

## Dynamics and predictions in the co-event interpretation

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2009 J. Phys. A: Math. Theor. 42 235303

(<http://iopscience.iop.org/1751-8121/42/23/235303>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.154

The article was downloaded on 03/06/2010 at 07:52

Please note that [terms and conditions apply](#).

# Dynamics and predictions in the co-event interpretation

Yousef Ghazi-Tabatabai<sup>1</sup> and Petros Wallden<sup>2</sup>

<sup>1</sup> Blackett Laboratory, Imperial College, London, SW7 2AZ, UK

<sup>2</sup> Raman Research Institute, Bangalore 560 080, India

Received 23 January 2009, in final form 26 March 2009

Published 22 May 2009

Online at [stacks.iop.org/JPhysA/42/235303](http://stacks.iop.org/JPhysA/42/235303)

## Abstract

Sorkin has introduced a new, observer independent, interpretation of quantum mechanics that can give a successful realist account of the ‘quantum micro-world’ as well as explaining how classicality emerges at the level of observable events for a range of systems including single time ‘Copenhagen measurements’. This ‘co-event interpretation’ presents us with a new ontology, in which a single ‘co-event’ is real. A new ontology necessitates a review of the dynamical and predictive mechanism of a theory, and in this paper we begin the process by exploring means of expressing the dynamical and predictive content of histories theories in terms of co-events.

PACS number: 03.65.Ta

## 1. Introduction

### 1.1. Opening remarks

Building on the work of Dirac [7] and Feynman [12, 13], the *histories approach* to quantum mechanics [17, 21, 25, 31, 38] is a reformulation and generalization of (Copenhagen) quantum mechanics in an observer independent framework. With a ‘spacetime history’ replacing the ‘state’ as the fundamental object of the theory, the dynamics can be rephrased as a ‘decoherence functional’ (or alternatively a ‘quantum measure’) on subsets of the ‘sample space’ of histories. Observer independence is then to be sought through the ‘embedding’ of the (Copenhagen) concept of experimentally observable events in the wider (though ill-defined) notion of ‘classical partitions’, which are typically associated with the property of decoherence (dynamical classicality).

However the need for an interpretation remains. Because the mathematical formalism of a histories theory is reminiscent of probability theory, a naive application of the (relatively) well-understood ‘one history is real’ classical interpretation may seem appropriate; however, it fails to explain gedanken-experimentally realizable theories based on the Kochen–Specker theorem [9].

The *consistent histories* interpretation [22, 23] is perhaps the most developed alternative. Identifying as ‘classical’ every dynamically classical (decoherent) partition, it allows us to simultaneously assign truth valuations to all propositions within a single *consistent set* (a classical partition), within the ‘context’ of that consistent set, in a manner satisfying Boolean logic. However, two propositions that do not participate in a common consistent set cannot be simultaneously (truth) evaluated.

Sorkin has proposed a new interpretation [9, 40, 41] that can assign truth valuations to all elements of the event algebra simultaneously. This is achieved by moving from the classical ‘one history is real’ to an interpretation based on ‘one co-event is real’; where a co-event is associated with a truth valuation map  $\phi : \mathfrak{A} \rightarrow \mathbb{Z}_2$  obeying certain restrictions.

A new ontology necessitates a review of the dynamical and predictive mechanism of a theory. How do our observations and predictions relate to the objects we have identified as (potentially) real? Can this innovation help us to develop a better understanding of ‘classicality’, or a more complete theory of measurement? In this paper, we make a start at answering such questions by exploring means of expressing the dynamical and predictive content of histories theories in terms of co-events.

## 1.2. Outline of this paper

In section 1.3 we introduce quantum measure theory, a formulation of the histories approach which phrases the dynamics in a fashion resembling a probability measure. In section 1.4 we introduce the co-event interpretation, reviewing its application to classical stochastic theories before turning to the multiplicative scheme. We are then in a position to set out and motivate the goals of this paper in section 1.5. In section 2 we discuss the possibility of ‘moving wholesale’ from histories theories to ‘co-event theories’, in which the dynamics is described by a probability measure on the space of potentially real co-events. Unfortunately this approach stalls, so in section 3 we focus our attention to experimentally falsifiable predictions, and via Cournot’s principle (section 3.2) attempt to explain these through *approximate preclusion* (section 3.3); which we apply in a manner consistent with literal strong Cournot (section 3.4), and then in a manner consistent with operational weak Cournot (section 3.5). We conclude in section 4; the appendix reviews some results regarding the principle classical partition.

## 1.3. Quantum measure theory

*1.3.1. The quantum measure.* Quantum measure theory [38] is a histories based attempt to approach quantum mechanics as a generalization of classical stochastic theories; phrasing the quantum dynamics as a generalization of the probability measure. We start with a brief review of quantum measure theory and refer to [34, 38–41] for more details.

A *histories theory* (sometimes referred to as a *generalized measure theory*) consists of a triple,  $(\Omega, \mathfrak{A}, \mu)$ , of a ‘sample space’ of histories, an event algebra and a measure. The sample space,  $\Omega$ , contains all the ‘fine-grained histories’ or ‘formal trajectories’ for the system e.g. for  $n$ -particle mechanics—classical or quantum—a history would be a set of  $n$  trajectories in spacetime, and for a scalar field theory, a history would be a field configuration on spacetime. In a classical theory  $\Omega$  is the usual measure theoretic sample space.

The event algebra,  $\mathfrak{A}$ , contains all the (unasserted) propositions that can be made about the system. We will call elements of  $\mathfrak{A}$  *events*, following standard terminology in the theory of stochastic processes. Now the power set,  $P\Omega$ , of  $\Omega$  is usually thought of as a Boolean algebra under union and intersection, however for our purposes it will be more useful to view  $P\Omega$  as a ring (or an algebra over  $\mathbb{Z}_2$ ), with symmetric difference playing the role of addition

$(A+B := (A \cup B) \setminus (A \cap B))$  and intersection playing the role of multiplication ( $AB := A \cap B$ ). In cases where  $\Omega$  is finite,  $\mathfrak{A}$  can be identified with the whole power set,  $\mathfrak{A} = P\Omega$ , when  $\Omega$  is infinite,  $\mathfrak{A}$  can be identified with a sub-ring of the power set,  $\mathfrak{A} \subset P\Omega$ . In both cases, the algebraic properties of  $\mathfrak{A}$  play a central role in the formulation of the co-event interpretation. In a classical theory  $\mathfrak{A}$  will typically be the  $\sigma$ -algebra (or  $\delta$ -algebra) of measurable sets.

Predictions about the system—the dynamical content of the theory—are to be gleaned, in some way or another, from a generalized measure  $\mu$ , a non-negative real function on  $\mathfrak{A}$ .  $\mu$  is the dynamical law and initial condition rolled into one. In a classical theory,  $\mu$  will be a probability measure. In a quantum theory,  $\mu$  will obey [34, 38]:

(1) *Positivity*

$$\mu(A) \geq 0, \tag{1}$$

(2) *Sum rule*

$$\mu(A \sqcup B \sqcup C) = \mu(A \sqcup B) + \mu(B \sqcup C) + \mu(A \sqcup C) - \mu(A) - \mu(B) - \mu(C), \tag{2}$$

(3) *Unitality*

$$\mu(\Omega) = 1, \tag{3}$$

where  $A, B, C \in \mathfrak{A}$  are disjoint elements of  $\mathfrak{A}$ , as indicated by the symbol ‘ $\sqcup$ ’ for disjoint union. Given a histories theory, we can construct the following series of symmetric set functions, which are sometimes referred to as the Sorkin hierarchy<sup>3</sup>:

$$I_1(X) \equiv \mu(X),$$

$$I_2(X, Y) \equiv \mu(X \sqcup Y) - \mu(X) - \mu(Y),$$

$$I_3(X, Y, Z) \equiv \mu(X \sqcup Y \sqcup Z) - \mu(X \sqcup Y) - \mu(Y \sqcup Z) - \mu(Z \sqcup X) + \mu(X) + \mu(Y) + \mu(Z),$$

and so on, where  $X, Y, Z$ , etc are disjoint. A histories theory is *of level k* (alternatively it is a *measure theory of level k*) if the sum rule  $I_{k+1} = 0$  is satisfied. It is known that this condition implies that all higher sum rules are automatically satisfied, namely  $I_{k+n} = 0$  for all  $n \geq 1$  [38].

A level 1 theory is thus one in which the measure satisfies the usual Kolmogorov sum rules of classical probability theory,  $\mathbf{P}(A \sqcup B) = \mathbf{P}(A) + \mathbf{P}(B)$ , classical Brownian motion being a good example. We refer to level 1 theories as *classical theories*. A level 2 theory is one in which the Kolmogorov sum rules may be violated but  $I_3$  is nevertheless zero. Any unitary quantum theory can be cast into the form of a generalized measure theory and its measure satisfies the condition  $I_3 = 0$ , which trivially implies equation (2). We refer to level 2 theories, therefore, as *quantum measure theories*. It is known that whenever a system can be described by standard (Copenhagen) quantum mechanics via the usual Hilbert space construction (through what we will call a *Hilbert space theory*), we can also find a quantum measure theory describing the system. We say that the histories theory is derived from the Hilbert space theory; the quantum measure defined using Feynman path integrals [26].

Note that the replacement of the Kolmogorov sum rule with equation (2) has significant consequences for the null set structure (a null set is a set of measure zero); in particular because of (destructive) interference negligible sets (subsets of null sets) are no longer null in general. Unfortunately we lack general results about the structure of null sets, which is a hinderance to the development of our interpretations of quantum measure theory. There is recent research activity in this area, for example work on the implications of the existence of ‘antichains’ of null sets [42].

<sup>3</sup> These are the generalized interference terms introduced in [38].

In what follows, unless explicitly stated otherwise, we shall assume all histories theories to be level 2. Further, because of the current uncertainty as to the correct general construction of  $\mu$  and  $\mathfrak{A}$  in the case of an infinite sample space, we shall unless specifically mentioned otherwise, assume that all sample spaces are finite and that  $\mathfrak{A} = P\Omega$ .

The existence of a quantum measure,  $\mu$ , is more or less equivalent [38] to the existence of a *decoherence functional*,  $D(\cdot, \cdot)$ , a complex function on  $\mathfrak{A} \times \mathfrak{A}$  satisfying [24, 25]:

- (i) Hermiticity:  $D(X, Y) = D(Y, X)^*$ ,  $\forall X, Y \in \mathfrak{A}$ ;
- (ii) Additivity:  $D(X \sqcup Y, Z) = D(X, Z) + D(Y, Z)$ ,  $\forall X, Y, Z \in \mathfrak{A}$  with  $X$  and  $Y$  disjoint;
- (iii) Positivity:  $D(X, X) \geq 0$ ,  $\forall X \in \mathfrak{A}$ ;
- (iv) Normalization:  $D(\Omega, \Omega) = 1$ .<sup>4</sup>

The quantum measure is related to the decoherence functional by

$$\mu(X) = D(X, X) \quad \forall X \in \mathfrak{A}. \tag{4}$$

The quantity  $D(X, Y)$  is interpretable as the quantum interference between two sets of histories in the case when  $X$  and  $Y$  are disjoint.

Finally, we end this section by introducing the important concept of *coarse graining* [25, 26], which is crucial to the study of subsystems and emergent classicality.

**Definition 1.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. We say that a histories theory  $(\Lambda, \mathfrak{A}_\Lambda, \mu_\Lambda)$  is a coarse graining of  $(\Omega, \mathfrak{A}, \mu)$  if:

- (i)  $\Lambda$  is a partition of  $\Omega$ ;
- (ii)  $\mathfrak{A}_\Lambda = P\Lambda$ ;
- (iii)  $\mu_\Lambda(A) = \mu(A) \forall A \in \mathfrak{A}_\Lambda$ .

We refer to  $(\Omega, \mathfrak{A}, \mu)$  as a fine graining of  $(\Lambda, \mathfrak{A}_\Lambda, \mu_\Lambda)$ . In this context we sometimes refer to the histories  $\gamma \in \Omega$  as fine-grained histories to distinguish them from the coarse-grained  $A \in \Omega_\Lambda$ .

**1.3.2. Classicality and observation.** We need to relate this framework to our (experimental) observations and to develop the ability to make falsifiable predictions. Now we envision that given a histories theory  $(\Omega, \mathfrak{A}, \mu)$ , some of the events in  $\mathfrak{A}$  may be ‘observable’, so that we can determine their truth or falsity through our experience. However not all the elements of  $\mathfrak{A}$  will in general be observable, leading us to enquire into the properties the set of observable events.

In ‘Copenhagen quantum mechanics’ described by a ‘Hilbert space theory’ we assume that the ‘observables’ are given; determined by an assumed measuring apparatus that is external to the system. Further, we can assign (Copenhagen) probabilities to each of these events; these probabilities constitute the predictions of the theory<sup>5</sup>. If a histories theory  $(\Omega, \mathfrak{A}, \mu)$  is derived from such a Copenhagen framework, then we inherit a set of ‘observable events’, each of which is an element of an ‘observable partition’ (of  $\Omega$ ) whose elements are all observable events. Further, it can be shown that such partitions are *decoherent*, so that the decoherence functional  $D$  associated with the measure obeys  $D_\Lambda(X, Y) = 0$  whenever  $X, Y \in \Lambda$  and  $X \neq Y$  [25]. It can be shown that this implies that the coarse-grained measure  $\mu_\Lambda$  obeys the axioms of a probability measure, and that  $\mu_\Lambda(A)$  is equal to the ‘Copenhagen probability’ of

<sup>4</sup> The normalization condition may turn out not to be necessary, but we include it because all the quantum measures we consider in this paper will satisfy it.

<sup>5</sup> Though as we shall see a ‘probability statement’ such as  $\mathbf{P}(A) = p$  is not in general falsifiable through a single trial; we must construct a methodology for testing such assertions by relating them to statements that are falsifiable.

the event  $A$ . These ‘probability statements’ regarding observable events form the predictive content of such theories.

However for a general quantum measure theory we do not have a fully developed framework for dealing with observable events. Further, one of the key themes of the histories approach is observer independence; observations are certainly not observer independent. This leads us to ‘embed’ the notion of (experimentally) observable events into the wider (but also at present ill-defined) notion of ‘classicality’, which we require to be observer independent. Thus we wish to distinguish some partitions as being ‘classical’ in some objective and observer independent manner; these *classical partitions* are to possess the properties we require of observable partitions, and are to be interpreted in the same manner. The idea is that the set of classical partitions will contain the set of observable partitions; thus the ‘classical’ behaviour of observed events is to be thought of an instance of an objective and observer independent phenomenon.

However, as alluded to above we are as yet uncertain what this broader phenomenon of classicality should be. Perhaps the simplest suggestion is that of *dynamical classicality*, or *decoherence*. As we saw above, ‘Copenhagen’ observable partitions are decoherent; further, when the measure is classical the whole sample space is decoherent. This leads us to the notion of dynamical classicality, which identifies classicality with decoherence; thus a partition  $\Lambda$  of  $\Omega$  is dynamically classical if and only if  $D(X, Y) = 0$  for all  $X, Y \in \Lambda$  such that  $X \neq Y$ . As before, it can be shown that this implies that the coarse-grained measure  $\mu_\Lambda$  obeys the axioms of a probability measure. The notion of *approximate decoherence*,  $D(X, Y) < \epsilon \ll 1$  for all  $X, Y \in \Lambda$  such that  $X \neq Y$ , is sometimes offered as an alternative to decoherence. In particular, in the study of environmental decoherence, observed ‘emergent classical behaviour’ is due to approximate decoherence [10, 18]. However for the purposes of this paper we will focus on the ‘full’ decoherence condition. Thus, given a general quantum measure theory, we will assume that we can identify the observable events, that each observable event participates in an observable partition and that all observable partitions are dynamically classical.

