $||\phi(f)x||^2$ is finite. We can view $\mathcal{H} = L^2(\mathbb{R}^d)$. Let $\varepsilon > 0$. There are two cases. If $\varepsilon > ||\phi(f)x||$ then for any $a, b \geq 0, ||\phi_a^b(f)x - \phi(f)x|| < \varepsilon$. Otherwise, if $||\phi(f)x|| \geq \varepsilon$, then there exists a $n \in \mathbb{R}$ such that

$$\frac{\int_{-n}^{n} d\lambda |\phi(f)x|^2}{||\phi(f)x||^2} > 1 - \frac{\varepsilon^2}{||\phi(f)x||^2}. \tag{4.31}$$

This follows from the fact that $I(n) = \int_{-n}^{n} d\lambda |\phi(f)x|^2$ is a monotonically increasing function and that $||\phi(f)x||^2$ is finite. Picking $a, b > n$, we have that

$$\frac{\int_{-a}^{b} d\lambda |\phi(f)x|^2}{||\phi(f)x||^2} > 1 - \frac{\varepsilon^2}{||\phi(f)x||^2} \implies \int_{\mathbb{R} \setminus [-a,b]} d\lambda |\phi(f)x|^2 < \varepsilon^2. \tag{4.32}$$

Hence,

$$\left\| \phi(f)x - \phi_a^b(f)x \right\| = \left\| \int_{\mathbb{R} \setminus [-a,b]} d\lambda \phi(f)x \right\| \leq \left(\int_{\mathbb{R} \setminus [-a,b]} d\lambda |\phi(f)x|^2 \right)^{\frac{1}{2}} < \varepsilon, \tag{4.33}$$

where we have used Cauchy–Schwarz for the penultimate inequality.
By an identical argument to 58, it is also a member of $\mathcal{E}$ (under the embedding). Unlike with the unbounded operator, we are considering resolutions of $[-a, b]$, and so the Appendix 10.7 of [61], where it is shown all resolutions of $\mathcal{R}$ lead to acausal update maps, does not apply directly. We do not prove an equivalent, instead showing that a natural coursegraining of $[-a, b]$ into N bins $B_n = [c_n, c_n + 1), n < N, B_N = [c_N, b]$, with $c_1 = -a$, which satisfy the definition of a resolution of $[-a, b]$, leads to an acausal update map. This does not prove that all $\phi_a^b(f)$ measurements are acausal, however it shows the existence of at least one, which is sufficient. The proof of this fact is left to Appendix 6.3. Hence, there are bounded linear operators in AQFT that are acausal, and so we have strengthened our conjecture: QFT violates the no-restriction hypothesis.

Chapter 5

Final remarks

5.1 Discussion

In our investigation, we have seen how each approach to foundations emphasises and de-emphasises fundamental questions. It is clear that, at least at this stage, none of the frameworks explored are able to answer every question. The device independent framework allows us to quickly make strong statements about the strength of non-classical behaviours such as non-locality and contexuality. However, it provides little in the way of answers regarding the dynamical or operational structure of the theory, and it seems unlikely that an experimental proposal could be born out of device independent research alone.

The GPT framework allows us to focus on the operational structure as the fundamental structure of a theory, and provides a complete, albeit often not natural, definition of a state and measurement. It has shown that many non-classical behaviours are generic, strengthen the results of the device independent framework, While it is possible to impose dynamics as maps between states, this gives little evidence of their form.

With the histories framework, we have no operational definition, states, or measurement, yet experiments feel far more tangible. Following the usual histories idea, that all measurements are measurements of position, and given that any quantum measure is defined over spacetime trajectories, it feels far more plausible that an experiment could be proposed[1]. Equally, the main object of study is the decoherence functional, which encodes the dynamics, placing them front and centre, in contrast to the other two approaches. However, there are still many open questions around the proper form of the decoherence functional, and the tie-ins to the device independent formalism imply a rich connection between constraints on correlations and constraints on the decoherence functional.

