arXiv:2603.11857v2 [quant-ph] 8 Sep 2026

ON CONTEXTUALITY AS A FEATURE OF LOGIC AND PROBABILITY THEORY

ASK ELLINGSEN

Abstract. In quantum mechanics, not everything that can be observed can be observed simultaneously. Observational data exhibits *contextuality* – a generalisation of nonlocality – if the result of an observation is necessarily dependent on which combination of observables was measured. This article gives a mathematical introduction to contextuality, emphasising its nature as a general feature of probability theory and logic, rather than of any particular quantum theory.

1. Introduction

One hundred years after the inception of quantum mechanics (QM), there still appears to be no universal consensus on what really makes a physical system *quantum* as opposed to *classical*. Certainly, some physical properties are hallmarks of QM – such as energy quantisation, uncertainty principles, and wave-particle duality – but none of these is universally present in all "quantum" systems.[1] Conversely, there seems to be no property which is uniquely classical, in the sense that it never holds for a "quantum" system. Rather, quantum mechanics regularly produces predictions that are compatible with some alternative "classical" explanation.[2]

However, if one were to attempt to single some quintessentially quantum feature – one of the best candidates might be the *simultaneous availability of information*. Namely, in a general quantum system, *not all information that may in principle be retrieved from the system may be retrieved at the same time*. If we conduct an experiment where we observe a certain feature of the system (say particle position, or spin in the z-direction), then there are other features (particle momentum, spin in the x-direction) that are *unobservable* in that experiment. This is reflected in the theory – in standard formulations of QM, no predictions are made about the outcomes of incommensurable observations (represented in the standard Hilbert space formalism by non-commuting operators).

This feature opens up an interesting possibility: The outcome of a quantum observation may in theory depend on the *context* in which it is observed. By a "context" is meant some subset of the set of all observations that one could in principle make. The key fact is that while each *individual* observation is possible, only certain *combinations* of observations are possible *simultaneously*.[3]

In a classical system, there is no theoretical restriction on which contexts are possible. Therefore, we remain assured that even if we choose to only ever observe certain combinations of features in conjunction with each other, whatever information we ignore is still well-defined, whether we check in on it or not. An observation made about a classical system simply reveals information that is "already there", and choosing a context merely means to ignore some of this information.

By contrast, if only certain subsets of information are simultaneously accessible even in theory, it is not obvious that these subsets of information "glue" to a whole, predetermined truth which is *independent* of the choice of which subset to access. A scenario of the second kind, where the context necessarily influences the outcome, is called *contextual*. A non-contextual theory is one where a measurement can be thought of as merely revealing pre-existing information.

Date: September 9, 2026.

[1]Indeed, not all quantum systems support notions of "energy", "waves", or "particles" at all. For example, where are the waves, the particles, or the energy in an (abstract) quantum bit?

[2]Indeed, if quantum mechanics did not behave deceivingly similarly to classical mechanics "most of the time", then classical physics would probably never have been very useful to begin with.

[3]It should be noted that unlike contextuality, the *uncertainty principle* does not fundamentally require this incompatibility. Namely, the statistical variances of multiple random variables can be large even when those random variables can be simultaneously evaluated.

Models of classical mechanics are usually non-contextual by construction. On the other hand, contextuality – and the closely related concept of *nonlocality* – have been generally accepted features of QM (or at least of the mathematical models thereof) since at least the work of Bell, Kochen and Specker in the 1960's [8, 9, 30]. Since then, multiple mathematical frameworks that describe some notion of contextuality have been proposed (see [15] for a fairly recent review of the current state of the art), but as of yet no one notion commands universal acclaim.

Having a readily available notion of contextuality general enough to be applicable in a multitude of scenarios seems desirable, as it would allow direct comparisons between notions of contextual behaviour in different formulations of quantum (and non-quantum) systems. Indeed, several existing notions of contextuality already apply beyond QM. Bell-type contextuality is a *probabilistic* obstruction to consistent, global value assignments; and Kochen-Specker type contextuality is a *logical* obstruction to the existence of any global value assignments at all.

This article is primarily intended as a review article attempting to introduce contextuality from this perspective – as a property of collections of probability distributions rather than of QM. In particular, I present no new results of my own, but in Section 5 I outline some ideas for future work.

In Section 2, we introduce the concept of contextuality with a thought experiment. In Section 3, we review how probability distributions are calculated in QM, and show how these distributions have the potential to be contextual. In Section 4, we outline the theory of *partial Boolean algebras*, and how these can be used to model the logical relationships between events in a contextual theory. We finish that section by giving a version of the Kochen-Specker theorem, which prohibits the existence of global value assignments for all high-dimensional enough quantum systems.

Some more modern approaches to contextuality employ the language of categories, sheaves, and topoi; and risk appearing hopelessly abstract to the uninitiated. In Section 5, we briefly mention two of the more prominent of these approaches: The sheaf-theoretical theory of *measurement scenarios* due to Abramsky and Brandenburger [1, 2, 3, 5, 6, 7] (which quickly found applications in the theory of quantum computation, see e.g. [32]); and the related *Bohrification* programme, which grew out of the topos-theoretical formulation of the Kochen-Specker theorem originally proposed by Butterfield and Isham [10, 16, 17, 19, 26, 27, 28, 29, 37]. A secondary goal of this article is to show that these formulations are well motivated by some natural modelling assumptions about contextual scenarios.

2. Alice and Bob Receive Letters

Remark 2.1. This section is largely based on [6], with some slight modifications – including making the thought experiment more "analogue" by phrasing it in terms of envelopes and pieces of paper; as well as emphasising the role of signalling distributions with an explicit example.

To illustrate how context-dependence may appear, let us perform a simple thought experiment in the form of a game. Two players, Alice and Bob, are situated in different rooms and unable to communicate. A referee, Nate,[4] presents Alice with two envelopes, marked A0 and A1, and Bob with envelopes marked B0 and B1. Alice and Bob are each allowed to choose one, and only one, of their envelopes to open. Inside each envelope is a piece of paper marked either 0 or 1. This procedure is repeated many times.

For each round of the game, Alice and Bob record their choice of envelope and the number they found inside. After the experiment is over, they are allowed to compare notes. Alice and Bob's goal is to decide if Nate placed the numbers inside the envelopes before Alice and Bob made their choices, or if he tampered with them afterward.

In each round of the experiment, there are $2^4 = 16$ possible **joint outcomes**, each labelled by a quadruple (x, y, a, b), where $a, b \in \{0, 1\}$ represent Alice and Bob's respective choices of envelopes, and $x, y \in \{0, 1\}$ the numbers they find inside. We call each envelope pair $(a, b) \in \{0, 1\}^2$ a **context**, and each pair $(x, y) \in \{0, 1\}^2$ a **result**.

Let $\mathbb{P}(x, y, a, b)$ be the total statistical probability of the outcome (x, y, a, b), and let

$$\mathbb{P}(a, b) = \sum_{x,y} \mathbb{P}(x, y, a, b) \tag{2.1}$$

[4] For *Nature*.

be the total probability that a and b were opened. We assume that $\mathbb{P}(a,b) > 0$ for all a, b (no context was left unexplored). Let

$$\mathbb{P}(x, y \mid a, b) = \frac{\mathbb{P}(x, y, a, b)}{\mathbb{P}(a, b)} \tag{2.2}$$

be the probability of the result (x, y), conditioned on the context (a, b).

With this notation in place, Alice and Bob's goal is to determine whether one of the following conditions fail:

(1) *(No tampering)* The envelopes are loaded before Alice and Bob make their choices. I.e., there is a well-defined *hidden probability distribution* $\mathbb{P}^h(x, x', y, y')$, describing the probability that Alice's two envelopes are loaded beforehand with the numbers x, x' respectively, and Bob's with y, y';

(2) *(Independence)* Alice and Bob's choices are not influenced by the preparation of the envelopes. I.e., the probability $\mathbb{P}(x, y, a, b)$ is given by a product

$$\mathbb{P}(x, y, a, b) = \mathbb{P}^h|_{a,b}(x, y)\mathbb{P}(a, b), \tag{2.3}$$

where

$$\mathbb{P}^h|_{0,0}(x, y) = \sum_{x', y'} \mathbb{P}^h(x, x', y, y') \tag{2.4}$$

is the *marginal probability* that the envelopes A0 and B0 contain the results (x, y), ignoring the contents of A1 and B1 – and similarly for all other envelope combinations (a, b).

