Two Forms of Inconsistency in Quantum Foundations

Jeremy Steeger
jsteeger@nd.edu

Nicholas Teh
nteh@nd.edu

December 4, 2017

Abstract

Recently, there has been some discussion of how Dutch Book arguments might be used to demonstrate the rational incoherence of certain hidden variable models of quantum theory (Feintzeig, 2015; Feintzeig and Fletcher, 2017; Wroński and Godziszewski, 2017). In this paper, we argue that the 'form of inconsistency' underlying this alleged irrationality is deeply and comprehensively related to the more familiar 'inconsistency' phenomenon of contextuality. Our main result is that the hierarchy of contextuality due to Abramsky and Brandenburger (2011) corresponds to a hierarchy of additivity/convexity-violations which yields formal Dutch Books of different strengths. We then use this result to provide a partial assessment of whether these formal Dutch Books can be interpreted normatively.

1 Introduction

Within the philosophical foundations of quantum theory, there is a long tradition of attempting to argue that the predictions of quantum theory cannot be accounted for by classical *hidden variable theories* **(HVTs)** of a certain kind. Very roughly speaking, such 'no-go' arguments take the following form:

> **(Standard)** There exists a quantum model whose empirical probabilities are inconsistent with the probabilities that can be reproduced by the relevant HVT.[1]

The development of this tradition has since resulted in a veritable cornucopia of models witnessing such inconsistencies. Thus, one important avenue of contemporary research (Abramsky and Brandenburger, 2011; Acín et al., 2012; Liang et al., 2011) seeks to impose order on this landscape by uncovering what we shall refer to as 'forms of inconsistency', i.e. high-level structures that provide a unifying explanation—and perhaps even a classification—of how these models give rise to such inconsistencies. The goal of this paper is to argue for a thoroughgoing and novel connection between two forms of inconsistency in quantum foundations, and to apply this result to several philosophical puzzles.

The first form, which has a long and distinguished pedigree, is the *Gluing Inconsistency* that is suggested by 'contextuality arguments' against certain HVTs, as exemplified by the Kochen-Specker

[1] We will give more detail in Sections 2 and 3 about the class of HVTs that we wish to consider. In brief: we will only consider *non-contextual* 'factorizable HVTs' in the sense of (Abramsky and Brandenburger, 2011, Section 8); thus, (Standard) arguments do not apply to Bohmian mechanics, which is both contextual and non-local.

theorem (Kochen and Specker, 1975). Its structural essence has recently been rigorously articulated by Abramsky and Brandenburger (2011) (henceforth **AB**), who not only generalize it to apply to a large class of physical theories, but also provide a three-tiered 'hierarchy of contextuality' which classifies the ways in which physical models can deviate from the classical probability assumed by an HVT.

The second form, which we call *Dutch Bookability*, is somewhat more recent and thus less explored; indeed, much of the work of this paper will be to show that it really is a 'form of inconsistency' in our sense. In its simplest incarnation, it arises in (Feintzeig, 2015) and (Feintzeig and Fletcher, 2017) and appears to differ from the first form in two ways. First, it is not presented as arising from an argument against an HVT, but rather against a *Generalized HVT* **(GHVT)**. Second, the 'inconsistency' in question stems from the fact that the GHVT is 'Dutch Bookable' in the sense that it violates a certain formal constraint. Feintzeig and Fletcher (2017) then make a move that has become standard in philosophical discussions concerning *classical* probability: they interpret this constraint normatively, i.e. as a rational constraint on the credences of an epistemic agent, and thus take its violation to provide a no-go argument against non-contextual GHVT models of quantum theory. For the sake of clarity, we will distinguish between the formal constraint and its normative intepretation by using 'Dutch Bookability' to refer to the violation of the constraint, and '*normative* Dutch Bookability' to refer to the normative interpretation of this violation.

Despite these apparent differences, we will argue that Gluing Inconsistency and Dutch Bookability are really two ways of understanding the same set of phenomena which fall under (Standard), because each form allows us to represent the empirical probabilities of a physical model and to classify the ways in which certain HVTs fail to reproduce these probabilities.[2] Our argument culminates in the result (Theorem 1) that, for a large class of physical models (including quantum models), AB's three-tiered hierarchy of Gluing Inconsistency corresponds to a three-tiered hierarchy of Dutch Bookability, which provides the aforementioned classification.

This result yields powerful applications to two puzzles raised by the recent exploration of the relationship between Dutch Books and quantum theory. First, the only connection between GHVT models of quantum theory and Dutch Bookability that has appeared in the literature (Feintzeig, 2015; Feintzeig and Fletcher, 2017) stems from an incredibly strong (indeed maximal!) violation of 'subadditivity' which leads to Dutch Books.[3] What explanation can be given of the strength of this violation? Our main result allows us to answer this question by explaining the strength of the violation in terms of AB's three-tiered classification of the contextuality exhibited by quantum (and non-quantum) models: the 'maximal Dutch Bookability' of the above violation derives from its being the strongest form of contextuality (exemplified in quantum theory by Kochen-Specker models), but there are also quantum models which exhibit weaker forms of Dutch Bookability corresponding to the second and third tier respectively, viz. the Hardy model (Hardy, 1993) and the Bell model (Bell, 1964).

Second, what room is there to interpret Dutch Bookability normatively, i.e. as a no-go argument against non-contextual GHVTs? While we do not provide a comprehensive answer to this question, our main result will turn out to undermine the applicability of the standard normative interpretation of Dutch Bookability to a large class of GHVT models of quantum theory (including the GHVT

[2] More carefully, we take this to be the minimal interpretation of Dutch Bookability; nothing yet that we have said militates against its further interpretation as providing a no-go argument against GHVTs.

[3] We note that in the standard Dutch Book literature, it is typical to discuss the violation of 'additivity' (both finite and countable), but we have not been able to find discussions of subadditivity-violation, let alone *maximal* violations of this kind!

versions of the Kochen-Specker and Hardy models).

We proceed as follows. Section 2 reviews relevant background concepts such as 'contextuality' and offers a non-technical account of the two forms of inconsistency. Section 3 then provides an elementary introduction to AB's Gluing Inconsistency and explains how the features of this framework can be represented within the probabilistic setting assumed by a GHVT. Section 4 presents our main result, viz. the three-tiered correspondence between Gluing Inconsistency and Dutch Bookability (with proofs relegated to the Appendix). Finally, Section 5 discusses the prospects for a normative interpretation of Dutch Bookability.

2 Contextuality and Inconsistency

2.1 Contextuality

The subject of 'contextuality' was introduced into quantum foundations in (Bell, 1966) and (Kochen and Specker, 1975), after which it has become common to characterize quantum theory as 'contextual'. In order to explain what this means, we now review several preliminary definitions. First, a (measurement) *context* is a set of 'co-measurable' or 'compatible' measurements, i.e. measurements that can be made jointly (thus, in the particular case of quantum theory, a 'context' refers to a set of commuting observables). Second, a *non-contextual* **(NC)** HVT is one whose 'response functions' (i.e. functions that specify the probabilities of measurement events conditional on the system's being in some particular state) are independent of any information about contexts.[4] Third, we define a *possibility assignment* as an assignment of either 1 or 0 to events. Following Spekkens (2005), we will use the term *outcome deterministic* **(OD)** to refer to an HVT whose response functions only make possibility assignments.

The characterization of quantum theory as 'contextual' then arises through the following 'quantum contextuality argument', which is a special case of (Standard): there exists a quantum model whose empirical possibilities are inconsistent with the response functions of an (NC, OD) HVT. We say that quantum theory is 'contextual' because such quantum models exist.

We now discuss how quantum contextuality suggests a geometric form of inconsistency that we shall call 'Gluing Inconsistency'. One classic way of demonstrating quantum contextuality is by means of a (discrete) Kochen-Specker model, i.e. a quantum model for which 'measurements' correspond to (a finite set of) projectors, 'maximal contexts' are given by sets of mutually orthogonal projectors that resolve to the identity, and the empirical possibilities are assumed to satisfy certain functional constraints.[5] The inconsistency between the empirical probabilities of a Kochen-Specker model and those generated by an (NC) possibilistic HVT can then be rendered geometrically, i.e. as the impossibility of consistently 'gluing together' orthonormal bases (corresponding to the maximal contexts) with the following possibility assignments: each orthonormal basis has exactly one basis vector that is assigned the possibility 1 and all other basis vectors are assigned the possibility 0. For instance, an impossible gluing scenario that corresponds to the Kochen-Specker model in (Cabello et al., 1996) is depicted in Fig. 2.1, where the vertices represent possible measurements,

[4] We note that Spekkens (2005) generalizes the above definitions of 'context' and 'non-contextual HVT' to apply to contexts of preparations and transformations of systems as well; however, we will only be considering measurement contexts in this paper.

[5] In particular, the assignment $h : \mathcal{P}(\mathcal{H}) \to \{0, 1\}$ of possibilities must respect the additivity of projectors; thus, h will be a 2-valued partial algebra homomorphism, as originally defined by Kochen and Specker (1975). In other words, we require that $h(\perp) = 0$ and for any set $\{P_i\}$ of mutually orthogonal projectors, $h\left(\bigvee_i P_i\right) = \sum_i h(P_i)$.

the rectangles represent maximal contexts, and the vertices are colored black for possibility 1, and white for possibility 0.

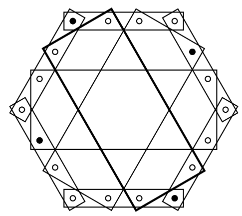


Figure 2.1: An inconsistent possibility assignment for Cabello's Kochen-Specker model.

Such models present us with a compelling image of inconsistency which admits of an intuitive parallel with the global geometric inconsistency of 'impossible figures' such as Penrose's Triangle (Penrose, 1992) and Escher's staircase. As we will see in Section 3.1, AB's account of Gluing Inconsistency provides both a precisification and a generalization of this core geometric intuition.

2.2 Gluing Inconsistency

Although the geometric intuition underlying Gluing Inconsistency has long been implicit in the quantum contextuality tradition, it was not until relatively recently that Abramsky and Brandenburger (2011) **(AB)** used a rigorous framework—viz. sheaf theory—to distill its essence and generalize it far beyond the confines of quantum contextuality.

The generalization offered by AB's framework is one that we shall refer to as *theory-independence*, because it allows us to consider models of a large class of theories, including *non-quantum* theories. Theory-independence can be motivated by noting that (Standard) arguments function by demonstrating an inconsistency between the 'structure of the empirical probabilities'—which happen to be generated by a quantum model—on the one hand, and a certain non-contextual HVT, on the other. This observation suggests that one might be able to devise theory-independent versions of such arguments, in which the empirical probabilities do not necessarily come from quantum theory.

In order to do so, AB introduce a formal generalization of 'the empirical probabilities of a quantum model', which they call an 'empirical model'. Roughly speaking, an empirical model is a family of (classical) probability distributions that satisfies a generalized no-signaling condition (cf. Section 3.1 for the details). *Gluing Inconsistency* is then defined as the failure of such a family of distributions to 'glue together' to form a global/joint distribution. Furthermore, since AB show that the existence of such joint distributions is equivalent to the existence of a certain classical HVT, this 'gluing failure' can be understood as a measure of the extent to which an empirical model deviates from classical probability.

