On propensity-frequentist models for stochastic phenomena; with applications to Bell's theorem

Tomasz Placek

Abstract

The paper develops models of statistical experiments that combine propensities with frequencies, the underlying theory being the branching space-times (BST) of Belnap (1992). The models are then applied to analyze Bell's theorem. We prove the so-called Bell-CH inequality via the assumptions of a BST version of Outcome Independence and of (non-probabilistic) No Conspiracy. Notably, neither the condition of probabilistic No Conspiracy nor the condition of Parameter Independence is needed in the proof. As the Bell-CH inequality is most likely experimentally falsified, the choice is this: contrary to the appearances, experimenters cannot choose some measurement settings, or some transitions, with SLR initials, are correlated, or both.

1 Introduction

The objective of this paper is twofold.[1] First, we will construct models that accommodate both propensity aspects and frequentist aspects of statistical experiments. A view underlying this task has it that on a given run of the experiment, a measurement is capable of producing alternative outcomes, the weights of such capacities being propensities. That is, the propensity of a transition from a measurement to one of its outcomes is the measure of how possible the transition is. The frequentist aspect consists in a desideratum that observed frequencies of results should reflect the underlying propensities. The second aim (and, to confess the raison d'être of this paper) is to analyze in the constructed models a particular variety of Bell's theorem, the so-called Bell-CH theorem.

[1] I would like to thank my audience, and in particular – Leszek Wroński. Support by MNiSW research grant 3165/32 is gratefully acknowledged.

One might ask at this point, however, why yet another analysis of Bell's theorem? The answer is that derivations of Bell's theorem have thus far been either fully informal, or have left some aspect of the reasoning informal. This resistance to formal treatment stems from the multifarious nature of Bell's theorem: it appeals to modal aspects (e.g., "other results might have occurred"), spatiotemporal aspects, with some rudiments of relativity ("space-like related events should not influence one another"), and probabilistic aspects ("the common past should screen off the correlations"). Thus, to adequately analyze Bell-type theorems, a framework is needed which includes all three features: it is modal, spatiotemporal, and probabilistic.

A brief survey of Bell-type derivations, with respect to the use of formal methods, reveals that physicists (incl. Bell, Mermin, and Shimony) have all done this informally. Starting with their (1999) paper, the so-called Budapest school of Rédei, Szabó and Hofer-Szabó has constructed rigorous models accounting for probabilistic aspects. Belnap and Szabó (1996) and Kowalski and Placek (1999) have presented models of the GHZS theorem (a non-probabilistic Bell-type theorem)[2] with modal aspects rigorously analyzed, but the spatiotemporal aspect left informal. Müller and Placek (1999–2002) have created models accommodating modal and probabilistic aspects of the Bell-Aspect set-up (Aspect et al., 1982), the spatiotemporal aspect being left informal.

In recent years, however, the branching space-times (BST), as conceived by Belnap (1992), has been developed in a series of papers (by Belnap and his collaborators) into a theory encompassing all the three aspects required in analyses Bell's theorems.[3] This is an axiomatic second order theory, with probabilities assignable to the so-called transitions of events; the theory has a class of models (the so-called Minkowskian Branching Structures) in which possible histories are isomorphic to the Minkowski space-time. The aim of this paper is thus to harness the formal machinery of BST in order to analyze the Bell-Aspect experiment, with the hope of determining, once and for all, what the theorem's implications actually are.

Although in the technical sections of this paper we work in the rigorous possible-worlds framework of BST, let us begin by asking how a statistical experiment is to be represented in a possible-worlds theory. As a preliminary illustration intending to convey the underlying idea, consider the statistical experiment of tossing a given coin. Let us suppose that, after a sufficiently

[2]See Greenberger, Horne, Shimony and Zeilinger (1990).

[3]After Belnap's (1992) paper, the essential development of the theory is reported in Belnap and Szabó (1996), Weiner and Belnap (2006), Belnap (2005), Müller (2002), Müller (2005), Müller, Belnap and Kishida (2008), Wroński and Placek (2009), and Placek and Wroński (2009).

many tosses, the experimenter estimates that the probability of heads, written as $p^{ex}(H)$, tends to 0.6. The experimenter can take two stances in interpreting what $p^{ex}(H) = 0.6$ means.

First, she might take the experiment as deterministic, i.e., claiming that, on each toss of the coin, there is exactly one possible outcome, either heads or tails, and the numbers 0.6 vs. 0.4 stem from there being a few *types of runs*, the particular numerical values following from the frequencies of types of runs within the set of all runs of the experiment. Using symbol Λ^H for the set of those types of runs, whose tokens yield Heads, and w_λ for the frequency of runs of type λ in the collection of all runs of the experiment, it must be that

(†) $\quad p^{ex}(H) = \sum_{\lambda \in \Lambda^H} w_\lambda,$

and analogously for the probability of tails. At an extreme, only two types of runs may be posited: one in which Heads are determined to occur, and the other in which Tails are determined to occur. In this case, experimental probabilities are identified with frequencies of corresponding types of runs. A deterministic model of the experiment will consist of a single history, with a chain of "toss, result" pairs, the former preceding the latter.

Second our experimenter can take an indeterministic stance while interpreting probability $p^{ex}(H) = 0.6$, acknowledging that a given run can have more than one possible outcome (of which, of course, exactly one occurs at a given time). Types of runs are then thought of as differentiated by propensities: if two runs belong to one type, for each possible result, the propensity of a result on one run must agree with the propensity of that result on the other run.[4] In other words, runs of a given type are uniform with respect to propensities. Note a peculiarity of the notion of run: it is a modally thick notion since, to represent a run of a given type, it is not sufficient to specify its (actual) outcome, but all the possible outcomes and propensities thereof. In what follows, we will put forward a certain representation of this notion, understood as modally thick.

For propensities to make contact with experimentally given probabilities, a faithfulness assumption is needed: the number of a given (type-level) result (say, Heads) produced in runs of a given type should approximate the propensity for the corresponding (token-level) result in a run[5] of the given type of runs. The experimental probability $p^{ex}(H)$ is then a weighted average

[4]Note that we are talking type-results here, like Heads. In the formal treatment we will have outcome tokens, and, in that context, it is more accurate to say that in two runs propensities of outcome-tokens agree.

[5]Equally well we could say "in every run" as we assumed that corresponding results attain the same propensity in every run of a given type.

of propensities for H, as obtaining in the types of runs:

$$p^{ex}(H) = \sum_{\lambda \in \Lambda} pr_{\lambda}(H) w_{\lambda}, \tag{1}$$

where Λ is the set of types of runs in our coin tossing experiment, $pr_{\lambda}(H)$ is the propensity for obtaining result H in any run of type $\lambda \in \Lambda$ and w_{λ} is the frequency of runs of type λ in the set runs of the experiment.

On this propensity approach, our model of coin tossing must have many histories, as it contains "forks" representing possible outcomes of tosses, and contains all the information concerning types of runs, and propensities for results. In short, it is like a branching structure stretching indefinitely long. We turn now to the task of defining propensity BST models for statistical experiments.

2 Propensity-Frequentist Models of Statistical Experiments

We are now going to explain propensity-frequentist BST models without providing any technical details on the BST theory. The required notions are explained in Section (4), hence, in order to fully grasp the models, the reader is advised, after having become acquainted with the technical notions of that section, to jump back to the present section.

For the present purpose of introducing propensity-frequentist models of statistical experiments, it suffices to know that BST is a possible-world theory, in which the role of possible worlds is played by histories. Chanciness is represented by branching of histories, so that, for instance, an initial event with two alternative possible outcomes is represented by two branching histories that share an initial event but contain different outcomes. Various types of events are definable in BST of which two are relevant for our purpose here: initial events (which are simply upper bounded subsets of histories) and outcome events (which are more complex set-theoretical constructions). There is a natural notion of consistency: two events are consistent iff there is a history in which the two occur. We also say that two events are space-like related (SLR) iff they are distinct and every element of one and every element of the other are consistent, but none is above the other. A pair consisting of an initial event and an outcome event above it is labeled a transition. Importantly, transitions are objects on which propensities are defined. Finally, we assume that histories are isomorphic to Minkowski space-time; as a result, spatiotemporal locations of events and the relation of being space-like related are definable. Technically, we work in the framework of the Minkowskian Branching Structures (MBS's), which form a class of BST models (for the explanation, see Section (4)).

Statistical experiment idealized Let us suppose that our statistical experiment consists of measuring $\tilde{A}$, which has possible outcomes $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k$. Be aware that we are talking types now: in our illustration, $\tilde{A}$ is a toss of a particular coin, and $\tilde{a}_1, \tilde{a}_2$ are heads up and tails up on this coin, respectively. To keep clear about the types vs. tokens distinction, we use the tilde to signify type-level notions. The statistical aspect has it that an experiment has runs, and in each run the measurement $\tilde{A}$ and some one of its outcomes, $\tilde{a}_i$, is instantiated. What makes runs is experiment-sensitive: in the simplest cases, instantiations of measurement $\tilde{A}$ is a good guide as to when a run begins. In more sophisticated cases, to be considered later, we need to appeal to some other phenomena. For simplicity, let us assume that on each run n of the experiment a token A^n of $\tilde{A}$ and a token $\mathbf{a}_i^n$ of a single $\tilde{a}_i$ occur. This is not a stringent requirement, as one can always add "a failure outcome" to the list of outcomes. More controversially, we assume that on each run n of the experiment, all outcomes $\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_k^n$ are in a sense possible. That is, all these outcomes are supposed to be in a BST model and situated in such a way that each sentence "It is possible that $\mathbf{a}_i^n$ will occur" is true if evaluated at A^n. On the other hand, we allow for some $\mathbf{a}_i^n$ to have propensity zero. This sounds like (and perhaps is) double talk, since, at a time, $\mathbf{a}_i^n$ is possible semantically speaking, but impossible propensity-wise. Unfortunately, we have to engage in this to satisfy a premise of our probabilistic machinery.

The aim of our next assumption is to circumvent the problem of estimating probabilities on the basis of finitely many runs of the experiment: we assume that the experiment has (countably) infinitely many runs. Clearly, as the experiment may finish at this or that stage, our BST models will have histories with a finite number of runs. Yet we take such histories to be irrelevant for our modeling of a statistical experiment. Thus, what we really require is a possible history with infinitely many runs of the experiment. We will define what it takes for the history of a BST model to adequately represent a statistical experiment.

An objection might be raised at this point that an adequate representation of a statistical experiment, with propensities and the idealization of infinitely many runs, should incorporate infinitely many possible histories, representing varying distributions of the measurement outcomes. Among these histories, there will be a majority (by the law of large numbers) of those in which frequencies of results agree with propensities, as well as a minority of those in which frequencies of results do not agree with propensities. Anything short of accounting for this aspect of probabilities does not deserve the name of a"model of statistical experiments, with propensities"—the objection continues. We agree, yet we leave this ambitious task for some future project. Since the import of this paper is negative, as it boils down to a claim that "a propensity-frequentist model for statistical experiment, subject to such-and-such conditions" is not possible, a simplified representation in terms of one history is enough. That is, if a single history of a simplified representation cannot have some desired feature, then none of the (possibly infinitely many) histories of a full-fledged representation can posses that feature, either.