Combining the independence assumption (2.3) with (2.2) yields the condition

$$\mathbb{P}(x, y \mid a, b) = \mathbb{P}^h|_{a,b}(x, y). \tag{2.5}$$

That is, the probability to see the result (x, y) *conditioned* on the context (a, b) is equal to the *marginal* probability that envelopes (a, b) were loaded with (x, y) to begin with. If *no* distribution $\mathbb{P}^h$ satisfying (2.5) exists, then Alice and Bob can be confident that there is some correlation between their choices of envelopes and the contents therein. That is, the result is dependent on the context, or **contextual**. If *some* $\mathbb{P}^h$ satisfying (2.5) exists, then Alice and Bob have no evidence of tampering.

Definition 2.2. Following [4], with notation as above, we will refer to an array of conditionals $\mathbb{P}(x, y \mid a, b)$ as a **measurement scenario**. A probability distribution $\mathbb{P}^h : \{0,1\}^4 \to [0,1]$ satisfying (2.5) will be called a **hidden global distribution** for the scenario. If a hidden distribution exists, the scenario is **non-contextual**; and otherwise **contextual**.

Remark 2.3. If a hidden distribution $\mathbb{P}^h$ satisfying (2.5) exists, it is typically not unique, since a probability distribution cannot in general be reconstructed from its marginals.

The 16 conditional probabilities $\mathbb{P}(x, y \mid a, b)$ can be organised into a 4×4 table. These tables can be neatly visualised through the use of *bundle diagrams*, first introduced in [6]. To draw a bundle diagram, one draws a *base graph* consisting of a vertex for each envelope A0, A1, B0, B1; and an edge connecting those vertices that belong to a common context. Over each vertex lies a *fibre* consisting of two points corresponding to the results $0, 1$. A joint outcome (x, y, a, b) corresponds to a line connecting points of neighbouring fibres (a "local section" of the bundle), and the conditional probability $\mathbb{P}(x, y \mid a, b)$ can be visualised as a "weighting" of the corresponding line.

Example 2.4 (Deterministic outcome)**.** Table 1a (fig. 1a) represents a scenario where the referee always loads all envelopes with the number 1. This scenario is non-contextual.

Example 2.5 (50/50)**.** Table 1b (fig. 1b) represents a scenario where the referee always loads all envelopes with the same number within a round, with a 50% chance of choosing 0 or 1 each time. This scenario is non-contextual.

Example 2.6 (Signalling)**.** Table 1c (fig. 1c) represents to a scenario where the referee loads Alice and Bob's envelopes with the result 1 if they choose the context $(0, 0)$ or $(1, 1)$, and with 0 if they choose $(0, 1)$ or $(1, 0)$. This scenario is contextual.

Example 2.7 (Popescu-Rohrlich box)**.** Table 1d (fig. 1d) represents a scenario where Alice and Bob always see the same result as each other (with a 50% chance of each) in the contexts $(0, 0)$, $(0, 1)$, and $(1, 0)$, but opposite results in the context $(1, 1)$. This scenario is the contextual **Popescu-Rohrlich (PR) box** [31].

TABLE 1. Probability tables for the scenarios described in Examples 2.4 – 2.10.

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	0	0	0	1
(0,1)	0	0	0	1
(1,0)	0	0	0	1
(1,1)	0	0	0	1

(A) The deterministic scenario from Example 2.4.

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	1/2	0	0	1/2
(0,1)	1/2	0	0	1/2
(1,0)	1/2	0	0	1/2
(1,1)	1/2	0	0	1/2

(B) The noncontextual 50/50-scenario from Example 2.5.

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	0	0	0	1
(0,1)	1	0	0	0
(1,0)	1	0	0	0
(1,1)	0	0	0	1

(C) The signalling scenario from Example 2.6.

	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	1/2	0	0	1/2
(0,1)	1/2	0	0	1/2
(1,0)	1/2	0	0	1/2
(1,1)	0	1/2	1/2	0

(D) The Popescu-Rohrlich box from Example 2.7.

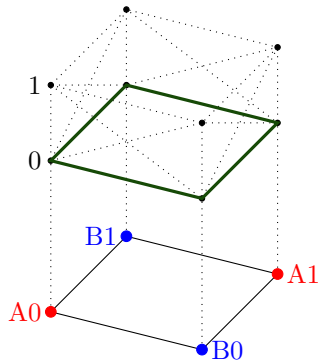
	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	1/2	0	0	1/2
(0,1)	3/8	1/8	1/8	3/8
(1,0)	3/8	1/8	1/8	3/8
(1,1)	1/8	3/8	3/8	1/8

(E) The CHSH scenario from Example 2.8.

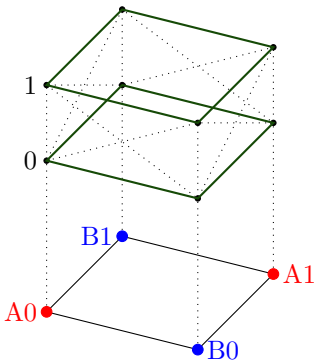
	(0,0)	(0,1)	(1,0)	(1,1)
(0,0)	1/2	0	0	1/2
(0,1)	1/2	0	0	1/2
(1,0)	1/2	0	0	1/2
(1,1)	1/4	1/4	1/4	1/4

(F) The logically contextual scenario from Example 2.10.

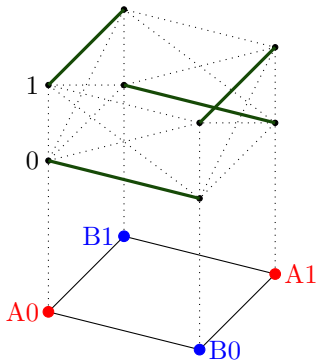
FIGURE 1. Bundle diagrams corresponding to the tables in Table 1. Darker colours and thicker lines correspond to higher probabilities. Joint outcomes that are assigned zero probability are shown as dotted lines.



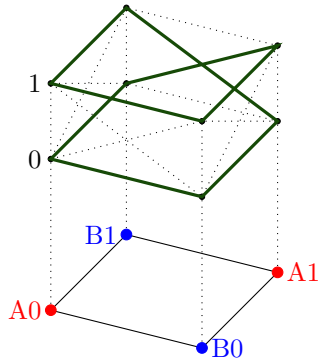
(A) Bundle diagram corresponding to Table 1a.



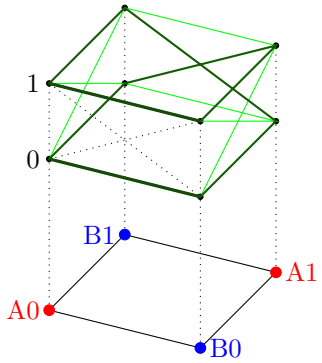
(B) Bundle diagram corresponding to Table 1b.



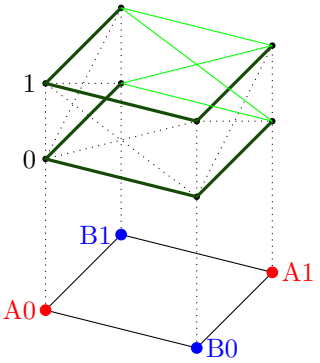
(C) Bundle diagram corresponding to Table 1c.



(D) Bundle diagram corresponding to Table 1d.



(E) Bundle diagram corresponding to Table 1e.



(F) Bundle diagram corresponding to Table 1f.

Example 2.8 (CHSH)**.** Table 1e (fig. 1e) represents a weakened version of the PR box from Example 2.7, where the outcomes that were possible in the PR box are the *most likely*, but other outcomes are also possible. We will refer to this as is the **Clauser-Horne-Shimony-Holt (CHSH) scenario**. It is contextual.

We claim that Examples 2.6 – 2.8 are all contextual, but they are so in increasingly subtle ways. In the bundle diagrams, a global assignment of results (x, x', y, y') corresponds to a closed, simple loop traversing each fibre exactly once. A probability table is non-contextual precisely if the corresponding bundle diagram can be built by adding together such simple loops of constant weight [6].

Examples 2.6 and 2.7 clearly fail this condition, as they do not contain any closed loops at all. Such distributions are called **strongly contextual**. They correspond to a situation where no global assignment of values is at all compatible with the data.

In Example 2.6, we see that there are "discontinuities" in the bundle diagram. This reflects that this example is **signalling**, meaning that Alice's result allows her to deduce what envelope Bob chose to open, and vice versa. Signalling is witnessed by *incompatible marginals*.

Definition 2.9. With notation as above, a measurement scenario is **no-signalling**, or satisfies **no-disturbance** at a if

$$\sum_y \mathbb{P}(x, y \mid a, 0) = \sum_y \mathbb{P}(x, y \mid a, 1). \tag{2.6}$$

Similarly for b. A measurement scenario is **signalling** if the no-signalling condition fails at any a or b.