Now given a single dynamically classical partition  $\Lambda$ , we can treat the coarse-grained histories theory  $(\Lambda, \mathfrak{A}_\Lambda, \mu_\Lambda)$  as a classical theory and apply our usual interpretation; identifying a single history as real. We are then able to simultaneously assign truth values (in  $\mathbb{Z}_2$ ) to every element in  $\Lambda$ . However, when the underlying theory  $(\Omega, \mathfrak{A}, \mu)$  is not itself classical, we may have various *incompatible* (or *incomparable*) partitions [22]; we are not able to simultaneously assign truth values to all events in two incompatible partitions using the classical ‘one history is real’ interpretation [10]. The *consistent histories* interpretation addresses this issue by interpreting every ‘classical proposition’ (an event that is an element of a dynamically classical partition) as *conditional* on the (‘largest’) dynamically classical partition in which it participates. Essentially, it is only meaningful to assign truth valuations to classical partitions, and this must be done within the context of a given dynamically classical (or *consistent*) partition; two propositions can only be simultaneously answered if they both participate in a common consistent partition [10, 22, 23].

Sorkin [40, 41] has introduced a new interpretation of quantum measure theory that can assign truth valuations to all elements of the event algebra simultaneously, rather than to individual consistent partitions. It is to this, *co-event interpretation* that we now turn.

#### 1.4. The co-event interpretation

*1.4.1. Co-events for classical theories.* When  $(\Omega, \mathfrak{A}, \mathbf{P})$  is a classical theory, the ‘standard interpretation’ identifies exactly one of the elements of  $\Omega$  as the ‘real history’,  $\gamma_r$ , which ‘actually occurs’. However the dynamics (in the form of the measure  $\mathbf{P}$ ) does not uniquely

identify  $\gamma_r$ ; the most we can say with complete confidence is that the dynamics should not *preclude* (rule out) the real history, which would occur were  $\gamma_r$  to be an element of a null set (a set of measure zero). We thus use the term *potential reality* in referring to a history which is not an element of a null set, and denote the set of such histories as the *space of potential realities*. To avoid confusion with the potential realities, we will henceforth refer to the real history  $\gamma_r$  as the *actually real history*.

Recalling that we can think of the event algebra  $\mathfrak{A}$  as the set of propositions that can be made about the system, note that knowledge of which potential reality is actually real allows us to give true/false answers to each of these propositions. More concretely, we say that an event is ‘true’ if it contains the actually real history  $\gamma_r$ , otherwise it is ‘false’. Identifying our space of truth valuations with  $\mathbb{Z}_2$  (where 1 is associated with ‘yes’/‘true’ and 0 with ‘no’/‘false’), we can use the history  $\gamma_r$  to construct a truth valuation map, or *co-event*:

$$\begin{aligned} \gamma_r^* : \mathfrak{A} &\rightarrow \mathbb{Z}_2, \\ \gamma_r^*(A) &= \begin{cases} 1 & \gamma_r \in A \\ 0 & \gamma_r \notin A. \end{cases} \end{aligned} \tag{5}$$

We can then phrase our requirement that no potential reality should lie in a null set in the language of truth valuation maps, so that  $\gamma \in \Omega$  is a potential reality whenever

$$\mathbf{P}(A) = 0 \Rightarrow \gamma^*(A) = 0, \tag{6}$$

a condition which we will refer to as *preclusion*. We will refer to co-events obeying equation (6) as *preclusive*.

Sorkin has suggested a subtle shift in thinking whereby the co-events  $\gamma^*$  should be regarded as real rather than the histories  $\gamma$  [9, 40, 41]. Thus given a classical histories theory  $(\Omega, \mathfrak{A}, P)$  our space of potential realities is now the set of preclusive  $\gamma^*$  where  $\gamma \in \Omega$ .

This shift in thinking is non-trivial because co-events have a natural algebraic structure which we can use to generalize them from the classical case. As previously noted,  $\mathfrak{A}$  is both a ring and an algebra over  $\mathbb{Z}_2$  with addition identified with symmetric difference and multiplication with intersection. As a field,  $\mathbb{Z}_2$  is also a ring (and could be thought of as an algebra over  $\mathbb{Z}_2$ ), inheriting addition and multiplication from  $\mathbb{Z}$ . When the sample space is finite, it can be shown that the co-events  $\gamma^*$  are precisely the (ring) homomorphisms from  $\mathfrak{A}$  to  $\mathbb{Z}_2$ .

**Lemma 1.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. Then*

$$\{\gamma^* | \gamma \in \Omega\} = \text{Hom}(\mathfrak{A}, \mathbb{Z}_2),$$

where  $\text{Hom}(\mathfrak{A}, \mathbb{Z}_2)$  is the set of (ring) homomorphisms from  $\mathfrak{A}$  to  $\mathbb{Z}_2$  excluding the zero map.

Thus the co-events  $\gamma^*$  can be characterized algebraically by the fact that they are (non-zero) homomorphisms, which is equivalent to their adherence to the following ‘rules’:

$$\gamma^*(A + B) = \gamma^*(A) + \gamma^*(B) \text{ (linearity),} \tag{7}$$

$$\gamma^*(AB) = \gamma^*(A)\gamma^*(B) \text{ (multiplicativity),} \tag{8}$$

for all  $A, B \in \mathfrak{A}$ .

**1.4.2. Co-events for quantum theories.** The co-events introduced in the previous section are tied to individual histories. We can generalize this structure to encompass a wider class of answering map.

**Definition 2.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory. Then a map

$$\phi : \mathfrak{A} \rightarrow \mathbb{Z}_2,$$

is a co-event if it maps the empty set to zero. We denote the space of co-events by  $\mathfrak{A}^*$ . A co-event is preclusive if

$$\mu(A) = 0 \Rightarrow \phi(A) = 0 \quad \forall A \in \mathfrak{A}.$$

It is clear that when the underlying histories theory is classical, the above definition of preclusiveness reduces to the one given in equation (6). Because of this, we can define the ‘classical’ co-events within this more general setting, describing any co-event (other than the zero map) as *classical* if it obeys linearity and multiplicativity (equations (7) and (8)). Lemma 1 means that the classical co-events are all the co-events of the form  $\gamma^*$  for some  $\gamma \in \Omega$ , as described above. Given a histories theory  $(\Omega, \mathfrak{A}, \mu)$ , we refer to the set of preclusive classical co-events as  $\mathcal{C}(\Omega, \mathfrak{A}, \mu)$ .

Perhaps the most obvious means of applying co-events to quantum measure theories would be to use the classical structure unaltered; thus given a histories theory  $(\Omega, \mathfrak{A}, \mu)$  our space of potential realities would be  $\mathcal{C}(\Omega, \mathfrak{A}, \mu)$ . However it is possible to show that classical co-events are an inadequate ontology for quantum dynamics; in particular the Kochen–Specker theorem [27], often cited as an obstruction to realism, and its proof by Asher Peres [32], can be used to construct a ‘gedanken-experimentally realizable’ system that cannot be described by any classical co-event [9]. This leads to

**Theorem 1.** There exists a ‘gedanken-experimentally realizable’ histories theory  $(\Omega, \mathfrak{A}, \mu)$  with a finite sample space such that

$$\mathcal{C}(\Omega, \mathfrak{A}, \mu) = \emptyset.$$

**Proof.** See [9]. □

The failure of classical co-events to adequately describe quantum measure theories has led to a search for alternatives. A number of co-event ‘schemes’ have been proposed, typically generalizing from classical co-events by omitting and/or relaxing one or both of equations (7) and (8). Many of these proposals, such as the *linear* [41], *quadratic* and *general polynomial* schemes, have been conclusively ruled out [19]. Other proposals, such as the *ideal* and *multiplicative* schemes remain as viable candidates [19, 40]. For various reasons [8, 19, 37], the multiplicative approach has been adopted as the current working model of co-event theory and it is to this scheme that we now turn.

**1.4.3. The multiplicative scheme.** Starting with a histories theory  $(\Omega, \mathfrak{A}, \mu)$ , a co-event  $\phi \in \mathfrak{A}^*$  is *multiplicative* if it is not the zero map and it obeys the multiplicativity condition (equation (8)):

$$\phi(AB) = \phi(A)\phi(B) \quad \forall A, B \in \mathfrak{A}.$$

An important feature of multiplicative co-events is their description in terms of filter theory:

**Lemma 2.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space and let  $\phi \in \mathfrak{A}^*$  be a multiplicative co-event. Then  $\phi^{-1}(1)$  is a principal filter.

**Proof.** Let  $A \in \phi^{-1}(1)$ . If  $A \subset B$  then  $AB = A$  and therefore  $1 = \phi(A) = \phi(AB) = \phi(A)\phi(B) = \phi(B)$ , which means that  $B \in \phi^{-1}(1)$ . Further, if  $A, B \in \phi^{-1}(1)$  then  $\phi(AB) = \phi(A)\phi(B) = 1$ , so  $A \cap B \in \phi^{-1}(1)$ ; thus  $\phi^{-1}(1)$  is a filter.

Now because  $\Omega$  and therefore  $\mathfrak{A}$  are finite  $\phi^{-1}(1)$  must contain a minimal element (under the partial order defined by set inclusion), which we will denote  $\phi^* \in \mathfrak{A}$ . Further  $\phi^*$  must be unique, otherwise if  $A \in \phi^{-1}(1)$  is also minimal then  $A \cap \phi^*$  is in the filter  $\phi^{-1}(1)$  and is contained in both  $A$  and  $\phi^*$ , contradicting the assumption that they are minimal. Therefore  $\phi^{-1}(1)$  is a principal filter and  $\phi^*$  its principal element.  $\square$

The association of a multiplicative co-event  $\phi$  with the principal element  $\phi^*$  of the related filter  $\phi^{-1}(1)$  can be described as a map:

$$* : \mathfrak{A}^* \rightarrow \mathfrak{A}, \quad * : \phi \mapsto \phi^*. \tag{9}$$

We can extend  $*$  to an involution on  $\mathfrak{A} \times \mathfrak{A}^*$  by defining

$$* : \mathfrak{A} \rightarrow \mathfrak{A}^*, \quad * : A \mapsto A^*, \tag{10}$$

where for all  $B \in \mathfrak{A}$

$$A^*(B) = \begin{cases} 1 & \text{if } A \subset B \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

It is easy to see that  $*$  is bijective, and that  $(\phi^*)^* = \phi$ ,  $(A^*)^* = A$ . We can therefore think of  $*$  as a duality and will call  $\phi^*$  the *dual* of  $\phi$  (and  $A^*$  the dual of  $A$ ).

Note that every classical co-event is also a multiplicative co-event, and that  $\{\gamma\}^* = \gamma^*$ , so our definition of the  $*$  map is consistent with our previous notation. Now the classical co-events  $\gamma^*$  are dual to single histories  $\gamma$ , which are the potential realities (given that they are preclusive) in classical theories. This suggests that we might think of a multiplicative co-event  $A^*$  as an *ontological coarse graining*; rather than one history  $\gamma$  (or co-event  $\gamma^*$ ) describing reality we now have a multiplicity of histories  $A$  (or the co-event  $A^*$ ), and only the properties common to all of these histories are true.

To generalize from (preclusive) classical to (preclusive) multiplicative co-events we have dropped the linear rule (equation (7)), thus increasing the number of ‘allowed’ co-events. Unfortunately we now have ‘too many’ co-events; in particular given a classical measure we could have non-classical multiplicative co-events, implying non-classical behavior of dynamically classical systems. Indeed, if the measure does not admit null sets then the dual of every event will be preclusive and multiplicative. We therefore need a further constraint that will restrict us to classical co-events given a classical measure. We say that a multiplicative co-event  $\psi$  *dominates* a multiplicative co-event  $\phi$  if

$$\phi(A) = 1 \Rightarrow \psi(A) = 1 \quad \forall A \in \mathfrak{A}.$$

A preclusive multiplicative co-event is *primitive* (among the set of preclusive multiplicative co-events) if it is not dominated by any other preclusive multiplicative co-event. We can also motivate primitivity using the natural order on filters. A preclusive multiplicative co-event  $\phi$  is *primitive* if and only if the associated filter  $\phi^{-1}(1)$  is maximal among the filters corresponding to preclusive multiplicative co-events. Primitivity can be thought of as a condition of ‘maximal detail’ or ‘finest graining’ consistent with preclusion. We are now in a position to fully specify the concept of potential reality in the multiplicative scheme

**Definition 3.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. Then we denote the set of primitive preclusive multiplicative co-events by  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ .

Comparing the multiplicative scheme with our previous results for classical theories, it is  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  which replaces  $\mathcal{C}(\Omega, \mathfrak{A}, \mu)$  as the space of potential realities. In this scheme individual histories (or equivalently classical co-events) may no longer be potentially real, and

thus cannot be expected to necessarily obey the logic describing reality. Notice that we use ‘standard’ (Boolean) logic in our reasoning concerning the co-events themselves.

Recall that classical co-events were discarded because we could find gedanken-experimentally realizable histories theories that no preclusive classical co-event could describe (theorem 1). We are therefore reassured by the following,

**Lemma 3.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a history theory with a finite sample space and let  $A \in \mathfrak{A}$  be non-negligible (i.e.  $A$  is not a subset of a null set). Then  $\exists \phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  such that  $\phi(A) = 1$ .*

**Proof.** Assume that  $\nexists \phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  such that  $\phi(A) = 1$ . Because  $A$  is non-negligible we know that  $A^*$  is preclusive. Then since  $A^*(A) = 1$ , by assumption  $A^*$  cannot be primitive, so there exists some preclusive  $\psi_1$  dominating  $A^*$  and by the definition of domination  $\psi_1(A) = 1$ . The same argument shows that  $\psi_1$  cannot be primitive, so there exists a preclusive  $\psi_2$  dominating  $\psi_1$  with  $\psi_2(A) = 1$ . Carrying on in this fashion we can find an infinite sequence of preclusive co-events  $\{\psi_i\}_{i=0}^\infty$  with  $\psi_0 = A^*$ , with  $\psi_i$  dominating  $\psi_{i-1}$  (and thus  $\psi_{i-n}$ ) and with  $\psi_i(A) = 1$ . However this contradicts our assumption that  $\Omega$ , and therefore  $\mathfrak{A}$  and  $\mathfrak{A}^*$ , are finite; hence the result.  $\square$

Note that  $\mu(\Omega) = 1$  in any histories theory, thus whenever  $\Omega$  is finite we know that  $\mathcal{M}(\Omega, \mathfrak{A}, \mu) \neq \emptyset$ . Further, we can show that primitivity is indeed successful in ensuring that given a classical measure our space of potential realities consists of classical co-events.

**Lemma 4.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a history theory with a finite sample space and a classical (level 1) measure, then  $\mathcal{M}(\Omega, \mathfrak{A}, \mu) = \mathcal{C}(\Omega, \mathfrak{A}, \mu)$ .*

**Proof.** Let  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$ . The Kolmogorov sum rule implies that the union,  $Z$ , of null sets in  $\mathfrak{A}$  is itself null, so the preclusivity of  $\phi$  implies  $\phi^* \not\subset Z$ . Thus  $\exists \gamma \in \phi^* \setminus Z$  such that  $\gamma$  is not an element of any null set. This means  $\gamma^*$  is preclusive, but then  $\gamma \in \phi^*$  means that either  $\gamma^*$  dominates  $\phi$ , or that  $\phi = \gamma^*$ . But by assumption  $\phi$  is primitive and so is not dominated by any preclusive co-event. Therefore  $\phi = \gamma^*$ , a homomorphism, and  $\phi \in \mathcal{C}(\Omega, \mathfrak{A}, \mu)$ . Further, any homomorphism is a multiplicative co-event, thus  $\mathcal{M}(\Omega, \mathfrak{A}, \mu) = \mathcal{C}(\Omega, \mathfrak{A}, \mu)$ .  $\square$

Now, as touched on in section 1.3.2, the histories approach typically assumes that dynamical classicality is an ‘emergent’ phenomena, so that classical histories theories are to be thought of as coarse grainings of ‘deeper’ quantum measure theories. Thus the ‘measure of the universe’ is level 2 whereas we access it through dynamically classical partitions of the sample space on which the induced measure is level 1. We would therefore like to see a stronger version on lemma 4 which ensured that (given  $(\Omega, \mathfrak{A}, \mu)$ ) all  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  ‘behave classically’ on ‘classical partitions’. Of course whether or not we can achieve such a result depends upon how we define these two terms, ‘classical behavior’ turns out to be the simpler of the two.