Turning next to spatiotemporal aspects, we assume the weak gravitational fields condition ensuring that Minkowski space-time is an adequate representation of the spatiotemporal aspects of the phenomena considered.

As for spatiotemporal locations, we naturally require that every possible outcome $\mathbf{a}_i^n$ of A^n belongs to the future of possibilities of A^n. We next assume the same spatiotemporal location of alternative outcomes, that is, the following statement:

Condition 1 *Suppose that outcome $\mathbf{a}_i^n$ occurred in spatiotemporal region R on run n; if on this run $\mathbf{a}_j^n$ $(i \neq j)$ occurred instead, it would occur in the same spatiotemporal region R.*

Note that the condition concerns possible outcomes of a measurement in a fixed run. As for possible outcomes of different runs, they are assumed to happen in different spatiotemporal regions: we assume the idealization that there is a minimal time interval between measurements in any two consecutive runs.

Condition 2 *There is $\Delta \in \Re$ such that for every run l of the experiment, if A^l occurs in spatiotemporal location L^l, then A^{l+1} occurs in a shifted by time $\epsilon \geqslant \Delta$ spatiotemporal location $L^{l+1} = T_\epsilon(L^l)$, where for spatiotemporal location $L \subset \Re^4$, $T_\epsilon(L) = \{\langle x_0 + \epsilon, x_1, x_2, x_3\rangle \mid \langle x_0, x_1, x_2, x_3\rangle \in L\}$. A sequence $A^1, A^2, \ldots$ that satisfies this condition will be called "shifted in time" sequence.*

A statistical experiment in which the above idealizing conditions are satisfied, will here be called an idealized statistical experiment.

MBS models for idealized statistical experiments The task of specifying a BST model with propensities for an idealized statistical experiment will be split into two parts. We first focus on the modal and spatiotemporal aspects of the phenomena. With the idealizations assumed above, these aspects can be adequately captured by MBS's. In the next part, we will add propensities to the picture, and with propensities there will come *types* of runs.

Definition 3 (MBS weakly represents) *In an MBS model, history σ weakly represents an idealized statistical experiment in which measurement event $\tilde{A}$ has k possible outcomes $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k$ iff*

1. *there is in σ a sequence $A^1, A^2, A^3, \ldots$ of shifted in time initial events,*
2. *for every $n \in \mathcal{N}$ there is k pairwise inconsistent, occurring in the same spatiotemporal region, outcomes $\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_k^n$.*
3. *for every $n \in \mathcal{N}$, A^n and the set $\{\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_k^n\}$ yields the set $\mathfrak{A}^n := \{A^n \to \mathbf{a}_i^n \mid 0 < i \leqslant k\}$ of transitions.*[6]
4. *for every history h, if A^n occurs in h, then for some $i \leqslant k$, $\mathbf{a}_i^n$ occurs in h as well.*

We say that an MBS model weakly represents the experiment in terms of history σ and the sequnce $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ of sets of transitions.

A central idea is that the experiment is analyzable in terms of transitions: these are transitions from a (token-level) measurement event to its possible outcomes (token-level as well). For simplicity, we assume here that only one measurement (type-level) is performed in the experiment. This limitation is removed in the definitions given in the Appendix. Observe the subtle interplay between the notion of run and the notion of history σ. By conditions (1) and (4), an infinite sequence of measurements and, after each measurement, one of its possible outcomes occur in history σ. Thus, σ codes a particular way in which the statistical experiment (idealized as infinitely long) can go. In and of itself, history σ (as any other history) does not contain any information as to what and where was/is/will be possible. Yet, this modal information is coded in the model, by means of various transitions. Accordingly, the notion of the run codes the measurement and its possible outcomes. We may thus say that the notion of the run is modally thick. Note finally that the model must have at least k^ω histories, i.e., uncountably many.

[6]In the formal treatment, these outcome events will be identified with scattered outcome events and accordingly the transitions will be from intial events to scattered outcomes.

MBS models with propensities for idealized statistical experiments A BST model can be equipped with probability measures, with each probability measure μ_T defined on some particular set T of basic transitions. Given that a probability measure is defined on a particular set of basic transitions, this measure imposes probabilities on larger (i.e., not necessarily basic) transitions. The probabilities in question are naturally interpreted as propensities, i.e., degrees of possibility of the transitions in question. Yet, in order for a BST model to have propensities for transitions, a μ_T probability must be defined on an appropriate set T of basic transitions—cf. Definition (19). In modeling an (idealized) statistical experiment we require that the relevant propensities be defined, which in turn means that the relevant μ-probabilities exist.

Definition 4 (MBS has propensities) *An MBS model $\mathfrak{M}$ has propensities for an idealized statistical experiment with measurement event $\tilde{A}$ and its k possible outcomes $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k$ iff $\mathfrak{M}$ represents (modally) the experiment in terms of some history σ, a sequence $A^1, A^2, A^3, \ldots$ of initial events, and the sequence $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ of sets of transitions and propensity is defined on every $t \in \bigcup_{n \in \mathcal{N}} \mathfrak{A}^n$.*[7]

Now, even if a BST model has propensities for all the relevant transitions, one might still be unable to relate the propensities to the experimentally estimated probabilities. What we need is a kind of faithfulness assumption, which requires that the numerical values of propensities be exhibited as frequencies of results in sufficiently long series of appropriate runs. Since we assumed that the set of all runs is countably infinite, we straightforwardly require that a numerical value of propensity for a result occurring in a run of a given type be equal to the limiting frequency of the corresponding results occurring in the type of runs in question.

To illustrate, consider the n-th run. The set $\mathfrak{A}^n$ of transitions has some propensity assignment, that is, for every transition $t \in \mathfrak{A}^n$, its propensity $pr(t)$ is defined. We first require that in history σ there is a non-empty subset Ω of the set of all runs that are "exactly like" the n-th run. We mean by this that a transition on the nth run and the corresponding transition on each run from Ω have the same propensity. Second, for any $i \leqslant k$, the frequency in Ω of transitions of the form $A^m \to \mathbf{a}_i^m$ must tend to propensity $pr(A^n \to \mathbf{a}_i^n)$.

How can it be interpreted that two sets $\mathfrak{A}_i$ and $\mathfrak{A}_k$ of transitions have the same propensity assignment? Uncontroversially, propensities stem from

[7] It follows from the Definition (19) of propensities in terms of causal probabilities, that pr behaves like probability. For instance, since $\mathfrak{A}^n$ is an exhaustive (see clause 3) set of alternative transitions, we have: $\sum_{t \in \mathfrak{A}^n} pr(t) = 1$.

some ontological stuff, so that if a measurement process on one run and a measurement process on another run are very much alike, the propensity assignment on the first run should agree with the propensity assignment on the other run. Now it is natural to reverse this story to use the same propensity assignments to define types of runs, as follows:

Definition 5 (types of runs) *Let $\mathfrak{M}$ be an MBS with propensities that weakly represents an idealized statistical experiment consisting in measuring $\tilde{A}$ with k possible outcomes $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k$, in terms of history σ and the sequence $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ of sets of transitions.*
The l-th run and the m-th run belong to the same type of runs iff for every $j \leqslant k$: $pr(A^l \to \mathbf{a}^l_j) = pr(A^m \to \mathbf{a}^m_j)$.

Putting the above ideas and observations together, we arrive at this definition of an MBS representing fully, i.e., modally, spatiotemporally, and probabilistically, an idealized statistical experiment:

Definition 6 (MBS model fully represents) *Let $\mathfrak{M}$ be an MBS model with propensities for an idealized statistical experiment consisting of measurement $\tilde{A}$, its k possible outcomes $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_k$, and their experimental probabilities $p^{ex}(\tilde{a}_1), \ldots, p^{ex}(\tilde{a}_k)$. Let the experiment be weakly represented in $\mathfrak{M}$ in terms of history σ and the sequence $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ of sets of transitions. We say that $\mathfrak{M}$ fully represents the experiment in terms of propensity assignment pr, history σ and the set $\mathfrak{A}^1, \mathfrak{A}^2, \mathfrak{A}^3, \ldots$ iff there is a partition Λ of the set $\mathcal{N}$ of natural numbers such that*

1. *for every $l, m \in \mathcal{N}$: $l, m \in \lambda$ for some $\lambda \in \Lambda$ iff for every $j \leqslant k$: $pr(A^l \to \mathbf{a}^l_j) = pr(A^m \to \mathbf{a}^m_j)$,*

2. *$p^{ex}(\tilde{a}_i) = \sum_\lambda w_\lambda pr(A^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)})$, where $\Theta : \Lambda \to \mathcal{N}$ is a "selection function" such that $\Theta(\lambda) \in \lambda$, and*
$$w_\lambda = lim_{n\to\infty} \frac{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda\}}{n},$$

3. *for every $\lambda \in \Lambda$ and $i \leqslant k$:*
$$pr(A^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)}) = \lim_{n\to\infty} \frac{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda \wedge \sigma \in H_{\langle \mathbf{a}_i^m \rangle}\}}{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda\}},$$
where $H_{\langle \mathbf{a}_i^m \rangle}$ is the set of histories in which $\mathbf{a}_i^m$ occurs.

To explain, the model has a history with an infinite sequence of initial events, representing (token-level) measurement events. Thus, by partitioning the set of natural numbers, we partition the set of initial events, each element of the partition being intuitively thought of as a type of runs.[8] Clause 1 ensures

[8]Note that Λ is at most countably infinite.

that Λ is the set of types of runs in the sense of Definition (5). Clause 2 is an Empirical Adequacy condition, as it draws a bridge between experimentally estimated probabilities and propensities (via weights). A value of the weight of a given type is equated to the limiting frequency of runs of this type in history σ. Observe that a type whose tokens appear only finitely many times in σ receives weight zero. The last clause postulates that history σ exhibits propensities of transitions in terms of limiting frequencies of corresponding outcomes, as occurring in runs of appropriate types.

Why might one want to postulate types of runs and propensities, which are unobserved and controversial entities, to account for a statistical experiment? The aim that comes to the fore in the context of quantum non-locality is to explain correlations between spatially distant results. Having seen that explaining such correlations in terms of a deterministic model is not viable, one may turn to models with propensities, as there seems to be more freedom in propensities than in strict determinism. The crucial task is then to derive correlated experimental probabilities from some "more basic" probabilities without correlations and from a distribution of types of runs in the experiment.