In other words, a scenario is no-signalling iff we have well-defined marginal distributions

$$\mathbb{P}(x \mid a) = \sum_y \mathbb{P}(x, y \mid a, 0) = \sum_y \mathbb{P}(x, y \mid a, 1), \tag{2.7}$$

and

$$\mathbb{P}(y \mid b) = \sum_x \mathbb{P}(x, y \mid 0, b) = \sum_x \mathbb{P}(x, y \mid 1, b), \tag{2.8}$$

That the presence of signalling expresses itself as a discontinuity in the bundle diagram is suggestive of the fact that this type of signalling is *locally detectable*. Alice and Bob can infer information about the other's choice from their own result.

Signalling distributions are not terribly exciting, as they require Alice and Bob to be able to exchange information. What is perhaps less immediately intuitive is that there is some wiggle room between non-contextual and signalling distributions. Namely, marginalising a probability distribution loses information about correlations between events. These correlations may contain contextual information even when Alice and Bob's locally available, marginal distributions do not. The slogan is that even if the distributions are *locally consistent* they may be *globally inconsistent* [6]. No-signalling contextual distributions can allow Alice and Bob to *coordinate* without *communicating*.

The PR box from Example 2.7 is no-signalling – meaning that no information is exchanged between Alice and Bob within a round; reflected in the lack of "discontinuities" in the bundle diagram. Nevertheless, the diagram still does not contain any closed *simple* loop (only a double loop), and so is strongly contextual. To see that no local assignment can be extended to a global assignment, assume for example that A0 contains 1. This leads to the following "liar's cycle":

$$\text{A0 contains 1} \implies \text{B0 contains 1} \implies \text{A1 contains 1} \implies \text{B1 contains 0} \implies \text{A0 contains 0};$$

an immediate contradiction. A similar contradiction follows from assuming that A0 contains 0, so no assignment of values to envelope A0 is consistent with *any* global assignment of values to all envelopes.

If the PR box is "qualitatively" contextual, the CHSH scenario from Example 2.8 is only *quantitatively* contextual. To derive a contradiction, one can assume that envelope A0 contains 0 and calculate the conditional probability that A1 contains 0. Passing via B0 gives a conditional probability of 3/4, while passing via B1 gives a conditional probability of only 3/8. We say that the CHSH scenario is **probabilistically**, or **weakly contextual**.

In between strong and weak contextuality lies **logical contextuality**. In a logically (but not strongly) contextual scenario, the bundle diagram may contain *some* closed simple loops, but not all edges belong to one. I.e., not every local section of the diagram is extendable to a global section.

Example 2.10. Table 1f (fig. 1f) shows a logically contextual, but not strongly contextual scenario. Namely, there *exist* global sections, corresponding to the scenarios that all envelopes contain the same number (these are the two closed simple loops in fig. 1f). But in the context $(A1, B1)$, the outcomes where Alice and Bob's envelopes contain different numbers (the "crossbars" in fig. 2.10) do not belong to any global section. Following the bundle diagram clockwise, we have the sequence of implications

$$\text{A1 contains 0} \implies \text{B0 contains 0} \implies \text{A0 contains 0} \implies \text{B1 contains 0.}$$

Similarly, if A1 contains 1, then B1 must also contain 1. Nevertheless, in the context $(A1, B1)$, half the time that A1 contains a 0, B1 contains a 1, and vice versa.

Remark 2.11. The three forms of contextuality *strong*, *logical*, and *weak* are strictly ordered by strength:

$$\text{Strong contextuality} \implies \text{Logical contextuality} \implies \text{Weak contextuality.}$$

However, *signalling* is a somewhat independent property. A signalling distribution is always weakly contextual, but may or may not be strongly or logically contextual.

The bundle diagrams from Fig. 1 suggest that contextuality is in some way a "topological obstruction" to the existence of global sections. This perspective is expanded upon in [4, 5, 6].

3. Quantum Probability

In this section we briefly review how probabilities are calculated in the Hilbert space based textbook approach to QM. The goal of this section is to illustrate how contextual collections of probability distributions may emerge from the familiar rules of quantum mechanics. For simplicity, we will only discuss pure states and sharp measurements.

3.1. **The Born rule.** One of the main predictive laws of QM is the *Born rule* [13, 14]. In its simplest form, the Born rule says that *events* a are represented by *projection operators* $P_a = P_a^2 = P_a^\dagger$ acting on a Hilbert space $\mathcal{H}$. A (pure) *state* of the system is represented by a normalised vector $\psi \in \mathcal{H}$, $\|\psi\| = 1$; and if one performs a measurement in which an event a represented by a projection P_a may or may not occur, then given that the system is in the state represented by ψ, the probability to observe a is

$$\mathbb{P}_\psi(a) = \langle \psi, P_a \psi \rangle \,. \tag{3.1}$$

Remark 3.1. Notice that the Born rule (3.1) is not a mathematical identity, but a *modelling assumption*, or *correspondence rule*. The quantity $\mathbb{P}_\psi(a)$ on the lhs is the *modelled probability* of the event (represented by) a occurring given that the system is in the state (represented by) ψ. The rhs is a number which can be calculated.

Remark 3.2. In the interest of simplicity, we only consider *projective* measurements and *pure* states. Alternatively, we may simply postulate the version of the Born rule given in (3.1) as a toy example, and study the properties of the resulting probability distributions.

The Born rule readily generalises to simultaneous measurements of multiple events $a_1, \ldots, a_n$, as long as the projection operators P_{a_i} *commute*. Namely, then $\prod_{i=1}^n P_{a_i}$ is a new projection operator, and the probability of simultaneously observing all the events a_i is given by

$$\mathbb{P}_\psi(a_1, \ldots, a_n) = \left\langle \psi, \left(\prod_{i=1}^n P_{a_i} \right) \psi \right\rangle . \tag{3.2}$$

We may think of it as though the projection $\prod_{i=1}^n P_{a_i}$ represents the event that *all* the events a_i individually occur.

We will call events a, b corresponding to non-commutative operators P_a, P_b **incommensurable**.[5] If any of the a_i are incommensurable, then not only does the product operator $\prod_{i=1}^n P_{a_i}$ ambiguously depend on the indexation of the events a_i; it may no longer even be a projection operator, and hence does not represent an event. The Born rule thus makes no claims about the probability that incommensurable events occur simultaneously.

[5] In the literature on contextuality, the term *incommeasurable* seems to be more common.

In QM, one often emphasises *observables* over events. These two viewpoints are equivalent, and in a sense "duals". An *observable property* of the system is postulated to be represented by a self-adjoint operator $T = T^\dagger$, defined on a dense subspace $\mathcal{D}_T$ of $\mathcal{H}$. Such an operator is also called an **observable**. By the spectral theorem, each observable T has a unique associated projection valued measure (PVM) $P^T : \mathscr{B}(\mathbb{R}) \to \mathrm{Proj}(\mathcal{H})$ – where $\mathscr{B}(\mathbb{R})$ is the σ-algebra of Borel sets on $\mathbb{R}$ and $\mathrm{Proj}(\mathcal{H})$ is the set of projections in $\mathcal{H}$ – such that

$$T = \int_{\mathbb{R}} \lambda \, \mathrm{d}P^T(\lambda). \tag{3.3}$$

If $U \subset \mathbb{R}$ is Borel measurable, then the event "T takes a value in U" – which we will denote symbolically as $(T \rightsquigarrow U)$ – is represented by the projection $P^T(U)$. The Born rule gives the probability of this event, assuming that the system is in a state $\psi \in \mathcal{D}_T$, as

$$\mathbb{P}_\psi(T \rightsquigarrow U) = \left\langle \psi, P^T(U)\psi \right\rangle; \tag{3.4}$$

and the expectation value of (the property represented by) T as

$$\mathbb{E}_\psi(T) = \langle \psi, T\psi \rangle . \tag{3.5}$$

In particular, the projection P_a corresponding to an event a is an observable, representing the "truth value" of a. The probability that a occurs agrees with the expectation value of the projection:

$$\mathbb{P}_\psi(a) = \langle \psi, P_a\psi \rangle = \mathbb{E}_\psi(P_a).$$

3.2. **Contexts in QM.** Since the Born rule only assigns probabilities to sets of commuting projections, it may in principle allow contextual distributions. If events are represented by projections $P_a \in \mathrm{Proj}(\mathcal{H})$, then **contexts** are subsets $C \subset \mathrm{Proj}(\mathcal{H})$ of mutually commuting projections.