---

[1]In fact, this has been done. Tests of higher order interference have been conducted, with only negative results so far [72]

Throughout this work we have commented on the possible validity of some of the results. For instance, while the device independent principles provide a clear operational definition for many sets of behaviours that have been elsewhere defined, it is the author's opinion that many are well motivated. We have discussed the issues of many of the information based principles like information causality and no trivial communication/computation, as well as macroscopic locality/non-contextuality. Far more work must be done if we are to accept these not just as convenient definitions of behaviour sets, but axioms used to exclude 'unreasonable' behaviours. Similarly, we have commented on strong positivity and the no-restriction hypothesis. SP, like the device independent principles, appears to exclude a large number of behaviours that we do not expect to exist, however work is still on-going to give it motivation [51]. While there is evidence that almost quantum and SPJQM ccould coincide, the branching issue is still unresolved, and the original results of Chapter 3, which show that SPJQM is itself closed under post-processing and may imply that SPJQM and $SPJQM_b$ coincide for (2 2 2), have not clarified the issue of a motivation for SP.

In sharp contrast, the no-restriction hypothesis seems to exclude many physical theories preemptively, and we have tried to argue that it would exclude what is often considered to be the most successful physical theory ever written down, quantum electrodynamics [73], and relativistic QFT more generally. While more work needs to be done to confirm this conjecture, our aim is also to push back somewhat on the often stated, yet rarely justified, statement that quantum theory (as a whole) satisfies the no-restriction hypothesis. It is perhaps unsurprising that statements like these exist, quantum foundations is a extremely broad field, covering effectively every physical theory we can imagine, and based on a number of frameworks and approaches that are not compatible in general, and many researchers stay away from fully relativistic quantum theory, "*However, since relativistic quantum-field theory has well-known technical difficulties, we will in subsequent sections formulate a standard nonrelativistic quantum model...*"-Aharonov *el al* [63].

It is to this end that we have aimed to state and prove new connections between three frameworks, and with relativistic quantum theory, as well as clear up what we believe to be misunderstandings in the literature caused by the disconnect between approaches.

5.2 Conclusion

We have presented a number of known and novel results from across three quantum foundations frameworks, emphasising the connections and tensions

between the approaches. We have strengthened the connection between the histories set SPJQM and device independent closure conditions by showing closure under post-processing. Further, we have hinted at a connection between the device independent principle 'local orthogonality' and the histories principle 'branching'. If further investigation of this connection bears out, then it would imply that any SPJQM is compatible with branching and that SPJQM=$\tilde{Q}$, again strengthening the connection between the device independent and histories frameworks. Finally, we have investigated and critiqued a number of principles in the device independent and GPT frameworks, suggesting that some principles should not be used to rule out sets of behaviours without further investigation. Namely, we have given evidence that QFT violates the no-restriction hypothesis, contradicting the implication that almost quantum should be abandoned for violating it.

Bibliography

[1] P. A. M. Dirac, *The Principles of Quantum Mechanics*. Clarendon Press, 1981. Google-Books-ID: XehUpGiM6FIC.

[2] J. Von Neumann, R. T. Beyer, and N. A. Wheeler, *Mathematical foundations of quantum mechanics*. Princeton: Princeton University Press, new edition ed., 2018. OCLC: on1004934776.

[3] S. Rivat, "Drawing Scales Apart: The Origins of Wilson's Conception of Effective Field Theories," Nov. 2021. arXiv:2111.03148 [hep-th, physics:physics].

[4] A. Einstein, B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?," *Physical Review*, vol. 47, pp. 777–780, May 1935.

[5] R. D. Sorkin, "Impossible Measurements on Quantum Fields," in *Directions in General Relativity: Proceedings of the 1993 International Symposium, Maryland, Vol. 2: Papers in honor of Dieter Brill*, pp. 293–305, Cambridge University Press, Feb. 1993.

[6] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, "Emergence of Noncontextuality under Quantum Darwinism," *PRX Quantum*, vol. 2, p. 030351, Sept. 2021. Publisher: American Physical Society.

[7] H. Everett, ""Relative State" Formulation of Quantum Mechanics," *Reviews of Modern Physics*, vol. 29, pp. 454–462, July 1957.