One of the most attractive features of AB's framework is that it clearly picks out certain features of empirical models which characterize the degree to which the model exhibits Gluing Inconsistency; henceforth, we will use the term 'contextual' to describe models with such features (we will continue to use the term 'quantum contextual' to describe quantum models that deviate from classical possibility). AB (2011) were thus able to provide the first comprehensive classification of contextuality in terms of a three-tiered hierarchy, where a higher tier (a stronger failure of gluing) implies a lower tier (a weaker failure of gluing), but not vice versa. As we will soon see, the top and middle tier of the hierarchy involve only empirical *possibilities*, whereas the lowest tier of this hierarchy is fully *probabilistic* in its characterization. In particular, the GHZ model and all Kochen-Specker models fall into the top tier, the Hardy model falls into the middle tier (but not the top), and the Bell model falls into the lowest tier (but not the middle). In sum, the 'AB hierarchy' encompasses all known quantum models that witness (Standard), and classifies the different strengths with which they deviate from classical probability.

2.3 Dutch Bookability

While the notion of Gluing Inconsistency has a straightforward relationship with the quantum contextuality tradition, the same cannot be said for our second form of inconsistency, viz. Dutch Bookability. But a relationship is present nonetheless, and is most clearly seen by dividing an account of Dutch Bookability into two parts, viz. (i) the partial algebra approach; and (ii) the Dutch connection.

(i) The partial algebra approach

The key formal notion that Feintzeig and Fletcher (2017) use to establish a link between Dutch Books and quantum foundations is that of a Weak Probability Space (WPS), i.e. a generalization of a measure space to a setting in which the 'set of events' is not required to have any algebraic structure. However, their actual use of a WPS to represent quantum models still requires *some* algebraic structure, because *co-measurable* WPS events (i.e. those that arise from measurements within the same context) are required to form an algebra. Thus, these WPS representations possess a *partial* algebraic structure in which certain algebraic relations model the co-measurability of events stemming from compatible measurements, and the absence of these algebraic relations models the non-co-measurability of events stemming from incompatible measurements—we shall call this the 'partial algebra approach'.

The partial algebra approach has a long history within quantum foundations that is not motivated by Dutch Books, but rather by the recognition that (because of non-commuting observables) quantum models will in general give rise to measurements that fall into distinct maximal contexts; in other words, partial algebras are meant to capture the very structure that gives rise to 'quantum contextuality'. This use of partial algebras goes at least as far back as the original Kochen-Specker paper (Kochen and Specker, 1975), which employs the notion of a partial Boolean algebra for this purpose. Since then, the partial algebra approach (and its associated measure theory) has been developed in various ways. For instance, one strand (Roumen, 2016; Staton and Uijlen, 2015; van den Berg and Heunen, 2010) focuses on developing the approach in a purely algebraic manner, wherein the events do not have to be realized by a particular collection of sets. This work has resulted in elegant (indeed, category-theoretic) constructions as well as powerful characterization results that are made available by the appeal to algebraic methods.

On the other hand, a different strand of the approach attempts to cleave very closely to the standard notion of a measure space, in which events are realized by sets. As such, it has to reckon with minutiae such as first specifying an ambient set Y that will serve as the 'sample space' in order to further specify the set of events (some subset of the power-set of Y) that carries the partial structure. For instance, one attempt to work this out in detail is provided by the notion of a 'generalized measure space' in (Gudder, 1973). And since a WPS is a weakening of this notion, it is helpful to situate it within this strand of the partial algebra approach.

(ii) The Dutch connection

The connection between Dutch Books and the partial algebra approach has an antecedent in the work of Fine (1982), who essentially suggests that one might be able to avoid (Standard) arguments against HVTs if one generalizes these HVTs to account for non-co-measurability by appealing to a partial algebraic framework for probability—call these Generalized HVTs (GHVTs). The connection with Dutch Books then emerges when Feintzeig (2015) and Feintzeig and Fletcher (2017) attempt to mount a Dutch Book argument against a particular implementation of Fine's suggestion. We describe this in two steps.

First, they take the relevant GHVTs to at least have the structure of a WPS and argue as follows:

> **(Dutch)** There exists a quantum model Q such that any WPS that reproduces Q's empirical probabilities will violate a formal constraint, which in turn shows that it will be susceptible to a (formal) Dutch Book.[6]

Regardless of whether we are discussing quantum or non-quantum models, we will refer to this type of constraint-violation as *Dutch Bookability.*

Second, they interpret the Dutch Book normatively, thereby taking it to provide a successful no-go argument against *non-contextual* GHVTs of a certain kind. The question of providing enough detail to this argument so that one can evaluate it is certainly an interesting one: how, for instance, should 'non-contextual' be understood for such GHVTs? However, our focus in this paper will be on investigating a far more fundamental puzzle concerning (Dutch); we will briefly consider its implications for the normative interpretation of Dutch Books in Section 5.

This more fundamental puzzle concerns the relationship between (Dutch) and (Standard), as well as their theory-independent generalizations. Admittedly, if one forgets about the 'partial algebra approach' and focuses only on the 'Dutch connection' part of the story, the investigation of such a relationship might seem ill-motivated: after all, Fine's suggestion was meant to sidestep (Standard) from the get-go, so why expect (Dutch) to have any connection with (Standard)? But a moment's reflection on 'the partial algebra approach' makes it clear that this appearance is misleading, because the *raison d'etre* of the partial algebra tradition is to represent facts about contextuality, and it is precisely such facts that allow quantum models to feature in (Standard) arguments. Indeed, the detailed argument for (Dutch) in (Feintzeig and Fletcher, 2017) supports this claim: there, Q is taken to be a Kochen-Specker model, and (as we will see in Section 4.1) the argument for (Dutch) turns on precisely those features that render the model Gluing Inconsistent. Based on this evidence, it is reasonable to conjecture that an even more comprehensive and general relationship holds between Gluing Inconsistency and a *theory-independent* version of Dutch Bookability.

In what follows, we will argue that this conjecture is true. First, we will develop the tools to represent AB's notion of an 'empirical model' in terms of a WPS. We will then use these tools to argue for the following formulation of the conjecture: AB's hierarchy of contextuality corresponds to a three-tiered hierarchy of violations of 'formal constraints', each of which implies the existence of a formal Dutch Book. As we will see, the question of how to specify these 'formal constraints' is somewhat subtle, since there are different choices that one could make, each having its own respective virtues. For instance, Feintzeig and Fletcher (2017) take the formal constraint to be a strong form of sub-additivity (which has a measure-theoretic flavor), whereas Wroński and Godziszewski (2017) re-interpret the constraint in terms of convexity (which has a more direct formal connection with

[6] In fact, several different kinds of formal constraints imply formal Dutch Books—we further clarify this point in Section 4.

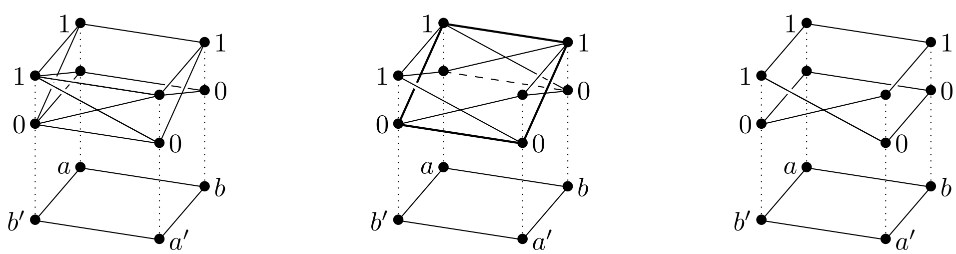
Dutch Books); Section 4 will treat both aspects. But regardless of which set of constraints one uses, the broader moral is the same: just like Gluing Inconsistency, Dutch Bookability is a 'form of inconsistency'—it represents the empirical probabilities of various physical models, and classifies the strength of their deviation from the classical probability assumed by an HVT.

3 Representing empirical models

3.1 Gluing Inconsistency: Empirical models

We now describe AB's (2011) framework for Gluing Inconsistency. Although Abramsky and collaborators (2011; 2015) deploy the full machinery of 'sheaves' and 'presheaves', we will not need to explicitly define these objects because we will only use their most elementary properties, which we will describe directly. Like AB, we will only discuss finite sets for the sake of simplicity.

We begin by defining an *empirical scenario* as a triple $(X, \mathcal{M}, O)$, where X is a set of measurements, $\mathcal{M}$ is the set of maximal contexts in the power-set $P(X)$, and O is the set of possible outcomes for each measurement (recall that a 'context' is a set of co-measurable/compatible measurements). We will also find it useful to refer to the set of all (possibly non-maximal) contexts, denoted $\mathcal{M}'$. Empirical scenarios can be represented by 'bundle diagrams': e.g. in Fig. 3.1, vertices of the 'base' represent $X = \{a, b, a', b'\}$, edges of the 'base' represent $\mathcal{M} = \{\{a, b\}, \{b, a'\}, \{a', b'\}, \{b', a\}\}$, and vertices over the 'base' represent $O = \{0, 1\}$.



(a) Bell model (b) Hardy model (c) PR box

Figure 3.1: Bundle diagrams for the support of e

Given a measurement scenario, we can define (measurement) events by means of the *event sheaf* $\mathcal{E}$, which maps each $U \subseteq X$ to the set of *events* $\mathcal{E}(U) := O^U$ (i.e. an event is a function from U to O), and which maps inclusions $U \subseteq U'$ to the obvious restriction $s \mapsto s|_U$, where $s \in \mathcal{E}(U')$.[7] When we want to emphasize the geometric character of events in the AB framework, we will refer to events as *sections* over U; if $U = X$, we say the sections are *global*, and if $U \neq X$, we say the sections are *local*. This geometric interpretation is illustrated in Fig. 3.1, where sections are represented as line segments over the base. For instance, in Fig. 3.1a, there are 4 local sections over the context $\{b', a'\}$.

We now highlight an obvious but important feature of events in the AB framework. Let $U \subseteq X$ be covered by $\{U_i\}$. We say that a family of sections $\{s_i \in \mathcal{E}(U_i)\}$ is *compatible* just in case

[7]Henceforth, we will use the symbol | to refer to both set-theoretic restriction and the sort of restriction specified here—in category-theoretic terms, the restriction along a morphism as specified by a presheaf (MacLane and Moerdijk, 1992, p. 25). The intended meaning will be clear from context.

$s_i|_{U_i \cap U_j} = s_j|_{U_i \cap U_j}$ for all i, j. The 'sheaf property' of $\mathcal{E}$ amounts to the following unremarkable fact: for every compatible family $\{s_i\}$, there there is a unique 'larger' section $s \in \mathcal{E}(U)$ that restricts to s_i on each U_i, viz. the function $s : U \to O$ that is constructed piecewise from $\{s_i\}$. In particular, when $U = X$ then the family $\{s_i\}$ gives rise to a unique global section (for instance, in Fig. 3.1b, the family of thick edges over the base glues together to form a unique global section).

Having defined events, we now proceed to define the probabilities of these events. Let a probability distribution on a finite set X be a function $d : X \to [0, 1]$ whose sum over elements of X is normalized to 1; we will use $\mathcal{D}(X)$ to denote the set of probability distributions on X. AB describe probabilities on the events of a measurement scenario by means of a 'pre-sheaf' $\mathcal{DE}$. We will only need two features of this object. First, $\mathcal{DE}$ maps a context $U \in \mathcal{M}'$ to the set of distributions $\mathcal{DE}(U)$ on the events (sections) over U; we will use $e_U \in \mathcal{DE}(U)$ to denote a distribution on this set.[8] Second, let $U, U' \in M'$ such that $U' \supset U$ and let A_s be the set of all $r \in \mathcal{E}(U')$ such that $r|_U = s$. $\mathcal{DE}$ defines a notion of 'distribution marginalization', viz. $e_{U'}|_U(s) := \sum_{r \in A_s} e_{U'}(r)$, which satisfies

$$e_U = e_{U'}|_U. \tag{3.1}$$

For the purposes of our discussion in Section 3.2, it will be important to note that AB's 'events' and 'distributions' constitute algebraic versions of set-theoretic measure spaces in which events over a context are mutually exclusive (thus, the framework has no need to represent conjunctions of these events).