Commutativity is a reflexive and symmetric, but generally non-transitive binary relation on the set of projections. We write

$$P_a \odot P_b \quad \text{if} \quad P_aP_b = P_bP_a. \tag{3.6}$$

A context is a subset $C \subset \mathrm{Proj}(\mathcal{H})$ such that $P_a \odot P_b$ for all $P_a, P_b \in C$.

Notice that contexts have the following property:

> **Specker's principle.** If $C \subset \mathrm{Proj}(\mathcal{H})$ is a set such that all $P_a, P_b \in C$ are *pairwise* commensurable, then they are *all* commensurable.

This principle, also called **coherence**, holds because commutativity is a binary relation. In higher generality, one can imagine situations where a collection of events may be pairwise commensurable without all being commensurable as a group.

We also note that the set of projections support versions of the *logical connectives and*, *or*, and *not* – but only between commensurable projection operators. If $P_a \odot P_b$, then the event *a and b* is represented by the projection P_aP_b; the event *a or b* is represented by $P_a + P_b - P_aP_b$; and the event *not a* is represented by $1 - P_a$.

Furthermore, the projection onto the trivial subspace 0 is always assigned probability 0 by the Born rule (3.1), so may be taken as representing an *impossible* event; and the identity projection $1 : \mathcal{H} \to \mathcal{H}$ always has probability 1, so it represents a *necessary* event (anything happens).

Projection operators are equivalent to linear subspaces, and these subspaces are partially ordered by inclusion. In terms of projections, the inclusion order of subspaces is equivalent to the order

$$P_a \leq P_b \quad \text{if} \quad P_aP_b = P_bP_a = P_a. \tag{3.7}$$

Under interpretation, if $P_a \leq P_b$, then *a implies b*.

In the classic paper [11], Birkhoff and von Neumann attempted to use the set of projections to provide a logical foundation for a probability theory where probabilities are calculated using the Born rule. In their approach, dubbed (Birkhoff-von Neumann) *quantum logic*, the order structure of the space of projections was assigned the highest importance.

Later, Kochen and Specker introduced another approach to quantum logic using the language of *partial Boolean algebras*, which placed more emphasis on the commensurability relation [30]. This will be the topic of the next section.

4. Logic, Probability, and Order

As the thought experiment involving Alice and Bob in Section 2 illustrates, contextuality is a general feature of logical relationships and probability distributions, not only of quantum mechanics. Hence it seems natural to attempt to look for a general formulation of contextuality in terms of probability distributions, rather than directly in terms of quantum mechanical notions.

In this section we will see that the standard Kolmogorovian formulation of probability theory using measure theory on σ-algebras (or Boolean algebras) of sets is not sufficiently general to support any notion of contextuality, since it assumes the existence of a *sample space* of global outcomes.

Remark 4.1. Boole himself (who did his work on probability long before the invention of set theory) actually recognised the necessity of additional assumptions to avoid Bell-like phenomena within his framework – see Section 11 of Chapter XIX in [12]. Later authors have also recognised the necessity of certain constraints to be satisfied in order for "local" probability distributions to "glue" – see e.g. Vorob'ev [38]. For an interesting discussion on the history, see the introduction to [25] (where the references to Boole and Vorob'ev can also be found). I thank the referee for this comment, as I was unaware of this connection between Boole's and Bell's works.

Furthermore, thanks to a famous theorem due to Stone (Theorem 4.9), the existence of a sample space can be inferred from the order theoretical structure describing the logical relationships between events, so in order to allow contextual probability distributions we must first allow a more general model of the events themselves. At the end of this section we explore the potential of *partial Boolean algebras* as one such model.

4.1. **What is probability?** The *probability* of an event is, in the most general sense of the word, simply a number indicative of some form of *confidence* in the event – either that it will occur, has occurred, or an estimate of how often the event can be expected to occur, on average. A probability *distribution* assigns probabilities to multiple events.

As a starting point for a naive mathematical model of probability, we may take a set $\mathcal{A}$, whose elements represent **events**. How one goes about assigning probabilities to events in $\mathcal{A}$ is a matter of modelling – it could be through statistical testing, plain guesswork, or otherwise – but generally, a mathematical model of probability is assumed to satisfy some version of the following assumptions:

(1) The probability is 0 that *nothing* happens, and 1 that *something* happens. That is, if $\bot \in \mathcal{A}$ is the **impossible event** and $\top \in \mathcal{A}$ is the **necessary event**, then

$$\mathbb{P}(\bot) = 0, \qquad \mathbb{P}(\top) = 1. \tag{4.1}$$

(2) Probability is *additive* with respect to the logical connectives *and* and *or*:

$$\mathbb{P}(a \text{ or } b) + \mathbb{P}(a \text{ and } b) = \mathbb{P}(a) + \mathbb{P}(b). \tag{4.2}$$

Remark 4.2. The above axioms are not intended as mathematically well-formed statements, only as intuitive "sketches" motivating the definitions making up a concrete model of probability. Do note however that the second axiom implies that the conjunctions "a and b" as well as "a or b" must necessarily themselves exist as *single events*, so that they can be assigned probabilities.

4.2. **Measure theoretic probability.** In the standard Kolmogorovian, measure theory based approach to modelling (classical) probability distributions, a core assumption is the existence of a **sample space** Ω (a set). As the set of events one takes a collection $\mathcal{A} \subset \mathscr{P}(\Omega)$ of subsets $a \subset \Omega$ ($\mathscr{P}(\Omega)$ is the *power set* of Ω). A point $\omega \in \Omega$ is an **outcome**. We say that a **occurs** (w.r.t. ω) if $\omega \in a$. Each outcome ω thus yields an evaluation map $\delta_\omega : \mathcal{A} \to \{0, 1\}$, defined by

$$\delta_\omega(a) = \chi_a(\omega) = \begin{cases} 1, & \text{if} \quad \omega \in a; \\ 0, & \text{if} \quad \omega \notin a. \end{cases} \tag{4.3}$$

In this way, each point (outcome) ω corresponds to a *global* assignment of simultaneous *truth values* $\delta_\omega(a) \in \{0, 1\}$ to all events $a \in \mathcal{A}$.

Sample spaces mimic an experimental scenario, where after each round of the experiment the outcome ω is definite, but beforehand is unpredictable. Morally at the very least, a probability distribution assigns a fundamental probability to each outcome ω, and the total probability of an event is the sum of the

probabilities of all outcomes that imply it. In the limit of infinitely many outcomes, this leads to the suggestive integral formula

$$\mathbb{P}(a) = \int_a \mathrm{d}\mathbb{P}(\omega). \tag{4.4}$$

Furthermore, if $a, b \in \mathcal{A}$ are events, then *both* a and b occur (w.r.t. ω) iff $\omega \in a \cap b$, and *at least one of them* occur iff $\omega \in a \cup b$. Hence it makes sense to interpret $a \cap b$ as the event a *and* b, and $a \cup b$ as a *or* b. It is then immediate from (4.4) that the additivity axiom

$$\mathbb{P}(a \cup b) + \mathbb{P}(a \cap b) = \mathbb{P}(a) + \mathbb{P}(b) \tag{4.5}$$

holds. We take $\bot = \varnothing$ as the impossible event (no outcome at all), and $\top = \Omega$ as the necessary event (any outcome); and normalise so that $\mathbb{P}(\Omega) = 1$.

A probability distribution constructed in this way can never be contextual, since it is constructed by summing over global truth value assignments. We may attempt to remedy this is by removing the sample space Ω, and work directly with the space of events $\mathcal{A}$.

4.3. **Boolean algebras.** For the normalisation and additivity axioms on a probability distribution to make sense, the event set $\mathcal{A}$ needs to carry some additional structure modelling the logical relationships between events. One popular way to equip $\mathcal{A}$ with such structure is by making it into a *Boolean algebra*.

Definition 4.3. Let $\mathcal{A}$ be a partially ordered set (poset). We say that $\mathcal{A}$ is a **lattice** if the meet $a \wedge b = \inf\{a, b\}$ and join $a \vee b = \sup\{a, b\}$ exist for all $a, b \in \mathcal{A}$. The lattice $\mathcal{A}$ is:

- **Distributive** if the meet and join satisfy the **distributive law**

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \qquad a, b, c \in \mathcal{A}. \tag{4.6}$$

- **Bounded** if there are elements $\top, \bot \in \mathcal{A}$ such that

$$\bot \leq a \leq \top, \qquad a \in \mathcal{A}; \tag{4.7}$$

- **Complemented** if it is bounded, and there for each $a \in \mathcal{A}$ exists $\neg a \in \mathcal{A}$ such that

$$a \wedge \neg a = \bot, \qquad a \vee \neg a = \top. \tag{4.8}$$

A **Boolean algebra** is a complemented, bounded, distributive lattice.