[8] F. Dowker and A. Kent, "On the Consistent Histories Approach to Quantum Mechanics," *Journal of Statistical Physics*, vol. 82, pp. 1575–1646, Mar. 1996. arXiv:gr-qc/9412067.

[9] N. Bodendorfer, "An elementary introduction to loop quantum gravity," July 2016. arXiv:1607.05129 [gr-qc].

[10] J. Bedford, "An Introduction to String Theory," June 2012. arXiv:1107.3967 [gr-qc, physics:hep-th, physics:physics].

[11] J. S. Bell, "Physics Physique Fizika 1, 195 (1964) - On the Einstein Podolsky Rosen paradox," *Physics Physique Fizika*, no. 3, pp. 193–200, 1964.

[12] C. M. Ferrera, R. Simmons, J. Purcell, D. Collins, and S. Popescu, "Classical-to-quantum non-signalling boxes," July 2023. arXiv:2303.17268 [quant-ph].

[13] M. Piani, M. Horodecki, P. Horodecki, and R. Horodecki, "On quantum non-signalling boxes," *Physical Review A*, vol. 74, p. 012305, July 2006. arXiv:quant-ph/0505110.

[14] V. B. Scholz and R. F. Werner, "Tsirelson's Problem," Dec. 2008. arXiv:0812.4305 [math-ph, physics:quant-ph].

[15] W. Slofstra, "Tsirelson's problem and an embedding theorem for groups arising from non-local games," *Journal of the American Mathematical Society*, vol. 33, pp. 1–56, Sept. 2019. arXiv:1606.03140 [math-ph, physics:quant-ph].

[16] B. S. Cirel'son, "Quantum generalizations of Bell's inequality," *Letters in Mathematical Physics*, vol. 4, pp. 93–100, Mar. 1980.

[17] S. Popescu and D. Rohrlich, "Quantum nonlocality as an axiom," *Foundations of Physics*, vol. 24, pp. 379–385, Mar. 1994.

[18] A. Kent, "Quantum nonlocal correlations are not dominated," Aug. 2013. arXiv:1308.5009 [quant-ph].

[19] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, "Proposed Experiment to Test Local Hidden-Variable Theories," *Physical Review Letters*, vol. 23, pp. 880–884, Oct. 1969.

[20] Z. Ji, A. Natarajan, T. Vidick, J. Wright, and H. Yuen, "MIP*=RE," Nov. 2022. arXiv:2001.04383 [quant-ph].

[21] M. Navascues, S. Pironio, and A. Acin, "A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations," *New Journal of Physics*, vol. 10, p. 073013, July 2008. arXiv:0803.4290 [quant-ph].

[22] F. Dowker, J. Henson, and P. Wallden, "A histories perspective on characterising quantum non-locality," *New Journal of Physics*, vol. 16, p. 033033, Mar. 2014. arXiv:1311.6287 [gr-qc, physics:quant-ph].

[23] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín, "Almost quantum correlations," *Nature Communications*, vol. 6, p. 6288, Feb. 2015. Number: 1 Publisher: Nature Publishing Group.

[24] M. Navascues and H. Wunderlich, "A glance beyond the quantum model," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, pp. 881–890, Mar. 2010. arXiv:0907.0372 [quant-ph].

[25] J. Henson and A. B. Sainz, "Macroscopic non-contextuality as a principle for Almost Quantum Correlations," *Physical Review A*, vol. 91, p. 042114, Apr. 2015. arXiv:1501.06052 [quant-ph].

[26] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, "Local orthogonality as a multipartite principle for quantum correlations," *Nature Communications*, vol. 4, p. 2263, Aug. 2013. Number: 1 Publisher: Nature Publishing Group.

[27] G. Brassard, H. Buhrman, N. Linden, A. A. Methot, A. Tapp, and F. Unger, "A limit on nonlocality in any world in which communication complexity is not trivial," *Physical Review Letters*, vol. 96, p. 250401, June 2006. arXiv:quant-ph/0508042.

[28] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, "Information causality as a physical principle," *Nature*, vol. 461, pp. 1101–1104, Oct. 2009. Number: 7267 Publisher: Nature Publishing Group.