We are now in a position to define the notion of an 'empirical model' that we earlier said would serve as a theory-independent generalization of a quantum model. Given an empirical scenario $(X, \mathcal{M}, O)$ and a pre-sheaf $\mathcal{DE}$, an *empirical model* $e := \{e_C\}_{C \in \mathcal{M}}$ is a family of distributions that satisfies the following compatibility condition: for all $C, C' \in \mathcal{M}$,

$$e_C|_{C \cap C'} = e_{C'}|_{C \cap C'}. \tag{3.2}$$

AB (2011) show that, by means of a standard prescription, any quantum model induces a corresponding empirical model. For instance, the probabilities of the Bell model define a family $\{e_C\}$ for which the compatibility condition (3.2) simply amounts to the statement that model's correlations satisfy the 'no-signaling' condition (for this reason, (3.2) is also referred to in the literature as a *generalized no-signaling* condition). Furthermore, AB show that the class of empirical models is much larger than just the quantum models: it includes various important *non-quantum* models which have been used as 'foils' in the foundations literature, such as the PR Box (Popescu, 2014) and generalized Kochen-Specker scenarios like the Specker Triangle (Liang et al., 2011).

The notion of an empirical model allows us to define a three-tiered *hierarchy of contextuality* that corresponds to three different strengths of 'gluing failure'. The first/strongest/top tier and the second/middle tier are purely 'possibilistic' in the sense that their definitions only use the coarse-graining of e's probabilities into 'possible' (non-zero probability) and 'impossible' (zero probability). We now define this sector of *possibilistic contextuality*. First, we will say that a global (event) section $s \in \mathcal{E}(X)$ is consistent with the support of an empirical model e just in case $s|_C \in \text{supp}(e_C)$ for all C; otherwise we will say that s is inconsistent with the support of e.

- Tier 1: e is *strongly contextual* iff all its global (event) sections are inconsistent with the support of e. Well-known quantum examples of strong contextuality are provided by all (quantum)

[8] Note that we could also have used the notation $\mathcal{D}(\mathcal{E}(U))$ to emphasize that this is the set of distributions on $\mathcal{E}(U)$.

Kochen-Specker models and the GHZ model. Well-known non-quantum examples of strong contextuality are provided by the PR Box and the Specker Triangle.[9]

- Tier 2: e is *logically contextual* iff at least one of its global (event) sections is inconsistent with the support of e. Clearly, logical contextuality implies strong contextuality, but not vice versa (and so logical contextuality is in this sense 'weaker' than strong contextuality). A well-known quantum example of logical contextuality is provided by Hardy's model (Hardy, 1993).

The geometric/topological essence of possibilistic contextuality can easily be visualized by means of 'bundle diagrams' of (event) sections in the support of e, as depicted in Fig. 3.1. In Fig. 3.1c we see that the PR Box is strongly contextual because no section over a maximal context can be extended to a global section of the bundle (indeed, the topological structure of this bundle is a combinatorial version of a 'Mobius strip', which has no global sections). Similarly, in Fig. 3.1b, we see that the Hardy model is logically contextual because there exists a section (the dotted edge) which cannot be extended to a global section of the bundle. However, the Hardy model is *not* strongly contextual, because it also contains local sections (e.g. the thick edges) that extend to global sections of the bundle.

The third and lowest tier of the hierarchy is a failure of gluing that is defined at the level of *probabilities*, and not merely at the level of possibilities:

- Tier 3: $e := \{e_C\}_{C \in \mathcal{M}}$ is *probabilistically contextual* iff there does not exist a unique global (or joint) distribution e_X that restricts to e_C on each maximal context C.

Clearly, any empirical model that is possibilistically contextual is also probabilistically contextual; however, as we are about to see, the converse does not hold.

A well-known example of probabilistic contextuality is the Bell model, whose bundle diagram is depicted in Fig. 3.1a. Notice that all of the bundle's sections extend to global sections, and thus it is not possibilistically contextual. But as Bell famously proved, the model is nonetheless probabilistically contextual, because the family of probability distributions (determined by Alice and Bob's different measurement settings) does not give rise to the joint probability distribution required for the existence of a local (NC) HVT. More generally, Abramsky and Brandenburger (2011) show that an empirical model e has a global distribution iff there exists an (NC) 'factorizable' HVT that reproduces the probabilities of e: this result thus constitutes a theory-independent generalization of (Standard).[10]

To sum up, the hierarchy of contextuality is:

$$\text{Strong} \implies \text{Logical} \implies \text{Probabilistic} \tag{3.3}$$

In the next subsection, we explain how the structural features that allow us to define this hierarchy can be represented within the WPS framework.

---

[9]It is known that the class of strongly contextual empirical models properly contains the class of AvN models, which in turn properly contains generalized Kochen-Specker models; see e.g. Abramsky et al. (2015).

[10]'Factorizability', which is generally taken to formalize Bell's notion of locality (Brunner et al., 2014), is defined as follows. Consider an HVT within the 'ontological models' framework, i.e. where a preparation of a system induces a distribution $h_\Lambda \in \mathcal{D}_R(\Lambda)$ over ontic states $\lambda \in \Lambda$ and ontic states induce response functions $h_C^\lambda \in \mathcal{D}_R\mathcal{E}(C)$ which encode the probabilities for observing the event s given that the system is in state λ after it is prepared and the measurements in C are performed on it. Furthermore, the distributions h_C^λ are required to be compatible and to reproduce the empirical probabilities, i.e. $e_C = \sum_{\lambda \in \Lambda} h_C^\lambda \cdot h_\Lambda(\lambda)$ for all $C \in \mathcal{M}$. Such an HVT is *factorizable* (Abramsky and Brandenburger, 2011, Section 8) iff, for any $\lambda \in \Lambda$ and for any $s \in \mathcal{E}(C)$ given any $C \in \mathcal{M}$, $h_C^\lambda(s) = \prod_{m \in C} h_C^\lambda|_{\{m\}}\left(s|_{\{m\}}\right)$.

3.2 Weak Probability Spaces: Representations of empirical models

Feintzeig and Fletcher (2017) assume that Generalized HVTs (GHVTs) will at least have the structure of a Weak Probability Space (WPS). Thus, in order to demonstrate that GHVTs are Dutch Bookable, they provide a method of representing a quantum Kochen-Specker model within the WPS framework. The main task of this section will be to obtain a theory-independent generalization of this method by constructing a WPS representation of an empirical model.

Recall that a probability space is a triple (Y, Σ, μ) where Σ is an algebra of events (viz. a σ-field on Y), and $\mu : \Sigma \to [0,1]$ is a probability measure.[11] A WPS is also a triple (Y, Σ, μ), but it generalizes a measure space by replacing the algebra of events with a mere set of events $\Sigma \subseteq P(Y)$ and by replacing the measure with a mere set function $\mu : \Sigma \to [0,1]$; we will nonetheless find it convenient to refer to this set function as a 'measure' when there is no danger of confusion.[12] While this definition of Σ is incredibly weak, remember that for our purposes (and in keeping with the partial algebra approach) it only functions as a blank template: when a WPS is used to represent to an empirical model, this template will need to be filled in the relevant co-measurability relations.

We now proceed to explain how events in AB's framework ('AB-events') can be *represented* in a WPS, and how the WPS events ('Σ-events') can then be constructed from a proper subset of these representations. For brevity, we will use 's_U' to denote an AB-event section in $\mathcal{E}(U)$.

Let e be an empirical model. An *event representation* of e, denoted $\mathcal{E}^\dagger$, consists of a sample space Y and an injective 'event transfer map'

$$\bar{E} : \bigcup_{U \subseteq X} \mathcal{E}(U) \to P(Y). \tag{3.4}$$

that satisfies the following analog of the 'sheaf' property for AB-events: for any $U \subseteq U' \subseteq X$,

$$\bar{E}(s_{U'}) = \bigcap_{\{s_U \,:\, s_{U'}|_U = s_U\}} \bar{E}(s_U) \neq \emptyset. \tag{3.5}$$

Recall that $s_{U'}$ is uniquely determined by the AB-events $\{s_U\}$ that it restricts to; (3.5) simply encodes this set-theoretically by representing $s_{U'}$ as the conjunction of $\{\bar{E}(s_U)\}$ (the intersections are required to be non-trivial because any of these events may be possible).

At this juncture, one could impose two further conditions to ensure that an event representation $\mathcal{E}^\dagger$ fully encodes the combinatorial structure of the event sheaf $\mathcal{E}$ of e, viz.

> **Strong Mutual Exclusivity** For any distinct $s, s' \in \mathcal{E}(U)$ and any $U \subseteq X$, $\bar{E}(s) \cap \bar{E}(s') = \emptyset$.
>
> **Exhaustiveness** For any $U \subseteq X$, $\bigcup_{s \in \mathcal{E}(U)} \bar{E}(s) = Y$.

Call an event representation satisfying these additional conditions a *combinatorial event representation*.[13] In what follows, we will prove our main result (Theorem 1) about the relationship between

[11]Recall that a (probability) measure is an additive function $\mu : \Sigma \to \mathbb{R}_{\geq 0}$ such that $\mu(\emptyset) = 0$ and $\mu(Y) = 1$. A σ-field Σ on Y is a collection of subsets of Y containing $\emptyset$ that is closed under complementation and countable unions and intersections.

[12]Cf. Definition 5 of (Feintzeig and Fletcher, 2017, p. 297). We alter the codomain of their definition to the unit interval, since we have no need of the additional generality.

[13]Every empirical model has a minimal combinatorial event representation; we construct one example in the Appendix.

contextuality and Dutch Bookability without assuming that the event representation is combinatorial (this degree of generality allows us to keep our treatment as close to that of Feintzeig and Fletcher (2017) as possible). However, we also show (Props. 1 and 2) that if one assumes the combinatorial structure, one can give a simpler argument for a stronger result, which has converses.

In order to avoid the clutter of repeatedly writing $\bar{E}$, we will introduce the notation $S_U := \bar{E}(s_U)$ for event representations, as well as $\mathcal{E}^\dagger(U) := \{\bar{E}(s_U) : s_U \in \mathcal{E}(U)\}$ for the set of event representations over $U \subseteq X$. This notation provides us with a useful mnemonic, because it reminds us that S_U is the set in $P(Y)$ that represents the AB-event s_U.

We note that there is a 'duality' between AB-events and their WPS representations: if $U \subset U'$ and $s_U = s_{U'}|_U$, then by (3.5), $S_U \supset S_{U'}$. Furthermore, just as the event sheaf $\mathcal{E}$ comes with a notion of restriction, $\mathcal{E}^\dagger$ possesses a dual notion of extension, i.e. $S_{U'}|^U = S_U$, where $s_U = s_{U'}|_U$. And just as every compatible family of AB-event sections $\{s_i \in \mathcal{E}(U_i)\}$ (where $\{U_i\}$ covers X) gives rise to a unique global section s_X, any compatible family $\{S_i \in \mathcal{E}^\dagger(U_i)\}$ (i.e. a family that satisfies $S_i|^{U_i \cap U_j} = S_j|^{U_i \cap U_j}$ for all i, j) has a unique intersection S_X such that $S_X|^{U_i} = S_i$ for all i.