**Definition 4.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space and let  $\Lambda$  be a partition of  $\Omega$ . Then we say that a co-event  $\phi \in \mathfrak{A}^*$  is classical on  $\Lambda$  if the restriction of  $\phi$  to the subalgebra of  $\mathfrak{A}$  generated by  $\Lambda$  is a homomorphism.*

As discussed in section 1.3.2, defining exactly what we mean by a ‘classical partition’ remains an open area of research. On the one hand given a histories theory  $(\Omega, \mathfrak{A}, \mu)$  we would like every  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  to be classical on every classical partition of  $\Omega$ , on the other hand we want to be able to regard classical partitions as ‘dynamically classical’ in some sense; at the very least partitions corresponding to experimentally observable alternatives should be

classified as classical. Perhaps the most obvious way forward would be to regard all decoherent partitions as classical, however this condition turns out to be too strong. At the very least, because we observe and reason with Boolean logic we require our potentially real co-events  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  to treat classical partitions classically. We call this condition *ontological classicality* or *classicality with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$*  to distinguish it from both dynamical classicality and the more general notion of a classical partition, which we may require to be stronger.

**Definition 5.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. We say that a partition  $\Lambda$  of  $\Omega$  is classical with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  if every  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  is classical on  $\Lambda$ .*

In the appendix, we demonstrate the existence of a *principle classical partition*  $\Lambda_C$  which all  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  treat classically; further any partition on which any  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  is classical is a sub-partition of  $\Lambda_C$ .

Relating this back to the dynamics, Sorkin has proposed a new dynamical condition, *preclusively separability* [37], that ensures classicality with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ . A partition  $\Lambda$  of  $\Omega$  is preclusively separable if for any null set  $Z \in \mathfrak{A}$  and any partition element  $A \in \Lambda$ , either  $Z \cap A = \emptyset$  or  $\exists$  a null set  $Z_A \subset A$  such that  $Z_A \supset A \cap Z$ .

The class of preclusively separable partitions includes all partitions in which any two histories belonging to distinct partition elements ‘end’ at different locations at the final time (or equivalently correspond to different final time projectors in their class operators, if the histories theory is derived from a Hilbert space theory). Thus in any histories theories derived from a Hilbert space theory, any partition corresponding to the alternative outcomes of a particular Copenhagen measurement (thus corresponding to ‘orthogonal projectors’) obeys preclusively separability [8, 37], so the fact that preclusively separable partitions are always classical with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  is a powerful result. This remains an open area of research.

*1.4.4. Co-events for coins.* We illustrate the use of co-events with a simple example. Consider a single toss of a coin; we assume that the outcome will be either ‘heads’ (‘h’) or ‘tails’ (‘t’), so our sample space is

$$\Omega = \{h, t\},$$

and our event algebra

$$\mathfrak{A} = P\Omega = \{\emptyset, \{h\}, \{t\}, \Omega\}.$$

We assume no interference between the two outcomes, so that our measure is classical

$$\begin{aligned} \mu(\emptyset) &= 0, \\ \mu(\{h\}) &= p, \\ \mu(\{t\}) &= 1 - p, \\ \mu(\Omega) &= 1. \end{aligned}$$

Then assuming that  $p \notin \{0, 1\}$  we have

$$\mathcal{C}(\Omega, \mathfrak{A}, \mu) = \{h^*, t^*\},$$

and the multiplicative co-events are

$$h^*, t^*, \Omega^*.$$

Note that  $\emptyset^*$  is not included because it is not a co-event (since no co-event maps  $\emptyset$  to unity). Now assuming that  $p \notin \{0, 1\}$  all three of the above co-events are preclusive, however both  $h^*$  and  $t^*$  dominate  $\Omega^*$ , while being themselves primitive. Then we have

$$\mathcal{M}(\Omega, \mathfrak{A}, \mu) = \mathcal{C}(\Omega, \mathfrak{A}, \mu),$$

as we would expect from lemma 4. Nevertheless, under non-classical measures non-classical co-events similar to  $h^*t^*$  are often primitive. It is thus instructive to delve further into the non-classicality of  $\Omega^*$ , which we can do by constructing its ‘truth table’:

	$\emptyset$	$\{h\}$	$\{t\}$	$\Omega$
$\Omega^*$	0	0	0	1

where each table entry is the truth valuation  $\Omega^*(A)$  of the relevant event  $A$ . We can rephrase this in ‘question and answer’ format:

Question	Answer
Does the coin show heads?	No
Does the coin show tails?	No
Does the coin show one of heads or tails?	Yes

Noting that  $\{h, t\}$  can be thought of as a (trivial) partition of  $\Omega$ , we could say that  $\Omega^*$  is not classical on the ‘classical partition’  $\{h, t\}$ . This is because  $\Omega$  has non-empty intersection with both partition elements.

### 1.5. The goals of this paper

As far as co-events go our predictive power is (so far) limited. Given an event algebra and a measure, our co-event approach identifies the primitive preclusive co-events (assuming a particular scheme) as the potential realities of our theory; however, we cannot in general reconstruct the measure from the set of potentially real co-events. Were we to ignore the measure and use only the potentially real co-events themselves (without knowledge of which co-event is actually real), we would be unable to make any predictive statements beyond ruling out as impossible those events that are precluded (mapped to zero) by all the potentially real co-events, and identifying as certain all those events that are mapped to unity. This may not be quite as restrictive as it appears, for by *conditioning* on known truth valuations of events (disregarding the co-events that disagree with these valuations) we may be able to narrow down the set of potentially real co-events to the point that such predictions are relatively powerful [37].

However there is reason to believe (depending on the definition of classicality we adopt) that conditioning on previous observations will not increase our power to predict other observable events; though conditioning on non-observable events may increase our ability to predict observable events and conditioning on observable events may increase our ability to predict non-observable events.

In any case, it is clear that such conditioning will not in general allow us to recover the *entirety* of either the measure or its predictive content, whatever definition of classicality we adopt; whereas we would like to be able to express the full dynamical and predictive content of a quantum measure theory in terms of the potentially real co-events rather than the histories. We would thus be able to ‘start with’ a set of potentially real co-events and make our predictions using only these co-events and structures defined in terms of them. In this paper, we will explore attempts to achieve this; we are motivated in our endeavor by three arguments.

Firstly, we are treating the primitive co-events themselves as the potential realities of the histories approach. The authors feel that we should be able to express all meaningful statements concerning a theory in terms of the theory's potential realities; even if *in practise* it is simpler to use the measure  $\mu$  defined on the event algebra, at least *in principle* this structure should 'emerge' from a 'deeper' expression of the dynamical and predictive content of the theory in terms of the co-events themselves.

Secondly, the presence of 'exactly null' sets in many gedanken experimental frameworks (such as the double slit experiment) depends upon idealized assumptions (such as point slits) that do not hold in practice. Regardless of our interpretation, co-events as they now stand cannot distinguish between events of large or small measure, so our use of null sets to approximate 'almost null' sets in idealized gedanken experiments requires justification. As mentioned above, at present the range of predictive statements we can make with the set of potentially real co-events is narrow; were we able to fully express the predictive content of quantum measure theory in terms of the potentially real co-events we may be able to better grapple with issues such as 'almost null sets' that are phrased in terms of probability rather than preclusion.

Thirdly, were we able to express in terms of co-events the predictive content of those histories theories derived from Hilbert space theories, we might then be in a position to extrapolate our findings and better understand the predictive content of a general quantum measure theory (in which, for example, there may be no 'external' measuring apparatus).

In what follows we will explore two recent attempts to rephrase in terms of co-events the dynamical and predictive content of quantum measure theories. We first consider attempts to move wholesale from the 'histories theory' structure to one of 'co-event theories' in which the dynamics is defined directly in terms of co-events (as anticipated in [40]). We then turn our attention to the important concept of 'approximate preclusion', and discuss the details of its implementation.

Unless specifically noted otherwise we will assume that all sample spaces are finite, and that all histories theories are derived from Hilbert space theories (we will continue to make the assumptions described above regarding observable events). Further, as previously stated, given a histories theory we will assume that we can identify the observable events, that each observable event participates in an observable partition and that all observable partitions are dynamically classical.

## 2. Dynamics on co-events

### 2.1. Placing a measure on the space of co-events

One approach, anticipated by Sorkin in [40], is to rephrase the dynamics (the quantum measure  $\mu$ ) in terms of the co-events themselves; for example, by placing a measure on the space of co-events. This would be a natural development, given that the co-events are the basic potential realities of the theory, and could lead to the rephrasing of the entirety of quantum mechanics in terms of co-events, bypassing our current need to construct dynamics upon the event algebra. We could then explain the rarity of low probability events (such as 'almost null' sets) by using the dynamics on our co-events to suppress (assign a low probability to) those co-events that value them to 1.

Perhaps the most obvious way to do this would involve literally 'moving' the dynamics from the event algebra onto the co-events, which can thus realize their status as the central objects of the theory. So far we have always begun with a histories theory  $(\Omega, \mathfrak{A}, \mu)$  from whence we have derived our allowed co-events  $\mathcal{S}(\Omega, \mathfrak{A}, \mu)$ . Since we are asserting that

$S(\Omega, \mathfrak{A}, \mu)$  rather than  $\Omega$  is the ‘true’ sample space of ‘potential realities’ it would seem more natural to place the co-events at the centre of our structure, moving from a histories theory  $(\Omega, \mathfrak{A}, \mu)$  to a full *co-event theory*  $(S, PS, \mathbf{P}_S)$ , where  $S$  is a set of (allowed) co-events, for example  $S = \mathcal{M}(\Omega, \mathfrak{A}, \mu)$ ,  $PS$  is the power set of  $S$ , and  $\mathbf{P}_S$  is a probability measure on  $PS$ . We have assumed a probability measure, for if we were to use a more general higher order measure we would not have gained anything from our adoption of co-events in place of histories, encountering the same interpretational difficulties that led us to co-events. Indeed, following that route might even lead us to adopt co-co-events, co-co-co-events and so on. Conversely, since we are assuming that ‘one co-event is real’, the use of a probability measure echoes the use of probability in classical stochastic theories in which we believe that ‘one history is real’.

In a sense this approach is born of the Bayesian approach to probability [5, 6, 33], which looks on probability as representing a degree of information concerning a system and thus ascribes meaning to individual probability statements which can therefore be treated as ‘logical’ predicates in their own right [5]. Following this philosophy we seek to express our incomplete information about the actual reality in terms of probability statements concerning the potential realities, much as we are accustomed to do in the classical theory. We can, therefore, think of  $\mathbf{P}_S(\{\phi\})$  as the probability that the co-event  $\phi$  is the actually real co-event; consequently  $\mathbf{P}_S(\{\phi\}) = 0$  could be interpreted to mean that  $\phi$  is not a potential reality. Indeed, we could even enforce our choice of  $S$  as the set of potential realities by extending the domain of  $\mathbf{P}_S$  to all of  $\mathfrak{A}^*$ , and constraining  $\mathbf{P}_S$  to be zero on all co-events outside  $S$ .

Nevertheless our observations are in terms of events in  $\mathfrak{A}$ , so we must relate this new structure back to our Copenhagen probabilistic predictions. As stated above we are focusing only on histories theories  $(\Omega, \mathfrak{A}, \mu)$  that are derived from a Hilbert space theory; we have assumed that in such theories we can identify the set of ‘observable events’, that each observable event is an element of an ‘observable partition’, and that  $\mu$  is classical on all such partitions. As noted above, if  $A$  is an observable event,  $\mu(A)$  can be interpreted as a probability and this probability agrees with the Copenhagen predicted probability for  $A$ . Then at the very least we want, for all observable  $A$ ,

$$\mu(A) = \begin{cases} 1 \\ 0 \end{cases} \Rightarrow \mathbf{P}_S(\{\phi \in S | \phi(A) = 1\}) = \begin{cases} 1 \\ 0 \end{cases}. \tag{12}$$

As we will discuss below, non-binary (i.e., not equal to 0 or 1) probabilities themselves require interpretation and can arguably be defined in terms of repeated trials and binary probabilities. However we will steer clear of this issue for now, and instead of using equation (12) to define probabilities through repeated trials we will instead use the stronger

$$\mathbf{P}_S(\{\phi \in S | \phi(A) = 1\}) = \mu(A), \tag{13}$$

for all observable events  $A$ . However it is not clear how ‘observable events’ should be defined in our quantum measure theory (see section 1.3.2). We sidestep this question by simply strengthening the condition once more to require:

$$\mathbf{P}_S(\{\phi \in S | \phi(A) = 1\}) = \mu(A) \quad \forall A \in \mathfrak{A}. \tag{14}$$

One difficulty with this condition is that in general  $\mu(A)$  can exceed 1. However, even when we are able to constrain  $\mu$  to take values in  $[0, 1]$ , equation (14) is a problematic condition for multiplicative co-events, as the following theorem shows.

**Theorem 2.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space, let  $S \subset \mathfrak{A}^*$  be a set of multiplicative co-events on  $\mathfrak{A}$ , so that*

$$\phi(AB) = \phi(A)\phi(B) \quad \forall A, B \in \mathfrak{A}, \quad \forall \phi \in S.$$

Further, let  $\mathbf{P}_S$  be a probability measure on  $S$  such that

$$\mathbf{P}_S(\{\phi \in S | \phi(A) = 1\}) = \mu(A) \quad \forall A \in \mathfrak{A}.$$

Then for any  $\phi \in S$ ,  $\mathbf{P}_S(\{\phi\}) = 0$  unless for all  $A, B, C \in \mathfrak{A}$

$$\phi(A + B + C) = \phi(A + B) + \phi(B + C) + \phi(C + A) + \phi(A) + \phi(B) + \phi(C). \quad (15)$$

As noted above, the use of a probability measure on the space of co-events echoes its use on histories in classical stochastic theories. Then following our usual interpretation that dynamically precluded ‘events’ do not occur, we conclude that the actually real co-event must obey equation (15). Now equation (15) is the defining property of the *quadratic (co-event) scheme* [19] we alluded to in section 1.4.2, a *quadratic co-event* being in general a co-event obeying equation (15) though not necessarily obeying linearity or indeed multiplicativity. Then since the quadratic scheme has been ruled out [19], this puts into question our idea of ‘moving’ the dynamics onto the co-events by constructing ‘co-event theories’ in place of histories theories, unless we are willing to abandon the multiplicative scheme.

Of course this approach is not entirely ruled out, to arrive at equation (14) we have made several simplifying assumptions so there may still be some scope for further investigation. For example, if we accept  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  as the space of potential realities of a histories theory  $(\Omega, \mathfrak{A}, \mu)$ , we might attempt to define all other elements of the theory in terms of the co-events  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$ . In particular, we could define events as maps from  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  to  $\mathbb{Z}_2$  using the ‘dual map’  $A[\phi] = \phi(A)$ . Since we may not be able to distinguish between all elements of  $\mathfrak{A}$  in this manner, we might perhaps choose to take equivalence classes of ‘distinguishable events’ ( $A \sim B$  if  $A[\phi] = B[\phi]$  for all  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$ ), and to attempt to define a measure on this set as a stepping stone to a ‘full co-event theory’. Such a structure is yet to be explored. Alternatively, we might explore the restrictions on  $\mu$  that would enable equation (14), or we might turn to another scheme.

## 2.2. Proof of theorem 2

In this section, we will prove theorem 2. We will need the following technical lemmas.