We interpret $\top$ as the necessary event, $\bot$ as the impossible event, $a \wedge b$ as the event a *and* b, $a \vee b$ as a *or* b, and $\neg a$ as *not* a. The partial ordering relation $\leq$ can be thought of as *implication*.

Remark 4.4. The partial order relation $\leq$ can be recovered from the binary operations $\wedge, \vee : \mathcal{A} \times \mathcal{A} \to \mathcal{A}$, via

$$a \leq b \quad \text{if} \quad a \wedge (a \vee b) = a. \tag{4.9}$$

Therefore, a Boolean algebra can be equivalently defined as a set $\mathcal{A}$ equipped with appropriate operations $(\wedge, \vee, \neg, \top, \bot)$.

A **Boolean algebra morphism** is a function $f : \mathcal{A} \to \mathcal{B}$ between Boolean algebras $\mathcal{A}$ and $\mathcal{B}$ such that

$$f(a \wedge b) = f(a) \wedge f(b), \qquad f(a \vee b) = f(a) \vee f(b), \qquad f(\top_{\mathcal{A}}) = \top_{\mathcal{B}}, \qquad f(\bot_{\mathcal{A}}) = \bot_{\mathcal{B}}. \tag{4.10}$$

A Boolean algebra morphism automatically also preserves complements: $f(\neg a) = \neg f(a)$. With these morphisms, Boolean algebras form a category **BA**. Two Boolean algebras are **isomorphic** if they are isomorphic in this category. Explicitly, an **isomorphism** of Boolean algebras is an invertible Boolean algebra morphism f whose inverse f^{-1} is also a Boolean algebra morphism. If $\mathcal{A}$ and $\mathcal{B}$ are isomorphic we write $\mathcal{A} \cong \mathcal{B}$. A **Boolean subalgebra** of a Boolean algebra $\mathcal{A}$ is a subset $\mathcal{S} \subset \mathcal{A}$ which is a Boolean algebra with the same $(\leq, \wedge, \vee, \neg, \top, \bot)$. A Boolean subalgebra is a subobject in the category theoretical sense.

Example 4.5. The **trivial**, or **terminal** Boolean algebra is the one-point set $1 = \{0\}$, and the **initial** Boolean algebra is the two-point set $2 = \{0, 1\}$, each equipped with the usual order.

Example 4.6. For any set Ω, the power set $\mathscr{P}(\Omega)$ is a Boolean algebra under the inclusion order. For example, if $\Omega = \varnothing$, then $\mathscr{P}(\Omega) \cong 1$ is the trivial Boolean algebra, and if $\Omega = 1 = \{0\}$ is a singleton, then $\mathscr{P}(\Omega) \cong 2$.

Definition 4.7. A **concrete** Boolean algebra is one that is isomorphic to a Boolean *subalgebra* of some power set $\mathscr{P}(\Omega)$.

A concrete Boolean algebra is one that can be realised as a set of subsets. *Stone's representation theorem* (Theorem 4.9) implies that *all* Boolean algebras are actually of this type. This means that the existence of global outcomes does not have to be assumed, but is a consequence of the abstract order-theoretic definition.

As a matter of notation, we employ the convention that a natural number n is identified with the set of its predecessors: $0 = \varnothing$ and $n = \{0, 1, \ldots, n-1\}$. In particular, $1 = \{0\}$ is the one-point set, and $2 = \{0, 1\}$ is the two-point set; and a truth valuation on $\mathcal{A}$ is a map $\omega : \mathcal{A} \to 2$. Also, if $\mathbf{C}$ is a category and $X, Y \in \mathbf{C}$, we write

$$\mathbf{C}(X, Y) = \mathrm{Hom}_{\mathbf{C}}(X, Y) = \{f : X \to Y\}$$

for the set of $\mathbf{C}$-morphisms from X to Y.

Lemma 4.8. *If $\mathcal{A}$ is a Boolean algebra, let $S(\mathcal{A}) = \mathbf{BA}(\mathcal{A}, 2)$ be the set of Boolean algebra morphisms $\omega : \mathcal{A} \to 2$.*[6] *The sets of the form $U_a = \{\omega \in S(\mathcal{A}) \mid \omega(a) = 1\}$ form a basis for a topology on $S(\mathcal{A})$ as a varies over $\mathcal{A}$.*

We call the thus defined topology the **Stone topology** on $S(\mathcal{A})$, and $S(\mathcal{A})$ the **Stone space**, or the **Stone spectrum** of the Boolean algebra $\mathcal{A}$.[7] The idea is that each $\omega \in \mathbf{BA}(\mathcal{A}, 2)$ forms a *point* of a topological space, which can be taken as a sample space for $\mathcal{A}$. Each event $a \in \mathcal{A}$ can then be identified with the set U_a of outcomes that assign a a value of 1. In this way we recover the standard picture of a Boolean algebra as a set of sets.

Theorem 4.9 (Stone)**.** *If $\mathcal{A} \in \mathbf{BA}$ is a Boolean algebra and $S(\mathcal{A})$ is its Stone spectrum, then the set $\mathrm{Cl}(S(\mathcal{A}))$ of clopen sets in $S(\mathcal{A})$ is a Boolean algebra under the inclusion order, and*

$$\mathrm{Cl}(S(\mathcal{A})) \cong \mathcal{A}. \tag{4.11}$$

Conversely, if X is any Stone space (totally disconnected Hausdorff space), its set of clopens $\mathrm{Cl}(X)$ is a Boolean algebra, and

$$S(\mathrm{Cl}(X)) \cong X, \tag{4.12}$$

i.e., the Stone space $S(\mathrm{Cl}(X))$ is homeomorphic to X.

Proof. See [35]. □

Corollary 4.10. *Every Boolean algebra is concrete.*

Stone's theorem says that the sample space Ω is implicit in the logical structure of the event set $\mathcal{A}$. As long as $\mathcal{A}$ carries a partial order making it into a Boolean algebra, we may always realise it as a Boolean algebra of subsets of some sample space Ω. Moreover, the construction explicitly identifies the sample space Ω with the set of global truth value assignments, or *global outcomes* $\omega : \mathcal{A} \to 2$.

Remark 4.11. We choose to focus on Boolean algebras in favour of σ-algebras as it simplifies the theory, but pause to note that a similar representation theorem is true for σ-algebras: the *Loomis-Sikorski theorem* [33, 36]. We will not discuss σ-algebras in any detail here, but note that the Loomis-Sikorski theorem comes with an additional conceptual caveat: While Stone's theorem says that any Boolean algebra is a *subalgebra* of a power set, the Loomis-Sikorski theorem says that any σ-algebra is a *quotient algebra* of a subalgebra of a power set. In particular, if $(X, \mathcal{A}, \mu)$ is a measure space, modding out by sets of μ-measure 0 gives an "abstract" σ-algebra that generally is not realisable as a σ-algebra of sets.

4.4. **Partial Boolean algebras.** Stone's theorem shows that if we assume that the event set $\mathcal{A}$ is equipped with a partial order making it into a Boolean algebra, then it can always be realised as a Boolean algebra of sets. But this means that events are represented by sets of global outcomes, which precludes the possibility of contextuality.

This is perhaps not too surprising from a modelling perspective, since the motivation for the definition of a Boolean algebra was precisely to guarantee that the events *a and b*, *a or b*, and *not a* are represented whenever a and b are. In a situation where the status of certain combinations of events are not simultaneously decidable, this motivation falls apart.

---

[6] Each $\omega \in S(\mathcal{A})$ is equivalent to an *ultrafilter* on $\mathcal{A}$ – namely the preimage $\omega^{-1}(1)$.

[7] We will generally opt for the term *Stone spectrum* when referring to the Stone space $S(\mathcal{A})$ of a particular Boolean algebra $\mathcal{A}$ (see e.g. [10, 37]), as the term *Stone space* usually refers to any compact totally disconnected Hausdorff space.

For this reason, one approach to contextual logic is through the use of *partial* Boolean algebras, originally introduced by Kochen and Specker [30]. The following formulation is adapted from [37].