To prepare ourselves for the definition of a WPS representation of e, let us consider how co-measurability and non-co-measurability relations can be faithfully transferred from the AB-events to the set of Σ-events that we now wish to construct. First, recall that in the AB framework, 'co-measurable events' are sections over sets of measurements that are jointly contained in a context, and 'non-co-measurable events' are sections over sets of measurements that are not. Thus, in order to transfer the structure of co-measurable events from the AB framework to the WPS framework, we will restrict the domain of $\bar{E}$ to AB-events over contexts, thereby defining the map $E : \bigcup_{U \in \mathcal{M}'} \mathcal{E}(U) \to P(Y)$. It is this map that we will now use to construct the event set Σ of a WPS representation.

For any context $U \in \mathcal{M}'$, let Σ_U denote the algebra generated by the collection of sets $\bigcup_{x \in U} \mathcal{E}^\dagger(x)$. The set of WPS events Σ is then defined as the union $\bigcup_{U \in \mathcal{M}'} \Sigma_U$, where Σ_U is treated as a collection of sets. Non-co-measurable Σ-events are thus elements of Σ which are not jointly contained in some Σ_U.

A *WPS representation* of e, denoted $e^\dagger$, is a triple $(\mathcal{E}^\dagger, \Sigma, \mu)$ such that (Y, Σ, μ) form a WPS that satisfies the following three conditions. First, we require that μ allows us to define a probability space for each algebra Σ_U:

> **Weak Classicality (WC)** Let $e^\dagger_U := \mu|_{\Sigma_U}$. For any context $U \in \mathcal{M}'$, $(Y, \Sigma_U, e^\dagger_U)$ is a probability space.

Second, in order to ensure that $e^\dagger$ accurately encodes the empirical probabilities of e, we require:

> **Empirical Consistency (EC)** For any context $U \subset C \in \mathcal{M}$ (where C is a maximal context) and any AB-event $s \in \mathcal{E}(U)$,

$$e^\dagger_U(S) = e_C|_U(s). \tag{3.6}$$

Third, since e treats events over contexts as mutually exclusive, we require:

> **Mutual Exclusivity (ME)** For any distinct AB-events $s, s' \in \mathcal{E}(U)$, $e^\dagger_U(S \cap S') = 0$.

This thus concludes our construction of a WPS representation $e^\dagger$ of an empirical model e.[14]

[14]For completeness, we include the following comment on the relationship between Feintzeig and Fletcher's *weak hidden variable representations* and our *WPS representations*. First, recall that Feintzeig and Fletcher directly

As one might expect, any WPS representation $e^\dagger$ will obey duals of the marginalization condition (3.1) and the compatibility condition (3.2) for e. To describe the dual of (3.1), let $U \subseteq U' \subseteq C$ and $s \in \mathcal{E}(U)$, let A_S be the set of $R \in \mathcal{E}^\dagger(U')$ such that $R|^U = S$, and define 'marginalization for extension' as $e^\dagger_{U'}|^U(S) := \sum_{R \in A_S} e^\dagger_{U'}(R)$. By (EC), it immediately follows that

$$e^\dagger_{U'}|^U = e^\dagger_U, \tag{3.7}$$

which is the dual of (3.1). The dual of (3.2) is simply the statement that WPS representations satisfy the following compatibility condition: for all maximal contexts $C, C' \in \mathcal{M}$,

$$e^\dagger_C|^{C \cap C'} = e^\dagger_{C'}|^{C \cap C'}. \tag{3.8}$$

We now have two ways of formalizing the empirical probabilities of a wide class of physical models: AB's sheaf-theoretic approach, on the one hand, and the above 'dual' WPS approach, on the other. In the next section, we will argue that AB's hierarchy of contextuality (or Gluing Inconsistency) corresponds to a hierarchy of Dutch Bookability in the WPS approach.

4 The hierarchies

We begin this section by situating our result with respect to previous work on hierarchies of constraint-violations leading to Dutch Books, albeit work that has only been done for *abstract* WPS-es, and not WPS representations of *physical* models. Recall that Feintzeig and Fletcher (2017) show that any WPS representation of a Kochen-Specker (i.e. a quantum) model will violate a constraint (in their terminology: the 'no finite null cover' (NFNC) condition), which amounts to a maximal violation of 'subadditivity' in a sense that we will make precise below. This of course raises the question of why this violation is maximal, and whether there might be a hierarchy of different violations enjoyed by quantum and other physical models.

Feintzeig (2015) and Wroński and Godziszewski (2017) have attempted to place some constraints on this question by reasoning abstractly about WPS-es. For instance, Feintzeig (2015) notes that

construct representations of 'quantum mechanical experiments', viz. triples $(\mathcal{H}, \psi, \mathcal{O}_n)$ where $\mathcal{H}$ is a Hilbert space, $\psi \in \mathcal{H}$ is a normalized vector, and $\mathcal{O}_n = \{P_1, \ldots, P_n\}$ is a finite set of projectors on $\mathcal{H}$. We also note that in Abramsky and Brandenburger (2011), an empirical model e is said to have a *quantum representation* if there is some $\mathcal{O}_n = X$ whose maximal subsets of commuting elements form the elements of $\mathcal{M}$, $O = \{0, 1\}$ (the spectrum of each P), and $e_C(s) = \langle \psi, P_s \psi \rangle$, where P_s is the projector on $\mathcal{H}$ corresponding to the section $s \in \mathcal{E}(C)$. Thus, every quantum mechanical experiment is the quantum representation of the appropriate empirical model e.

In Feintzeig and Fletcher (2017), a *weak hidden variable representation* of a quantum mechanical experiment $(\mathcal{H}, \psi, \mathcal{O}_n)$ is a weak probability space (Y, Σ, μ) and a map $E' : \mathcal{O}_n \to \Sigma$ satisfying the following two conditions: $\mu(E'(P)) = \langle \psi, P\psi \rangle$ for all $P \in \mathcal{O}_n$, and for all orthogonal $P_i, P_j \in \mathcal{O}_n$, $E'(P_i) \cap E'(P_j) \in \Sigma$ and $\mu(E'(P_i) \cap E'(P_j)) = 0$. In order to derive a connection between these representations and additivity-violations, Feintzeig and Fletcher require that these spaces satisfy an analog of (WC): for every set $Q \subseteq \mathcal{O}_n$ of mutually orthogonal projectors spanning $\mathcal{H}$, and letting Σ_Q be the algebra formed by Q, $\Sigma_Q \subseteq \Sigma$ and $(Y, \Sigma_Q, \mu|_Q)$ is a probability space.

Now let e be the empirical model induced by $(\mathcal{H}, \psi, \mathcal{O}_n)$. It is clear that by (WC), (EC), and (ME), any WPS representation $e^\dagger$ will also be a weak hidden variable representation of $(\mathcal{H}, \psi, \mathcal{O}_n)$ where $E'(P) = E(P \mapsto 1)$ for all $P \in \mathcal{O}_n$. However, not every weak hidden variable representation of $(\mathcal{H}, \psi, \mathcal{O}_n)$ will be a WPS representation of e; the former are blind to the structure of the event sheaf and so need not satisfy condition (3.5). We note that condition (3.5) imposes the minimal additional structure needed in order to derive interesting connections between sheaf-theoretic inconsistency and violations of additivity; these are established by Theorem 1. To go further and establish an *equivalence* between these two notions of inconsistency, one also needs to assume strong mutual exclusivity and exhaustiveness for $\bar{E}$.

even if an abstract WPS does not maximally violate subadditivity, it could still exhibit a weaker violation of subadditivity that gives rise to a formal Dutch Book. This leads him to conjecture that violating subadditivity is the lowest tier of a hierarchy of additivity-violating conditions that give rise to Dutch Books, because any abstract WPS that satisfies subadditivity will not be Dutch Bookable—if this were true, then it would leave conceptual room for a WPS representation of a physical model to both offer a non-classical alternative to HVTs, and yet avoid Dutch Books. However, Wroński and Godziszewski (2017) show that this conjecture is false by constructing an example of an abstract WPS that is Dutch Bookable but still satisfies subadditivity. More importantly, they adapt an elementary convex analysis result from Paris (2001) to argue that an abstract WPS is Dutch Bookable iff it can be modeled by classical probability (i.e. iff it has a 'classical extension', cf. Section 4.1 below).

We concur with Wroński and Godziszewski (2017) that this exchange demonstrates the futility of simultaneously using WPS-es as a non-classical framework for GHVTS and accepting (formal) non-Dutch Bookability as a constraint on one's theorizing. However, for all this talk of an abstract hierarchy, it could still be the case that no moderately interesting class of physical models gives rise to Dutch Books except by maximally violating subadditivity! In other words, it might be the case that as regards WPS representations of *physical models*, no discussion of a hierarchy is necessary beyond the original Dutch Book violation demonstrated by Feintzeig and Fletcher (2017).

In what follows, we will show that this is not the case by arguing for a correspondence between AB's hierarchy of contextuality and a hierarchy of 'formal constraint'-violations that lead to Dutch Bookability. Since Feintzeig (2015) emphasizes formal constraints with a measure-theoretic flavor, viz. additivity conditions, and the approach of Wroński and Godziszewski (2017) suggests convexity constraints, we will discuss both aspects in Sections 4.1 and 4.2 respectively. It will be an immediate consequence of our discussion that there are physical models which violate subadditivity, but not maximal subadditivity (or NFNC); and that there are physical models which violate additivity, but not subadditivity.

4.1 Contextuality and additivity-violation

We now introduce the relevant additivity conditions that will be violated in our hierarchy. These violations are always witnessed by the 'defect of subadditivity' for some collection V of Σ-events,

$$\mathfrak{a}(V) := \mu\left(\bigcup V\right) - \sum_{a \in V} \mu(a). \tag{4.1}$$

A WPS violates subadditivity just in case there is some collection V of Σ-events such that $\mathfrak{a}(V) > 0$, and it maximally violates subadditivity if $\mathfrak{a}(V) = 1$ for some such collection. It violates additivity just in case there is some collection V of *disjoint* Σ-events such that $\mathfrak{a}(V) \neq 0$.

In order to construct a hierarchy of additivity-violation, we will also need the notion of an 'extension' of a WPS $W \equiv (Y, \Sigma, \mu)$, as well as the more refined notion of a monotonic extension. An *extension* of W is a WPS (Y, Σ', μ'), where Σ' contains the algebra generated by Σ, and $\mu'|_\Sigma = \mu$.[15] A *monotonic extension* of W is an extension (Y, Σ', μ') of W for which $\mu' : \Sigma' \to [0, 1]$ is a monotonic function, i.e. if $A \subseteq B$, then $\mu'(A) \leq \mu'(B)$.[16]

[15] A classical extension is just an extension where μ' is a probability measure.

[16] Note that μ' is really a 'fuzzy measure' in the sense of (Murofushi and Sugeno, 1989). There are also deeper theoretical reasons for imposing monotonicity on an extension: when we describe the space of 'probabilistic but non-possibilistic' contextuality, imposing subadditivity on the extension and taking $P(Y)$ to be the domain of μ' yields

In the case of *abstract* WPS-es (which are not constrained to be representations of empirical models) it is easy to construct a hierarchy of additivity-violation that is tantalizingly similar to AB's hierarchy of contextuality. First, it is clear that a maximal violation of subadditivity implies a violation of subadditivity. Second, if a WPS violates subadditivity, then any of its monotonic extensions must violate additivity. This simple observation provides a heuristic for the kind of hierarchy that one might expect for WPS representations of empirical models: if the analogy holds, we expect that strong contextuality will correspond to a maximal violation of subadditivity; logical contextuality will correspond to a violation of subadditivity; and probabilistic contextuality will correspond to a violation of additivity for monotonic extensions of $e^{\dagger}$. The following theorem shows that this expectation is realized.