**Lemma 5.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space and let  $\phi$  be a (not necessarily multiplicative) co-event. Then*

$$\phi(A \sqcup B \sqcup C) = \phi(A \sqcup B) + \phi(B \sqcup C) + \phi(C \sqcup A) + \phi(A) + \phi(B) + \phi(C), \quad (16)$$

$\forall$  disjoint  $A, B, C \in \mathfrak{A}$  if and only if

$$\phi(A' + B' + C') = \phi(A' + B') + \phi(B' + C') + \phi(C' + A') + \phi(A') + \phi(B') + \phi(C'), \quad (17)$$

$\forall A', B', C' \in \mathfrak{A}$ .

**Proof.** First note that equation (16) is simply the restriction of equation (17) to disjoint sets, so that we immediately see that equation (16)  $\Leftarrow$  equation (17).

To prove the converse we consider sets  $A_1, A_2, A_3 \in \mathfrak{A}$  and break down their union  $\bigcup_{i=1}^3 A_i$  into disjoint components by defining

$$\begin{aligned} A_1 A_3 A_3 &= A_1 \cap A_2 \cap A_3 \\ \overline{A_i A_j} &= (A_i \cap A_j) \setminus (A_1 A_3 A_3) \\ \overline{A_i} &= A_i \setminus (\overline{A_i A_j} \sqcup \overline{A_i A_k} \sqcup A_1 A_3 A_3). \end{aligned} \quad (18)$$

Then

$$\bigcup_i A_i = \left( \bigsqcup_i \overline{A_i} \right) \sqcup \left( \bigsqcup_{i < j} \overline{A_i A_j} \right) \sqcup A_1 A_2 A_3.$$

Assuming equation (16), we can use this notation to decompose the terms in equation (17). The left-hand side becomes

$$\begin{aligned} \phi(A_1 + A_2 + A_3) &= \phi \left( \left( \bigsqcup_i \overline{A_i} \right) \sqcup A_1 A_2 A_3 \right) \\ &= \sum_{i < j} \phi(\overline{A_i} \sqcup \overline{A_j}) + \sum_i \phi(\overline{A_i} \sqcup A_1 A_2 A_3). \end{aligned}$$

Similarly turning our attention to the right-hand side, for  $i, j, k$  distinct we find that

$$\begin{aligned} \phi(A_i + A_j) &= \phi(\overline{A_i} \sqcup \overline{A_i A_k} \sqcup \overline{A_j} \sqcup \overline{A_j A_k}) \\ &= \phi(\overline{A_i} \sqcup \overline{A_j}) + \phi(\overline{A_i} \sqcup \overline{A_i A_k}) + \phi(\overline{A_j} \sqcup \overline{A_j A_k}) \\ &\quad + \phi(\overline{A_j} \sqcup \overline{A_i A_k}) + \phi(\overline{A_i} \sqcup \overline{A_j A_k}) + \phi(\overline{A_i A_k} \sqcup \overline{A_j A_k}) \end{aligned}$$

and

$$\begin{aligned} \phi(A_i) &= \phi(\overline{A_i} \sqcup \overline{A_i A_j} \sqcup \overline{A_i A_k} \sqcup A_1 A_2 A_3) \\ &= \phi(\overline{A_i} \sqcup \overline{A_i A_j}) + \phi(\overline{A_i} \sqcup \overline{A_i A_k}) + \phi(\overline{A_i} \sqcup A_1 A_2 A_3) \\ &\quad + \phi(\overline{A_i A_j} \sqcup \overline{A_i A_k}) + \phi(\overline{A_i A_j} \sqcup A_1 A_2 A_3) + \phi(\overline{A_i A_k} \sqcup A_1 A_2 A_3). \end{aligned}$$

Comparing these it is easy to see that the left- and right-hand sides of equation (17) are equal, hence the result.  $\square$

We now introduce some notation. First, given a histories theory  $(\Omega, \mathfrak{A}, \mu)$  and a co-event  $\phi \in \mathfrak{A}^*$  we define

$$Q_{ABC}(\phi) = \phi(A \sqcup B \sqcup C) + \phi(A \sqcup B) + \phi(B \sqcup C) + \phi(C \sqcup A) + \phi(A) + \phi(B) + \phi(C), \tag{19}$$

where  $A, B, C \in \mathfrak{A}$  are disjoint events. Then lemma 5 tells us that  $\phi \in \mathfrak{A}^*$  is quadratic if and only if  $Q_{ABC}(\phi) = 0$  for all disjoint  $A, B, C \in \mathfrak{A}$ .

Now though co-events take values in  $\mathbb{Z}_2$ , probability measures take values in  $\mathbb{R}$ . It will be useful to algebraically combine the images of (the values returned by) co-events and probability measures, which we can achieve by thinking of  $\mathbb{Z}_2$  as a subset of  $\mathbb{R}$ ; thus given a co-event  $\phi \in \mathfrak{A}^*$  we can construct the related map:

$$\begin{aligned} \tilde{\phi} : \mathfrak{A} &\rightarrow \mathbb{R} \\ \tilde{\phi}(A) &= \begin{cases} 0 & \phi(A) = 0 \\ 1 & \phi(A) = 1. \end{cases} \end{aligned}$$

Using this map we can define a real-valued analogue of the function  $Q_{ABC}$ :

$$\begin{aligned} R_{ABC} : \mathfrak{A}^* &\rightarrow \mathbb{R} \\ R_{ABC}(\phi) &= \tilde{\phi}(A \sqcup B \sqcup C) - \tilde{\phi}(A \sqcup B) - \tilde{\phi}(B \sqcup C) - \tilde{\phi}(C \sqcup A) + \tilde{\phi}(A) + \tilde{\phi}(B) + \tilde{\phi}(C), \end{aligned}$$

where  $A, B, C \in \mathfrak{A}$  are disjoint events. We can show that if  $\phi$  obeys the multiplicative rule then  $R_{ABC}$  is either 0 or 1, regardless of our choice of  $A, B, C$ .

**Lemma 6.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space and let  $\phi \in \mathfrak{A}^*$  be multiplicative. Then for all disjoint  $A_1, A_2, A_3 \in \mathfrak{A}$  we have

$$R_{A_1 A_2 A_3}(\phi) \in \{0, 1\}.$$

**Proof.** Since  $\phi$  is multiplicative, for any  $A \in \mathfrak{A}$  we know that  $\phi(A) = 1$  if and only if  $\phi^* \subset A$ . Then fixing our three events  $A_1, A_2, A_3$  we have four cases:

- (1)  $\phi^* \not\subset A_1 \sqcup A_2 \sqcup A_3$ ; then  $\phi(A_1 \sqcup A_2 \sqcup A_3) = \phi(A_i \sqcup A_j) = \phi(A_k) = 0$  for all  $i, j, k$ , so that  $R_{A_1 A_2 A_3}(\phi) = 0$ .
- (2)  $\phi^* \subset A_1 \sqcup A_2 \sqcup A_3, \phi^* \not\subset A_i \sqcup A_j$  for any  $i, j$ ; then  $\phi(A_1 \sqcup A_2 \sqcup A_3) = 1$  but  $\phi(A_i \sqcup A_j) = \phi(A_k) = 0$  for all  $i, j, k$ , so that  $R_{A_1 A_2 A_3}(\phi) = 1$ .
- (3)  $\phi^* \subset A_i \sqcup A_j$  for some  $i, j$ , but  $\phi^* \not\subset A_k$  for any  $k$ . Without loss of generality we can assume that  $i, j = 1, 2$ , then  $\phi(A_1 \sqcup A_2 \sqcup A_3) = 1, \phi(A_1 \sqcup A_2) = 1$  and  $\phi(A_k) = 0$  for all  $k$ . Further,  $\phi^* \not\subset A_2 \sqcup A_3$  otherwise  $\phi^* \in A_2$ , similarly  $\phi^* \not\subset A_3 \sqcup A_1$ , so  $\phi(A_2 \sqcup A_3) = \phi(A_3 \sqcup A_1) = 0$ . Therefore  $R_{A_1 A_2 A_3}(\phi) = 1 - 1 = 0$ .
- (4)  $\phi^* \subset A_i$  for some  $i$ ; without loss of generality assume that  $i = 1$ . Then  $\phi^* \not\subset A_2, A_3$ , since these sets are disjoint to  $A_1$ . Then  $\phi(A_k) = \delta_{1k}, \phi(A_1 \sqcup A_2) = \phi(A_1 \sqcup A_3) = 1, \phi(A_2 \sqcup A_3) = 0$ , and finally  $\phi(A_1 \sqcup A_2 \sqcup A_3) = 1$ . Thus  $R_{A_1 A_2 A_3}(\phi) = 1 - 1 - 1 + 1 = 0$ .

Hence the result. □

Finally, we can relate  $R_{ABC}$  and  $Q_{ABC}$ .

**Lemma 7.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space and let  $\phi \in \mathfrak{A}^*$  be a multiplicative co-event. Then if  $A, B, C \in \mathfrak{A}$  are disjoint

$$R_{ABC}(\phi) = 0 \Rightarrow Q_{ABC}(\phi) = 0.$$

**Proof.** First, denote by  $X \text{ Mod } 2$  the  $\mathbb{Z}_2$  element corresponding to the integer  $X \text{ mod } 2$ , where  $X$  is an integer in  $\mathbb{R}$ . Then  $\phi(A) = \tilde{\phi}(A) \text{ Mod } 2$  for any  $A \in \mathfrak{A}$ , so noting that  $+$  and  $-$  are equivalent in  $\mathbb{Z}_2$ , for any disjoint  $A, B, C \in \mathfrak{A}$  we have

$$\begin{aligned} R_{ABC}(\phi) \text{ Mod } 2 &= \tilde{\phi}(A \sqcup B \sqcup C) \text{ Mod } 2 - \tilde{\phi}(A \sqcup B) \text{ Mod } 2 \\ &\quad - \tilde{\phi}(B \sqcup C) \text{ Mod } 2 - \tilde{\phi}(C \sqcup A) \text{ Mod } 2 \\ &\quad + \tilde{\phi}(A) \text{ Mod } 2 + \tilde{\phi}(B) \text{ Mod } 2 + \tilde{\phi}(C) \text{ Mod } 2, \\ &= \phi(A \sqcup B \sqcup C) + \phi(A \sqcup B) + \phi(B \sqcup C) + \phi(C \sqcup A) \\ &\quad + \phi(A) + \phi(B) + \phi(C), \\ &= Q_{ABC}(\phi). \end{aligned}$$

Hence the result. □

We are now in a position to prove the theorem.

**Proof of theorem 2.** By assumption our probability measure  $\mathbf{P}_S$  obeys equation (14), then the sum rule obeyed by the quantum measure (equation (2)) gives us:

$$\begin{aligned} \mathbf{P}_S(\{\phi|\phi(A \sqcup B \sqcup C) = 1\}) &= \mathbf{P}_S(\{\phi|\phi(A \sqcup B) = 1\}) + \mathbf{P}_S(\{\phi|\phi(B \sqcup C) = 1\}) \\ &\quad + \mathbf{P}_S(\{\phi|\phi(C \sqcup A) = 1\}) - \mathbf{P}_S(\{\phi|\phi(A) = 1\}) \\ &\quad - \mathbf{P}_S(\{\phi|\phi(B) = 1\}) - \mathbf{P}_S(\{\phi|\phi(C) = 1\}), \end{aligned} \tag{20}$$

for all disjoint  $A, B, C \in \mathfrak{A}$ . Now because  $\mathbf{P}_S$  is a probability measure we can use the Kolmogorov sum rule to decompose its valuation on any set into a sum of its valuations on the elements of that set. Thus if we define  $p_\phi = \mathbf{P}_S(\{\phi\})$  we get

$$\begin{aligned} \mathbf{P}_S(\{\phi | \phi(A) = 1\}) &= \sum_{\phi(A)=1} p_\phi \\ &= \sum_{\phi \in S} p_\phi \tilde{\phi}(A), \end{aligned} \tag{21}$$

for all  $A \in \mathfrak{A}$ . Putting this back into equation (20) we get

$$\begin{aligned} \sum_{\phi \in S} p_\phi \tilde{\phi}(A \sqcup B \sqcup C) &= \sum_{\phi \in S} p_\phi \tilde{\phi}(A \sqcup B) + \sum_{\phi \in S} p_\phi \tilde{\phi}(B \sqcup C) + \sum_{\phi \in S} p_\phi \tilde{\phi}(C \sqcup A) \\ &\quad - \sum_{\phi \in S} p_\phi \tilde{\phi}(A) - \sum_{\phi \in S} p_\phi \tilde{\phi}(B) - \sum_{\phi \in S} p_\phi \tilde{\phi}(C), \end{aligned} \tag{22}$$

for all disjoint  $A, B, C \in \mathfrak{A}$ , which we can rewrite as

$$0 = \sum_{\phi \in S} p_\phi R_{ABC}(\phi) \quad \forall \text{ disjoint } A, B, C \in \mathfrak{A}. \tag{23}$$

Now fix  $A, B, C$  and let  $Z_{ABC} \subset S$  be the set of co-events  $\phi$  such that  $R_{ABC} = 0$ . Further denote by  $\bar{Z}_{ABC}$  the set of co-events such that  $R_{ABC} \neq 0$ , so that  $S = Z_{ABC} \sqcup \bar{Z}_{ABC}$ . Now by assumption every  $\phi \in S$  is multiplicative, so by lemma 6  $\phi \in \bar{Z}_{ABC} \Rightarrow R_{ABC}(\phi) = 1$ . Then equation (23) gives us

$$\begin{aligned} 0 &= \sum_{\phi \in Z_{ABC}} p_\phi R_{ABC}(\phi) + \sum_{\phi \in \bar{Z}_{ABC}} p_\phi R_{ABC}(\phi), \\ &= \sum_{\phi \in \bar{Z}_{ABC}} p_\phi. \end{aligned} \tag{24}$$

Noting that  $p_\phi \geq 0$ , this means  $p_\phi = 0$  for all  $\phi \in \bar{Z}_{ABC}$ . Since  $A, B, C$  are arbitrary disjoint sets, we conclude that  $p_\phi = 0$  unless  $R_{ABC}(\phi) = 0$  for all disjoint  $A, B, C$ . But by lemma 7  $R_{ABC}(\phi) = 0$  implies that  $Q_{ABC}(\phi) = 0$ , which by lemma 5 implies that  $\phi$  is quadratic. This completes our proof.  $\square$

### 3. Approximate preclusion

Rather than tackling the whole of the dynamics, an alternate approach is to focus directly on the predictive content of a histories theory. Thus instead of, or as a step towards, expressing  $\mu$  in terms of the theory's allowed co-events, we would primarily be concerned with the probability statements that are the (experimentally falsifiable) predictions of our theory. As before, we seek to be able to 'start' with a set of potentially real co-events, and to make our experimental predictions in terms of these co-events and structures defined in terms of them.

Our strategy will be to examine more closely the meaning (or interpretation) of probability *as applied to (experimentally) falsifiable predictions*; thus we look to understand a dynamical statement  $\mu(A) = q$  or  $\mathbf{P}(B) = p$  through its (experimentally) falsifiable implications. We will begin by focusing on classical theories, where we can draw from a long history of thought upon the matter to pick a suitable interpretation of probability; we then seek to generalize this interpretation in terms of the potentially real co-events in such a manner that can be extended to quantum histories theories.

This strategy requires a somewhat rigorous approach to the interpretation of probability, for we are not simply attempting to justify pre-existing statistical practice, but to construct a predictive framework for a new ontology. In a world of co-events, what do we mean by  $\mu(A) = q$ , or even  $\mathbf{P}(A) = p$ ? Since co-events have been connected to the measure through the preclusion of null sets, a natural approach is to seek an interpretation of probability based on this notion. We begin with a simple example that will guide us through this process.