Leaving aside the task of constructing deterministic BST models for Bell-type phenomena, as they have been discussed elsewhere, we focus here on indeterministic models with propensities. We will first analyze a simplified EPR set-up, to finally turn to Bell's theorem, deriving the so called Bell-Clauser-Horne inequality (in short, Bell-CH), cf. Clauser and Horne (1974).[9]

3 Simple EPR Explained Away

A peculiar feature of Bell-type experiments is that chanciness (or its appearance) occurs at two levels: on each measurement device, there is a selection of a measurement setting (e.g., direction of polarization), and each measurement has more than one possible outcome. As a warm-up, we will focus first on a simplified experiment, however, in which no selection of measurement settings occur. This is, in essence, as set-up with fixed polarizations, for which there is a hidden variable model—see Bell (1987); here we will call this "simple EPR." The description of the setup is as follows: there are two space-like related measurement events $\tilde{A}$ and $\tilde{B}$ (type-level), with alternative possible results (type-level) $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_I$ of $\tilde{A}$ and $\tilde{b}_1, \tilde{b}_2, \ldots, \tilde{b}_J$ of $\tilde{B}$. Every

[9]An important step towards this formula was the argument of Clauser, Holt, Shimony, and Horne (1969).

combination $\tilde{a}_i, \tilde{b}_j$ is possible. Results are correlated, i.e.,

$$p^{ex}(\tilde{a}_i) \cdot p^{ex}(\tilde{b}_j) \neq p^{ex}(\tilde{a}_i \text{ and } \tilde{b}_j).$$

We will hereby construct an MBS, in which some history fully represents this experiment. Yet, as we are troubled by the (experimentally observed) correlations above, we attempt the following explanation: the runs of this experiment are not uniform as perhaps the state of the source of particles vary, or the measurement process might not exactly be alike on every run. Yet, one may consider a division of the set of runs into types of exactly alike runs. Our article of faith, motivated solely by our uneasiness with distant correlations, is that in each set of exactly alike runs there are no distant correlations, i.e., probabilities factor. In the language of the Bell literature, we envisage some "hidden variable" parametrizing hidden states of pairs of particles involved in the experiment—"hidden" as it is not accounted by quantum mechanics. The hidden variable can take on a number of values, and each joint probability mentioned above, if conditioned on each value of the variable, factors into single probabilities (conditional on the value). Now, on the assumption of indeterminism, a measurement in a given run has many possible outcomes, and the idea of propensity comes up naturally.Furthermore, drawing on an argument sketched in the last section, if some runs are exactly alike, the propensity of corresponding transitions in these runs should be the same. By this train of reasoning, we end up with the task of constructing an MBS, which yields experimentally observed distant correlations (the results of the existence of many types of runs) with such values of propensities that in each run the distant results are uncorrelated. In other words, a crucial part of the task is to show that it is it possible to recover experimentally estimated probabilities, including the correlations above, on the assumption that for every type λ of runs there is no correlation, i.e., for every $i \leqslant I$ and every $j \leqslant J$ we have, schematically,

$$pr(a_i^\lambda) \cdot pr(b_j^\lambda) = pr(a_i^\lambda \text{ and } b_j^\lambda)? \tag{2}$$

0ur MBS $\mathfrak{M}$ should fulfill these desiderata:

Desiderata (†)

1. There is in $\mathfrak{M}$ a history σ containing a sequence $A^1 \cup B^1, A^2 \cup B^2, A^3 \cup B^3, \ldots$ of temporally shifted initial events and for every $n \in \mathcal{N}$: A^n and B^n are SLR;

2. for every $n \in \mathcal{N}$, there is $I \in \mathcal{N}$ pairwise inconsistent, occurring in the same spatiotemporal region, (scattered) outcomes $\mathbf{a}_1^n, \mathbf{a}_2^n, \ldots, \mathbf{a}_I^n$ such

that A^n is below each $\mathbf{a}_i^n$ $(i \leqslant I)$,[10] and each such outcome is SLR to B^n;

3. for every $n \in \mathcal{N}$, there is $J \in \mathcal{N}$ pairwise inconsistent, occurring in the same spatiotemporal region, (scattered) outcomes $\mathbf{b}_1^n, \mathbf{b}_2^n, \dots, \mathbf{b}_J^n$ such that B^n is below each $\mathbf{b}_j^n$ $(i \leqslant J)$, and each such outcome is SLR to A^n;

4. if $A^n \cup B^n$ occurs in history h, then for some $i \leqslant I, j \leqslant J$, both $\mathbf{a}_i^n$ and $\mathbf{b}_j^n$ occur in h;

5. $\mathbf{a}_i$ and $\mathbf{b}_j$ are consistent, for every $i \leqslant I$ and $j \leqslant J$;

6. for any history h and any n, A^n occurs in h iff B^n occurs in h, i.e., $H_{[A^n]} = H_{[B^n]}$.

From conditions (1), (2) and (3) it immediately follows that for every $n \in \mathcal{N}$, there are two sets of transitions, induced by A^n and B^n, respectively, namely $\mathfrak{A}^n := \{A^n \to \mathbf{a}_I^n \mid 0 < i \leqslant I\}$ and $\mathfrak{B}^n := \{B^n \to \mathbf{b}_j^n \mid 0 < j \leqslant J\}$. With clause (4) added, they also imply that for every $n \in \mathcal{N}$ there is a third set $\mathfrak{C}^n := \{A^n \cup B^n \to \mathbf{a}_i^n \cup \mathbf{b}_j^n \mid 0 < i \leqslant I \text{ and } 0 < j \leqslant J\}$ of transitions from initials to (scattered) outcomes–see Fact (23).

Having accepted the "enumerating" convention: $(A \cup B)^n := A^n \cup B^n$ and $(\mathbf{a}_i^n \cup \mathbf{b}_j)^n := \mathbf{a}_i^n \cup \mathbf{b}_j^n$, it is immediate to see that an MBS $\mathfrak{M}$ which fulfills the above desiderata weakly represents our statistical experiments in terms of history σ and the sequence $\mathfrak{C}^1, \mathfrak{C}^2, \mathfrak{C}^3, \dots$ of sets of transitions.

Observe that, in the terminology developed in Section (5), the stipulations above amount to saying that for every $n \in \mathcal{N}$, $\mathfrak{A}^n$ and $\mathfrak{B}^n$ is a set of exclusive alternative transitions of the same spatiotemporal location, and each such pair induces a third set $\mathfrak{C}^n$ of alternative transitions of the same spatiotemporal location. Further, given that A^n and B^n occur in exactly the same histories, the transitions of the last set are also exclusive—cf. facts of section (5).

We next need to equip our model with propensities, defined for every relevant transition, that is, for every transition from $\mathfrak{A}^n$, $\mathfrak{B}^n$ and $\mathfrak{C}^n$, for every $n \in \mathcal{N}$. Given the facts of section (5), it suffices to postulate that some particular causally significant sets, called sets of past causal loci (*pcl*), are eligible for producing a causal probability space. We require that the numerical values of these propensities recover experimentally estimated probabilities and that with respect to these propensities, every transition from $\mathfrak{A}^n$ be probabilistically independent from every transition from $\mathfrak{B}^n$.

[10] In the sense of $A^n <_{\forall\exists} \mathbf{a}_i^n$.

The existence of a model fully representing the experiment requires a set Λ of types of runs, with each $\lambda \in \Lambda$ having associated weight w_λ. Empirical Adequacy (clause 3 of Definition (6)) then dictates:

$$p^{ex}(\tilde{a}_i \text{ and } \tilde{b}_j) = \sum_{\lambda \in \Lambda} w_\lambda \cdot pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)} \cup \mathbf{b}_j^{\Theta(\lambda)}) \tag{3}$$

for every $n \in \mathcal{N}, \lambda \in \Lambda, i \leqslant I$, and $j \leqslant J$, where $\Theta : \Lambda \to \mathcal{N}$ is such that $\Theta(\lambda) \in \lambda$.

What is, however, a proper formulation of the condition that with respect to propensities, every transition from $\mathfrak{A}^n$ be probabilistically independent from every transition from $\mathfrak{B}^n$? It turns out that, thanks to desiderata (2), (3), (5) and (6) imposed on our MBS, the assumptions of Definition (32) of a special case of probabilistic independence are satisfied, and this independence condition amounts to the satisfaction of these equations:

$$pr(A^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)}) pr(B^{\Theta(\lambda)} \to \mathbf{b}_j^{\Theta(\lambda)}) = pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)} \cup \mathbf{b}_j^{\Theta(\lambda)}). \tag{4}$$

Above Θ is a "selection function"— for every type of runs it chooses a representative run of this type. Formally, $\Theta : \Lambda \to \mathcal{N}$ and $\Theta(\lambda) \in \lambda$.

Moreover, $\mathfrak{A}^n$ and $\mathfrak{B}^n$ satisfy the assumptions of Fact 29, which in turn allows us to rewrite the above set of equations as follows:

$$\begin{split}(\sum_{l \leqslant J} pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)} \cup \mathbf{b}_l^{\Theta(\lambda)}))(\sum_{m \leqslant I} pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \\ \mathbf{a}_m^{\Theta(\lambda)} \cup \mathbf{b}_j^{\Theta(\lambda)})) = pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)} \cup \mathbf{b}_j^{\Theta(\lambda)}),\end{split} \tag{5}$$

We will now argue that there is a BST model with finitely many types of runs and propensities that fully represents the statistical experiment described, and in this model space-like related transitions are independent propensity-wise. That is, there is a finite partition Λ of $\mathcal{N}$ such that for every $\lambda \in \Lambda$, its weight w_λ is non-zero and the propensity assignment satisfies Equations (3) and (5). Our task is to find weights w_λ and propensities pr for any set of experimental probabilities p^{ex} that can arise in this experiment.

To introduce a shorter notation, let us put:

$$q_{ij} := p^{ex}(\tilde{a}_i \text{ and } \tilde{b}_j) \quad p_{ij\lambda} := pr(A^{\Theta(\lambda)} \cup B^{\Theta(\lambda)} \to \mathbf{a}_i^{\Theta(\lambda)} \cup \mathbf{b}_j^{\Theta(\lambda)}) \cdot w_\lambda$$

Then we rewrite the above two sets (3) and (5) of equations as:

$$q_{ij} = \sum_{\lambda \in \Lambda} p_{ij\lambda} \quad \text{and} \quad (\sum_{j \leqslant J} p_{kj\lambda}) \cdot (\sum_{i \leqslant I} p_{im\lambda}) = w_\lambda p_{km\lambda}. \tag{6}$$

We assume that q_{ij}'s are given and solve for $p_{ij\lambda}$'s and w_λ's.

We will argue that there are always quasi-deterministic solutions of Equations (6), in the sense that every probability $pr(A^n \cup B^n \to \mathbf{a}_i^n \cup \mathbf{b}_j^n)$ is either zero or one. To see that this is indeed the case, we assume that every probability of the form above is either zero or one, and then define partition Λ of natural numbers into at most $I \times J$ subsets λ_{ij} $(i \leqslant I, j \leqslant J)$ such that:[11]

$$\lambda_{ij} := \{n \in \mathcal{N} \mid pr(A^n \cup B^n \to \mathbf{a}_i^n \cup \mathbf{b}_j^n) = 1.\}$$

Then we may easily check that we have these solutions for Eqs. 6:

$$p_{ij\lambda} = q_{ij} \text{ if } \lambda = \lambda_{ij} \text{ and 0 otherwise} \tag{7}$$

$$\omega_\lambda = q_{ij} \text{ if } \lambda = \lambda_{ij} \text{ and 0 otherwise .} \tag{8}$$

Observe that with these solutions, types of runs work deterministically: if we are in nth run and the run is of type λ_{ij}, the propensity for the transition $(A^n \cup B^n \to \mathbf{a}_i^n \cup \mathbf{b}_j^n)$ is one, and the propensity for a transition to any other outcome of $A^n \cup B^n$ is zero. In the hidden-variable language, this is a determinate-value (or deterministic) hidden variable model.