Theorem 1. *Let e be an empirical model and let $e^{\dagger}$ be any of its WPS representations.*

1. *e is strongly contextual only if $e^{\dagger}$ maximally violates subadditivity.*
2. *e is logically contextual only if $e^{\dagger}$ violates subadditivity.*
3. *e is probabilistically contextual only if any monotonic extension of $e^{\dagger}$ violates additivity.*

In fact, if one is willing to make the stronger assumption that $e^{\dagger}$ is *combinatorial*, meaning that its event representation satisfies Strong Mutual Exclusivity and Exhaustiveness (cf. Section 3.2 for the definitions), then the main ideas in the proof simplify dramatically, and one obtains the result that there is an *equivalence* between each tier of contextuality and a corresponding condition that witnesses additivity-violation. We now proceed to state these conditions and the equivalence result.

Given that strong and logical contextuality only concern the possibilistic data of an empirical model e, we expect the corresponding (sub)additivity-violations to depend only on the possibilistic data of $e^{\dagger}$. To characterize such violations, we will introduce the notion of an *additive cover*, viz. a collection $A \subseteq \Sigma$ of mutually disjoint sets that is additive (so $\mathfrak{a}(A) = 0$) and which covers the sample space Y.

Definition 1. $e^{\dagger}$ *strongly $\mathcal{V}$-violates subadditivity* just in case a collection V of measure-zero sets in $\mathcal{V} \subseteq \Sigma$ covers *all* the non-measure-zero elements of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathfrak{a}(V) = 1$).

Note that V witnesses a maximal violation of subadditivity.

Definition 2. $e^{\dagger}$ *logically $\mathcal{V}$-violates subadditivity* just in case the collection V of measure-zero sets in $\mathcal{V} \subseteq \Sigma$ covers *some* non-measure-zero element a of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathfrak{a}(V \cup A \setminus \{a\}) > 0$).

Here, $V \cup A \setminus \{a\}$ witnesses a (possibly non-maximal) violation of subadditivity. Finally, in order to characterize additivity-violations stemming from the probabilistic data of $e^{\dagger}$, we turn our attention to monotonic extensions.

Definition 3. $e^{\dagger}$ *$\mathcal{V}$-violates additivity* just in case the algebra generated by $\mathcal{V}$ contains a collection V of disjoint sets that violates additivity in any monotonic extension of $e^{\dagger}$ (and so $\mathfrak{a}(V) \neq 0$ in these extensions).

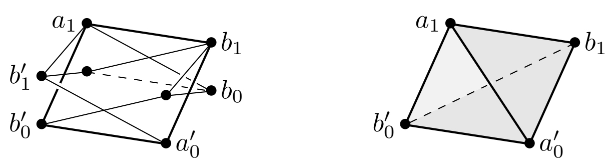
In the following statement of the equivalence between contextuality and additivity-violation, the above conditions are witnessed by the set $\mathcal{V}_{\mathcal{M}}$ of Σ-events corresponding to AB-events over maximal contexts (i.e. $\mathcal{V}_{\mathcal{M}} := \{S : S \in \mathcal{E}^{\dagger}(C), C \in \mathcal{M}\}$).

a Caratheodory 'outer measure', which gives us control over when we can construct genuine measures on algebras within $P(Y)$.

Proposition 1. *Let e be an empirical model and let $e^{\dagger}$ be any of its combinatorial WPS representations.*

1. *e is strongly contextual iff $e^{\dagger}$ strongly $\mathcal{V}_{\mathcal{M}}$-violates subadditivity.*
2. *e is logically contextual iff $e^{\dagger}$ logically $\mathcal{V}_{\mathcal{M}}$-violates subadditivity.*
3. *e is probabilistically contextual iff $e^{\dagger}$ $\mathcal{V}_{\mathcal{M}}$-violates additivity.*

The proof of Theorem 1 involves some set-theoretic yoga, which is relegated to the Appendix, but it is possible to briefly convey an intuitive idea of how the argument works. First, observe that Y contains two measure-zero collections of Σ-events which do not represent any AB-events, viz. the collection of 'contradictory events' D_1 and the collection of 'non-existent outcomes' D_2 (cf. Eqns (5.1) and (5.2) of the Appendix for the full definitions). In order to translate the definitions of strong/logical/probabilistic contextuality into a WPS setting, we thus excise $\bigcup(D_1 \cup D_2)$ from Y to obtain Z, which represents all the properties that we care about. The translation is then carried out by relying on a key lemma (Lemma 1), which tells us that any $z \in Z$ is contained in some S_X (representing a global section s_X). For instance, this lemma lets us translate 'logical contextuality' into the statement that some subset of Z is contained in a collection of measure zero sets, which (in conjunction with other collections of sets encoding the properties of e) lets us construct the relevant sub-additivity-violating collection.



(a) Hardy model (b) 3-simplex

Figure 4.1

In the case where $e^{\dagger}$ D_3-violates subadditivity (i.e. where e is possibilistically contextual), there is also a beautiful topological interpretation that mirrors the bundle diagrams of Fig. 3.1. For instance, consider $e^{\dagger}$ for the Hardy model. Fig. 4.1a is a 'nerve diagram' of the Σ-events in the support of μ, meaning that n-simplices represent n-intersections of Σ-events. Notice that diagram is essentially the bundle diagram 3.1b, except that the AB-event $a \mapsto 0$ has been replaced by the Σ-event (or 0-simplex) $a_0 := E(a \mapsto 0)$, the AB-event $\{a \mapsto 0, b \mapsto 0\}$ with the Σ-event (or 1-simplex) $a_0 \cap b_0$, and so on. The relevant violation of subadditivity can then be understood as stemming from the fact that the dashed 1-simplex is not the edge of a 3-simplex (the WPS-analogue of a global section) whose edges are in the support of μ.[17]

Two important corollaries immediately follow from Thm. 1. First, since Abramsky and Brandenburger (2011) have shown that the Hardy model is logically contextual but not strongly contextual, and that the Bell model is probabilistically contextual but not logically contextual, our result shows that there are quantum models which 'inhabit' lower tiers of the hierarchy of additivity-violation. It also extends the original NFNC-violation result of Feintzeig and Fletcher (2017) by showing that

[17]Clearly, this suggests a cohomological formulation of additivity-violation, although we have not the room to treat this aspect here.

e is strongly contextual	$\Longrightarrow$	e is logically contextual	$\Longrightarrow$	e is probabilistically contextual	$\Longleftrightarrow$	e has no NC, factorizable HVT
$\Downarrow$		$\Downarrow$		$\Downarrow$		$\Downarrow$
$e^\dagger$ maximally violates subadditivity	$\Longrightarrow$	$e^\dagger$ violates subadditivity	$\Longrightarrow$	any monotonic extension of $e^\dagger$ violates additivity	$\Longleftrightarrow$	$e^\dagger$ is Dutch-bookable

Figure 4.2: The hierarchy of contextuality (top) and the hierarchy of Dutch Bookability from additivity-violation (bottom).

non-quantum models like the PR Box have WPS representations which maximally violate subadditivity.

To state the second corollary, we will need a final notion of 'extension' that has already been discussed in (Feintzeig, 2015; Wroński and Godziszewski, 2017): A *classical extension* of W is an extension (Y, Σ', μ') that is a probability space, i.e. Σ' is an algebra, and μ' is additive. The corollary is that e is contextual (either strongly, or logically, or probabilistically) only if its WPS representations do not have a classical extension. Furthermore, Thm. 2 of (Wroński and Godziszewski, 2017) (which is in turn a minor adaptation of a result of (Paris, 2001)) states that an abstract WPS is Dutch Bookable iff it does not have a classical extension; thus, if e is contextual then $e^\dagger$ is Dutch Bookable. The complete set of relations between our two hierarchies is shown in Fig. 4.2 below.

We note that the result of Paris (2001) invoked by Wroński and Godziszewski (2017) is essentially one about 'convexity', which lies at the heart of the formal concept of a Dutch Book. Thus, to complete this circle of ideas, the next section reviews formal Dutch Books and proceeds to describe the hierarchy of contextuality in terms of a hierarchy of convexity-violation.

4.2 Dutch Bookability and convexity-violation

The violation of the formal 'No Dutch Books' constraint and its standard normative interpretation is well-trodden territory (both mathematically and philosophically), so we shall be brief in our summary of both. Our main aim is to stress that since formal Dutch Bookability is a statement about convexity-violation, it follows that AB's hierarchy of contextuality can be straightforwardly understood as a hierarchy of convexity-violation (cf. Prop. 2 below).[18] In order to separate this point from normative issues, we will defer our review of the standard normative interpretation of formal Dutch Books to the end of this subsection.

Let (Y, Σ, μ) be a WPS. To succintly state the standard definition of a formal Dutch Book (adapted to WPS-es), it will be useful to define the (finite) set of functions $\mathbb{V} := \{\ V_y = \chi_{(\ .\ )}(y) : \Sigma \to \{0,1\}\ \}$, where $y \in Y$ and $\chi_A : Y \to \{0,1\}$ is the characteristic function of $A \in \Sigma$. Each function $V_y \in \mathbb{V}$ can be identified with the 'elementary/atomic event' specified by an element y of the sample space Y; we note that in general, such 'events' will of course not be Σ-events. We say

[18]Abramsky et al. (2017) note that every empirical model decomposes into a convex combination of a noncontextual empirical model and a strongly contextual one, and they use the weight of the latter component as a quantitative measure of contextuality. In what follows, we highlight the qualitative links between convexity-violation and AB's original hierarchy.

that a WPS violates a 'No Dutch Books' constraint (or: is Dutch Bookable) iff there exists a function $s : \Sigma \to \mathbb{R}$ and a collection of sets $\mathcal{V} \subset \Sigma$ such that for any $V_y \in \mathbb{V}$,

$$\sum_{A \in \mathcal{V}} s(A) \cdot (V_y(A) - \mu(A)) < 0. \tag{4.2}$$

It is also common to leave (Y, Σ) implicit and simply say that the set-function μ is Dutch Bookable.