### 3.1. The classical coin

Consider a coin described by classical stochastic dynamics. We assume that if we throw the coin we will either have a ‘heads’ or a ‘tails’ result, we will further assume that all such throws are in some sense ‘equivalent’ so that the probability of ‘heads’ is always  $p$  and the probability of ‘tails’ always  $1 - p$ . But what exactly do we mean by this?

One approach to its interpretation is to emphasize the role of probability in representing our knowledge or expectations about a system, so in this case we would for example say that  $p = 1/2$  if we had no information or view on the outcome of a coin toss [5, 6, 33]. An alternate approach focuses on the practical testing of such assertions, which takes place through a set of repeated, independent trials [15, 28]. In the latter framework a probability assertion  $\mathbf{P}(\text{heads}) = 1/2$  is a statement, not about a single trial, but about a theoretical ensemble of trials, or a prediction regarding a hypothesized set of future trials. Although the former approach is valid and useful, for example in the theory of decision making [1], in light of our difficulties in constructing a probability measure on the space of co-events (section 2) we will focus on the latter approach, which is more in accordance with the predictive and experimental nature of our field, and as we shall see lends itself naturally to the concept of preclusion. We begin by discussing how we might test an assertion  $\mathbf{P}(\text{heads}) = p$  in practice.

We test a theory, or a statement within a theory, by testing its experimentally falsifiable predictions. But when we believe there is one realized outcome (the actually real co-event, or the real history if we using the naive interpretation) we face the problem that unless  $p \in \{0, 1\}$ , the statement  $\mathbf{P}(\text{heads}) = p$  cannot be falsified by either outcome ‘heads’ or ‘tails’. We can however make falsifiable statements about sequences of trials.

The frequentist approach, as articulated by Friedman [15], interprets a probability statement  $\mathbf{P}(\text{heads}) = p$  by identifying it with the asymptotic relative frequency of the outcome ‘heads’ in an infinite sequence of repeated trials; under the requirement that the sequence of trials conforms to a ‘randomness’ criterion [15, 44]. Indeed, if we had a theoretical ensemble of infinitely many trials, the proportion of the trials resulting in heads would be exactly<sup>6</sup>  $p$ ; alternatively the assertion that ‘the proportion of heads is not  $p$ ’ is precluded by the measure on the ensemble, and herein lies our link with co-events.

However asymptotic properties of a sequence cannot be determined using a finite number of elements of the sequence and an infinite ensemble of trials is not realizable in terms of an experiment, whereas we seek to phrase probability in terms of experimentally falsifiable predictions. Further, the co-event is meant to be ‘real’, so it is not meaningful to say it precludes an imagined event in a theoretical ensemble that is not part of the ‘actual’ event algebra on which the co-event is defined.

We are thus pushed into defining probabilities based upon finite trials, and will focus on one particular technique of testing probability assertions through a sequence of repeated trials; the standard statistical technique of hypothesis testing [43]. We present a simple version of hypothesis testing that will suffice for the purposes of this paper.

<sup>6</sup> More precisely it would be  $p$  with probability 1.

We begin with our assertion, or *hypothesis*,  $\mathbf{P}(\text{heads}) = p$  and repeat our trial  $n$  times, resulting in a history  $\gamma$  consisting of  $n$  ordered ‘heads’ or ‘tails’ outcomes (we will ignore the issue of the ‘randomness’ of this sequence for now). We will denote by  $H$  (or  $H(\gamma)$ ) the number of these trials that result in a ‘heads’ outcome, so that the proportion of heads is  $H/n$ ; to compare this ‘statistic’ [43] with the implications of our hypothesis we assume the product measure (which we shall also denote by  $\mathbf{P}$ ) on the sequence of trials. Using  $N_H$  to denote the event that the number of heads in the realized history is  $H$ , we define the *cumulative probability* of  $N_H$  to be

$$\mathbb{CP}(N_H) = \sum_{m=1}^H \mathbf{P}(N_m). \quad (25)$$

Thus  $\mathbb{CP}(N_k)$  is the probability that the number of heads in the realized history is less than or equal to  $k$ . For the purposes of this paper we will say that we *reject the hypothesis at the  $\epsilon$  level* if  $\mathbb{CP}(N_{H(\gamma)}) < \epsilon$ , where  $0 < \epsilon \leq 1$  (typically  $\epsilon \ll 1$ ). If  $\gamma$  is such that  $\mathbb{CP}(N_{H(\gamma)}) \geq \epsilon$  we *fail to reject the hypothesis at the  $\epsilon$  level*<sup>7</sup>. Note crucially that we have implicitly *chosen* to organize the potential outcomes according to the number (or equivalently the proportion) of head outcomes, based on our hypothesis; the importance of this choice will become clear below.

This technique (or a more sophisticated version thereof [43]) can be used in the testing of scientific theories;  $\epsilon$  is then chosen to represent the ‘degree of certainty’ we wish to test our theory to. However, assuming  $p \notin \{0, 1\}$  no history (no sequence of outcomes) is actually precluded by our hypothesis; strictly speaking the hypothesis has no falsifiable implications for the realized outcomes, thus no realized history can falsify the hypothesis.

To justify our hypothesis testing technique and to place it on a more rigorous footing, we turn to *Cournot’s principle* [4, 35]; which has long been used to connect ‘the mathematical formalism of probability to the empirical world’ [16].

### 3.2. Cournot’s principle

Crudely speaking, in one way or another Cournot’s principle ‘rules out’ events of small probability. This concept was certainly known to Bernoulli, who asserted that high probability can be treated as a ‘moral certainty’ [2]. Cournot himself made the connection with physics, arguing that events of small probability may be mathematically possible but ‘physically impossible’ [4]. Among many other references, the principle is used or alluded to by Levy [29, 30], Markov [35], Borel [3] and Kolmogorov [28], under the name ‘Principle B’. More recently it has been applied by Goldstein *et al* to statistical mechanics [20] and Bohmian mechanics [11] and by Galvan to quantum mechanics [16]. For a more detailed discussion of the history of Cournot’s principle see [35].

However it was the French mathematician Maurice Frechet who coined the term ‘principe de Cournot’, which has come into English as ‘Cournot’s principle’, or the ‘Cournot principle’ [35]. Frechet distinguished between strong and weak forms of Cournot’s principle [14]; Shafer describes these two formulations as follows [35]:

The strong form refers to an event of small or zero probability that we single out in advance of a single trial: it says the event will not happen on that trial. The weak form says that an event with very small probability will happen very rarely in repeated trials. Some authors, including Levy, Borel and Kolmogorov, adopted the

<sup>7</sup> This is essentially a ‘one-tailed test’; while both one- and two-tailed tests would be equally valid in this context (though they would correspond to different values of  $\epsilon$ ), we will find the one-tailed test simpler to analyse.

strong principle. Others, including Chuprov and Frechet himself, preferred the weak principle.

Kolmogorov's statement of strong Cournot (his 'Principle B') as concerns the probability  $\mathbf{P}(E)$  of an event  $E$  occurring in an experiment  $C$  is [28] (as translated by Sahfer and Vovk [36]):

If  $\mathbf{P}(E)$  is very small, one can be practically certain that when  $C$  is carried out only once, the event  $E$  will not occur at all.

These statements contain some ambiguities, for example, in Kolmogorov's 'Principle B', what exactly do we mean by 'practically certain'? To apply Cournot's principle to co-events, we will have to be more rigorous. We begin by giving a more precise specification of the strong Cournot principle:

*Literal strong Cournot.* Events of probability less than  $\epsilon \ll 1$  do not occur, for some  $\epsilon > 0$ .

Where there is no danger of confusion with other variations of the strong Cournot principle, we will refer to literal strong Cournot simply as the 'strong Cournot principle' or 'strong Cournot'. This literal specification of the strong Cournot principle is in some ways a 'maximal' interpretation of the Cournot principle, and as such may appear to be 'neat' or 'simple'; it would also allow us to deal with probability 'objectively', and in the absence of observation, by grounding it in the ontology. However it is problematic in the case of classical stochastic theories. Firstly what do we mean by a small probability, are we to take  $\epsilon$  as a constant of nature? Secondly there is nothing to prevent the probability of every single history in the repeated trial being much smaller than the probability of the events we wish to preclude, in which case a literal application of strong Cournot would rule out all single histories, and thus any reality. In fact an example of such a situation is provided by our coin, given a sufficiently large number of trials.

Since literal strong Cournot is not appropriate for classical theories, we instead present a 'minimal' version of weak Cournot that allows us to perform the hypothesis testing described above, but little else.

*Operational weak Cournot.* In a repeated trial, an event of probability less than  $\epsilon \ll 1$ , identified ahead of the repeated trial, will not occur, for some  $\epsilon > 0$ . Only a single event can be considered for a given repeated trial.

Thus we can only make one prediction for each repeated trial. Where there is no danger of confusion with other variations of the weak Cournot principle, we will refer to operational weak Cournot simply as the 'weak Cournot principle' or 'weak Cournot'. Note that our reference to repeated trials is essentially semantic; a single trial can be considered as a special, ' $n = 1$ ', case of a repeated trial, and a general ' $n$  times' repeated trial can be considered as a single trial in which sequences of length  $n$  are thought of as single outcomes. We insist on the term 'repeated trial' to emphasize the link between operational weak Cournot and hypothesis testing<sup>8</sup>.

Operational weak Cournot has the advantage that it fits in well with our actual methods of falsifying theories, such as the statistical hypothesis testing discussed in section 3.1. When we make the hypothesis  $\mathbf{P}(\text{heads}) = p$ , there may be certain outcomes that would convince us that this hypothesis has been falsified; for example we may reconsider the assertion  $\mathbf{P}(\text{heads}) = 1/2$

<sup>8</sup> Indeed, we regard operational weak Cournot as a variation on Shafer's 'weak principle' precisely because of the shared emphasis on repeated trials. We note however, that in its insistence that certain 'low probability' events *will not* occur (rather than *occurring rarely*, operational weak Cournot bears some resemblance to Shafer's strong principle'.

were we to toss our coin a million times only to find a heads outcome for every toss. We can turn this around by saying that *those outcomes that would falsify a theory are precluded by it*. Of course there remains some ambiguity in the choice of  $\epsilon$ , however this ambiguity is inherent and has not been introduced by our adoption of weak Cournot; whatever interpretation of probability we adopt we would have to choose  $\epsilon$  we use to falsify our theories (of course this  $\epsilon$  need not be unique). In a sense this identification takes the arbitrariness of our choice of  $\epsilon$  out of our theory and into the ‘meta’ level on which we compare and reject theories. The ‘meta’ level is always present and by using it to give us  $\epsilon$  we have avoided the addition of ‘new’ ambiguity.

Applying this to our coin (section 3.1), we start with our hypothesis  $\mathbf{P}(\text{heads}) = p$ , then in the context of a ‘possible’ experiment (or  $n$ -fold repeated trial) we *choose* to consider the proportion of heads that would be realized in such an experiment, and thus single out the event<sup>9</sup>  $L$  that the realized history  $\gamma$  will obey  $\mathbb{C}\mathbb{P}(H(\gamma)) < \epsilon$ . Since the probability of this event is less than  $\epsilon$ , using operational weak Cournot we are justified in asserting that it will not occur (we are assuming the hypothesis). The non-occurrence of  $L$  then becomes a falsifiable prediction of our hypothesis; and once we perform the experiment, the occurrence of the event  $L$  would falsify the hypothesis.

However, given the realized history  $\gamma$  (once the experiment has been carried out), if  $n$  is sufficiently large we will always be able to find an event that has occurred (i.e., it contains the realized history) yet is assigned probability  $< \epsilon$  by the hypothesis; for example, as mentioned above the event that the realized history  $\gamma$  will occur will itself be of probability  $< \epsilon$  for sufficiently large  $n$ . Because of this, to rule out the problems faced by literal strong Cournot we have phrased operational weak Cournot so as to allow conclusions to be drawn concerning only the single<sup>10</sup> falsifiable prediction (implied by our hypothesis) singled out before the experiment and being tested by it.

In this way operational weak Cournot provides us with a practical method of relating probability statements to experimental measurements, though we have had to adopt a ‘minimalist’ and perhaps ‘restrictive’ interpretation of probability, which might for example restrict our ability to treat probability statements as logical predicates [5]. However in the classical theory this has no impact on either our ontology or the ‘meaning’ we attribute to dynamics, for we typically assume a deeper deterministic theory operating at a more fundamental level. This will not hold in the quantum case, which therefore will require us to take our chosen interpretation of probability more seriously.

### 3.3. Approximate co-events

In one way or another Cournot’s principle rules out events of small probability. Because this is a straightforward generalization of the concept of preclusion, which rules out null sets, we can easily express it in terms of co-events by introducing the concept of *approximate preclusion*. For the sake of clarity we will henceforth refer to preclusion itself as *exact preclusion*, thus an exactly preclusive co-event (which we will sometimes shorten to an *exact co-event*) obeys

$$\mu(A) = 0 \Rightarrow \phi(A) = 0$$

whilst an *approximately preclusive co-event* is given by

<sup>9</sup> This is an event in the event algebra related to the sample space of all possible sequences of outcomes in our repeated trial.

<sup>10</sup> Each such prediction is an element of the event algebra related to the sample space of all possible sequences of outcomes in our repeated trial. Thus, if we wish to make multiple predictions in the context of a single experiment, we can combine them using the logical (Boolean) operations in the event algebra to yield a single event.

**Definition 6.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory. Given  $\epsilon > 0$  a co-event  $\phi_\epsilon$  is approximately preclusive at the  $\epsilon$  level if

$$\mu(A) < \epsilon \Rightarrow \phi_\epsilon(A) = 0.$$

We say that  $\phi_\epsilon$  is an approximately preclusive co-event, or simply an approximate co-event.

Note that this definition holds for a general histories theory, not simply in the classical case, and thus might be used to generalize Cournot's principle.

Though the introduction of approximate co-events is a fundamental change in the theory, its implications for the 'internal structure' of the 'allowed' co-events is less than radical. To be precise, we have effectively altered the precluded events, including 'almost null' as well as null sets. We will refer to events of measure less than  $\epsilon$  as 'approximately precluded' or ' $\epsilon$ -null' sets, with ' $\epsilon$ -negligible' sets defined similarly. However other than this we have made no changes, and have not altered the algebraic structure of the co-events, and so can define many of the same concepts and prove many of the same theorems that we could for exactly preclusive co-events. In particular we can define multiplicativity, domination and primitivity based upon approximate preclusion in exactly the same way we defined these concepts for exact preclusion, leading to a notion of 'approximate co-event schemes' and in particular an approximate multiplicative scheme, which we shall henceforth assume. Further, we can prove 'approximate co-event versions' of many of the results we have established for the exact multiplicative scheme, for example the following lemma, which shall make use of below, is analogous to lemma 3.

**Lemma 8.** Let  $(\Omega, \mathfrak{A}, \mu)$  be a history theory with a finite sample space and let  $A \in \mathfrak{A}$  be non-negligible (i.e.  $A$  is not a subset of a null set). Then there exists a primitive approximately preclusive multiplicative co-event  $\phi_\epsilon$  such that  $\phi_\epsilon(A) = 1$ .

**Proof.** See proof of lemma 3. □

In a classical theory, how we interpret  $\phi_\epsilon$  depends on the version of Cournot's principle we are following. Weak Cournot means that  $\phi_\epsilon$  are essentially theoretical tools, used to phrase (experimentally falsifiable) predictions in terms of preclusion. In this case  $\epsilon$  will be taken from our 'meta' level choice of  $\epsilon$  used to falsify a given theory, and may not be the same for every system. On the other hand, strong Cournot means that  $\phi_\epsilon$  are the potential realities, in which case we may consider  $\epsilon$  as a constant of nature. Note that our basic ontology remains unchanged, the 'actually real' co-event remains our description of reality and its internal structure (or logic) is still given by the multiplicative rule<sup>11</sup>. What we have done is to alter the set of potential realities *given* a measure, thus essentially we have altered the role and meaning of the measure.