Definition 4.12. A **partial Boolean algebra** consists of a set $\mathcal{A}$ together with:

- A reflexive and symmetric binary **commensurability** relation $\odot \subseteq \mathcal{A} \times \mathcal{A}$;
- Elements $\bot, \top \in \mathcal{A}$;
- A (total) unary operation $\neg : \mathcal{A} \to \mathcal{A}$;
- (Partial) binary operations $\wedge, \vee : \odot \to \mathcal{A}$;

such that every set $C \subset \mathcal{A}$ of pairwise commensurable elements $a \odot b$ is contained in a set $C \subset C' \subset \mathcal{A}$ of pairwise commensurable elements, on which the above operations and elements define a Boolean algebra structure.[8]

A Boolean algebra is a partial Boolean algebra where the commensurability relation $\odot$ is total. We will therefore refer to Boolean algebras also as *total* Boolean algebras. A morphism of partial Boolean algebras $\mathcal{A}$ and $\mathcal{B}$ is a map $f : \mathcal{A} \to \mathcal{B}$ such that $f(a) \odot f(b)$ whenever $a \odot b$, and which preserves the operations wherever defined. Partial Boolean algebras equipped with these morphisms form a category, which we denote **pBA** (following [1]).

Remark 4.13. While a total Boolean algebra can be equivalently defined as a partially ordered set, the same is not true for *partial* Boolean algebras, since the commensurability condition $\odot$ is generally not transitive.

The **contexts** of a partial Boolean algebra $\mathcal{A}$ are sets of mutually commensurable events. It is part of the definition that each context is included in some total Boolean subalgebra, so we generally restrict attention to contexts that are themselves total Boolean algebras. If $C \subset \mathcal{A}$ is a total Boolean subalgebra, we can construct its Stone space $S(C)$. The assignment of Stone spectra to Boolean algebras defines a contravariant functor,[9] which we call the **Stone presheaf**

$$S : \mathbf{BA}^{\mathrm{op}} \to \mathbf{Stone}, \tag{4.13}$$

where **Stone** is the category of Stone spaces. Restricting this presheaf to the total Boolean subalgebras $C \subset \mathcal{A}$ of a given partial Boolean algebra $\mathcal{A}$ yields an outcome space $S(C)$ for every context C, but no *global* outcome space for the whole set $\mathcal{A}$.

Example 4.14. If $\mathcal{H}$ is a Hilbert space, the set $\mathrm{Proj}(\mathcal{H})$ of projections on $\mathcal{H}$ is a partial Boolean algebra, where $P_a \odot P_b$ if $P_a P_b = P_b P_a$. Meets, joins, and complements are given by

$$P_a \wedge P_b = P_a P_b, \qquad P_a \vee P_b = P_a + P_b - P_a P_b, \qquad \neg P_a = 1 - P_a.$$

The Stone construction provides an outcome space $S(C)$ for each set of commuting projections, but not for every projection at once.

Remark 4.15. The definition of a partial Boolean algebra takes commensurability to be a binary relation. This means that a partial Boolean algebra by definition satisfies Specker's principle. This suggests that while partial Boolean algebras can be useful for quantum mechanics, they cannot function as a setting for formulations of more general, *incoherent* theories of contextual logic and probability theory.

4.5. **The Kochen-Specker Theorem.** The problem of hidden variables in quantum mechanics asks whether observations made in quantum mechanics can be explained by a "hidden" theory where all quantum observables are given simultaneous, context-independent values. The Kochen-Specker theorem answers this question strongly in the negatory.[10]

Given a collection of partially overlapping "contexts" $\mathscr{C} = \{C_i\}_i$, each which is a Boolean algebra, we can build from it a *partial* Boolean algebra $\mathcal{A}$. Namely, we have the following theorem:

Theorem 4.16 (van den Berg-Heunen)**.** *The category* **pBA** *is complete and co-complete.*

Proof. See [37]. □

[8] In particular, we must have $\bot, \top \in C'$, and $\neg a \in C'$ for each $a \in C'$.

[9] That S is a *contravariant functor* (i.e., a *presheaf*) means that every Boolean algebra morphism $f : \mathcal{A} \to \mathcal{B}$ corresponds to a continuous function in the opposite direction, $S(f) = f^* : S(\mathcal{B}) \to S(\mathcal{A})$.

[10] This section is largely based on the papers [10, 37].

Corollary 4.17. *If $\mathscr{C} = \{C_i\}_i$ is a directed set of Boolean algebras, there exists a unique (up to isomorphism) partial Boolean algebra*

$$\mathcal{A} = \operatorname{colim} \mathscr{C}. \tag{4.14}$$

What this means is (by definition of $\mathcal{A}$ being a colimit), that if we have a collection of truth valuations (Boolean algebra morphisms) $\omega_i : C_i \to 2$, one for every context $C_i \in \mathscr{C}$, and if these *agree on overlaps* in the sense that

$$\omega_C|_{C_i \cap C_j} = \omega_{C_j}|_{C_i \cap C_j}, \qquad C_i, C_j \in \mathscr{C}; \tag{4.15}$$

then there is a unique morphism of partial Boolean algebras $\omega : \mathcal{A} \to 2$ making the diagram

$$\begin{array}{ccc} & & \mathcal{A} \\ & \nearrow & \vdots\,\omega \\ & & \downarrow \\ C_i & \xrightarrow[\omega_i]{} & 2 \end{array} \tag{4.16}$$

commute for every i.

Furthermore, a partial Boolean algebra is characterised precisely as the colimit of its total Boolean subalgebras (contexts).

Theorem 4.18. *If $\mathcal{A}$ is a partial Boolean algebra, and $\mathscr{C}(\mathcal{A})$ is the collection of total Boolean subalgebras of $\mathcal{A}$, ordered by inclusion, then*

$$\mathcal{A} = \operatorname{colim} \mathscr{C}(\mathcal{A}). \tag{4.17}$$

In particular, a global truth valuation (partial Boolean algebra morphism) $\omega : \mathcal{A} \to 2$ exists if and only if there is a collection of compatible local truth valuations $\omega_i : C_i \to 2$, defined on the total Boolean subalgebras of $\mathcal{A}$. The Kochen-Specker theorem says that for a sufficiently high-dimensional quantum system, *no* such collection of valuations can exist.

Theorem 4.19 (Kochen-Specker)**.** *Let $\mathcal{A} = \operatorname{Proj}(\mathbb{C}^3)$, considered as a partial Boolean algebra. If $\mathcal{B}$ is a total Boolean algebra and $\omega : \mathcal{A} \to \mathcal{B}$ is a partial Boolean algebra morphism, then $\mathcal{B} = 1$ is the trivial Boolean algebra.*

Proof. See [10, 30]. □

Corollary 4.20. *There exists no morphism of partial Boolean algebras* $\operatorname{Proj}(\mathbb{C}^3) \to 2$.

In other words, it is impossible to assign all events (projections) on $\operatorname{Proj}(\mathbb{C}^3)$ consistent truth values. As $\operatorname{Proj}(\mathbb{C}^3)$ is included as a subalgebra of $\operatorname{Proj}(\mathcal{H})$ for any higher-dimensional Hilbert space $\mathcal{H}$, the problem only gets worse in higher dimension. The Kochen-Specker theorem strongly obstructs the existence of any hidden variable model for quantum mechanics.

5. Outlook

This article can only hope to give a small glimpse into the field of contextuality. Much work has been done on the subject, and much is not yet fully understood (by anyone, and much less than that by this particular author). Let us therefore finish with a brief outlook.

We have not covered the topic of *weak* contextuality beyond the illustrating example in Section 2. As long as we are content with using a partial Boolean algebra $\mathcal{A}$ as our event set, it seems natural to generalise the notion of probability distributions from Boolean algebras to partial Boolean algebras; as a collection of probability measures $\mu_C : C \to [0, 1]$ defined on total Boolean subalgebras C.

If $\mathcal{A}$ is not embeddable into a total Boolean algebra, then there is no hope of realising the μ_C as marginals of a global distribution, as there are no global outcomes to assign probabilities in the first place. However, even if $\mathcal{A}$ *is* a total Boolean algebra (or is embeddable into one), it may be that the μ_C cannot be extended to a global distribution. Then the μ_C exhibit weak, but not strong contextuality.

A conceptual difference between strong and weak contextuality as loosely defined above is that if a collection of local outcomes $\omega_i : C_i \to 2$ agree on overlaps, then they can always be glued to a global truth valuation. The presence of strong contextuality, which rules out global truth valuations, hence also rules out the existence of a compatible family of *local* truth valuations covering the space of events. This is a familiar scenario from topology and geometry, where local functions may always be glued to global functions if they agree on the overlaps of their domains.