[29] N. Linden, S. Popescu, A. J. Short, and A. Winter, "No quantum advantage for nonlocal computation," *Physical Review Letters*, vol. 99, p. 180502, Oct. 2007. arXiv:quant-ph/0610097.

[30] W. Van Dam, *Nonlocality and communication complexity*. PhD Thesis, University of Oxford, 1999.

[31] M. Plávala, "General probabilistic theories: An introduction," Aug. 2021. arXiv:2103.07469 [quant-ph].

[32] L. Hardy, "Quantum Theory From Five Reasonable Axioms," Sept. 2001. arXiv:quant-ph/0101012.

[33] G. Ludwig, *An Axiomatic Basis for Quantum Mechanics*. Berlin, Heidelberg: Springer, 1985.

[34] S. N. Filippov, S. Gudder, T. Heinosaari, and L. Leppäjärvi, "Operational Restrictions in General Probabilistic Theories," *Foundations of Physics*, vol. 50, pp. 850–876, Aug. 2020.

[35] A. B. Sainz, Y. Guryanova, A. Acín, and M. Navascués, "Almost quantum correlations violate the no-restriction hypothesis," *Physical Review Letters*, vol. 120, p. 200402, May 2018. arXiv:1707.02620 [quant-ph].

[36] T. Gonda, R. Kunjwal, D. Schmid, E. Wolfe, and A. B. Sainz, "Almost Quantum Correlations are Inconsistent with Specker's Principle," *Quantum*, vol. 2, p. 87, Aug. 2018.

[37] V. J. Wright and S. Weigert, "General Probabilistic Theories with a Gleason-type Theorem," *Quantum*, vol. 5, p. 588, Nov. 2021. Publisher: Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften.

[38] V. J. Wright, "Intermediate determinism in general probabilistic theories," *Journal of Physics A: Mathematical and Theoretical*, vol. 55, p. 464002, Nov. 2022. Publisher: IOP Publishing.

[39] P. Janotta and R. Lal, "Generalized Probabilistic Theories Without the No-Restriction Hypothesis," *Physical Review A*, vol. 87, p. 052131, May 2013. arXiv:1302.2632 [quant-ph].

[40] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, "Limits on nonlocal correlations from the structure of the local state space," *New Journal of Physics*, vol. 13, p. 063024, June 2011.

[41] R. D. Sorkin, "Quantum Mechanics as Quantum Measure Theory," *Modern Physics Letters A*, vol. 09, pp. 3119–3127, Oct. 1994. arXiv:gr-qc/9401003.

[42] R. D. Sorkin, "Quantum Measure Theory and its Interpretation," June 1997. arXiv:gr-qc/9507057.

[43] J. Hartle, "Generalized Quantum Mechanics," Oct. 2021. arXiv:2110.11268 [gr-qc, physics:hep-th, physics:quant-ph].

[44] F. Dowker, "The birth of spacetime atoms as the passage of time," *Annals of the New York Academy of Sciences*, vol. 1326, no. 1, pp. 18–25, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.12542.

[45] F. Dowker, "Causal Set Quantum Gravity and the Hard Problem of Consciousness," May 2023. arXiv:2209.07653 [gr-qc, physics:hep-th, physics:physics].

[46] F. Dowker, S. Johnston, and R. D. Sorkin, "Hilbert spaces from path integrals," *Journal of Physics A: Mathematical and Theoretical*, vol. 43, p. 275302, June 2010.

[47] S. Surya, "The causal set approach to quantum gravity," *Living Reviews in Relativity*, vol. 22, p. 5, Dec. 2019. arXiv:1903.11544 [gr-qc].

[48] G. R. Rocha, D. Rickles, and F. J. Boge, "A Brief Historical Perspective on the Consistent Histories Interpretation of Quantum Mechanics," Mar. 2021. arXiv:2103.05280 [physics, physics:quant-ph].

[49] H. F. Dowker and J. J. Halliwell, "Quantum mechanics of history: The decoherence functional in quantum mechanics," *Physical Review D*, vol. 46, pp. 1580–1609, Aug. 1992. Publisher: American Physical Society.

[50] R. D. Sorkin, "Notes on a histories formulation of quantum mechanics.," 2011.