Moving from classical stochastic theories to quantum mechanics and the multiplicative scheme, we are no longer treating a single history as real<sup>12</sup> so our objections to strong Cournot may no longer be relevant. This leads us to question whether it may be possible to achieve strong Cournot in the context of quantum measure theory and to take  $\phi_\epsilon$  literally. One advantage of such an approach lies in our practise of treating some events of small probability as null. In section 1.5, we raised concerns regarding the preclusion of 'approximately null' sets, arguing that true null sets are rare; in fact, due to experimental inaccuracies the events we are characterizing as null will in general be only of small probability. Approximate preclusion

<sup>11</sup> We assume that the multiplicative scheme holds for all histories theories, in a classical theory it 'happens' to coincide with the classical scheme.

<sup>12</sup> More precisely we are no longer expecting our primitive multiplicative co-events to be classical.

is tailor made to address such concerns, in particular if we are able to adopt strong Cournot we can preclude such sets directly.

If strong Cournot, as expressed by approximate co-event, is to be applicable to a general quantum measure theory, it must certainly make sense in the context of classical theories. In the following section, we therefore use approximate preclusion to explore the application of strong Cournot to hypothesis testing in a classical theory; returning to the example of our simple coin (section 3.1). In what follows we assume that all co-events are multiplicative.

### 3.4. Can we achieve strong Cournot?

*3.4.1. Approximate preclusion for a coin.* As before we assume that if we throw the coin we will either have a ‘heads’ (*h*) or a ‘tails’ (*t*) result, we will further assume that all such throws are in some sense ‘equivalent’ so that the probability of ‘heads’ is always  $p$  and the probability of tails always  $1 - p$ . Thus in the histories formalism, a single throw corresponds to the sample space:

$$\Omega = \{h, t\}. \tag{26}$$

In the language of our hypothesis testing technique (section 3.1), our hypothesis is that  $\mathbf{P}(\{h\}) = p$ , so that the classical measure  $\mathbf{P}$  is

$$\mathbf{P}(\{h\}) = p, \quad \mathbf{P}(\{t\}) = 1 - p, \quad \mathbf{P}(\{h, t\}) = 1. \tag{27}$$

An ‘experiment’ consisting of  $n$  trials (which we assume to be independent) has the sample space

$$\Omega = \{h, t\}^n, \tag{28}$$

of histories (ordered sequences of outcomes)  $\gamma = a_1, \dots, a_n$ , where  $a_i \in \{h, t\}$ . The corresponding event algebra is  $\mathcal{P}\Omega$  and since we have assumed that our trials are independent we use the product measure, which by abuse of notation we shall also denote by  $\mathbf{P}$ .

Now we have previously denoted the number of ‘heads’ outcomes in a history  $\gamma$  by  $H(\gamma)$ , so that the proportion of ‘heads’ is  $H(\gamma)/n$ . Then we have

$$\mathbf{P}(\{\gamma\}) = p^{H(\gamma)}(1 - p)^{n-H(\gamma)}. \tag{29}$$

We are less interested in the individual histories than in the proportion of heads, which, as discussed above, should reflect the underlying probability  $p$  of a heads outcome in a single trial. We will denote by  $N_H$  the event that the number of heads realized in the sequence is  $H$ . Then we have

$$\mathbf{P}(N_H) = \binom{n}{H} p^H (1 - p)^{n-H}. \tag{30}$$

Further, we will label by  $L_H$  and  $G_H$  the events that the number of heads in the realized sequence is less than or equal to or greater than  $H$ , respectively. Then we get

$$\begin{aligned} \mathbf{P}(L_H) &= \sum_{m \leq H} \mathbf{P}(m), \\ &= \mathbb{C}\mathbf{P}(N_H). \end{aligned} \tag{31}$$

Thus if we performed such an experiment and realized a proportion of heads corresponding to a small  $\mathbf{P}(L_H) < \epsilon$  we would reject our hypothesis  $\mathbf{P}(h) = p$  (at the  $\epsilon$  level). Turning this around, given the assumption  $\mathbf{P}(h) = p$  we wish to preclude all events  $L_H$  (and  $N_H$ ) with  $\mathbf{P}(L_H) < \epsilon$  for some small  $\epsilon$ .

We can make this more precise; since the measure  $\mathbf{P}$  is classical and non-zero everywhere, and since  $L_m$  is a proper subset of  $L_{m+1}$ , we can see that  $\mathbf{P}(L_m)$  is monotonic in  $m$ . Thus

defining  $H_\epsilon$  as the greatest  $H$  such that  $\mathbf{P}(L_H) < \epsilon$ , we have  $\mathbf{P}(L_{H_\epsilon+m}) \geq \epsilon \forall m > 0$ ; in particular  $\mathbf{P}(L_{H_\epsilon}) < \epsilon \leq \mathbf{P}(L_{H_\epsilon+1})$ . Thus, following the argument of section 3.1 we would like to say that  $L_{H_\epsilon}$  is precluded while  $L_{H_\epsilon+1}$  is not. Since  $\mathbf{P}(L_H) > 0$  in all cases, we turn to approximate preclusion.

Adopting strong Cournot, we fix  $\epsilon$  and assume that events of measure less than  $\epsilon$  never occur. Then we can immediately rule out  $L_{H_\epsilon}$ , indeed  $\phi_\epsilon(L_{H_\epsilon}) = 0$  for all approximately preclusive co-events  $\phi_\epsilon$ . However, this does not hold for  $L_{H_\epsilon+1}$ , indeed since  $\mathbf{P}(L_{H_\epsilon+1}) > \epsilon$  by lemma 8 there exists a primitive approximate co-event  $\phi_{H_\epsilon}$  mapping it to unity. In other words,

$$\phi_{H_\epsilon}(L_{H_\epsilon}) = 0, \quad \phi_{H_\epsilon}(L_{H_\epsilon+1}) = 1. \tag{32}$$

Thus every potential reality precludes  $L_{H_\epsilon}$ , but not every potential reality precludes  $L_{H_\epsilon+1}$ . This is enough to allow us to use the hypothesis testing technique outlined in section 3.1, thus enabling us to deal with (experimentally falsifiable) predictions in terms of our approximate co-events. Since approximate preclusion is defined in terms of a general histories theory, this seems to raise the possibility of a predictive approximate co-event framework for a general quantum measure theory. However, though widening the scope of preclusion by moving from exact to approximate co-events has introduced useful features to our theory, we must make sure that it has not also introduced problems.

*3.4.2. The failure of strong Cournot.* The above construction seems promising, it seems that our shift from exact to approximate preclusion has succeeded in introducing new and useful features to co-event theory; however, we must check that it has not inadvertently introduced new problems as well. In dealing with exactly preclusive co-events we can show that (under certain conditions) classical measures always imply classical outcomes meaning that the co-events will necessarily behave classically (see section 1.4.3); do the approximate co-events always ‘behave’? Such concerns lead to three objections to the construction we outlined above.

- (1) *The status of single histories.* The first question to be raised regards the status of single histories. In a system obeying non-classical dynamics we may be happy with single histories not being realized, however in this system with its classical dynamics we could still find all single histories to be ruled out for a sufficiently large number of trials. This raises interpretational issues, for example if we take  $p = 1/2$ ,  $\epsilon = 10^{-3}$  and  $n = 10$  we find all single histories ruled out. However if  $n = 9$  all single histories are allowed. We have several problems here, firstly we may not be comfortable with the non-classical behavior of the  $n = 10$  system. Secondly there is the question of the value of  $\epsilon$ , which leads to classical outcomes in the  $n = 9$  case but not the  $n = 10$  case. Finally it seems odd that one additional trial will ‘disallow’ classical behaviour, particularly given that a single trial behaves classically in and of itself.

Note that we do not have a causality problem here; although in the  $n = 10$  system no single history is realized this will not become apparent in the first trial. The two outcomes of the first trial can be thought of as a coarse-grained partition of the  $n = 10$  sample space and as such each outcome has probability greater than  $\epsilon$ . This will also hold in general, any single history in the  $n = 10 - m$  system (where  $1 \leq m \leq 9$ ) corresponds to a coarse-grained event in the  $n = 10$  system that has probability greater than  $\epsilon$  and so is not precluded.

Further note that this problem of systems that behave classically becoming non-classical in a repeated trial can occur in the histories framework without the use of co-events. We could take the view that no system is truly classical, but rather that at

the level of observable events we have an emergent classicality based on environmental decoherence. Adopting this view, observed ‘classical’ behaviour is due to approximate decoherence [37, 8], and repeated trials of an approximately decohering system (such as a ‘quantum coin’ with interference  $\epsilon$  between heads and tails) may lead to non-classical outcomes, notably the reappearance of quantum coherence [37].

- (2) *The problems of assigning ontology to the approximate co-events.* The second question to be raised concerns the range of possible co-events ‘allowed’ by approximate preclusion. We may be able to construct allowed co-events  $\phi_\epsilon$  such that some observable questions have no definite answer. In other words our framework allows potential realities that if realized would imply observable anhomomorphisms.

For example,  $\{L_{H_\epsilon}, G_{H_\epsilon}\}$  is a partition of  $\Omega$ , both elements of which are observable (since all events are observable in this classical system). Then using lemma 8 there will be at least one primitive approximate co-event  $\phi_{G_\epsilon}$  mapping  $G_{H_\epsilon}$  to 1, and  $\phi_{G_\epsilon}$  will map  $L_{H_\epsilon}$  to 0 (as do all approximate co-events) because  $\mathbf{P}(L_{H_\epsilon}) < \epsilon$ . However, although we can ‘cherry-pick’ the co-event  $\phi_{G_\epsilon}$  to treat the partition  $\{G_{H_\epsilon}, L_{H_\epsilon}\}$  classically, we do not know if  $\phi_{H_\epsilon}$  will map  $G_\epsilon$  to one and so treat this partition classically. More generally we do not have any guarantee that all the allowed co-events will treat this partition classically; in fact, we can find primitive approximate co-events that map both elements of the observable partition  $\{G_{H_\epsilon}, L_{H_\epsilon}\}$  to zero.

We can find subsets of histories  $T \subset L_{H_\epsilon}$  and  $P \subset G_{H_\epsilon}$  such that  $\mathbf{P}(T), \mathbf{P}(P) < \epsilon$  but  $\mathbf{P}(T \sqcup P) > \epsilon$ . Now  $\mathbf{P}(T \sqcup P) > \epsilon$  means that by lemma 8 some subset  $C$  of  $T \sqcup P$  will be the dual of a primitive approximate co-event  $C^*$ . Further both  $T$  and  $P$  are approximately precluded so  $C$  cannot be a subset of either; therefore  $C$  has non-empty intersection with both  $L_{H_\epsilon}$  and  $G_{H_\epsilon}$ , which means

$$C^*(L_{H_\epsilon}) = 0 \quad C^*(G_{H_\epsilon}) = 0. \tag{33}$$

For an explicit example consider the case  $p = 1/2, n = 10^3$  and  $\epsilon = 10^{-3}$ , then  $H_\epsilon = 450$  since  $\mathbf{P}(L_{450}) < \epsilon$  whereas  $\mathbf{P}(L_{451}) \geq \epsilon$ . Now  $L_{451} = L_{450} \sqcup N_{451}$ , so there is some (not necessarily unique) subset  $S \subseteq N_{451}$  such that  $\mathbf{P}(S \sqcup L_{450}) \geq \epsilon$  whereas  $\mathbf{P}((S \setminus \{\gamma\}) \sqcup L_{450}) < \epsilon$  for any  $\gamma \in S$ . We can think of constructing  $S$  by adding fine-grained histories from  $N_{451}$  to  $L_{450}$  one by one until the measure is greater than  $\epsilon$ ; thus we cannot reduce the set  $S \sqcup L_{450}$  without its measure falling below  $\epsilon$ . In fact, because  $p = 1/2$  (and the measure is classical) every single history  $\gamma$  contributes the same amount,  $p^n = 2^{-10^3}$ , to the probability of any event containing  $\gamma$ , so that  $\mathbf{P}(S) = p^n |S|$ . Further, since the measure is classical we know that  $\epsilon \leq \mathbf{P}(S) + \mathbf{P}(L_{450}) < \epsilon + p^n$ . Therefore,

$$p^{-n}(\epsilon - \mathbf{P}(L_{450})) \leq |S| < p^{-n}(\epsilon - \mathbf{P}(L_{450})) + 1, \tag{34}$$

so that  $S$  could be any subset of  $N_{451}$  of cardinality:

$$|S| = \text{Int}\left(\frac{\epsilon - \mathbf{P}(L_{450})}{p^n}\right) \approx 1.4 \times 10^{297}, \tag{35}$$

where  $\text{Int}(x)$  denotes the least integer which is greater than or equal to  $x$ . Then the set  $S \sqcup L_{450}$  has measure greater than or equal to  $\epsilon$ , and so is not a subset of any  $\epsilon$ -null set (since the measure is classical). Further, we cannot reduce our set without the remainder being  $\epsilon$ -negligible. Therefore  $S \sqcup L_{450}$  is the base of an approximate co-event  $(S \sqcup L_{450})^*$  that maps both  $G_{H_\epsilon}$  and  $L_{H_\epsilon}$  to zero.

- (3) *The inconsistency of multiple partitions.* Finally, in the examples we have looked at so far, we considered a single partition  $\Omega = G_{H_\epsilon} \sqcup L_{H_\epsilon}$  which was treated non-classically

by some allowed co-events. Intuitively, both elements of this partition correspond to ‘meaningful’ propositions that the proportion of heads in the observed history is less than or equal to  $(L_{H_\epsilon})$ , or greater than  $(G_{H_\epsilon})$ ,  $H_\epsilon/n$ . However, we did not attach any particular ‘meaning’ to our ‘problem co-event’  $C^*$ , indeed as a proposition the related event  $C$  is ‘pathological’ in that it would be difficult to express it as an ‘English sentence’ as we were able to do for  $L_{H_\epsilon}$  and  $G_{H_\epsilon}$  above. This might lead one to speculate that we could perhaps focus on partitions and co-events corresponding to ‘good’, ‘meaningful’ or ‘useful’ propositions, whilst ignoring partitions and co-events corresponding to ‘bad’ or ‘pathological’ propositions; indeed we were able to find the ‘good’ approximate co-event  $\phi_{G_\epsilon}$  that was classical on the ‘good’ partition  $\Omega = G_{H_\epsilon} \sqcup L_{H_\epsilon}$ .

However the situation is not so simple; in the above we focused on a single partition, whereas there are many ‘good’ partitions, each with its associated ‘good’ co-events. Unfortunately the co-events that treat one partition classically will not in general be classical on another partition, as the following example shows.

When the number of trials of our coin is even,  $n = 2m$ , we can partition the sample space  $\Omega$  into even and odd coarse grainings as follows. Given a single history  $\gamma = a_1, a_2, \dots, a_{2m}$  (where  $a_i \in \{h, t\}$ ) we can form the *even* and *odd* histories  $E(\gamma) = a_2, a_4, a_6, \dots, a_{2m}$  and  $O(\gamma) = a_1, a_3, a_5, \dots, a_{2m-1}$ . Likewise we can form the even and odd coarse grainings  $\Omega_E = \{E(\gamma) | \gamma \in \Omega\}$  and  $\Omega_O = \{O(\gamma) | \gamma \in \Omega\}$ , which inherit their associated measures from  $\Omega$ , and it is easy to see that  $\Omega = \Omega_E \sqcup \Omega_O$ . Building on this, we can treat both  $\Omega_E$  and  $\Omega_O$  in the same way that we previously treated  $\Omega$  by looking at the number of heads in the even and odd trials,  $H^E$  and  $H^O$ , respectively. Given an  $\epsilon$  we can go on to define  $H_\epsilon^E, H_\epsilon^O$ , and the subsets  $G_{H_\epsilon^E}, G_{H_\epsilon^O}$  and  $L_{H_\epsilon^E}, L_{H_\epsilon^O}$  of the two distributions. Then, following the analysis above, we can define the co-events  $\phi_{G_\epsilon^E}, \phi_{G_\epsilon^O}$  corresponding to the questions ‘is the number of heads in the even distribution greater than or equal to  $H_\epsilon^E$ ?’ and ‘is the number of heads in the odd distribution greater than or equal to  $H_\epsilon^O$ ?’ Now as before we have

$$\phi_{G_\epsilon^E}^E(L_{H_\epsilon^E}) = 0 \quad \phi_{G_\epsilon^E}^E(G_{H_\epsilon^E}) = 1, \tag{36}$$

however now we also have

$$\phi_{G_\epsilon^E}^E(L_{H_\epsilon^O}) = 0 \quad \phi_{G_\epsilon^E}^E(G_{H_\epsilon^O}) = 0, \tag{37}$$

with  $\phi_{G_\epsilon^O}^O$  showing similar behaviour.