It is a well-known fact that condition (4.2) is a statement about convexity-violation. More precisely: let $\mathrm{Conv}(\mathbb{V})$ be the convex hull of $\mathbb{V}$, i.e. the set of all convex combinations of functions in $\mathbb{V}$. For convenience, we will say that μ is $\mathbb{V}$-convex iff $\mu \in \mathrm{Conv}(\mathbb{V})$. Then, as originally noted by de Finetti (1974) (for finite $\mathbb{V}$) μ is not Dutch Bookable iff μ is $\mathbb{V}$-convex.[19]

With this background in place, it is easy to see that AB's hierarchy of contextuality can be interpreted as a hierarchy of obstructions to $\mathbb{V}$-convexity. For $\mathcal{V} \subseteq \Sigma$, we say that a WPS representation $e^\dagger$ $\mathcal{V}$-violates $\mathbb{V}$-convexity so long as $\mu|_{\mathcal{V}}$ cannot be written as a convex sum of $V_y|_{\mathcal{V}}$ for $V_y \in \mathbb{V}$; this will constitute the lowest tier of the hierarchy. In order to describe the higher tiers, let us introduce the Boolean addition operator $\oplus$, i.e. $1 \oplus 1 = 1$, $0 \oplus 1 = 1 \oplus 0 = 1$, and $0 \oplus 0 = 0$ (logical sums of functions are conducted pointwise). We will say that $e^\dagger$ logically $\mathcal{V}$-violates $\mathbb{V}$-convexity iff $\mathrm{supp}(\mu|_S)$ is not a logical sum of $V_y|_{\mathcal{V}}$ for $V_y \in \mathbb{V}$; clearly, in such a case, no convex combination of V_y will yield μ. In this case, however, it may still be that

$$\mathrm{supp}(\mu|_{\mathcal{V}}) = f + \bigoplus_{V_y \in \mathbb{V}'} V_y|_{\mathcal{V}} \tag{4.3}$$

for some set $\mathbb{V}' \subseteq \mathbb{V}$ and some function $f : \mathcal{V} \to \{0, 1\}$. If (4.3) cannot be satisfied for $\mathcal{V}$, then we will say that $e^\dagger$ strongly $\mathcal{V}$-violates $\mathbb{V}$-convexity. Given these definitions, one immediately obtains the following correspondence between contextuality and convexity-violation (we sketch a proof in the Appendix):

Proposition 2. *For any empirical model e and any of its combinatorial WPS representations $e^\dagger$,*

1. *e is strongly contextual iff $e^\dagger$ strongly $\mathcal{V}_{\mathcal{M}}$-violates $\mathbb{V}$-convexity.*
2. *e is logically contextual iff $e^\dagger$ logically $\mathcal{V}_{\mathcal{M}}$-violates $\mathbb{V}$-convexity.*
3. *e is probabilistically contextual iff $e^\dagger$ $\mathcal{V}_{\mathcal{M}}$-violates $\mathbb{V}$-convexity.*

By de Finetti's result, the probability measure μ of $e^\dagger$ violates $\mathbb{V}$-convexity iff it is Dutch Bookable; thus, this hierarchy of violations of $\mathbb{V}$-convexity is a hierarchy of conditions witnessing Dutch Bookability.

The above formal definition of Dutch Bookability (Eq. 4.2) comes with a standard normative interpretation that we will now review. Following Hajek (2008), we note that the standard interpretation takes as given a sample space Y of possible states of affairs and identifies the function $\mu : \Sigma \to [0, 1]$ with the credences (or subjective degrees of belief) of an agent. Thus, a function $V_y \in \mathbb{V}$ gives yes/no answers to the set of questions {'Does event A occur given that the world is in state y?' : $A \in \Sigma$ }, and thereby determines an agent's *ideal* degrees of belief in A given that the world is in state y. A normative Dutch Book Argument then proceeds as follows: consider a betting game in which a bookie assigns stakes $s_i \in \mathbb{R}$ to every event $A_i \in \Sigma$. For non-negative

[19] See also Theorem 2 of Paris (2001).

stakes, we assume the agent is willing to buy into the 'fair' s_i-dollar bet on A_i for $\mu(A_i) \cdot s_i$ dollars. The bookie then checks the outcomes specified by the state of the world $V_y(A_i)$ and proceeds to pay out $V_y(A_i) \cdot s_i$ dollars. We also consider reverse bets: for negative s_i, the bookie pays the agent $\mu(A_i) \cdot |s_i|$, but the agent must pay back $V_y(A_i) \cdot s_i$ when the bookie comes to collect. It turns out that condition (4.2) is satisfied (equivalently: the agent's credences μ are Dutch Bookable) precisely when the agent will always lose money in this betting game. Thus, the argument concludes that if μ is Dutch Bookable, the agent's credences must be irrational. As emphasized by Hajek (2008), this putative irrationality is supposed to be 'internal' to the agent's degrees of belief (as opposed to an inconsistency between the agent's beliefs and the 'external world', e.g. the empirical probabilities of a physical model) and is in principle detectable by 'a priori' means alone.[20] In the next and final section, we take up the question of whether this normative interpretation can be intelligibly ascribed to (formally) Dutch Bookable WPS representations.

5 Conclusion

The results of this paper show that Dutch Bookability has the following minimal interpretation: it is a 'form of inconsistency' in the sense that it characterizes and classifies the ways in which the empirical probabilities of WPS representations deviate from the classical probability assumed by an HVT. We now end with a partial assessment of whether—as per the suggestion of Feintzeig and Fletcher (2017)—WPS representations can be used as a probabilistic framework for *non-contextual* GHVTs, and whether Dutch Bookability can in turn be used to demonstrate the incoherence of credences which are identified with the 'measure' μ of such a GHVT.

At minimum, any (WPS representation-based) GHVT should specify the possible states of affairs according to the theory, as well as the 'response functions' that specify the probabilities of outcomes of measurements given some particular state of affairs. Furthermore, from the discussion in Section 4.2, it is clear that in order for Dutch Bookability to be interpreted normatively, i.e. as yielding a Dutch Book Argument, these elements will need to be elaborated on as follows. First, (G1) the GHVT's set of possible states of affairs will be taken to be the sample space Y. Second, (G2) the response function associated with a state $y \in Y$ will be given by $V_y \in \mathbb{V}$, and is thus *outcome-deterministic* (OD) (i.e. takes values in $\{0, 1\}$) and *non-contextual* (i.e. is fully specified by the state and so does not depend on information about the co-measurability of events). One can relate these features of a GHVT to the bundle diagrams of Fig. 3.1: each V_y corresponds to some global section of a bundle, which specifies the outcomes for measurements if the world is in state y. Finally, recall from the definition of a WPS representation of an empirical model that (G3) the empirical probabilities are encoded in the WPS 'measure' μ, which the Dutch Book Argument then identifies with the credences of a hypothetical agent.

Given (G1)–(G3) and our result that Dutch Bookability is a 'form of inconsistency' (cf. Thm. 4.1 and Prop. 1), it now follows that if an empirical model e is *possibilistically* contextual, a GHVT based on e's WPS representations must have OD response functions (or: atomic states of affairs) which are inconsistent with the empirical probabilities encoded in an agent's credences μ—indeed, we have demonstrated that it is precisely this fact which gives rise to (formal) Dutch Bookability in the possibilistic case. But then it is clear that the standard normative interpretation cannot be given to Dutch Bookability: for recall that the standard interpretation is only supposed to ascribe

[20] We will not need to further explore the nature of this irrationality, although we note that it is the subject of vigorous debate in the general Dutch Book literature, and refer the reader to (Hajek, 2008) for further discussion.

'internal' irrationality to the agent's credences, whereas in the case at hand, the inconsistency is between the agent's credences (which are based on the empirical probabilities of a physical model) and what the GHVT takes to be the objective structure of the world (viz. the response functions V_y). Thus, here it is not the credences which are at fault (since they match the empirical data) but rather the structure of the states of affairs which are assumed by a standard Dutch Book Argument.

Thus, we conclude that an adequate understanding of Dutch Bookability as a 'form of inconsistency' undermines its normative interpretation with respect to WPS representations of quantum theory. We emphasize, however, that the scope of our present argument is limited to *possibilistically* contextual empirical models; for all we have said, it is still conceivable that the standard normative interpretation might be given to the Dutch Bookability of WPS representations of *probabilistically* contextual empirical models. In other words, if there were a physical theory whose only contextual models were *probabilistically* contextual, one might be able to use a standard Dutch Book Argument to demonstrate the incoherence of WPS representations of the theory's contextual models.[21] It is an open question whether any such physical theory exists, but one thing is certain: such a theory would not be quantum theory.[22]

Appendix

We begin by reviewing some notation. First, for AB-event sections, we will use a subscript to keep track of the set $U \subseteq X$ over which a section lies, i.e. $s_U \in \mathcal{E}(U)$ (where X is the measurement set of an empirical scenario). For instance, s_X will denote a particular global section. Second, we will use $S_U := \bar{E}(s_U)$ to simplify our notation for representations of AB-events (as sets within the WPS sample space Y); we will also use $\mathcal{E}^\dagger(U) := \{\bar{E}(s_U) : s_U \in \mathcal{E}(U)\}$ to denote the set of representations of AB-events over U. Third, in what follows, we will always use C and C' to refer to *maximal* contexts, i.e. $C, C' \in \mathcal{M}$.

In order to relate the different tiers of contextuality to properties of our WPS representations, we will want to excise the following two collections of sets from Y:

$$D_1 := \{E(s) \cap E(s') : s, s' \in \mathcal{E}(x) \text{ for some } x \in X\} \,. \tag{5.1}$$

$$D_2 := \left\{ Y - \bigcup_{s \in \mathcal{E}(x)} E(s) : x \in X \right\}. \tag{5.2}$$

D_1 represents 'contradictory events', whereas D_2 represents 'measurements with non-existent outcomes'; neither of these types of 'events' is represented within the AB framework, and by (WC),(EC)

[21] The question of what one would need to assume about a GHVT in order to intelligibly interpret probabilistic contextuality in terms of a Dutch Book Argument is a subtle one that we cannot explore here. We merely note that some conceptions of 'non-contextual hidden variable theories' are strong enough to rule out such an interpretation. For instance, on the 'ontological models' framework advanced by Spekkens (2005), the empirical probabilities can be reproduced by an NC factorizable HVT iff they can be reproduced by an NC outcome-deterministic HVT (see (Kunjwal, 2015) for a discussion of this aspect of 'Fine's theorem'). In such a case, our argument against interpreting possibilitic contextuality in terms of normative Dutch Books would seem to extend straightforwardly to probabilistic contextuality (once one takes into account the probability of preparing the system in some particular state).

[22] One candidate for such a theory *might* be the semi-classical fragment of the gravity dual in the ER=EPR correspondence (Maldacena and Susskind, 2013)—while this model can reproduce the empirical probabilities of the Bell model, quantum corrections are needed in order to obtain strongly contextual (e.g. GHZ) phenomena (Susskind, 2016).

and (ME), every set in these collections is a measure-zero Σ-event. Thus, it is the result of the excision $Z := Y - \bigcup(D_1 \cup D_2)$ which will contain the properties of interest to us. In the course of our analysis, we will frequently have reason to consider the intersection of Z with some set $B \in P(Y)$; we thus introduce the notation:

$$\tilde{B} := Z \cap B \tag{5.3}$$

Let e be an empirical model, and let $e^\dagger$ be any one of its WPS reprsentations. We will now prove a key lemma that will be used repeatedly to help translate the 'contextuality' of e into the properties of $e^\dagger$.

Lemma 1. *If $z \in Z := Y - \bigcup(D_1 \cup D_2)$, then $z \in \bar{E}(s_X)$ for some $s_X \in \mathcal{E}(X)$.*

Proof.

We begin by arguing that for all $x \in X$, any $z \in Z$ is contained in exactly one $S_x \in \mathcal{E}^\dagger(x)$. Let us define the union

$$A_x := \bigcup_{S_x \in \mathcal{E}^\dagger(x)} S_x$$

for an arbitrary $x \in X$. Since D_2 contains A_x^c and D_2 is disjoint from Z (by the definition of Z), it is clear that $Z \cap A_x^c = \emptyset$, and so $A_x \supseteq Z$. In other words, any $z \in Z$ is contained in the union A_x. To further see that any $z \in Z$ is contained in exactly one S_x out of this union, we simply note that since D_1 is disjoint from Z,

$$\tilde{S}_x \cap \tilde{S}'_x = \emptyset \quad \text{for any } S, S' \in \mathcal{E}^\dagger(x), \tag{5.4}$$

where we recall that $\tilde{S} := S \cap Z$.