[51] F. Dowker and H. Wilkes, "An Argument for Strong Positivity of the Decoherence Functional," Sept. 2021. arXiv:2011.06120 [gr-qc, physics:hep-th, physics:quant-ph].

[52] E. Joos, H. D. Zeh, C. Kiefer, D. Giulini, J. Kupsch, and I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory*. Berlin, Heidelberg: Springer, 2003.

[53] F. Dowker, V. Havlíček, C. Lewandowski, and H. Wilkes, "A "problem of time" in the multiplicative scheme for the n-site hopper," *Journal of Physics A: Mathematical and Theoretical*, vol. 50, p. 455303, Nov. 2017. arXiv:1706.03793 [quant-ph].

[54] D. Craig, F. Dowker, J. Henson, S. Major, D. Rideout, and R. D. Sorkin, "A Bell Inequality Analog in Quantum Measure Theory," Apr. 2006.

[55] B. Lang, T. Vertesi, and M. Navascues, "Closed sets of correlations: answers from the zoo," Feb. 2014. arXiv:1402.2850 [quant-ph].

[56] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, "Non-local correlations as an information theoretic resource," *Physical Review A*, vol. 71, p. 022101, Feb. 2005. arXiv: quant-ph/0404097.

[57] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. Vértesi, "Closed sets of nonlocal correlations," *Physical Review A*, vol. 80, p. 062107, Dec. 2009. Publisher: American Physical Society.

[58] P. Boes and M. Navascues, "Composing decoherence functionals," *Physical Review A*, vol. 95, p. 022114, Feb. 2017. arXiv:1609.09723 [gr-qc, physics:quant-ph].

[59] B. Lang, T. Vertesi, and M. Navascues, "Closed sets of correlations: answers from the zoo," Feb. 2014. arXiv:1402.2850 [quant-ph].

[60] S. Beigi and A. Gohari, "Monotone Measures for Non-Local Correlations," *IEEE Transactions on Information Theory*, vol. 61, pp. 5185–5208, Sept. 2015. arXiv:1409.3665 [quant-ph].

[61] E. Albertini and I. Jubb, "Are Ideal Measurements of Real Scalar Fields Causal?," June 2023. arXiv:2306.12980 [gr-qc, physics:hep-th, physics:math-ph, physics:quant-ph].

[62] G. Chiribella and X. Yuan, "Bridging the gap between general probabilistic theories and the device-independent framework for nonlocality and contextuality," *Information and Computation*, vol. 250, pp. 15–49, Oct. 2016.

[63] Y. Aharonov, S. Popescu, and D. Rohrlich, "On conservation laws in quantum mechanics," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 118, p. e1921529118, Jan. 2021.

[64] L. Lami, "Non-classical correlations in quantum mechanics and beyond," Mar. 2018. arXiv:1803.02902 [math-ph, physics:quant-ph].

[65] L. Lami, B. Regula, R. Takagi, and G. Ferrari, "Framework for resource quantification in infinite-dimensional general probabilistic theories," *Physical Review A*, vol. 103, p. 032424, Mar. 2021. arXiv:2009.11313 [math-ph, physics:quant-ph].

[66] L. Borsten, I. Jubb, and G. Kells, "Impossible measurements revisited," *Physical Review D*, vol. 104, p. 025012, July 2021. arXiv:1912.06141 [gr-qc, physics:hep-th, physics:quant-ph].

[67] H. Bostelmann, C. J. Fewster, and M. H. Ruep, "Impossible measurements require impossible apparatus," *Physical Review D*, vol. 103, p. 025017, Jan. 2021. arXiv:2003.04660 [gr-qc, physics:hep-th, physics:math-ph, physics:quant-ph].

[68] D. M. T. Benincasa, L. Borsten, M. Buck, and F. Dowker, "Quantum information processing and relativistic quantum fields," *Classical and Quantum Gravity*, vol. 31, p. 075007, Mar. 2014. Publisher: IOP Publishing.

[69] C. J. Fewster and R. Verch, "Quantum fields and local measurements," July 2020. arXiv:1810.06512 [gr-qc, physics:math-ph, physics:quant-ph].