For an explicit example consider the case  $p = 1/2, m = 10^3$  and  $\epsilon = 10^{-3}$ ; thus the even and odd distributions are similar to the example considered above. Then  $H_\epsilon^E = H_\epsilon^O = 450$ , and as before every single history  $\gamma$  contributes the same amount,  $p^{2m} = 2^{-2 \times 10^3}$ , to the probability of any event containing  $\gamma$ ; thus the measure of any event  $S$  is given by  $\mathbf{P}(S) = p^{2m} |S|$ . From this we can see that an event  $S$  is a subset of an approximately precluded set  $T$  if and only if  $S$  itself is approximately precluded, thus an approximate co-event  $\phi_\epsilon$  is preclusive if and only if its dual  $\phi_\epsilon^*$  has measure greater than or equal to  $\epsilon$ ; then if  $\phi_\epsilon$  is preclusive

$$\epsilon \leq \mathbf{P}(\phi_\epsilon^*) = p^{2m} |\phi_\epsilon^*|,$$

so that

$$|\phi_\epsilon^*| \geq \epsilon p^{-2m} = 10^{-3} \times 2^{2 \times 10^3}.$$

Thus  $\phi_\epsilon$  is primitive if and only if  $|\phi_\epsilon^*| = \text{Int}(\epsilon p^{-2m})$ , where  $\text{Int}(x)$  denotes the least integer that is greater than or equal to  $x$ . Now let  $\gamma_E = a_1, a_2, \dots, a_{2m}$  be the history

defined by the outcomes

$$a_i = \begin{cases} h & i \text{ even} \\ t & i \text{ odd.} \end{cases}$$

Then  $E(\gamma_E) = hhhh \dots$  and  $O(\gamma_E) = tttt \dots$ , so  $\gamma_E \in L_{H_\epsilon^o} \cap G_{H_\epsilon^E}$ . Now it is easy to see that  $\mathbf{P}(G_{H_\epsilon^E}) > \epsilon$ , so that  $|G_{H_\epsilon^E}| > \epsilon p^{-2m}$  and thus we can find an event  $C \subset G_{H_\epsilon^E}$  of cardinality  $\text{Int}(\epsilon p^{-2m})$  that contains the history  $\gamma_E$ . But then setting  $\phi_{G_\epsilon^E}^E = C^*$ , we have found a primitive approximately preclusive co-event satisfying equations (36) and (37). A similar construction can be made for  $\phi_{G_\epsilon^o}^o$ .

This is in fact a coarse graining problem (as are most of the problems in the multiplicative scheme) reminiscent of the interpretational problems of the consistent histories approach. Essentially we can recover the probabilities in the measure (or decoherence functional) but we are also recovering the interpretational problems in the sense that different coarse-grained partitions have become incompatible. In consistent histories each question ‘makes sense’ in one decoherent partition, but may not have a classical answer in other decoherent partitions. In our case every approximate co-event (which corresponds in a natural way to a question via its dual) will treat at least one partition classically (for example the partition consisting of its dual and the complement thereof) but may not treat other partitions classically, even though these partitions are dynamically classical (in that they decohere). However this feature is more of an issue for approximate preclusion than for consistent histories, since, as our coin example shows, approximate preclusion may have difficulties even when the underlying fine-grained histories obey classical dynamics, a problem not shared with consistent histories.

We can gain further insight into these problems by applying the machinery of the ‘principle classical partition’ introduced in the appendix. Essentially, even though the measure is classical we find non-classical behaviour at the level of the approximate co-events. This might suggest that our observable partitions are finer grained than the principle classical partition, and indeed while we are assured that when using exact preclusion a classical measure will lead to a fully fine-grained principle classical partition (every non-null singleton set is an element of the partition), we do not have such a guarantee when using approximate preclusion.

It is instructive, therefore, to calculate the principle classical partition for our repeated trial; to provide an explicit example we set  $p = 1/2$ ,  $n = 10^3$  and  $\epsilon = 10^{-3}$  as before. Then every history  $\gamma \in \Omega$  is of equal probability  $\mathbf{P}(\{\gamma\}) = 2^{-10^3}$ , and the Kolmogorov sum rule means that every set  $S \subset \Omega$  of trials has a probability determined solely by its cardinality,  $\mathbf{P}(S) = 2^{-10^3} |S|$ . Then if we set  $m = \text{Int}(2^{10^3} \epsilon)$  (where  $\text{Int}(x)$  denotes the least integer which is greater than or equal to  $x$ ), we see that  $\mathbf{P}(S) \geq \epsilon$  for every set of cardinality greater than or equal to  $m$  and that  $\mathbf{P}(S) < \epsilon$  for every set of cardinality less than  $m$ . Thus if  $S_m$  is of cardinality  $m$ , it is not itself approximately precluded, however every subset of  $S_m$  is approximately precluded. Thus  $S_m^*$  is a primitive approximately preclusive co-event. Since  $m > 2$ , given any two trials  $\gamma_1, \gamma_2 \in \Omega$  we can find a primitive approximately preclusive co-event  $\phi_\epsilon = S_m^*$  whose dual contains both trials,  $\gamma_1, \gamma_2 \in S_m$ . In other words, if we fix  $\gamma_1$ , given any  $\gamma_2 \in \Omega$  we can find a primitive  $\phi_\epsilon^{\gamma_2}$  whose dual contains both  $\gamma_1$  and  $\gamma_2$ . But then the transitive closure of the duals of all these primitive approximate co-events is the whole sample space  $\Omega$  and the principle classical partition is simply  $\{\Omega\}$ .

As a final comment on strong Cournot, we note that the value of  $\epsilon$  used in this section may be considered ‘large’, in that we can find experimentally observable events with smaller probability. Following this line of thought we might attempt to evade the above objections to strong Cournot by interpreting the results outlined above as placing a ‘cap’ on the level of  $\epsilon$  (which in strong Cournot we are taking to be a constant of nature). Thus we would

conclude that the value of  $\epsilon$  must be small compared to the probabilities of observable events (including experimentally achievable repeated trials)<sup>13</sup>. There are however objections to this argument. Firstly, such an approach would be unlikely to address the questions surrounding the status of single histories; in particular, with regard to an anticipated theory of quantum gravity, the individual ‘histories of the universe’ may not have ‘large’ measure. Secondly, and more seriously, it is our use of  $\epsilon$  to preclude observable events which allows us to discuss probability in the language of co-events, the very reason we turned to Cournot’s principle. Decreasing the value of  $\epsilon$  decreases our ability to falsify probability hypotheses, so a value of  $\epsilon$  so small as to avoid the problems outlined above would not give approximate preclusion any significant advantage over exact preclusion in the discussion of dynamics. Thus even if we adopted strong Cournot with such a small  $\epsilon$ , we would in any case be driven to some other mechanism (such as weak Cournot) to deal with dynamics and prediction. For these reasons, the authors prefer to discard strong Cournot in favour of weak Cournot.

### 3.5. Quantum operational weak Cournot

Although strong Cournot may be more philosophically satisfying, in particular with regard to co-event theory, the arguments of section 3.4 above conclusively demonstrate the failure of its application to multiplicative co-events due to violations of observable classicality. As in classical stochastic theories, this leads us to fall back on weak Cournot. Though perhaps less philosophically satisfying than the strong variety, it may be the ‘best we can do’, at least at the present time.

Our application of operational weak Cournot to quantum histories theories is similar in some ways to its classical application (section 3.2). We again evoke the ‘meta-level’ process by which we falsify theories to give us  $\epsilon$ , which consequently is no longer considered as a constant of nature and may be different for different systems. Our restriction of approximate preclusion to events singled out ahead of a repeated trial avoids the problems of multiple partitions and single histories encountered above. Essentially, we depart from strong Cournot by introducing a split between the ontology and the predictive content of the theory, whereas in strong Cournot both are described by a primitive approximate co-event we now propose:

*Ontology:* the potential realities are primitive exactly preclusive (multiplicative) co-events. In the example of the classical coin above, since the measure is classical, these primitive co-events correspond to single fine-grained histories.

*Predictions:* in an experiment consisting of repeated trials, an experimentally observable event  $A$  singled out in advance, of measure less than  $\epsilon$ , will not occur. We can of course phrase this in terms of approximate co-events, replacing the phrase ‘of measure less than  $\epsilon$ ’ by ‘precluded by all approximate co-events  $\phi_\epsilon$ ’. However this is tautological, and the use of approximate co-events is now rather vacuous since the ontology is given by primitive exactly preclusive co-events.

This formalism aims to allow us to use the hypothesis testing technique described in section 3.1 to test ‘Copenhagen predictions’ (discussed in section 1.5) through experiments consisting of repeated trials, for both classical and non-classical systems. There are, however, several objections we can make.

Firstly, the restriction of our predictive ability to such experiments may obstruct the application of weak Cournot beyond the Copenhagen framework, and the need to single out our predictions in advance of an experiment raises questions regarding observer independence;

<sup>13</sup> Repeated trials in which  $n$  is not so large that the repeated trial is experimentally unrealizable in practise.

a hallmark of the histories approach. Thus even if weak Cournot is successful in justifying the Copenhagen predictive mechanism from the perspective of quantum measure theory it is doubtful that it will assist us to develop a wider conception of the predictive content of a general quantum measure theory as hoped for in section 1.5. Simply by choosing to focus solely on (experimentally falsifiable) predictions we have constructed a formalism in which we cannot attach any ('physical') meaning to non-predictive dynamical statements. While such an 'instrumentalist' view of dynamics may be appealing to some [8], the authors feel that the dynamics should express something about the structure of reality, even in the absence of observation.

Had literal strong Cournot been applicable it would have avoided such issues, since Cournot's principle would have been applied at the ontological level and thus its implications would have been observer independent. More importantly this split between the ontology and the predictive mechanism separates us not only from strong Cournot but also from the application of weak Cournot in classical physics. In a classical theory we typically assume an underlying deterministic dynamics; the dynamical statements of this 'true theory' would then have observer independent meaning, allowing probability to be considered as an instrumentalist phenomenon without any implication for the interpretation of this 'true dynamics'.

Further, in a classical stochastic theory our statement of weak Cournot is made in terms of the event algebra and the measure, which are both defined in terms of the histories; the potential realities of the theory. In quantum measure theory our potential realities are primitive exactly preclusive multiplicative co-events, however our formulation of weak Cournot is still expressed in terms of events and the measure defined on them, or alternatively in terms of approximate co-events. Thus contrary to the goals we set out in section 1.5 we have not been able to express the predictive content of quantum measure theory in terms of its potential realities.

Finally, by adopting the above formulation of weak Cournot we have not made headway in our attempt to deal with 'almost null' sets (section 1.5). Given our practice of treating 'almost null' sets as 'exactly null' by assuming idealized conditions our inability to address this issue is a cause for concern.

#### 4. Conclusion

Our aim in this paper has been to rephrase the dynamical and predictive content of a quantum histories theory in terms of (multiplicative) co-events, or at the very least to demonstrate that this is possible in principle even if in practice it may be simpler to work with event algebras. In particular, we were concerned that due to experimental inaccuracies the events that we are characterizing as null will in general be only of small probability.

Our first, and potentially most appealing proposal was to shift the dynamics wholesale onto the co-events, 'completing' the co-event program by moving from 'histories theories' to 'co-event theories' in which we could use co-events and the dynamics defined on them to deal directly with physical systems. This would have allowed us to deal with events of small probability by assigning small probabilities to the co-events that found those events to be true. Unfortunately this approach has stalled.

Our second proposal was approximate preclusion. Following strong Cournot we aimed to explain the emergence of probability from preclusion by altering the status of the measure with regard to the allowed co-events. This would have allowed us to deal with events of small probability by directly precluding them. Again, this approach has not met with success.

Finally we have been driven to adopt weak Cournot, which can perhaps be used to justify the (existing) Copenhagen experimental framework from the vantage point of a quantum

histories theory. However we do not manage to express our predictions in terms of the ‘potentially real’ co-events themselves, and make no headway in extending our understanding of a histories theory’s predictive content beyond the narrow Copenhagen framework. Further, we are left in a quandary regarding the application of preclusion in practice. As we have previously pointed out the events that we are assuming to be null, for example in the double slit experiment, are only found to have zero measure following the assumption of an idealized system that we know cannot hold in practice. Thus for our results to have application to the real world an alternate solution to this problem must be found.

### Acknowledgments

The authors thank Fay Dowker, Rafael Sorkin and Sumati Surya for many insightful discussions on the co-event interpretation. The authors further thank Fay Dowker, Anna Gustavsson and Rafael Sorkin for their comments on drafts of this paper. This work was partly supported by the Royal Society-British Council International Joint Project 2006/R2. YGT was supported by an STFC studentship.

### Appendix A. The principle classical partition

For the purposes of this appendix, where there is no scope for confusion we will sometimes abbreviate ‘classical with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ ’ to ‘classical’. We can extend the notion of fine graining to a partial order on the partitions of  $\Omega$ . We say that a partition  $\Lambda$  of  $\Omega$  is a *fine graining of*, or equivalently is *less than*, a partition  $\Theta$  of  $\Omega$  if  $\Theta$  can be regarded as a partition of  $\Lambda$ . Equivalently we could say that  $\Theta$  is a *coarse graining*, or is *greater than*, the partition  $\Lambda$ . It is easy to see that fine graining induces a partial order on the space of partitions of  $\Omega$ .

Restricting to classical partitions, we can show that this partial order contains a unique minimal element, the ‘finest-grained’ classical partition, which we call the *principle classical partition* (with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ ).

**Theorem 3.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. Then there exists a classical partition (with respect to  $\mu$ )  $\Lambda$  of  $\Omega$  that is minimal among the classical partitions of  $\Omega$ . Furthermore  $\Lambda$  is unique.*

To prove this we first need a technical lemma:

**Lemma 9.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. Then the partition  $\Lambda$  of  $\Omega$  is classical with respect to  $\mu$  if and only if for every  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  there exists  $A_i \in \Lambda$  such that  $\phi^* \subset A_i$ .*

**Proof.** First assume that  $\Lambda = \{A_i\}_{i=1}^n$  is a classical partition of  $\Omega$ . Then by assumption  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  will behave classically on  $\Lambda$ , which means it will act as a homomorphism; obeying linearity (equation (7)) and multiplicativity (equation (8)) when restricted to the subalgebra generated by  $\Lambda$ . Then

- (1) Assume that  $\phi(A_i) = 0 \forall i$ ; linearity then implies that  $\phi(\Omega) = 0$ . Then multiplicativity implies that  $\phi$  is the zero map, which is excluded from  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ , contradicting our assumptions.
- (2) Assume  $\phi$  maps more than one element of  $\Lambda$  to one; without loss of generality  $\phi(A_1) = \phi(A_2) = 1$ . Noting that  $A_i$  are disjoint (as elements of a partition), multiplicativity implies  $\phi(\emptyset) = \phi(A_1 A_2) = \phi(A_1)\phi(A_2) = 1$ , contradicting our assumption that  $\phi$  is a co-event.