We now know that for any $z \in Z$, there is a collection $\{S_x\}_{x \in X}$ such that each S_x contains z. Since any collection of AB-events $\{s_x\}_{x \in X}$ defines a global section s_X (by the sheaf property of $\mathcal{E}$), and the corresponding S_X is in the intersection of $\{S_x\}$ (by Eq. (3.5)), we conclude that $z \in Z$ is contained in S_X.

□

We now use this lemma to prove Theorem 1.

Proof of Theorem 1.

(1) e strongly contextual $\Longrightarrow$ $e^\dagger$ maximally violates subadditivity

In order to show the existence of a collection of sets that maximally violate subadditivity, it will be convenient to define D_3, the collection of measure-zero sets S_C each of which contains some representation S_X of a global section. More formally:

$$D_3 := \{S_C \text{ corresponding to some } s_C : \mu(S_C) = 0 \text{ and } S_C \supset S_X \text{ corresponding to some } s_X\}.$$

Our strategy will be to use strong contextuality to show that the union of sets in D_3 contains Z, thus providing—in conjunction with D_1 and D_2—the collection of sets which maximally violates subadditivity.

We now translate 'strong contextuality' into a feature of the collection D_3. Recall that strong contextuality means that there are no global sections in the support of e; in other words, for any global section s_X, there exists a section $s_C := s_X|_C$ such that $e_C(s_C) = 0$. Thus, by (EC) and condition (3.5), strong contextuality implies that any S_X is contained in some measure-zero S_C; hence such an S_C belongs in D_3. We can now combine this fact with Lemma 1 to deduce that—since any $z \in Z$ is contained in some S_X—strong contextuality implies that $\bigcup D_3 \supset Z$.

In light of this last WPS characterization of strong contextuality, we will take our violation-inducing collection to be $\mathcal{V}_3 := D_1 \cup D_2 \cup D_3$. Since $Y = \bigcup \mathcal{V}_3$ and $\mu(Y) = 1$ by (WC), we can immediately compute that $\mathfrak{a}(\mathcal{V}_3) = \mu(\bigcup V_3) - \sum_{A \in \mathcal{V}_3} \mu(A) = 1$, i.e. $\mathcal{V}_3$ yields a maximal violation of subadditivity.

(2) e logically contextual $\Longrightarrow$ $e^\dagger$ violates subadditivity

Our strategy here is similar, viz. we will find a collection of sets that covers Y and use this to show that the collection violates subadditivity. However, our previous argument will need to be supplemented, because logical contextuality only guarantees the existence of *one* section $s^* \in \mathcal{E}(C)$ in the support of e which does not extend to a global section in the support of e. Thus, we will only be able to show that the union of sets in D_3 contains a subset of Z; in order to construct a collection that completely contains Z, we will need to introduce a new collection that we call D_4.

From the definition of logical contextuality, we know that for all global sections s_X that restrict to s^*, there exists a distinct maximal context C' such that the restriction $s_X|_{C'} = s_{C'}$ is not in the support of $e_{C'}$, i.e. $e_{C'}(s_{C'}) = 0$. Thus, by (Empirical Consistency) and condition (3.5), $S_{C'}$ is a measure-zero set that contains any such S_X, i.e. it is an element of D_3.

Combining this fact with (3.5) and Lemma 1 now lets us argue that $\bigcup D_3$ contains $\tilde{S^*} := S^* \cap Z$. To see this, we first note that by (3.5) and Lemma 1, $S^* = \bigcup_{S_X : S_X|^C = S^*} S_X$. And since we have just argued that any S_X that extends to S^* is contained in $S_{C'}$, it follows that $\bigcup D_3 \supset \tilde{S^*}$.

However, to complete our argument, we will need to generate a collection whose union contains all of Z. In order to do so, we first define another collection $D_4 := \{S_C : S_C \neq S^*\}$ (i.e. sets that represent sections over C which are not s^*), and note that by (EC), any set in D_4 must have measure less than one because—by hypothesis—$\mu(S^*) \neq 0$. Supplementing D_3 with D_4 then allows us to construct the desired collection: $\bigcup (D_3 \cup D_4) \supset Z$, because by Lemma 1 and (3.5), any $z \in Z$ is contained in precisely one S_C.

We can now produce our subadditivity collection by defining $\mathcal{V}_4 := D_1 \cup D_2 \cup D_3 \cup D_4$, and noting that $Y = \bigcup \mathcal{V}_4$. Since the sets in D_1, D_2, D_3 are all of measure zero and the sets in D_4 are of measure less than one, it follows that $\mathfrak{a}(\mathcal{V}_4) = \mu(\bigcup V_4) - \sum_{A \in \mathcal{V}_3} \mu(A) > 0$. Thus, $\mathcal{V}_4$ violates subadditivity.

(3) e probabilistically contextual $\Longrightarrow$ any monotonic extension of $e^\dagger$ violates additivity

Note that if e is possibilistically (i.e. either strongly or logically) contextual, then any monotonic extension of $e^\dagger$ will violate subadditivity, and so it will also violate additivity. Thus, we will only need to consider the case in which e is probabilistically, but not possibilistically, contextual. In this case, $e^\dagger$ satisfies subadditivity, and its monotonic extensions which are *not* subadditive are not additive by default; thus, we need only consider the monotonic extensions of $e^\dagger$ which satisfy subadditivity. As such, we will now argue that if e is probabilistically contextual, then any subadditive and monotonic extension of $e^\dagger$ will contain a collection of sets $\mathcal{V}$ which violates additivity.

By probabilistic contextuality and (EC), there must exist some $S^* \in \mathcal{E}^\dagger(C)$ such that $\mu(S^*)$

cannot be recovered by marginalization from any μ', i.e.

$$\mu(S^*) \neq \sum_{\{S_X \in \mathcal{E}^\dagger(C):\ S_X|^C = S^*\}} \mu'(S_X). \tag{5.5}$$

The non-marginalization equation (5.5) bears a resemblance to the additivity-violation condition; however, there is insufficient information to show that it is defined on a disjoint collection $\{S_X\}$ whose union yields S^*. In order to obtain a related collection of sets that satisfies these two properties, we define $\mathcal{V} := \{\tilde{S}_X : S_X|^C = S^*\}$. We then observe that by Lemma 1, this is a disjoint collection; and that by Lemma 1 and (3.5), $\bigcup \mathcal{V} = \tilde{S}^*$, which relates the collection to the LHS of (5.5).

In order to relate the measures of the sets used in (5.5) to the measures of the sets in $\mathcal{V}$, we will need to establish two further facts, viz. $\mu'(\tilde{S}_X) = \mu'(S_X)$ and $\mu'(\bigcup \mathcal{V}) = \mu'(\tilde{S}^*) = \mu(S^*)$. We do so by arguing that $\mu'(\tilde{A}) = \mu'(A)$ for any $A \in \Sigma'$. Let $\bar{A} := A \cap Z^c$. To obtain the result, note that $\mu'(Z^c) = 0$ by the subadditivity of μ' and so $\mu'(\bar{A}) = 0$ by the monotonicity of μ'. Then observe that by subadditivity, $\mu'(A) = \mu'(\tilde{A} \cup \bar{A}) \leq \mu'(\tilde{A})) + \mu'(\bar{A}) \leq \mu'(\tilde{A})$, and again by monotonicity, $\mu'(\tilde{A}) \leq \mu'(A)$. Thus the result follows.

By means of the above facts, we can now use (5.5) to compute that $\mu'(\bigcup \mathcal{V}) \neq \sum_{A \in \mathcal{V}} \mu'(A)$, thus showing that $\mathcal{V}$ witnesses an additivity-violation for (Y, Σ', μ').

□

In preparation for the proofs of Propositions 1 and 2, we provide an illustration of a simple combinatorial WPS representation. Let e be the empirical model of the Specker triangle, where $X = \{a, b, c\}$, $O = \{0, 1\}$, and $\mathcal{M} = \{\{a, b\}, \{b, c\}, \{a, c\}\}$.[23] A simple combinatorial WPS representation $e^\dagger$ is obtained by letting the sample space Y contain a unique point for each global section $s_X \in \mathcal{E}(X)$ and no additional points. This WPS can be graphically represented by the cube in Figure 5.1. In this figure, the points exhaust Y, edges represent events over maximal contexts, and faces represent atomic events. For example, the shaded face is $\bar{E}(c \mapsto 0) = \{a_0b_0c_0, a_0b_1c_0, a_1b_0c_0, a_1b_1c_0\}$, and the thick edges represent the set $\{S : S \in \mathcal{E}^\dagger(\{a, b\})\}$.[24]

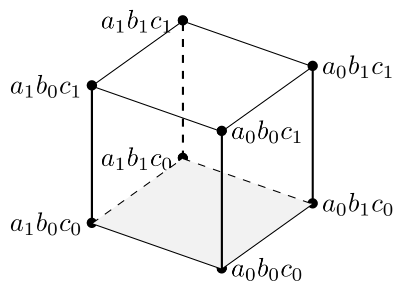


Figure 5.1: A combinatorial WPS-representation of the Specker triangle.

Before proving Proposition 1, we recall that $\mathcal{V}_\mathcal{M} := \{S_C : S_C \in \mathcal{E}^\dagger(C), C \in \mathcal{M}\}$ and that an 'additive cover' is collection of disjoint sets $A \subset \Sigma$ that covers the sample space such that $\mathfrak{a}(A) = 0$.

[23] This is all the data we will need for our example, but see (Liang et al., 2011) for the complete model.

[24] Feintzeig and Fletcher (2017) discuss a variety of types of weak probability spaces that satisfy certain additional conditions, but not all of these will admit of such combinatorial models. The model illustrated, for example, fails to be a Generalized Probability Space (GPS) representation of the Specker triangle because there are disjoint Σ-events whose union is not a Σ-event.

Proof of Proposition 1.

(1) e strongly contextual $\iff$ $e^\dagger$ *strongly* $\mathcal{V}_\mathcal{M}$*-violates subadditivity*

Recall that $e^\dagger$ strongly $\mathcal{V}_\mathcal{M}$-violates subadditivity iff the collection V of measure-zero sets in $\mathcal{V}_\mathcal{M}$ covers *all* the measure-nonzero elements of an additive cover $A \subseteq \mathcal{V}$. By (Exhaustiveness), we can set $A = \{S_C : S_C \in \mathcal{E}^\dagger(C)\}$ for some $C \in \mathcal{M}$. It is clear that $e^\dagger$ strongly $\mathcal{V}_\mathcal{M}$-violates subadditivity iff $\bigcup V = Y$. We will now show that this holds iff e is strongly contextual.

Recall that strong contextuality means that there are no global sections in the support of e; in other words, for any global section s_X, there exists a section $s_C := s_X|_C$ such that $e_C(s_C) = 0$.

Thus, by (EC), (3.5), and (Strong Mutual Exclusivity), every S_X is contained in some measure-zero S_C iff e is strongly contextual. By (Exhaustiveness), $\bigcup_{S_X \in \mathcal{E}^\dagger(X)} S_X = Y$, and by condition (3.5), no S_X is trivial. Thus, $\bigcup V = Y$ iff e is strongly contextual.