We have just proved that there is propensity assignment that fulfills clause (3) of our Definition (6) of what it means that a BST model fully represents a statistical experiment. As for the remaining clauses, it is rather straightforward to construct an MBS model that satisfies the desiderata (†)—they guarantee that the model modally and spatiotemporally represents the statistical experiment. We need further ensure that each transition of the form $A^n \rightarrowtail \mathbf{a}_i$ or $B^n \rightarrowtail \mathbf{b}_j$ involves only finitely many finitely splitting points—this will guarantee that the transition is eligible for propensity assignment—see Section (4). As postulated above, the model has at most $I \times J$ types of runs. We need to take care that these types be distributed in history σ, in accordance with clause 2 of Definition (6). This is simple if every ω_λ is rational (which in turn depends on whether every $p^{ex}(\tilde{a}_i \text{ and } \tilde{b}_j)$ is rational) and it is only a little more complex if the above numbers are irrational. Finally, to satisfy clause 4 of the Definition, one may require that in a run of type λ_{ij} outcomes $\mathbf{a}_i$ and $\mathbf{b}_j$ occur in history σ. For scarcity of space, we leave the explicit construction of the model to the reader. For the same reason, we do not enter here into a more exciting topic, namely whether it is possible to construct a full fledged indeterministic model with uncorrelated propensities for the experiment, where "full fledged" means that all (or, most) relevant transitions have neither propensity zero nor propensity one.

[11] The partition has less than $I \times J$ elements if $q_{ij} = 0$ for some $i \leqslant I, j \leqslant J$.

4 BST, MBS, and Causal Probability Spaces

The aim of Belnap's (1992) theory of branching space-times (BST) is to combine objective indeterminism and relativity. A BST model is a non-empty partially ordered set W subject to some postulates (listed below). W is called 'Our World' and interpreted as the set of all possible point events. A partial ordering $\leqslant$ on W is interpreted as a pre-causal order between point events.[12] Typically Our World W has many possible scenarios, this last idea being rendered by a technical notion of history:

Definition 7 *A set $h \subseteq W$ is upward-directed iff $\forall e_1, e_2 \in h\ \exists e \in h$ such that $e_1 \leqslant e$ and $e_2 \leqslant e$.*
A set h is maximal with respect to the above property iff $\forall g \in W$ such that $h \subsetneq g$, g is not upward-directed.
A subset h of W is a history iff it is a maximal upward-directed set.
For histories h_1 and h_2, any maximal element in $h_1 \cap h_2$ is called a choice event for h_1 and h_2.

A BST model is then defined as follows (for more information about BST, see Belnap (1992)):

Definition 8 *$\langle W, \leqslant\rangle$, where W is a nonempty set and $\leqslant$ is a partial ordering on W, is a model of BST if and only if it meets the following requirements:*

1. *The ordering $\leqslant$ is dense.*
2. *W has no maximal elements with respect to $\leqslant$.*
3. *Every lower bounded chain in W has an infimum in W.*
4. *Every upper bounded chain in W has a supremum in every history that contains it.*
5. *For any lower bounded chain $O \subseteq h_1/h_2$ there exists an event $e \in W$ such that e is maximal in $h_1 \cap h_2$ and $\forall e' \in O:\ e < e'$ (this is the Prior Choice Principle, abbreviated as PCP).*

A consequence of PCP is that every two histories overlap and have at least one choice point. The postulates also imply that the relation $\equiv_e$ of undividedness of histories at event e, (defined by $h \equiv_e h'$ iff $\exists e' :\ e < e' \wedge e' \in h \cap h'$) is an equivalence relation on the set $H_{(e)} := \{h \in Hist \mid e \in h\}$. Thus, $\equiv_e$ induces a partition Π_e of the set $H_{(e)}$; we write $\Pi_e\langle h\rangle$ (where $h \in H_{(e)}$) for a unique

[12] We read '$e_1 \leqslant e_2$' as 'e_2 is in a possible future of e_1,' or as 'e_1 can causally influence e_2.'

element of partition Π_e to which h belongs. And we say that event e is finitely splitting if $card(\Pi_e)$ is larger than one but finite. For I a subset of a history, one may define undividedness in I: $h \equiv_I h'$ iff $h \equiv_e h'$ for every $e \in I$. This is an equivalence relation on the set $H_{[I]} := \{h \in Hist \mid I \subseteq h\}$, so it induces a partition of it, written as Π_I.

BST allows one to define a few concepts of events (apart from point events) and relations between them. For our purposes, the most important are initial events and scattered outcome events, and the relation of being space-like related.

Definition 9 (types of events and SLR) *An initial event is an upper bounded subset of a history; an outcome chain event is a lower bounded non-empty chain in W; a scattered outcome event is a set of outcome chain events all of which overlap some one history.*
Two point events are space-like related (SLR) if they are distinct, share a history, and are incomparable by the pre-causal ordering; $A \subseteq h_1$ and $B \subseteq h_2$ are SLR iff every element of A and every element of B are SLR; Initial event A and scattered outcome $\mathbf{O} = \{O_\delta\}_{\delta \in \Delta}$ are SLR iff A and O_δ are SLR, for every $\delta \in \Delta$.

An important class of structures, as they play causal and probabilistic roles, are transitions. A transition is a pair consisting of an initial event and an outcome event located properly above it. We will only use two kinds of transitions:

Definition 10 (transitions) *A transition from initial event to scattered outcome event is a pair $\langle I, \mathbf{O} \rangle$, where I and $\mathbf{O}$ are, resp., an initial event and a scattered outcome event such that $I <_{\forall\exists} \mathbf{O}$, where the last condition means that $\forall e \in I\ \exists O \in \mathbf{O} \forall x \in O\ (e < x)$.*
A basic transition is a pair $\langle e, H \rangle$, where e is a choice event and $H \in \Pi_e$.
We will denote transitions by $I \rightarrowtail \mathbf{O}$ and $e \rightarrowtail H$, resp.

A way of looking at EPR phenomena is that a combinatorially allowable history, say one with spin + on the left and spin + on the right, is not possible. A BST property used to analyze this phenomenon is called Modal Funny Business (MFB) and defined as follows:

Definition 11 (MFB) *Two initial events A and B, and two histories h_A, h_B such that $A \subseteq h_A$ and $B \subseteq h_B$ constitute a case of modal funny business (MFB) iff (1) A SLR B and (2) $\Pi_A\langle h_A \rangle \cap \Pi_B\langle h_B \rangle = \emptyset$.*

In what follows, we will always assume No Modal Funny Business.

We need a few "propositional" notions, especially the so-called occurrence propositions, defined in terms of sets of histories. We say that an event occurs in history h iff h belongs to the occurrence proposition for this event.

Definition 12 (propositions) *The occurrence proposition for point event* e *is the set* $H_{(e)} := \{h \in Hist \mid e \in h\}$*; the occurrence proposition for initial event* I *is the set* $H_{[I]} := \{h \in Hist \mid I \subseteq h\}$*; the occurrence proposition for outcome chain event* O *is the set* $H_{\langle O \rangle} := \{h \in Hist \mid O \cap h \neq \emptyset\}$*; the occurrence proposition for scattered outcome event* $\mathbf{O}$ *is the set* $H_{\langle \mathbf{O} \rangle} := \bigcap_{O \in \mathbf{O}} H_{\langle O \rangle}$.

Occurrence propositions permit one to define generally the notion of consistency. Two objects, X and Y are consistent if their occurrence propositions intersect non-emptily.

We turn next to causal notions. A novelty of Belnap's (2005) analysis is that he asks for causes of transitions. Given indeterminism, the best one can get as a causal analysis of transition $I \rightarrowtail \mathbf{O}$ is the notion of a factor that keeps $\mathbf{O}$ possible rather than prohibit its occurrence. The events relevant to this process are choice events at which some history in which I occurs branches from every history in which $\mathbf{O}$ occurs. These events are called *causal loci* of a transition; some particular basic transitions, whose initials are *past* causal loci, are called originating causes, or *causae causantes* of a transition considered. Belnap proves that they satisfy a version of Mackie's INUS condition. The definitions are as follows:

Definition 13 (pcl and cc) *Let* $I \rightarrowtail \mathbf{O}$ *be a transition from initial event to scattered event.*
Event e *belongs to the set* $cl(I \rightarrowtail \mathbf{O})$ *of cause-like loci of the transition iff* $\exists h \in H_{[I]}\ h \perp_e H_{\langle \mathbf{O} \rangle}$.
Event e *belongs to the set* $pcl(I \rightarrowtail \mathbf{O})$ *of past causal loci of the transition iff* $e \in cl(I \rightarrowtail \mathbf{O})$ *and* $\exists O \in \mathbf{O} \forall x \in O\ e < x$.
Basic transition $t = e \rightarrowtail H$ *belongs to the set* $cc(I \rightarrowtail \mathbf{O})$ *of causae causantes of the transition iff* $e \in pcl(I \rightarrowtail \mathbf{O})$ *and* $H = \Pi_e\langle h \rangle$ *for some history* $h \in H_{\langle \mathbf{O} \rangle}$.

An important theorem, proved by Belnap (2002), says that if there is no MFB, then every element of the set of cause-like loci of a transition to a scattered outcome is in the past of some element of the scattered outcome. We put this theorem in the following form:

Theorem 14 (NB's theorem) *If there is no MFB in a BST model, then for any transition* $I \rightarrowtail \mathbf{O}$ *to scattered outcome,* $cl(I \rightarrowtail \mathbf{O}) <_{\forall\exists} \mathbf{O}$.

In our modeling of statistical experiments, we assume that the underlying space-times are Minkowski's. Thus, we need a class of BST models, in which histories are isomorphic to the Minkowski space-time. These are the so-called Minkowskian Branching Structures (MBS's).[13]

For our present purposes, it suffices to say that in an MBS point events have spatiotemporal locations; we write $s(e)$ for spatiotemporal location of point event e (for brevity, in what follows we abbreviate this term to "st-location"). In a natural way, we can use st-locations of point events to define the concept st-location for initial events, outcome chain events, scattered outcomes, etc.

4.1 Causal Probability Spaces

A (classical) probability space is a triple $\langle A, F, \mu \rangle$, where A is a non-empty set called a sample space, F_E is a Boolean σ-algebra of subsets of A (called event algebra), and μ is a normalized to unity, countably additive measure on F. For Müller's (2005) concept of causal probability space (the generalized version), consider a finite set E of finitely splitting points such that minimal elements of E are consistent (since E is finite, its every element is above some minimal element of E). The elements of a causal probability space $\langle A_E, F_E, \mu_E \rangle$ associated with set E are constructed as follows:

1. Consider the complete set of alternative basic transitions that have initials in E, i.e., set $\tilde{T}_E = \{t \in \mathcal{T}r \mid i(t) \in E)\}$, where $\mathcal{T}r$ is the set of basic transitions and $i(t)$ is the initial of transition t.