However, it is possible to find local *probability distributions* $\mu_i : C_i \to [0,1]$ which agree on overlaps and cover the entire space $\mathcal{A}$ of events, but nevertheless do not glue to any global distribution. This is because the right notion of "restriction" for probability distributions is *marginalisation*, which loses information about correlations. Conversely, it is possible that there is *no* global set of correlations that restrict to a given set of "local" correlations.

In conclusion, while strong contextuality appears as the lack of a global family of compatible local sections, weak contextuality can appear as the failure of a global family of compatible local sections to glue. The slogan is that *weak contextuality is the failure of the presheaf of local probability distributions to be a sheaf.*

In future work, I would like to address this idea in more detail. It is my belief that a sheaf-theoretical formulation of probability theory is not only the right arena to study contextuality in, but arguably is as much as natural and rich a framework for probability theory as it has proven to be for topology, geometry, and algebra. Possibly, much of this could already be understood by experts in point-free approaches to probability theory; but less so by the quantum mechanics community. I hope this article can be of some help in connecting these communities.

Finally, let us mention some interesting approaches to contextuality that we have not had time to discuss at any length here. As discussed in the introduction, two of the main approaches relying on the theory of sheaves and topoi to capture the idea of contextuality as "local to global incompatibility of information" are:

- The sheaf-theoretical approach to contextuality pioneered by Abramsky and Brandenburger in [4], de-emphasises the role of Boolean algebras, and instead studies general value assignments to sets of commensurable observables. Emphasis is placed on the potential use of cohomological methods to observe contextuality. See [2, 4, 5, 6, 7].
- The sheaf-theoretic approach of Abramsky et al was preceded, and at least in part inspired by, the topos-theoretical formulation of the Kochen-Specker theorem by Isham and Butterfield [16, 17, 26, 29]; from which eventually grew the ambitious programme of *Bohrification*. See [18, 19, 20, 21, 26, 27]. The formulation of the Kochen-Specker theorem used in this article is based on [10, 37], exploiting (and contributing to) this perspective.

There is far too much ongoing research related to contextuality to list all relevant directions here, but here are at least two other honourable mentions:

- Fritz and others have explored category theoretical foundations of probability theory, including the introduction of *Markov categories*. The basic idea is to treat probability as a theory of stochastic processes, or "noisy maps", which can be modelled as arrows in an appropriate category. See [22, 23, 24].
- In the literature, one also finds the notion of *Spekkens contextuality*, which is another formalisation of contextuality which can be applied also to non-quantum scenarios; and which precedes Abramsky Brandenburger contextuality by a few years. I will make no attempt to recount Spekkens' formalisation of contextuality here – instead I refer to his quite readable original paper [34]. It is however worth noting that despite Spekkens' definition of contextuality making no mention of sheaves, and being overall quite different in nature from Abramsky and Brandenburger's definition, Wester has shown that the two definitions coincide for a large subclass of scenarios [39] (see however Example 1 of Section 6 of the above reference for an example of when they differ).

The literature relevant to contextuality is vast and difficult to overview. I hope this article can at least open a window to a small corner of it.

Acknowledgements

I would like to thank Politecnico di Milano (Polimi) for organising the intensive period on quantum mathematics, and for inviting me to contribute to this special issue – with special thanks to Alessandro Olgiati for his tireless organisational work. I am very grateful to Douglas Lundholm and Tore Ellingsen for thoroughly proofreading the manuscript and giving helpful suggestions, comments and corrections. I also thank Lasse Svensson, Filippa Getzner, and Ella Wiiand for additional corrections, comments, and good questions. I especially want to thank the anonymous referee for their constructive feedback, interesting remarks, and additional references that greatly improved the final version of this article. Finally, I gratefully

acknowledge financial support from Polimi, from C.F. Liljewalchs Stipendiestiftelse, and from the Center for Interdisciplinary Mathematics at Uppsala University.

Declarations and statements

Research funding. This research was funded in part by the Center for Interdisciplinary Mathematics (CIM) at Uppsala University.

Conflict of interest. The author declares that he has no competing interests regarding the publication of this paper.

Availability of data. There are no data associated with the research in this paper.

References

[1] Samson Abramsky and Rui Soares Barbosa. "The Logic of Contextuality". In: *29th EACSL Annual Conference on Computer Science Logic (CSL 2021)*. Ed. by Christel Baier and Jean Goubault-Larrecq. Vol. 183. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021, 5:1–5:18. ISBN: 978-3-95977-175-7. DOI: 10.4230/LIPIcs.CSL.2021.5. URL: https://drops.dagstuhl.de/entities/document/10.4230/LIPIcs.CSL.2021.5 (visited on 02/26/2026).

[2] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. "Contextual Fraction as a Measure of Contextuality". In: *Phys. Rev. Lett.* 119.5 (Aug. 4, 2017), p. 050504. DOI: 10.1103/PhysRevLett.119.050504. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.050504 (visited on 02/28/2026).

[3] Samson Abramsky and Adam Brandenburger. "The Sheaf-Theoretic Structure Of Non-Locality and Contextuality". en. In: *New Journal of Physics* 13.11 (Nov. 2011). arXiv:1102.0264 [quant-ph], p. 113036. ISSN: 1367-2630. DOI: 10.1088/1367-2630/13/11/113036. URL: http://arxiv.org/abs/1102.0264 (visited on 02/25/2026).

[4] Samson Abramsky and Adam Brandenburger. "The Sheaf-Theoretic Structure Of Non-Locality and Contextuality". In: *New J. Phys.* 13.11 (Nov. 28, 2011), p. 113036. ISSN: 1367-2630. DOI: 10.1088/1367-2630/13/11/113036. arXiv: 1102.0264[quant-ph]. URL: http://arxiv.org/abs/1102.0264 (visited on 02/25/2026).

[5] Samson Abramsky, Shane Mansfield, and Rui Soares Barbosa. "The Cohomology of Non-Locality and Contextuality". In: *Electron. Proc. Theor. Comput. Sci.* 95 (Oct. 1, 2012), pp. 1–14. ISSN: 2075-2180. DOI: 10.4204/EPTCS.95.1. URL: http://arxiv.org/abs/1111.3620v2 (visited on 02/25/2026).

[6] Samson Abramsky et al. "Contextuality, Cohomology and Paradox". In: *LIPIcs, Volume 41, CSL 2015* 41 (2015), pp. 211–228. ISSN: 1868-8969. DOI: 10.4230/LIPIcs.CSL.2015.211. arXiv: 1502.03097[quant-ph]. URL: http://arxiv.org/abs/1502.03097 (visited on 02/25/2026).

[7] Rui Soares Barbosa et al. "Continuous-Variable Nonlocality and Contextuality". In: *Commun. Math. Phys.* 391.3 (May 1, 2022), pp. 1047–1089. ISSN: 1432-0916. DOI: 10.1007/s00220-021-04285-7. URL: https://doi.org/10.1007/s00220-021-04285-7 (visited on 02/28/2026).

[8] J. S. Bell. "On the Einstein Podolsky Rosen paradox". In: *Physics Physique Fizika* 1 (3 1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195. URL: https://link.aps.org/doi/10.1103/PhysicsPhysiqueFizika.1.195.

[9] John S. Bell. "On the Problem of Hidden Variables in Quantum Mechanics". In: *Rev. Mod. Phys.* 38.3 (July 1, 1966), pp. 447–452. DOI: 10.1103/RevModPhys.38.447. URL: https://link.aps.org/doi/10.1103/RevModPhys.38.447 (visited on 02/28/2026).

[10] Benno van den Berg and Chris Heunen. *Extending obstructions to noncommutative functorial spectra*. July 10, 2014. DOI: 10.48550/arXiv.1407.2745. arXiv: 1407.2745[math]. URL: http://arxiv.org/abs/1407.2745 (visited on 02/25/2026).

[11] Garrett Birkhoff and John Von Neumann. "The Logic of Quantum Mechanics". In: *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution*. Ed. by C. A. Hooker. Dordrecht: Springer Netherlands, 1975, pp. 1–26. ISBN: 978-94-010-1795-4. DOI: 10.1007/978-94-010-1795-4_1. URL: https://doi.org/10.1007/978-94-010-1795-4_1 (visited on 02/28/2026).

[12] George Boole. *An investigation of the laws of thought: on which are founded the mathematical theories of logic and probabilities*. Walton and Maberly, 1854.

[13] Max Born. "Statistical Interpretation of Quantum Mechanics". In: *Science* 122.3172 (Oct. 14, 1955), pp. 675–679. DOI: 10.1126/science.122.3172.675. URL: https://www.science.org/doi/abs/10.1126/science.122.3172.675 (visited on 02/28/2026).