(2) e logically contextual $\iff$ $e^\dagger$ *logically* $\mathcal{V}_\mathcal{M}$*-violates subadditivity*

Recall that $e^\dagger$ *logically* $\mathcal{V}$*-violates subadditivity* iff the collection V of measure-zero sets in $\mathcal{V} \subseteq \Sigma$ covers *some* non-measure-zero element a of an additive cover $A \subseteq \mathcal{V}$ (and so $\mathfrak{a}(V \cup A \setminus \{a\}) > 0$).

Suppose e is logically contextual; this means there is some section $s^* \in \mathcal{E}(C)$ for some $C \in \mathcal{M}$ which is in the support of e but which does not extend to a global section in the support of e. By condition (3.5), for each global section s_X that restricts to s^*, the associated WPS-event S_X is contained in some $S_C \in V$. Again by (3.5), the union of all these S_X is equal to $E(s^*) = S^*$; thus $\bigcup V$ contains S^*. Let $A = \{S_C : S_C \in \mathcal{E}^\dagger(C)\}$ and note that A covers Y (by Exhaustiveness) and contains S^*. Thus, $(A \setminus \{S^*\}) \cup V$ yields the desired logical violation of subadditivity, viz. $\mathfrak{a}(V \cup A \setminus \{S^*\}) > 0$.

Now suppose e is not logically contextual; this means *all* sections s_C in the support of e extend to global sections s_X in the support of e. By (3.5), any cover of a measure-nonzero S_C must cover all of the global sections to which it extends; that is, it must cover all S_X such that $S_X|^C = S_C$. Since at least one of these S_X is in the support of $e^\dagger$, by (EC) and (Strong Mutual Exclusivity), any cover using elements of $\mathcal{V}_\mathcal{M}$ must contain some measure-nonzero element. Thus no logical violation of subadditivity can be constructed.

(3) e probabilistically contextual $\iff$ $e^\dagger$ $\mathcal{V}_\mathcal{M}$*-violates additivity*

Recall $e^\dagger$ $\mathcal{V}_\mathcal{M}$-violates additivity iff the algebra generated by $\mathcal{V}_\mathcal{M}$ contains some collection V of disjoint sets that violates additivity in any monotonic extension of $e^\dagger$ (and so $\mathfrak{a}(V) \neq 0$ in these extensions).

First, note that $\mathcal{E}^\dagger(X)$ is a collection of disjoint sets that generates the algebra Σ' that is generated by $\mathcal{V}_\mathcal{M}$. Now note that if e is *not* probabilistically contextual, then (by definition) there exists some $e_X \in \mathcal{DE}(X)$. This e_X induces a monotonic extension (Y, Σ', μ') such that, by (EC) and (WC), all elements in Σ' marginalize appropriately. That is: for any $S \in U$ for any $U \subseteq X$,

$$\mu(S) = \sum_{\{S_X \in \mathcal{E}^\dagger(X):\ S_X|^U = S\}} \mu'(S_X). \tag{5.6}$$

Since the sets in $\mathcal{E}^\dagger(X)$ are all disjoint and generate the algebra, there is no collection of disjoint sets in this algebra that yield a violation of additivity. So there is an additive, monotonic extension of $e^\dagger$.

Now suppose that e is probabilistically contextual. By (EC), in any monotonic extension (Y, Σ', μ'), there must exist some $S^* \in \mathcal{E}^\dagger(C)$ such that $\mu(S^*)$ cannot be recovered by marginalization from μ'. Let $V = \{S_X \in \mathcal{E}^\dagger(X) : \; S_X|^C = S^*\}$, and note that V is a collection of disjoint sets such that $\mathfrak{a}(V) \neq 0$.[25] □

We now turn to Proposition 2. Its proof makes use of the following consequence of condition (3.5), (Strong Mutual Exclusivity), and (Exhaustiveness): letting $\mathbb{V}|_{\mathcal{V}_\mathcal{M}} := \{V|_{\mathcal{V}_\mathcal{M}} : V \in \mathbb{V}\}$, there is a bijection

$$v : \mathcal{E}(X) \to \mathbb{V}|_{\mathcal{V}_\mathcal{M}} :: s_X \mapsto V^{s_X}, \tag{5.7}$$

where for $C \in \mathcal{M}$, $V^{s_X}(S_C)$ is 1 if $S_X|^C = S_C$ and 0 otherwise. Additionally, if we let $\mathcal{D}_A\mathcal{E}(X)$ be the presheaf containing *all* the distributions on X, then there is a bijection

$$d : \mathcal{D}_A\mathcal{E}(X) \to \mathrm{Conv}(\mathbb{V}|_{S_\mathcal{M}}) :: e_X \mapsto \sum_{s_X \in \mathcal{E}(X)} e_X(s_X) \cdot v(s_X). \tag{5.8}$$

By means of these two maps, we can demonstrate the equivalence of the hierarchy of contextuality and the hierarchy of convexity-violation.

Proof of Proposition 2.

(1) e strongly contextual $\iff$ $e^\dagger$ *strongly* $\mathcal{V}_\mathcal{M}$*-violates* $\mathbb{V}$*-convexity*

($\Longrightarrow$) Suppose towards a contradiction that (4.3) is satisfied for some $\mathbb{V}'$ and some f. Pick a $V \in \mathbb{V}'$; then $v^{-1}(V|_{\mathcal{V}_\mathcal{M}}) = s_X$ must be in the support of e, which yields a contradiction.

($\Longleftarrow$) Suppose towards a contradiction that $s_X \in \mathcal{E}(X)$ is in the support of e. But then the set of $V \in \mathbb{V}$ which restricts to $v(s_X)$ satisfies (4.3) for an appropriate choice of f, which yields a contradiction.

(2) e logically contextual $\iff$ $e^\dagger$ *logically* $\mathcal{V}_\mathcal{M}$*-violates* $\mathbb{V}$*-convexity*

($\Longrightarrow$) Suppose that $\mathrm{supp}(\mu|_{\mathcal{V}_\mathcal{M}})$ is the logical sum of the elements of $\mathbb{V}' \subseteq \mathbb{V}$. Then every local section in the support of e extends to a global section in the pre-image of v over $\mathbb{V}'$, which yields a contradiction.

($\Longleftarrow$) Suppose that all local sections in the support of e extend to global sections; then the logical sum of the image of v over these global sections yields $\mathrm{supp}(\mu|_{\mathcal{V}_\mathcal{M}})$, which yields a contradiction.

(3) e probabilistically contextual $\iff$ $e^\dagger$ $\mathcal{V}_\mathcal{M}$*-violates* $\mathbb{V}$*-convexity*

($\Longrightarrow$) Suppose otherwise, viz. suppose that some convex combination of $\mathbb{V}$ yields $\mu|_{\mathcal{V}_\mathcal{M}}$. It follows that $d^{-1}(\mu|_{\mathcal{V}_\mathcal{M}})$ is a distribution e_X that marginalizes to each e_C, which yields a contradiction.

[25] These considerations are sufficient to demonstrate the stronger result that e is probabilistically contextual if and only if *any* (monotonic or non-monotonic) extension of $e^\dagger$ contains a subset in Σ' that violates additivity. We restrict our focus to monotonic extensions to emphasize the connection to Theorem 1.

($\Longleftarrow$) Suppose otherwise, viz. suppose that there is some $e_X \in \mathcal{DE}(X)$. Then $d(e_X)$ yields an element of Conv($\mathbb{V}$) which is equal to $\mu|_{\mathcal{V}_{\mathcal{M}}}$; contradiction.

□

References

S. Abramsky, R. Soares Barbosa, K. Kishida, R. Lal, and S. Mansfield. Contextuality, cohomology and paradox. *ArXiv e-prints*, February 2015.

Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, 2011.

Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. *ArXiv e-prints*, 2015.

Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. The contextual fraction as a measure of contextuality. *Physical Review Letters*, 119:050504, 2017.

A. Acín, T. Fritz, A. Leverrier, and A. Belén Sainz. A combinatorial approach to nonlocality and contextuality. *ArXiv e-prints*, December 2012.

John S. Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

John S. Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447–452, 1966.

Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Reviews of Modern Physics*, 86(2):419, 2014.

Adán Cabello, José M. Estebaranz, and Guillermo García-Alcaine. Bell-Kochen-Specker theorem: a proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996.

Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*, volume 1. Wilely, London, 1974.

Benjamin Feintzeig. Hidden variables and incompatible observables in quantum mechanics. *The British Journal for the Philosophy of Science*, 66:905–927, 2015.

Benjamin H. Feintzeig and Samuel C. Fletcher. On noncontextual, non-Kolmogorovian hidden variable theories. *Foundations of Physics*, 47:294–315, 2017.

Arthur Fine. Joint distributions, quantum correlations, and commuting observables. *Journal of Mathematical Physics*, 23:1306–1310, 1982.

Stanley P Gudder. Generalized measure theory. *Foundations of Physics*, 3(3):399–411, 1973.

Alan Hajek. Dutch Book arguments. In Paul Anand, Prasanta Pattanaik, and Clemens Puppe, editors, *The Oxford Handbook of Rational and Social Choice*. Oxford University Press, 2008.

Lucien Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Physical Review Letters*, 71(11):1665, 1993.

Simon Kochen and E P Specker. The problem of hidden variables in quantum mechanics. In *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution*, pages 293–328. Springer, 1975.

Ravi Kunjwal. Fine's theorem, noncontextuality, and correlations in Specker's scenario. *Physical Review A*, 91(2):022108, 2015.

Yeong-Cherng Liang, Robert W Spekkens, and Howard M Wiseman. Specker's parable of the overprotective seer: A road to contextuality, nonlocality and complementarity. *Physics Reports*, 506(1):1–39, 2011.

Saunders MacLane and Ieke Moerdijk. *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*. Springer Science & Business Media, Dordrecht, 1992.

J. Maldacena and L. Susskind. Cool horizons for entangled black holes. *Fortschritte der Physik*, 61(9):781–811, 2013. ISSN 1521-3978. doi: 10.1002/prop.201300020. URL `http://dx.doi.org/10.1002/prop.201300020`.

Toshiaki Murofushi and Michio Sugeno. An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure. *Fuzzy Sets and Systems*, 29(2):201–227, 1989.

J B Paris. A note on the Dutch Book method. In *Second International Symposium on Imprecise Probabilities and their Applications*, 2001. URL `http://www.sipta.org/isipta01/proceedings/`.

Roger Penrose. On the cohomology of impossible figures. *Leonardo*, 25(3):245–247, 1992.

Sandu Popescu. Nonlocality beyond quantum mechanics. *Nature Physics*, 10(4):264–270, 2014.

F. Roumen. Cohomology of Effect Algebras. *ArXiv e-prints*, February 2016.

Robert W Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5):052108, 2005.

Sam Staton and Sander Uijlen. Effect algebras, presheaves, non-locality and contextuality. In M. Halldórsson, K. Iwama, N. Kobayashi, and B. Speckmann, editors, *Proceedings, Part II, of the 42nd International Colloquium on Automata, Languages, and Programming*, volume 9135, pages 401–413. Springer-Verlag, 2015.

Leonard Susskind. Er=epr, ghz, and the consistency of quantum measurements. *Fortschritte der Physik*, 64(1):72–83, 2016. ISSN 1521-3978. doi: 10.1002/prop.201500094. URL `http://dx.doi.org/10.1002/prop.201500094`.

B. van den Berg and C. Heunen. Noncommutativity as a colimit. *ArXiv e-prints*, March 2010.

Leszek Wroński and Michał Tomasz Godziszewski. Dutch Books and nonclassical probability spaces. *European Journal for Philosophy of Science*, 7:267–284, 2017.