2. Identify the sample space A_E with the set of all maximally consistent subsets of $\tilde{T}_E$,

3. Since the case is finite, take for event algebra F_E the set-theoretical algebra of subsets of A_E,

4. Define a normalized to unity measure μ_E on F_E.

To explain clause 2 above, two basic transitions $e_1 \rightarrowtail H_1$ and $e_2 \rightarrowtail H_2$ are consistent iff $H_1 \cap H_2 \neq \emptyset$. As was mentioned before, there are finististic assumptions concerning the construction of causal probability spaces. For the record, we capture them here in this definition:

Definition 15 (eligible for producing a causal probability space) *$E \subset W$ is eligible for producing a causal probability space if E is a finite set of finitely splitting points and the minimal elements of E are consistent.*

[13]They were introduced by Placek (2000a), then rigorously defined for finitary cases by (Müller, 2002), and defined generally by Wroński and Placek (2009).

Many causal probability spaces may "live" in a given BST model, and a set of basic transitions may belong to many event algebras, each associated with a different subset E of W. Thus, a question arises of how to represent an element of a "smaller" probability space in a "larger" probability space so that one could subsequently postulate that an element of the smaller probability space and its representative in the larger probability space are ascribed the same value by the two probability measures. That is, let E and E' be sets of splitting points eligible for producing causal probability spaces, with $\tilde{T}_E$ and $\tilde{T}_{E'}$ their complete sets of alternative basic transitions, resp., and such that $\tilde{T}_E \subseteq \tilde{T}_{E'}$.

Müller defines representability only if $E \subseteq E'$ and requires that the larger set $\tilde{T}_{E'}$ differs from the smaller set $\tilde{T}_E$ only by transitions located above those of $\tilde{T}$, i.e., if $e' \in E'/E$, then $\forall e: \ e \in E \rightarrow e < e'$. Since we cannot fulfill this last postulate, we need to generalize the concept. We first state when one event algebra represents some other event algebra. Next, we define an auxiliary notion of rep of an element of one sample space in an event algebra. With the help of this notion we finally state what a representative of an element of one event algebra in some other event algebra is.

Definition 16 (representability) *Let $E \subseteq W$ and $E' \subseteq W$ be sets of splitting points, each eligible for the construction of a causal probability space. F_E is said to be representable in $F_{E'}$ iff $E \subseteq E'$, and if $e' \in E'/E$, then ($\forall e \in E$: $e < e'$ or e' SLR E).*

For $a \in A_E$, the rep $r_{E'}(a)$ of a in $F_{E'}$ is: $r_{E'}(a) := \{a' \in A_{E'} \mid a \subseteq a'\}$.

For $x \in F_E$, the representative $R_{E'}(x)$ of x in $F_{E'}$ is:

$$R_{E'}(x) := \bigcup \{z \in F_{E'} \mid \exists a \in A_E \ a \in x \wedge z = r_{E'}(a)\}.$$

Having the notion of representative, we may state the Marginal Property, which we assume in the rest of this paper. Marginal Property requires that an element of one event algebra and its representative in some other event algebra are assigned the same numerical value by the probability measures defined on the two event algebras:

Definition 17 (marginal property) *Let $\langle A_E, F_E, \mu_E \rangle$ and $\langle A_{E'}, F_{E'}, \mu_{E'} \rangle$ be causal probability spaces with F_E being representable in $F_{E'}$. Then if $x' \in F_{E'}$ is the representative of $x \in F_E$,*

$$\mu_{E'}(x') = \mu_E(x).$$

Recall that an element of the sample space of a causal probability space is a maximal consistent set of basic transitions. One might have a foreboding that basic transitions should be independent. With this intuition it is tempting to postulate that the probability of an atom be the product of probabilities of its elements — basic transitions. As Nature can go her own unaccountable ways, we had better not assume this postulate, but merely single out those causal probability spaces that satisfy this condition.

Definition 18 (uncorrelated probability space) *Let $\langle A_E, F_E, \mu_E \rangle$ be a causal probability space. We say that it is uncorrelated if for every $a \in A_E$:*

$$\mu_E(A) = \prod_{x \in a} \mu_{\{i(x)\}}(x),$$

where $i(x)$ is the initial event of basic transition x.

Note thats a causal probability space is produced out of a set of basic transitions, subject to some requirements. Can we use it to assign probabilities to other BST structures, in particular to (non-basic) transitions? It seems utterly reasonable that if probability is defined for a certain transition, it should be derivable from probabilities of its causae causantes. This is a desideratum of the 'nothing but causae causantes' that Weiner and Belnap (2006) assumed (their Postulate 4-4):

> When $pr(I \rightarrowtail \mathbf{a}_i)$ is defined, nothing counts except nature-given stochastic features of its *causae causantes*—including the possibility that one may need to take into account not only probabilities of individual *causae causantes*, but also probabilities of certain sets of them, taken as operating jointly. [...]

In line with the tradition, we will call single-case objective probabilities of transitions propensities, and write them as $pr()$. Allowing that causae causantes may work together, the desideratum above boils down to this definition of propensities:

Definition 19 (propensities) *A propensity of transition $I \rightarrowtail \mathbf{O}$, if defined, is the following:*

$$pr(I \rightarrowtail \mathbf{O}) = \mu_{pcl((I \rightarrowtail \mathbf{O})}(I \rightarrowtail \mathbf{O}).$$

5 Facts About Structures Occurring in EPR-Like Experiments

The purpose of this section is to justify our (seemingly loose) talk of propensities by relating it to hard facts about the so-called causal probability spaces of BST. The justification relies on showing that structures occurring in our BST models allow for defining causal probability spaces, exactly of the sort that assign propensities to the transitions considered. In our analysis of statistical experiments, we have typically considered finitely many transitions from one initial event to scattered outcome events, where the scattered outcome events are pairwise inconsistent. A set of such transitions represents a measurement event with its alternative possible outcome events (both kinds of events are token-level). We define:

Definition 20 (exclusive alternative transitions) *Let $Tr = \{I \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ be a set of transition from initial event I to scattered outcome events $\mathbf{a}_i$. We say that Tr is a set of exclusive alternative transitions of the same st-location iff*

1. $H_{\langle \mathbf{a}_i \rangle} \cap H_{\langle \mathbf{a}_j \rangle} = \emptyset$ *if* $i \neq j$; *alternative outcomes*
2. $\forall h \exists j : \ h \in H_{[I]} \rightarrow h \in H_{\langle \mathbf{a}_j \rangle}$, *exclusiveness*
3. *all* $\mathbf{a}_i$*'s have the same st-location.* *same st-location*

For each transition $A \rightarrowtail \mathbf{a}_i$, its set $pcl(A \rightarrowtail \mathbf{a}_i) =: pcl_i$ is consistent. We will always additionally assume that each pcl_i is finite and its every element is finitely splitting. We aim to build a causal probability space associated with $E := \bigcup_{i \leqslant I} pcl_i$, but for this to be viable, the minimal elements of E should be consistent. The fact below testifies that indeed this is the case.

Fact 21 *Let $Tr = \{I \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ be a set of exclusive alternative transitions of the same st-location such that each $pcl_i := pcl(I \rightarrowtail \mathbf{a}_i)$ is finite and its elements finitely splitting. Then minimal elements in $pcl := \bigcup_{i \leqslant I} pcl_i$ are consistent.*

Proof. Assume to the contrary that the set of minimal elements of pcl is not consistent. Then for every $h \in H_{[I]}$ there is a minimal element $e \in pcl$ such that $e \notin h$. Accordingly, one may take an arbitrary $h \in H_{[I]}$ and some $i \leqslant I$ such that $e \in pcl(I \rightarrowtail \mathbf{a}_i)$ but $e \notin h$ and e is minimal in pcl. Pick now some $h' \in H_{\langle \mathbf{a}_i \rangle}$; clearly $e \in h'$. By applying PCP to e, h, and h', we have $\exists e' : \ e' < e \wedge h \perp_{e'} h'$, from which $h \perp_{e'} H_{\langle \mathbf{a}_i \rangle}$ follows. Accordingly,

$e' \in pcl(I \rightarrowtail \mathbf{a}_i)$. Hence $e' < e$ contradicts the assumption that e is a minimal element of pcl. $\square$

We will now argue that for a set $Tr = \{I \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ of exclusive alternative transitions of the same st-location, such that $pcl := \bigcup_{i \leqslant I} pcl(I \rightarrowtail \mathbf{a}_i)$ is eligible for producing a causal probability space, the sample space A_{pcl} is exactly the set of all sets of causae causantes of transitions from Tr.

Fact 22 *Let $Tr = \{I \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ be a set of exclusive alternative transitions of the same st-location such that pcl $:= \bigcup_{i \leqslant I} pcl(I \rightarrowtail \mathbf{a}_i)$ is eligible for producing a causal probability space. Then the sample space A_{pcl} of the probability space $\langle A_{pcl}, F_{pcl}, \mu_{pcl}\rangle$ is: $A_{pcl} = \{cc(I \rightarrowtail \mathbf{a}_i) \mid i \leqslant I\}$.*

Proof: To the left first. Supppose for reductio that a certain $cc(I \rightarrowtail \mathbf{a}_i)$ is not a member of A_{pcl}, which means that it is not a *maximally* consistent subset of $\tilde{T}_{pcl}$ (recall that every set of causae causantes is consistent). Accordingly, there is $t \in \tilde{T}_{pcl}$, $t = e \rightarrowtail H$, such that $t \notin cc(I \rightarrowtail \mathbf{a}_i)$ and $cc(I \rightarrowtail \mathbf{a}_i) \cup \{t\}$ is consistent $(\star)$. If $e \in pcl(I \rightarrowtail \mathbf{a}_i)$, then either t belonged or was inconsistent with the set $cc(I \rightarrowtail \mathbf{a}_i)$. It thus must be that $e \in pcl(I \rightarrowtail \mathbf{a}_j)$ for some $j \neq i$, $j \leqslant I$. Accordingly, $e <_{\exists} \mathbf{a}_j$. But it cannot be that $(\dagger)$ $e <_{\exists} \mathbf{a}_i$: if $(\dagger)$ were true, $H_{\langle \mathbf{a}_i \rangle} \subseteq H^*$ for the same $H^* \in \Pi_e$, and, accordingly, $e \in pcl(A \rightarrowtail \mathbf{a}_i)$. Since $\mathbf{a}_i$ and $\mathbf{a}_j$ have the same spatiotemporal st-location, for $\neg(e <_{\exists} \mathbf{a}_i)$ to obtain, e and $\mathbf{a}_i$ must be inconsistent, from which it follows that t and $cc(I \rightarrowtail \mathbf{a}_i)$ are inconsistent. Contradiction with $(\star)$.