[70] I. Jubb, "Causal state updates in real scalar quantum field theory," *Physical Review D*, vol. 105, p. 025003, Jan. 2022. Publisher: American Physical Society.

[71] "Spectral radius - Encyclopedia of Mathematics."

[72] 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 *AIP Conference Proceedings*, pp. 200–207, 2009. arXiv:0811.2068 [quant-ph].

[73] R. P. Feynman, *QED: the strange theory of light and matter*. Alix G. Mautner memorial lectures, Princeton, N.J: Princeton University Press, 1985.

Chapter 6

Appendix

6.1 $\Sigma_T = \Sigma_S$

We must show that

$$\sum_{\bar{A} \in \bar{X}} \sum_{A' \in S(\bar{A})} |A'\rangle = \sum_{X' \in T(\bar{X})} \sum_{A' \in X'} |A'\rangle \tag{6.1}$$

i.e. $T[\bar{X}] = S[\bar{X}]$ where the square bracket sets denote (super)set of atoms that appear in the usual event subsets. In set notation

$$S[\bar{X}] = \left\{ A \text{ atom of } 2^\Omega | A_{ij} = \bar{A}_{ij} \text{ for } i, j > 1, \bar{A} \in \bar{X}, A_{11} = 1 \text{ or } 2 \text{ if } \bar{\vec{x}}_1 = 1, \right.$$
$$\left. \text{otherwise } A_{11} = \bar{A}_{11} + 1 \right\}, \tag{6.2}$$

and

$$T[\bar{X}] = \left\{ A \text{ atom of } 2^\Omega | A \in X \in \tilde{\mathcal{O}}, \bar{X} = (\bar{a}|\bar{\vec{x}}), X = (\vec{a}|\vec{x}), \vec{x} = \bar{\vec{x}}, \right.$$
$$\vec{a} = \bar{\vec{a}} \text{ if } \vec{x}_1 \neq 1, \text{ otherwise } \vec{a}_i = \bar{\vec{a}}_i \text{ for } i > 1 \text{ and}$$
$$\left. \vec{a}_1 = \bar{\vec{a}}_1 + 1 \text{ if } \vec{a}_1 > 1 \text{ otherwise } \vec{a}_1 = 1 \text{ or } 2 \right\} \tag{6.3}$$

where it should be understood that $\bar{X}$ may denote a fine or coarse-grained event. In the case of a coarse-grained event $\bar{X}$ such that the first party is ignored, both sets are trivially equal, as they contain only the atoms of $X = \bar{X}$. In the remaining cases, fine-grained or coarse-raining but not ignoring the first party, the following derivation holds.

Let $A \in S[\bar{X}]$. Then there are two cases.

Case 1: $A_{11} = 1, 2$ and $\exists \bar{A} \in \bar{X}$, with $A_{ij} = \bar{A}_{ij} j, i \neq 1$. Let $X \in \tilde{\mathcal{O}}$ with $\vec{x} = \bar{\vec{x}}, \vec{a}_i = \bar{\vec{a}}_i$ for $i > 1$ and $\vec{a}_1 = A_{\vec{x}_1 1}$. Then $A \in X$ and so $A \in T[\bar{X}]$.

Case 2: $A_{11} > 2$ and $\exists \bar{A} \in \bar{X}$ with $A_{ij} = \bar{A}_{ij} j, i \neq 1$ and $\bar{A}_{11} + 1 = A_{11}$. Let $X \in \tilde{\mathcal{O}}$ with $\vec{x} = \bar{\vec{x}}, \vec{a}_i = \bar{\vec{a}}_i$ for $i > 1$ and $\vec{a}_1 = \bar{\vec{a}}_1 + 1$. Then $A \in X$ and so $A \in T[\bar{X}]$. Hence $S \subseteq T$

The converse follows similarly, and so $T = S$.