Therefore there exists  $j$  such that  $\phi(A_i) = \delta_{ij}$ . But then  $\phi^* \subset A_j$ .

To prove the converse, note that because  $A_i$  are disjoint every element of the subalgebra generated by  $\Lambda$  is of the form  $B_I = \bigsqcup_{i \in I} A_i$  for some indexing set  $I$ . Now let  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$ , then by assumption there exists  $A_k \in \Lambda$  such that  $\phi^* \subset A_k$ . Then for all  $B_I$  in the subalgebra generated by  $\Lambda$  we have

$$\phi(B_I) = \begin{cases} 1 & k \in I \\ 0 & k \notin I. \end{cases}$$

Thus it is easy to see that  $\phi$  is a homomorphism on the subalgebra generated by  $\Lambda$ . □

We are now in a position to prove the theorem.

**Proof of theorem 3.** *Existence:* the partition  $\{\Omega\}$  is always classical.

*Minimality:* because  $\Omega$  is finite the set of partitions thereof, and in particular the set of classical partitions thereof, is also finite. It is easy to see that the ‘fine graining of partitions’ defines a partial order in the set of classical partitions of  $\Omega$ , which must therefore contain a minimal element  $\Lambda$ .

*Uniqueness:* we enumerate the primitive preclusive multiplicative co-events by setting  $\mathcal{M}(\Omega, \mathfrak{A}, \mu) = \{\phi_i\}_{i=1}^m$  where  $m = |\mathcal{M}(\Omega, \mathfrak{A}, \mu)|$ . Now let  $\Lambda_A = \{A_i\}_{i=1}^{n_A}$  and  $\Lambda_B = \{B_i\}_{i=1}^{n_B}$  be distinct partitions that are both minimal among the set of partitions of  $\Omega$  that are classical with respect to  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$ . Then for every  $\phi_i \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  lemma 9 gives us  $A_{j_i} \in \Lambda_A$  and  $B_{k_i} \in \Lambda_B$  such that  $\phi_i^* \subset A_{j_i}$  and  $\phi_i^* \subset B_{k_i}$ . But then  $\phi_i^* \subset A_{j_i} \cap B_{k_i}$  for all  $i$ , so that by lemma 9 the partition  $\Lambda_{AB} = \{A_j \cap B_k\}$  is classical and is finer than both  $\Lambda_A$  and  $\Lambda_B$ , contradicting our assumption. Therefore  $\Lambda_A = \Lambda_B$ , and the principle classical partition is unique. □

In what remains of this appendix, we will enumerate the primitive preclusive multiplicative,  $\mathcal{M}(\Omega, \mathfrak{A}, \mu) = \{\phi_i\}_{i=1}^m$ , where  $m = |\mathcal{M}(\Omega, \mathfrak{A}, \mu)|$ . If none of the duals of the co-events ‘overlap’, so that  $\phi_i^* \cap \phi_j^* = \emptyset$  for  $i \neq j$ , then it is easy to see (lemma 10) that the principle classical partition is given by

$$\{\phi_i^* | \phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)\} \sqcup \{\{\gamma\} | \gamma \notin \phi^* \text{ for any } \phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)\}.$$

However the duals of co-events in  $\mathcal{M}(\Omega, \mathfrak{A}, \mu)$  will in general overlap, leading to a generalization of the above. If  $\Lambda$  is a classical partition we already know (lemma 9) that given  $\phi_i \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  we can find  $A_{m_i} \in \Lambda$  such that  $\phi_i^* \subset A_{m_i}$ . Now assume that  $\phi_i^* \cap \phi_j^* \neq \emptyset$  for some  $i \neq j$ , then  $\phi_j^* \cap A_{m_i} \neq \emptyset$ , so using lemma 9  $\phi_j^*$  is also a subset of  $A_{m_i}$ . Repeating this argument, if we can find  $\phi_k^*$  that has non-empty intersection with  $\phi_j^*$ , then  $\phi_k^* \subset A_{m_i}$  even if  $\phi_k^* \cap \phi_i^* = \emptyset$ . Thus we really want to ‘group together’  $\phi_i^*$  that are related by some sort of transitive closure of intersection; we achieve this by the use of an equivalence relation.

First we define a relation,  $\sim_\cap$ , on the set  $\mathcal{M}^*(\Omega, \mathfrak{A}, \mu) := \{\phi^* | \phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)\}$  as follows:

$$\phi_i^* \sim_\cap \phi_j^* \Leftrightarrow \phi_i^* \cap \phi_j^* \neq \emptyset.$$

This relation is reflexive and symmetric, thus its transitive closure,  $\approx_\cap$ , is an equivalence relation, which we will call *intersection equivalence*. We can therefore partition  $\mathcal{M}^*(\Omega, \mathfrak{A}, \mu)$  using  $\approx_\cap$  to give us the set  $\mathcal{M}_\cap^*(\Omega, \mathfrak{A}, \mu)$ , and will denote by  $\Phi^*(\phi_i^*)$  the equivalence class containing  $\phi_i^*$ . Finally, we can ‘merge’ the elements of the equivalence class  $\Phi^*(\phi_i^*)$  to give us the set

$$\Phi_i^{F*} := \bigcup_{\phi^* \in \Phi^*} \phi^*. \tag{A.1}$$

It is easy to see that  $\Phi_i^{F*} \in \mathfrak{A}$  (since it is the union of elements of  $\mathfrak{A}$ ), so its dual  $\Phi_i^F := \Phi_i^{F**}$  is a co-event. This leads us to define:

$$\begin{aligned} \mathcal{M}_F^*(\Omega, \mathfrak{A}, \mu) &:= \{ \Phi_i^{F*} \mid \Phi^* \in \mathcal{M}_\cap^*(\Omega, \mathfrak{A}, \mu) \}, \\ \mathcal{M}_F(\Omega, \mathfrak{A}, \mu) &:= \{ \Phi_i^F \mid \Phi^* \in \mathcal{M}_\cap^*(\Omega, \mathfrak{A}, \mu) \}. \end{aligned} \tag{A.2}$$

$\Phi_i^{F*}$  (or their duals) are sometimes referred to as *fat co-events*. The fat co-event corresponding to  $\phi_i^*$  is denoted by  $\Phi^{F*}(\phi_i^*)$ . It is clear that  $\Phi_i^{F*}$  are disjoint, however they do not necessarily form a cover of  $\Omega$ . Thus if we define  $Z$  to be the set of singleton sets of histories that are not an element of any  $\Phi_i^{F*}$  (or equivalently any  $\phi_i^*$ ) we can form the partition:

$$\Lambda_C = \mathcal{M}_F^*(\Omega, \mathfrak{A}, \mu) \sqcup Z, \tag{A.3}$$

of  $\Omega$ . It is easy to see that  $\Lambda_C$  is the principle classical partition with respect to  $\mu$ .

**Lemma 10.** *Let  $(\Omega, \mathfrak{A}, \mu)$  be a histories theory with a finite sample space. Then  $\Lambda_C$  as defined above is the principle classical partition with respect to  $\mu$ .*

**Proof.** Every  $\phi \in \mathcal{M}(\Omega, \mathfrak{A}, \mu)$  is an element of an equivalence class  $\Phi^*(\phi) \in \mathcal{M}_\cap^*(\Omega, \mathfrak{A}, \mu)$ , and thus is a subset of the fat co-event  $\Phi^{F*}(\phi) \in \Lambda_C$ . Then by lemma 9  $\Lambda_C$  is a classical partition with respect to  $\mu$ .

Now let  $\Lambda = \{A_i\}$  be a classical partition (with respect to  $\mu$ ). In the above we showed that if  $\phi_i^* \subset A_m$  then  $\phi_j^* \subset A_m$  whenever  $\phi_i^* \cap \phi_j^* \neq \emptyset$ . Then using an inductive argument it is easy to see that  $\phi_k^* \subset A_m$  whenever  $\phi_k^* \approx_\cap \phi_i^*$ , so that  $\Phi^{F*}(\phi_i^*) \subset A_m$ . Further, for any element  $\{\gamma\}$  of  $Z$  we can find  $A_m$  such that  $\{\gamma\} \subset A_m$ , thus  $\Lambda$  can be considered a partition of  $\Lambda_C$ . But since  $\Lambda$  was arbitrary,  $\Lambda_C$  is less than or equal to every classical partition; and so must be the principle classical partition.  $\square$

The principle classical partition helps us to further develop the notion of classicality for the multiplicative scheme, and more generally the histories approach. Note the contrast with the idea of classicality in consistent histories (see for example [22]), in which the absence in general of a ‘principle decoherent partition’ allows the possibility of incompatible but ‘equally valid’ classical interpretations (or ‘frameworks’ [22]) of a given system.

Further, note that there is no way to distinguish between two intersectionally equivalent co-events using the classical partitions. Thus if we make the assumption that ‘observable events’ are elements of classical partitions, this suggests that no sequence of measurements could allow us to distinguish between the distinct realities described by two co-events within the same intersectional equivalence class. In this way the fat co-events may be a useful practical rule, representing the ‘finest-grained’ information we could possibly discover using experimental techniques; a state of affairs we might think of as conforming to a *co-event uncertainty principle*.

## References

- [1] Anand P 1993 *Foundations of Rational Choice Under Risk* (Oxford: Oxford University Press)
- [2] Bernoulli J 1713 *Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola gallice scripta de ludo pilae reticularis* Basileae, impensis Thurnisiorum, fratrum
- [3] Borel E 1909 *Elements de la theorie des probabilites* (Paris: Gauthier-Villars)
- [4] Cournot A A 1843 *Exposition de la thorie des chances et des probabilits* (Paris: Hachet)
- [5] Cox R T 1961 *Algebra of Probable Inference* (Baltimore: The Johns Hopkins University Press)
- [6] de Finetti B 1930 *Fondamenti Logici del Ragionamento Probabilistico* (Bollettino dell’Unione Matematica Italiana) 9:258–61
- [7] Dirac P A M 1933 The Lagrangian in quantum mechanics *Phys. Z. Sowjetunion* **3** 64–72

- [8] Dowker F 2008 Private communication
- [9] Dowker F and Ghazi-Tabatabai Y 2007 The Kochen–Specker theorem revisited in quantum measure theory arXiv:0711.0894
- [10] Dowker F and Kent A 1996 On the consistent histories approach to quantum mechanics *J. Stat. Phys.* **82** 1575–646
- [11] Durr D, Goldstein S and Zanghi N 1992 Quantum equilibrium and the origin of absolute uncertainty *J. Stat. Phys.* **67** 843
- [12] Feynman R P 1948 Space-time approach to nonrelativistic quantum mechanics *Rev. Mod. Phys.* **20** 367
- [13] Feynman R P 1965 The development of the space-time view of quantum electrodynamics, Nobel Lecture, [http://nobelprize.org/nobel\\_prizes/physics/laureates/1965/feynman-lecture.html](http://nobelprize.org/nobel_prizes/physics/laureates/1965/feynman-lecture.html)
- [14] Frechet M 1951 Rapport general sur les travaux du colloque de calcul des probabilites *Congres International de Philosophie des Sciences, Paris, 1949; IV: Calcul des Probabilites* ed R Bayer (Paris: Hermann) pp 3–21
- [15] Friedman C 1999 The frequency interpretation in probability *Adv. Appl. Math.* **23** 234–54 (doi: 10.1006/aama.1999.0653)
- [16] Galvan B 2008 A generalization of the born rule arXiv:0806.4935 v1
- [17] Gell-Mann M and Hartle J B 1990 *Complexity, Entropy and the Physics of Information, SFI Studies in the Science of Complexity* vol III (Reading, MA: Addison-Wesley)
- [18] Gell-Mann M and Hartle J B 1993 Classical equations for quantum systems *Phys. Rev. D* **47** 3345–3382
- [19] Ghazi-Tabatabai Y in preparation
- [20] Goldstein S 2001 Boltzmann’s approach to statistical mechanics *Chance in Physics: Foundations and Perspectives (Lecture Notes in Physics)* J Bricmont *et al* (eds) (Berlin: Springer Verlag) p574 (Preprint arXiv.org/abs/cond-mat/0105242)
- [21] Griffiths R B 1984 Consistent histories and the interpretation of quantum mechanics *J. Stat. Phys.* **36** 219–72
- [22] Griffiths R B 1996 Consistent histories and quantum reasoning *Phys. Rev. A* **54** 2759–74
- [23] Griffiths R B 1998 Choice of consistent family, and quantum incompatibility *Phys. Rev. A* **57** 1604–18
- [24] Hartle J B 1991 The quantum mechanics of cosmology *Quantum Cosmology and Baby Universes: Proc. 1989 Jerusalem Winter School for Theoretical Physics* ed S Coleman, J B Hartle, T Piran and S Weinberg (Singapore: World Scientific)
- [25] Hartle J B 1995 Space-time quantum mechanics and the quantum mechanics of space-time *Proc. Les Houches Summer School on Gravitation and Quantizations, Les Houches, France, 6 Jul–1 Aug 1992* ed J Zinn-Justin and B Julia (Amsterdam: North-Holland)
- [26] Isham C J and Linden N 1994 Quantum temporal logic and decoherence functionals in the histories approach to generalized quantum theory *J. Math. Phys.* **35** 5452–76
- [27] Kochen S and Specker E P 1967 The problem of hidden variables in quantum mechanics *J. Math. Mech.* **17** 59–83
- [28] Kolmogorov A N 1933 *Grundbegriffe der Wahrscheinlichkeitsrechnung* (Berlin: Springer)  
Kolmogorov A N 1933 *Foundations of the Theory of Probability* (New York: Chelsea) (English transl.)
- [29] Levy P 1925 *Calcul des probabilites* (Paris: Gauthier-Villars)
- [30] Levy P 1937 *Theorie de l’addition des variables aleatoires* (Paris: Gauthier-Villars)
- [31] Omnes R 1988 Logical reformulation of quantum mechanics: 1. Foundations *J. Stat. Phys.* **53** 893–932
- [32] Peres A 1991 Two simple proofs of the Kochen–Specker theorem *J. Phys. A: Math. Gen.* **24** L175–8
- [33] Ramsey F P 1931 Truth and probability *The Foundations of Mathematics and Other Logical Essays* (London: Routledge and Kegan Paul) pp 156–99
- [34] Salgado R B 2002 Some identities for the quantum measure and its generalizations *Mod. Phys. Lett. A* **17** 711–28
- [35] Shafer G 2005 From Cournot’s principle to market efficiency *Working Paper no. 15, The Game-Theoretic Probability and Finance Project* <http://www.probabilityandfinance.com>
- [36] Shafer G and Vovk V 2005 The origins and legacy of Kolmogorov’s ‘Grundbegriffe’ *Working Paper no. 4, The Game-Theoretic Probability and Finance Project* <http://www.probabilityandfinance.com>
- [37] Sorkin R D 2008 Private communication
- [38] Sorkin R D 1994 Quantum mechanics as quantum measure theory *Mod. Phys. Lett. A* **9** 3119–28
- [39] Sorkin R D 1997 Quantum measure theory and its interpretation *Quantum Classical Correspondence: Proc. 4th Drexel Symposium on Quantum Nonintegrability, 8–11 September 1994, Philadelphia, PA* ed D H Feng and B-L Hu (Cambridge, MA: International Press) pp 229–51
- [40] Sorkin R D 2007 An exercise in ‘anhomomorphic logic’ *J. Phys.* (special volume) (arXiv:quant-ph/0703276)
- [41] Sorkin R D 2007 Quantum dynamics without the wave function *J. Phys. A: Math. Gen.* **40** 3207–22
- [42] Surya S and Wallden P 2008 Quantum covers in quantum measure theory arXiv:0809.1951
- [43] Tucker H G 1962 *An Introduction to Probability and Mathematical Statistics* (New York: Academic)
- [44] von Mises R 1957 *Probability Statistics and Truth* (New York: Dover)