In the opposite direction, A_{pcl} is the set of maximally consistent subsets of $\tilde{T} = \{t \in \mathcal{B} \mid i(t) \in pcl)\}$, where $\mathcal{B}$ is the set of basic transitions and $i(t)$ is the initial event of basic transition t. Clearly, every $cc(I \rightarrowtail \mathbf{a}_i)$ is a maximally consistent subset of $\tilde{T}$. It remains to be seen that there is no maximal consistent subset of $\tilde{T}$ that is not $cc(I \rightarrowtail \mathbf{a}_i)$ for some $i \leqslant I$. By the condition of exclusiveness, every element of $\tilde{T}$ belongs to some $cc(I \rightarrowtail \mathbf{a}_i)$ for $i \leqslant I$. It remains to be seen that no element of A_{pcl} can have two elements, $t_1 : e_1 \rightarrowtail H_1$ and $t_2 := e_2 \rightarrowtail H_2$ such that, for some $i \neq j \leqslant I$:

$$t_1 \in cc(I \rightarrowtail \mathbf{a}_i) \text{ but } t_1 \notin cc(I \rightarrowtail \mathbf{a}_j) \text{ and } t_2 \notin cc(I \rightarrowtail \mathbf{a}_i) \text{ but } t_2 \in cc(I \rightarrowtail \mathbf{a}_j).$$

It cannot be that $e_1 \in pcl(I \rightarrowtail \mathbf{a}_j)$; otherwise $t_1 \in cc(I \rightarrowtail a_j)$ or t_1 is inconsistent with t_2. Similarly, $e_2 \notin pcl(I \rightarrowtail \mathbf{a}_i)$. Clearly, $e_1 <_{\exists} \mathbf{a}_i$. Yet, $\mathbf{a}_i$ and $\mathbf{a}_j$ are in the same st-location. But it cannot be that $e_1 <_{\exists} \mathbf{a}_j$, since then e_1 would belong to $pcl(I \rightarrowtail \mathbf{a}_j)$. It must thus be that e_1 and $H_{\langle \mathbf{a}_j \rangle}$ are inconsistent, and hence, (since $e_1 <_{\exists} \mathbf{a}_i$) $H_{\langle \mathbf{a}_i \rangle} \cap H_{\langle \mathbf{a}_j \rangle} = \emptyset$. Hence, t_1 is inconsistent with t_2, which means that the two cannot belong to any one element of A_{pcl}.

□

To represent a joint experiment in which two measurements occur in separate "stations," we need two sets of exclusive alternative transitions of the same st-location. Each set represents a measurement in one station. It is a welcome consequence that such two sets induce a third set of exclusive alternative transitions of the same st-location, naturally representing the joint measurement and its possible joint outcomes:

Fact 23 *Let $Tr_a := \{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ and $Tr_b := \{B \rightarrowtail \mathbf{b}_j \mid i \leqslant J\}$ be sets of exclusive alternative transitions of the same st-location such that $\mathbf{a}_i \cup \mathbf{b}_j$ is a scattered outcome for every $i \leqslant I, j \leqslant J$. Then $Tr_{ab} := \{A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j \mid i \leqslant I, j \leqslant J\}$ is a set of exclusive alternative transitions of the same st-location.*

Proof: Immediate. □

The induced set has an interesting property: the sets of past causal loci of, resp., Tr_a, Tr_b and the induced set TR_{ab}, are nicely related:

Fact 24 *For initial events A, B, scattered outcomes $\mathbf{a}, \mathbf{b}$, and $\mathbf{a} \cup \mathbf{b}$ such that $A <_{\forall\exists} \mathbf{a}$, $B <_{\forall\exists} \mathbf{b}$:*

$$pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b}) \subseteq pcl(A \rightarrowtail \mathbf{a}) \cup pcl(B \rightarrowtail \mathbf{b}).$$

Proof. From the order relations assumed in the premises, $A \cup B <_{\forall\exists} \mathbf{a} \cup \mathbf{b}$. Let $e \in pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})$. Then for some $h \in H_{A \cup B}$: $h \perp_e H_{\langle \mathbf{a} \cup \mathbf{b} \rangle}$ and $e <_{\exists} \mathbf{a} \cup \mathbf{b}$, and hence $e <_{\exists} \mathbf{a}$ or $e <_{\exists} \mathbf{b}$. Suppose the former (the other case is symmetric). Clearly, $h \in H_{[A]}$. Pick now an arbitrary $h' \in H_{\langle \mathbf{a} \cup \mathbf{b} \rangle}$; it follows that $h' \in H_{\langle \mathbf{a} \rangle}$ and $h \perp_e h'$, so $h \perp_e \Pi_e \langle h' \rangle$. Since $e <_{\exists} \mathbf{a}$ and $h' \in H_{\langle \mathbf{a} \rangle}$, it follows that $H_{\langle \mathbf{a} \rangle} \subseteq \Pi_e \langle h' \rangle$. Accordingly, $h \perp_e H_{\langle \mathbf{a} \rangle}$, which proves that $e \in pcl(A \rightarrowtail \mathbf{a})$.

□

Fact 25 *Let $A \rightarrowtail \mathbf{a}$ and $A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{a}$ be transitions from initials to scattered outcomes. Let $H_{[A]} \subseteq H_{[B]}$. Then $pcl(A \rightarrowtail \mathbf{a}) \subseteq pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})$.*

Proof: Let $e \in pcl(A \rightarrowtail \mathbf{a})$. Then for some $h \in H_{[A]}$: $h \perp_e H_{\langle \mathbf{a} \rangle}$. By the premise ($H_{[A]} \subseteq H_{[B]}$), $h \in H_{[A \cup B]}$. Since $H_{\langle \mathbf{a} \cup \mathbf{b} \rangle} \subseteq H_{\langle \mathbf{a} \rangle}$, it follows that $h \perp_e H_{\langle \mathbf{a} \cup \mathbf{b} \rangle}$. Thus $e \in pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})$.

□

Combining Facts (24) and (25) above, we arrive at the following:

Fact 26 *For initial events A, B, scattered outcomes $\mathbf{a}, \mathbf{b}$, and $\mathbf{a} \cup \mathbf{b}$ such that $A <_{\forall\exists} \mathbf{a}$, $B <_{\forall\exists} \mathbf{b}$, and $H_{[A]} = H_{[B]}$:*

$$pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b}) = pcl(A \rightarrowtail \mathbf{a}) \cup pcl(B \rightarrowtail \mathbf{b}).$$

This fact has a consequence for sets of causaes causantes:

Fact 27 *For initial events A, B, scattered outcomes $\mathbf{a}, \mathbf{b}$, and $\mathbf{a}\cup\mathbf{b}$ such that $A <_{\forall\exists} \mathbf{a}$, $B <_{\forall\exists} \mathbf{b}$, and $H_{[A]} = H_{[B]}$:*

$$cc(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b}) = cc(A \rightarrowtail \mathbf{a}) \cup cc(B \rightarrowtail \mathbf{b}).$$

As a hint for the proof, note that if $e <_{\exists} \mathbf{a}\cup\mathbf{b}$, then $\Pi_e\langle\mathbf{a}\cup\mathbf{b}\rangle$ is identical to $\Pi_e\langle\mathbf{a}\rangle$ or to $\Pi_e\langle\mathbf{b}\rangle$, or to both, depending on whether $e <_{\exists} \mathbf{a}$, or $e <_{\exists} \mathbf{b}$, or both. In the opposite direction, if $e <_{\exists} \mathbf{a}$, then $\Pi_e\langle\mathbf{a}\rangle$ is identical to $\Pi_e\langle\mathbf{a}\cup\mathbf{b}\rangle$, as $\mathbf{a}\cup\mathbf{b}$ is a scattered outcome by assumption.

Suppose now that we have three sets of exclusive alternative transitions: $Tr_a := \{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$, $Tr_b := \{B \rightarrowtail \mathbf{b}_j \mid i \leqslant J\}$ and $Tr_{ab} := \{A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j \mid i \leqslant I, j \leqslant J\}$ (the induced set). Suppose further that their past causal loci yield causal probability spaces, associated respectively with $E_1 := \bigcup_{i\leqslant I} pcl(A \rightarrowtail \mathbf{a}_i)$, $E_2 := \bigcup_{j\leqslant J} pcl(B \rightarrowtail \mathbf{b}_j)$, and $E := \bigcup_{i\leqslant I, j\leqslant J} pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$. We would like to represent elements of the event algebra associated with E_i $i \in \{1, 2\}$ in the event algebra associated with E. But is E eligible for producing a causal probability space, if each E_i is eligible for producing a causal probability space? And are the conditions of representability satisfied? Under a rather special condition of $H_{[A]} = H_{[B]}$, this is the case indeed.

We need to prove that (1) $E_i \subseteq E$ and (2) $e \in E/E_i \rightarrow (e >_{\forall} E_i \vee e\ SLR\ E_i)$, (3) the minimal elements of E are consistent if minimal elements of E_i are consistent, and (4) E is finite and its elements are finitely splitting, if E_i is finite and its elements are finitely splitting.

Given that $H_{[A]} = H_{[B]}$, (1) is a consequence of Fact (25). (4) follows from Fact (24). For (3), suppose that the set of minimal elements of E_1 is a subset of history h; then $A \subseteq h$. By $H_{[A]} = H_{[B]}$ we have $B \subseteq h$. By the exclusiveness condition on Tr_b, every minimal element of E_2 must be in h. Finally, (2) is proved in this Fact:

Fact 28 *Suppose that there are three sets of exclusive alternative transitions: $Tr_a := \{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$, $Tr_b := \{B \rightarrowtail \mathbf{b}_j \mid i \leqslant J\}$ and (the induced set) $Tr_{ab} := \{A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j \mid i \leqslant I, j \leqslant J\}$ such that $H_{[A]} = H_{[B]}$. Then: if $e \in \bigcup_{i \leqslant I, j \leqslant J} pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j) := E$, but $e \notin \bigcup_{i \leqslant I} pcl(A \rightarrowtail \mathbf{a}_i) = E_1$, then $e >_{\forall} E_1 \vee e\ SLR\ E_1$.*

Proof: Given Fact 26, the premises imply that $e \in pcl(B \rightarrowtail \mathbf{b}_j$ for some $j \leqslant J$, but (†)$\forall i \leqslant I : e \notin pcl(A \rightarrowtail \mathbf{a}_i)$. For reductio, let $e < e'$ for some $e' \in E_1$. It follows that $e' \in pcl(A \rightarrowtail \mathbf{a}_k)$ for some $k \leqslant I$. Since $e < e'$, it must be that $e \in pcl(A \rightarrowtail \mathbf{a}_k)$. Contradiction with (†). □

Representatives. To investigate representability in these settings, let us assume that $Tr_a := \{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ and $Tr_b := \{B \rightarrowtail \mathbf{b}_j \mid i \leqslant J\}$ are sets of exclusive alternative transitions of the same st-location, such that $H_{[A]} = H_{[B]}$ and that the two sets induce $Tr_{ab} = \{A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j \mid i \leqslant I, j \leqslant J\}$, yet another set of exclusive alternative transitions of the same st-location. Consider now four sets of splitting points, with which we wish to associate causal probability spaces:

$C := pcl(A \rightarrowtail \mathbf{a}_i)$ for a fixed $i \leqslant I$

$D := \bigcup_{i \leqslant I} pcl(A \rightarrowtail \mathbf{a}_i)$,

$E := \bigcup_{i \leqslant I, j \leqslant J} pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$

$G := pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$ for fixed $i \leqslant I, j \leqslant J$.