6.2 Atomic vectors

For every basic measurement x_i, we selected one outcome $(a_i'|x_i)$, and define for all the other outcomes $(a_i|x_i)$ etc

$$\mathcal{H}^{(a_i|x_i)} = \{|\vec{a}'', \vec{x}''\rangle \ (\vec{a}''|\vec{x}'') \in \tilde{\mathcal{O}}, \ (a_i|x_i) \subset (\vec{a}''|\vec{x}'')\} \text{ and } \mathbb{E}^{(a_i|x_i)} = \text{Proj}(\mathcal{H}^{(a_i|x_i)})$$
(6.4)

where we note that local orthogonality implies that $\mathbb{E}^{(a_i|x_i)}\mathbb{E}^{(a_i''|x_i)} = \delta_{a_i a_i''}\mathbb{E}^{(a_i|x_i)}$ and then define $\mathbb{E}^{(a_i'|x_i)}$ such that

$$\sum_{(a_i|x_i) \in x_i} \mathbb{E}^{(a_i|x_i)} = \mathbb{I}$$
(6.5)

i.e. such that the projectors form a complete set and orthogonality of the projectors follows. Let $\vec{a} \in \mathfrak{u}_{\vec{x}}$ be an outcome with $\vec{a} = (a_1, a_2, ...a_n) = a_1 \cap a_2... \cap a_n$ and define

$$\mathbb{E}^{(\vec{a}|\vec{x})} = \mathbb{E}^{(a_1|x_1)}\mathbb{E}^{(a_2|x_2)}...\mathbb{E}^{(a_n|x_n)}.$$
(6.6)

Define the new vectors

$$|\gamma\rangle = \prod_{(a_i|x_i) \in I(\gamma)} \mathbb{E}^{(a_i|x_i)} |\Omega\rangle$$
(6.7)

where $I(\gamma)$ is the set of all outcomes $(a_i|x_i)$ such that $\gamma = (a_1|x_1) \cap (a_2|x_2) \cap ...(a_n|x_n)$. Finally, we need to show 1 of SPJQM.

$$\sum_{\gamma \in (\vec{a}|\vec{x})} |\gamma\rangle = \mathbb{E}^{(\vec{a}|\vec{x})} \sum_{(\vec{a}'|\vec{x}) \in \vec{x}} |\vec{a}', \vec{x}\rangle = |\vec{a}, \vec{x}\rangle$$
(6.8)

satisfying $(\vec{a}|\vec{x}) \in \vec{x}$. By the sum rule 1 of $\tilde{Q}$, we can immediately extend this to all fine and coarsegrained physical measurement outcomes in $\tilde{\mathcal{O}}$. Finally, we let Eq 6.7 define vectors for all non physical events and we are done.

6.3   Acausal bounded field operator measurements

By Section 6.4 of [61], for an ideal measurement with resolution $\mathcal{R}$,

$$R_t = \bigcup_{n \in I} B_n \cap (B_n + t) \tag{6.9}$$

must equal $\emptyset$ or $[-a, b]$ for every $t \in \mathbb{R}$ if the update map is to be causal. Let $m = \text{textmin}(c_{n+1} - c_n)$ where $c_{N+1} = b$.

$$B_n \cap (B_n + t) = \begin{cases} \begin{cases} [c_n, c_{n+1}) \cap [c_n + t, c_{n+1} + t) & \text{if n}<\text{N} \\ [c_N, b] \cap [c_N + t, n + t] & \text{otherwise} \end{cases} & \text{if } 0 \leq t < m \\ 0 & \text{otherwise} \end{cases} \tag{6.10}$$

$$= \begin{cases} \begin{cases} [c_n + t, c_{n+1}) & \text{if n}<\text{N} \\ [c_N + t, b] & \text{otherwise} \end{cases} & \text{if } 0 \leq t < m \\ 0 & \text{otherwise} \end{cases} \tag{6.11}$$

Then, for $0 < t < m$, we have

$$R_t = \left(\bigcup_{n \in I} [c_n + t, c_{n+1}) \right) \cup [c_N + t, b] \neq \emptyset \tag{6.12}$$

we can see, assuming $N > 1$, by picking $\tau = c_2 + t/2$ that $\tau \in [-a, b]$ and $\tau \notin R_t$. Hence, $R_t \neq [-a, b]$ or $\emptyset$ for $0 < t < m$. By Claim 6.1 in [61], $\mathcal{E}_{\phi_a^b(f), \mathcal{R}}$ is acausal.