[14] Max Born. "Quantum mechanics of collision processes (1)." In: ().

[15] Costantino Budroni et al. "Kochen-Specker contextuality". In: *Rev. Mod. Phys.* 94.4 (Dec. 19, 2022), p. 045007. DOI: 10.1103/RevModPhys.94.045007. URL: https://link.aps.org/doi/10.1103/RevModPhys.94.045007 (visited on 02/28/2026).

[16] J. Butterfield and Chris J. Isham. "A Topos Perspective on the Kochen-Specker Theorem II. Conceptual Aspects and Classical Analogues". In: *International Journal of Theoretical Physics* 38.3 (Mar. 1, 1999), pp. 827–859. ISSN: 1572-9575. DOI: 10.1023/A:1026652817988. URL: https://doi.org/10.1023/A:1026652817988 (visited on 02/28/2026).

[17] J. Butterfield and Chris J. Isham. "Topos Perspective on the Kochen–Specker Theorem IV. Interval Valuations". In: *International Journal of Theoretical Physics* 41.4 (Apr. 1, 2002), pp. 613–639. ISSN: 1572-9575. DOI: 10.1023/A:1015276209768. URL: https://doi.org/10.1023/A:1015276209768 (visited on 02/28/2026).

[18] Andreas Döring. *Spectral presheaves as quantum state spaces*. (Visited on 02/28/2026).

[19] Andreas Döring and Chris Isham. ""What is a Thing?": Topos Theory in the Foundations of Physics". In: *New Structures for Physics*. Ed. by Bob Coecke. Berlin, Heidelberg: Springer, 2011, pp. 753–937. ISBN: 978-3-642-12821-9. DOI: 10.1007/978-3-642-12821-9_13. URL: https://doi.org/10.1007/978-3-642-12821-9_13 (visited on 02/28/2026).

[20] Andreas Döring and Chris J. Isham. *A topos foundation for theories of physics: I. Formal languages for physics*. URL: https://pubs.aip.org/aip/jmp/article-abstract/49/5/053515/911423/A-topos-foundation-for-theories-of-physics-I (visited on 02/28/2026).

[21] Andreas Döring and Chris J. Isham. *A topos foundation for theories of physics: II. Daseinisation and the liberation of quantum theory*. URL: https://pubs.aip.org/aip/jmp/article-abstract/49/5/053516/911247/A-topos-foundation-for-theories-of-physics-II?redirectedFrom=fulltext (visited on 02/28/2026).

[22] Tobias Fritz. "A synthetic approach to Markov kernels, conditional independence and theorems on sufficient statistics". In: *Advances in Mathematics* 370 (Aug. 26, 2020), p. 107239. ISSN: 0001-8708. DOI: 10.1016/j.aim.2020.107239. URL: https://www.sciencedirect.com/science/article/pii/S0001870820302656 (visited on 02/28/2026).

[23] Tobias Fritz and Eigil Fjeldgren Rischel. "Infinite products and zero-one laws in categorical probability". In: *Compositionality* Volume 2 (2020) (Aug. 11, 2020). ISSN: 2631-4444. DOI: 10.32408/compositionality-2-3. URL: https://compositionality.episciences.org/13509 (visited on 02/28/2026).

[24] Tobias Fritz et al. "Representable Markov categories and comparison of statistical experiments in categorical probability". In: *Theoretical Computer Science* 961 (June 15, 2023), p. 113896. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2023.113896. URL: https://www.sciencedirect.com/science/article/pii/S0304397523002098 (visited on 02/28/2026).

[25] Richard D. Gill. "Gull's Theorem Revisited". In: *Entropy* 24.5 (May 2022), p. 679. ISSN: 1099-4300. DOI: 10.3390/e24050679. URL: https://www.mdpi.com/1099-4300/24/5/679 (visited on 08/05/2026).

[26] J. Hamilton, Chris J. Isham, and J. Butterfield. "Topos Perspective on the Kochen-Specker Theorem III. Von Neumann Algebras as the Base Category". In: *International Journal of Theoretical Physics* 39.6 (June 1, 2000), pp. 1413–1436. ISSN: 1572-9575. DOI: 10.1023/A:1003667607842. URL: https://doi.org/10.1023/A:1003667607842 (visited on 02/28/2026).

[27] Chris Heunen, Nicolaas P. Landsman, and Bas Spitters. *Bohrification*. July 16, 2010. DOI: 10.48550/arXiv.0909.3468. arXiv: 0909.3468[quant-ph]. URL: http://arxiv.org/abs/0909.3468 (visited on 02/25/2026).

[28] Chris Heunen et al. "The Gelfand spectrum of a noncommutative C*-algebra: A topos-theoretic approach". In: *Journal of the Australian Mathematical Society* 90.1 (Feb. 2011), pp. 39–52. ISSN: 1446-8107, 1446-7887. DOI: 10.1017/S1446788711001157. URL: https://www.cambridge.org/core/journals/journal-of-the-australian-mathematical-society/article/gelfand-spectrum-of-a-noncommutative-calgebra-a-toposttheoretic-approach/9C7D009CFB824F0B6CA6109D60EAB7C7 (visited on 02/25/2026).

[29] Chris J. Isham and J. Butterfield. "Topos Perspective on the Kochen-Specker Theorem: I. Quantum States as Generalized Valuations". In: *International Journal of Theoretical Physics* 37.11 (Nov. 1, 1998), pp. 2669–2733. ISSN: 1572-9575. DOI: 10.1023/A:1026680806775. URL: https://doi.org/10.1023/A:1026680806775 (visited on 02/28/2026).

[30] Simon Kochen and E. P. Specker. "The Problem of Hidden Variables in Quantum Mechanics". In: *Ernst Specker Selecta*. Ed. by Gerhard Jäger et al. Basel: Birkhäuser, 1990, pp. 235–263. ISBN: 978-3-0348-9259-9. DOI: 10.1007/978-3-0348-9259-9_21. URL: https://doi.org/10.1007/978-3-0348-9259-9_21 (visited on 02/28/2026).

[31] Sandu Popescu and Daniel Rohrlich. "Quantum nonlocality as an axiom". In: *Foundations of Physics* 24.3 (1994), pp. 379–385.

[32] Robert Raussendorf. "Contextuality in Measurement-based Quantum Computation". In: *Physical Review A* 88.2 (Aug. 19, 2013), p. 022322. ISSN: 1050-2947, 1094-1622. DOI: 10.1103/PhysRevA.88.022322. arXiv: 0907.5449 [quant-ph]. URL: http://arxiv.org/abs/0907.5449 (visited on 08/07/2026).

[33] Roman Sikorski et al. *Boolean algebras*. Vol. 2. Springer, 1969.

[34] R. W. Spekkens. "Contextuality for Preparations, Transformations, and Unsharp Measurements". In: *Physical Review A* 71.5 (May 31, 2005), p. 052108. DOI: 10.1103/PhysRevA.71.052108. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.052108 (visited on 08/07/2026).

[35] M. H. Stone. "The Theory of Representation for Boolean Algebras". In: *Transactions of the American Mathematical Society* 40.1 (1936), pp. 37–111. ISSN: 0002-9947. DOI: 10.2307/1989664. URL: https://www.jstor.org/stable/1989664 (visited on 03/01/2026).

[36] Terence Tao. *245B notes 4: The Stone and Loomis-Sikorski representation theorems (optional)*. What's new. Jan. 12, 2009. URL: https://terrytao.wordpress.com/2009/01/12/245b-notes-1-the-stone-and-loomis-sikorski-representation-theorems-optional/ (visited on 03/01/2026).

[37] Benno Van Den Berg and Chris Heunen. "Noncommutativity as a Colimit". In: *Appl Categor Struct* 20.4 (Aug. 2012), pp. 393–414. ISSN: 0927-2852, 1572-9095. DOI: 10.1007/s10485-011-9246-3. URL: http://link.springer.com/10.1007/s10485-011-9246-3 (visited on 02/25/2026).

[38] Nikolai Nikolaevich Vorob'ev. "Consistent families of measures and their extensions". In: *Theory of Probability & Its Applications* 7.2 (1962), pp. 147–163.

[39] Linde Wester. "Almost Equivalent Paradigms of Contextuality". In: *Electronic Proceedings in Theoretical Computer Science* 266 (Feb. 27, 2018), pp. 1–22. ISSN: 2075-2180. DOI: 10.4204/EPTCS.266.1. URL: http://arxiv.org/abs/1708.00843v2 (visited on 08/09/2026).