As explained in Definition (19), propensity of $A \rightarrowtail \mathbf{a}_i$ is determined by measure μ_C, taken on the set of causae causantes of the transition, that is:[14]

$$pr(A \rightarrowtail \mathbf{a}_i) = \mu_C(cc(A \rightarrowtail \mathbf{a}_i)).$$

Given that $pcl(A \rightarrowtail \mathbf{a}_i)$ is eligible for producing a causal probability space (for every $i \leqslant I$), does D the same property? Clearly, $C \subseteq D$ and D is a finite set of finitely splitting points, if each $pcl(A \rightarrowtail \mathbf{a}_i)$ is. So D has minimal elements. Moreover, the minimal elements are consistent—cf. Fact (19). Is then the event algebra F_C representable in F_D? Obviously, $C \subseteq D$ and if $e \in D/C$, then $e >_\forall C$; otherwise for some e': $e < e'$ and $e' \in pcl(A \rightarrowtail \mathbf{a}_k))$, where $i \neq k$, which entails $e \in pcl(A \rightarrowtail \mathbf{a}_k))$. Contradiction with $e \in D/C$.

Finally, by Fact (22), $cc(A \rightarrowtail \mathbf{a}_i) \in A_C$, i.e., the singleton of $cc(A \rightarrowtail \mathbf{a}_i)$ is an atom in event algebra associated with C. This singleton is also an atom in event algebra associated to D, i.e., $cc(A \rightarrowtail \mathbf{a}_i) \in A_D$. For by the condition of alternative outcomes, adding to $cc(A \rightarrowtail \mathbf{a}_i)$ a transition from $cc(A \rightarrowtail a_k)/cc(A \rightarrowtail \mathbf{a}_i)$ (where $k \neq i$) would destroy consistency. Thus,

$$\mu_D(cc(A \rightarrowtail \mathbf{a}_i)) = \mu_C(cc(A \rightarrowtail \mathbf{a}_i)) = pr(A \rightarrowtail \mathbf{a}_i).$$

Let us next turn to representability of events from F_D in the event algebra F_E. By Fact (26), and the usual assumptions, F_E is eligible for producing a causal probability space. By Fact (25), $D \subseteq E$. And by Fact (28), if $e \in E/D$, then $e >_\forall D \vee e$ SLR D. Thus, F_D is representable in F_E. It is then immediate to see that the rep of $cc(A \rightarrowtail \mathbf{a}_i) \in A_D$ in F_E is the set: $\{cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j) \mid j \leqslant J\}$. We thus have:

$$pr(A \rightarrowtail \mathbf{a}_i) = \mu_C(cc(A \rightarrowtail \mathbf{a}_i)) = \mu_D(cc(A \rightarrowtail \mathbf{a}_i)) = \\ \mu_E(\bigcup \{cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j) \mid j \leqslant J\}) = \sum_{j \leqslant J} \mu_E(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)), \quad (9)$$

[14]Since the singleton of $cc(A \rightarrowtail \mathbf{a}_i)$ belongs to F_C, and not $cc(A \rightarrowtail \mathbf{a}_i)$ itself, we should write $\mu_C(\{cc(A \rightarrowtail \mathbf{a}_i)\})$; to avoid eyestrain, we neglect the curly brackets.

where the last equation follows from the fact that the singleton of each $cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$ is an atom of F_E. Consider finally set G: clearly, $G \subseteq E$, so if E is eligible for producing a causal probability space, so is G. By an argument exactly like the one at the beginning of this paragraph, F_G is representable in F_E. And clearly, $\{(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)\} \in F_G$ has the representative $\{(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)\} \in F_E$. Hence:

$$\mu_E(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)) = \mu_G(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)) = pr(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j). \tag{10}$$

The last equation follows from the Definition (19) of propensities. Combining equations (9) and (10) above, we get this rule on calculating propensities in models satisfying rather special conditions:

Fact 29 *Let each* $\{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ *and* $\{B \rightarrowtail \mathbf{b}_j \mid j \leqslant J\}$ *be a set of exclusive alternative transitions of the same st-location and* $\mathbf{a}_i$ *consistent with* $\mathbf{b}_j$ *for every* $i \leqslant I, j \leqslant J$. *Let each* $E_i := pcl(A \rightarrowtail \mathbf{a}_i)$ *be eligible for producing a causal probability space and* $H_{[A]} = H_{[B]}$. *Then*

$$pr(A \rightarrowtail \mathbf{a}_i) = \sum_{j \leqslant J} pr(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j). \tag{11}$$

6 Towards an Adequate Statement of Probabilistic Funny Business

Having seen how to calculate propensities in causal probability spaces satisfying a rather particular condition, we now turn towards formulating a condition to the effect that two transitions, with SLR initials, are probabilistically independent, or (in the BST terminology), do not constitute probabilistic funny business (PFB). For our search of a proper condition, it is instructive to note that under the particular circumstances assumed in the last section, sets of causae causantes of two transitions do not overlap.

Fact 30 *Let each* $\{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ *and* $\{B \rightarrowtail \mathbf{b}_j \mid j \leqslant J\}$ *be a set of exclusive alternative transitions of the same st-location such that every* $\mathbf{a}_i$ *is consistent with every* $\mathbf{b}_j$. *Then*

$$pcl(A \rightarrowtail \mathbf{a}_i) \cap pcl(B \rightarrowtail \mathbf{b}_j) = \emptyset.$$

Proof: Assume to the contrary that there is $e \in pcl(A \rightarrowtail \mathbf{a}_i) \cap pcl(B \rightarrowtail \mathbf{b}_j)$, which entails $e <_\exists \mathbf{a}_i \wedge e <_\exists \mathbf{b}_j$. There are now two cases:
(i) $H_{\langle \mathbf{a}_i \rangle} \perp_e H_{\langle \mathbf{b}_j \rangle}$, from which it follows that $\mathbf{a}_i$ and $\mathbf{b}_j$ are inconsistent, which contradicts the assumption.

(ii) there is some history h such that $h \notin H_{\langle \mathbf{a}_i \rangle}$, $h \notin H_{\langle \mathbf{b}_j \rangle}$, $h \perp_e H_{\langle \mathbf{a}_i \rangle}$ and $h \perp_e H_{\langle \mathbf{b}_j \rangle}$. But, since $e \in pcl(A \rightarrowtail \mathbf{a}_i)$ and $e \in h$, it must be that $A \subset h$, and then by Exclusiveness, $h \in H_{\langle \mathbf{a}_k \rangle}$ for some $k \leqslant I$. By the condition of the same st-location of Definition 20: $e <_\exists \mathbf{a}_k$, and hence $H_{\langle \mathbf{a}_k \rangle} \perp_e H_{\langle \mathbf{b}_j \rangle}$, from which inconsistency of $\mathbf{a}_k$ and $\mathbf{b}_j$ follows, contradicting the assumption.

The immediate consequence is no overlap of the respective sets of causae causantes:

Fact 31 *Let each* $\{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ *and* $\{B \rightarrowtail \mathbf{b}_j \mid j \leqslant J\}$ *be a set of exclusive alternative transitions of the same st-location such that every* $\mathbf{a}_i$ *is consistent with every* $\mathbf{b}_j$*. Then*

$$cc(A \rightarrowtail \mathbf{a}_i) \cap cc(B \rightarrowtail \mathbf{b}_j) = \emptyset.$$

What would be a proper statement of the probabilistic independence of two transitions $A \rightarrowtail \mathbf{a}$ and $B \rightarrowtail \mathbf{b}$, each belonging to a different set of exclusive alternative transitions of the same st-location, subject to an extra condition that $H_{[A]} = H_{[B]}$? By Facts (27) and (31), if causal probability space associated with $E = pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$ is uncorrelated and $H_{[A]} = H_{[B]}$, the following obtains:

$$\begin{aligned}
&\prod_{e \in pcl(A \rightarrowtail \mathbf{a}_i)} \mu_{\{e\}}(e \rightarrowtail \Pi_e\langle \mathbf{a}_i \rangle) \cdot \prod_{e \in pcl(B \rightarrowtail \mathbf{b}_j)} \mu_{\{e\}}(e \rightarrowtail \Pi_e\langle \mathbf{b}_j \rangle) = \\
&\prod_{e \in pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)} \mu_{\{e\}}(e \rightarrowtail \Pi_e\langle \mathbf{a}_i \cup \mathbf{b}_j \rangle) = \\
&\qquad \mu_{pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)}(cc(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)).
\end{aligned} \tag{12}$$

Combining this equation with the formula for propensities and Equation (11), we get this definition of a special case of probabilistic independence (or of No PFB):

Definition 32 (probabilistic independence, special case) *Let* $A \rightarrowtail \mathbf{a}$ *and* $B \rightarrowtail \mathbf{b}$ *be transitions from initial events to scattered outcome events and* $\mathbf{a}$ *be consistent with* $\mathbf{b}$ *and* $H_{[A]} = H_{[B]}$*. Then* $A \rightarrowtail \mathbf{a}$ *and* $B \rightarrowtail \mathbf{b}$ *are probabilistically independent iff*

$$pr(A \rightarrowtail \mathbf{a}) \cdot pr(B \rightarrowtail \mathbf{b}) = pr(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b}).$$

The immediate consequence of this definition and Equation 12 is the following fact:

Fact 33 *Let $A \rightarrowtail \mathbf{a}$ and $B \rightarrowtail \mathbf{b}$ be transitions from initial events to scattered outcome events and $\mathbf{a}$ be consistent with $\mathbf{b}$ and $H_{[A]} = H_{[B]}$. Then $A \rightarrowtail \mathbf{a}$ and $B \rightarrowtail \mathbf{b}$ are probabilistically independent if the probability measure $\mu_{pcl((A\cup B\rightarrowtail \mathbf{a}\cup\mathbf{b})}$ is uncorrelated.*

This definition, although useful in analysis of an EPR-style set-up of Section (3), is not applicable to the Bell-Aspect experiment, because of the assumption $H_{[A]} = H_{[B]}$. In the experiment, a measurement event in one wing may occur together with each of a few measurement events in the other wing. Thus, we need to remove this assumption. As a consequence of relaxing the above condition, we need something like a "conditional propensity," since we will now investigate, let's say, the propensity of transition $A \rightarrowtail \mathbf{a}_i$, on the supposition that B occurs. It is not immediately clear how this concept can be related to causal probability spaces, since B is not even a transition, so there is no room for it in an event algebra of a causal probability space.

The propensity of $I \rightarrowtail \mathbf{O}$ conditional on initial J (written as $pr(I \rightarrowtail \mathbf{O} \mid J))$ should be derived from the set of only those elements of the set $\tilde{T}_{pcl(I\rightarrowtail\mathbf{O})}$ of the alternative basic transitions that do not prohibit the occurrence of J. In other words, we need to take into account only those elements of $\tilde{T}_{pcl(I\rightarrowtail\mathbf{O})}$ that are consistent with J. Given No MFB, every transition from $\tilde{T}_{pcl(I\rightarrowtail\mathbf{O})}$, whose initial is not below some element of J is consistent with J. As for $t \in \tilde{T}_{cc(I\rightarrowtail\mathbf{O})}$, whose initial is below some element of J, it is consistent with J only if it belongs to the set $cc(I \rightarrowtail \mathbf{O})$ of causae causantes and J is consistent with $\mathbf{O}$. Now, if all inconsistent alternative basic transitions with the same initial $e <_\exists J$ are removed from the set $\tilde{T}_{cc(I\rightarrowtail\mathbf{O})}$, in the "pruned" set $\tilde{T}$ there will only be a single basic transition with initial e, namely one belonging to $cc(I \rightarrowtail \mathbf{O})$. Clearly, a basic transition unaccompanied by its alternatives in the "pruned" set $\tilde{T}_{cc(I\rightarrowtail\mathbf{O})}$ has no probabilistic consequences. Hence, in constructing a casual probability space for propensity conditional on J, we remove from $pcl(I \rightarrowtail \mathbf{O})$ all its elements that lie below some element of J. Then in the usual way, we build a sample space out of maximal consistent subsets of the set of basic transitions, whose initials lie in the pruned set.

From these considerations, the following recipe emerges for the construction of a causal probability space for conditional propensity $pr(I \rightarrowtail \mathbf{O} \mid J)$:

Consider set $E = \{x \in pcl(I \rightarrowtail \mathbf{O}) \mid x \not<_\exists J\}$. Construct the complete set $\tilde{T}$ of alternative basic transitions that have initials in E. As the sample space A_E take the set of all maximally consistent subsets of $\tilde{T}$. Then in a familar way define the event algebra F_E and probability measure μ_E. The sought probability space is now the triple $\langle A_E, F_E, \mu_E\rangle$. Conditional propensity is then defined as follows:

Definition 34 *For transition $I \rightarrowtail \mathbf{O}$ from initial I to scattered outcome $\mathbf{O}$, its propensity conditional on initial J, $pr(I \rightarrowtail \mathbf{O} \mid J)$, where J is consistent with $\mathbf{O}$, is the following:*

$$pr(I \rightarrowtail \mathbf{O} \mid J) = \mu_E(cc_J(I \rightarrowtail \mathbf{O})),$$

where $E = \{x \in pcl(I \rightarrowtail \mathbf{O} \mid x \not<_\exists J\}$ and $t \in cc_J(I \rightarrowtail \mathbf{O})$ iff $t \in cc(I \rightarrowtail \mathbf{O})$ and $i(t) \not<_\exists J$.

Next we have these facts relevant to conditional propensities:

Fact 35 *Let $A \rightarrowtail \mathbf{a}$ and $B \rightarrowtail \mathbf{b}$ be transitions to scattered outcomes such that $\mathbf{a}$ and $\mathbf{b}$ are consistent. Then:*

$$\textit{if } t \in cc_B(A \rightarrowtail \mathbf{a}), \textit{ then } t \in cc(A \cup B \rightarrow \mathbf{a} \cup \mathbf{b}), \tag{13}$$

where $cc_B(A \rightarrowtail \mathbf{a}) = \{t \in cc(A \rightarrowtail \mathbf{a}) \mid i(t) \not<_\exists B\}$.

Proof: For the implication above to fail, there must be $t = \{e\} \rightarrowtail H$ such that for some $h \in H_{[A]}$: $h \perp_e H_{\langle \mathbf{a} \rangle}$ and $e \not<_\exists B$, but (†) for every $h' \in H_{[A]}$ such that $h' \perp_e H_{\langle \mathbf{a} \rangle}$, $h' \notin H_{[B]}$. Since $\mathbf{a}$ is consistent with $\mathbf{b}$, e and B are consistent as well. Given this consistency, $e \not<_\exists B$ means: $\forall x : (x \in B \rightarrow (e \text{ SLR } x \vee x \leqslant e))$. Consider now two subsets that partition B: $B_1 = \{x \in B \mid x \text{ SLR } e\}$ and $B_2 = \{x \in B \mid x \leqslant e\}$. Clearly, e SLR B_1. By No MFB, there is $h^* \in \Pi_e\langle h \rangle \cap H_{[B_1]}$. Hence $B_1 \subseteq h^*$ and (since $e \in h^*$) $B_2 \subseteq h^*$, so $B \subseteq h^*$. From $e \in h^*$, it also follows that $A \subseteq h^*$. And since $h^* \equiv_e h$, it must be that $h^* \perp H_{\langle \mathbf{a} \rangle}$. Contradiction with (†). □

The immediate consequence of this implication and Fact (24)[15] is that, for $cc_B(A \rightarrowtail \mathbf{a}) = \{t \in cc(A \rightarrowtail \mathbf{a}) \mid i(t) \not<_\exists B\}$ and symmetrically for cc_A, the following is true:

$$cc_B(A \rightarrowtail \mathbf{a}) \cup cc_A(B \rightarrowtail \mathbf{b}) = cc(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b}). \tag{14}$$

Accordingly, if we frame the notion of probabilistic independence in terms of conditional propensities as defined above, and require further that $cc_B(A \rightarrowtail \mathbf{a}) \cap cc_A(B \rightarrowtail \mathbf{b}) = \emptyset$, we get that $A \rightarrowtail \mathbf{a}$ and $B \rightarrowtail \mathbf{b}$ are probabilistically independent if the relevant causal probability measure is uncorrelated—see the fact below. This motivates the following definition of probabilistic funny business.

[15] Additionally one needs to observe that if $e \in pcl(A \cup B \rightarrowtail \mathbf{a}_i \cup \mathbf{b}_j)$, then $\neg(e <_\exists A \cup B)$.

Definition 36 (probabilistic funny business) *Transitions $A \to \mathbf{a}$ and $B \to \mathbf{b}$ (where A and B are initial events and $\mathbf{a}$, $\mathbf{b}$ scattered outcomes) are a case of probabilistic funny business (PFB) iff*

1. *A SLR B;*
2. *$\mathbf{a} \cup \mathbf{b}$ is a scattered outcome;*
3. *$cc(A \to a) \cap cc(B \to b) = \emptyset$ (causal separation);*
4. *$pr(A \to \mathbf{a} \mid B) \cdot pr(B \to \mathbf{b} \mid A) \neq pr(A \cup B \to \mathbf{a} \cup \mathbf{b})$.*

With this definition, we have a desired fact: if the basic transitions are uncorrelated, then (if two transitions are causally separated, then they do not constitute a case of PFB).[16]

Theorem 37 *Let $A \to \mathbf{a}$ and $B \to \mathbf{b}$ be causally separated transitions from initial events to scattered outcomes such that A SLR B and $\mathbf{a} \cup \mathbf{b}$ is a scattered outcome. Then if the measure $\mu_{pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})}$ is uncorrelated, $A \to \mathbf{a}$ and $B \to \mathbf{b}$ are not a case of PFB.*

Proof. Assume this notation: $G = pcl(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})$, $F = \{e \in pcl(A \rightarrowtail \mathbf{a}) \mid e \not<_\exists B\}$, $E = \{e \in pcl(B \rightarrowtail \mathbf{b}) \mid e \not<_\exists A\}$, $cc_A(B \rightarrowtail \mathbf{b}) = \{t \in cc(B \rightarrowtail \mathbf{b}) \mid i(t) \not<_\exists A\}$, and $cc_B(A \rightarrowtail \mathbf{a}) = \{t \in cc(A \rightarrowtail \mathbf{a}) \mid i(t) \not<_\exists B\}$. our theorem follows from this sequence of equations:

$$\mu_G(cc(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})) \stackrel{1}{=} \prod_{t \in cc(A \cup B \rightarrowtail \mathbf{a} \cup \mathbf{b})} \mu_{(i(t))}(t) \stackrel{2}{=}$$
$$\prod_{t \in cc_A} \mu_{(i(t))}(t) \cdot \prod_{t \in cc_B} \mu_{(i(t))}(t) \stackrel{3}{=} \mu_F(cc_A) \cdot \mu_E(cc_B),$$

where $\stackrel{1}{=}$ and $\stackrel{3}{=}$ hold because basic transitions are uncorrelated, and $\stackrel{2}{=}$ because of Fact (35) (cf. Equation (14)) and causal separation. □

Now, in rather special circumstances, which however are most likely satisfied in the set-up of Bell's theorem, causal separation of transitions is true.

Fact 38 *Let $\mathfrak{A} = \{A \rightarrowtail \mathbf{a}_i \mid i \leqslant I\}$ be a set of exclusive transitions of the same st-locations, and $B \rightarrowtail \mathbf{b}$ be a transition to a scattered outcome such that $\mathbf{a}_i$ is consistent with $\mathbf{b}$ for every $i \leqslant I$. Then $A \rightarrowtail \mathbf{a}_i$ is causally separated from $B \rightarrowtail \mathbf{b}$ for every $i \leqslant I$.*

[16] This is an analogue of Müller's (2005) Theorem 1.

Proof. For $A \rightarrowtail \mathbf{a}_i$ and $B \rightarrowtail \mathbf{b}$ *not* to be causally separated, there must be some $e \in pcl(A \rightarrowtail \mathbf{a}_i) \cap pcl(B \rightarrowtail \mathbf{b})$, from which it follows that for some $h_1 \in H_{[A]} : h_1 \perp_e H_{\langle \mathbf{a}_i \rangle}$. By exclusiveness condition, for some outcome event $\mathbf{a}_k$ of some transition from $\mathfrak{A}$: $h_1 \in H_{\langle \mathbf{a}_k \rangle}$. Hence e is consistent with $\mathbf{a}_k$. Hence, by same st-location, since $e <_\exists \mathbf{a}_i$: $e <_\exists \mathbf{a}_k$. We will now argue that, contrary to the premise of the fact, $\mathbf{a}_k$ and $\mathbf{b}$ are inconsistent. Pick a "witness" history $h^* \in H_{\langle \mathbf{a}_i \rangle} \cap H_{\langle \mathbf{b} \rangle}$. Observe that for $h \in H_{\langle \mathbf{a}_k \rangle}$, since $e <_\exists \mathbf{a}_k$, it must be that $h \equiv_e h_1$, and hence $h \perp_e h^*$, whereas for $h \in H_{\langle \mathbf{b} \rangle}$, (since $e <_\exists \mathbf{b}$) it must be that $h \equiv_e h^*$. Contradiction. $\square$

6.1 Parameter Independence

A crucial premise in Bell's theorem is that an outcome of a measurement performed in one wing is to be independent of what is measured in the other wing, this being the content of Parameter Independence (PI). The present framework affords it a precise reading: the (conditional) propensity of a transition $A \rightarrowtail \mathbf{a}$ is the same, no matter whether conditioned on event B or on B', where each B and B' are SLR to A. But is Parameter Independence true in the present framework?

The answer is that, given certain assumptions, naturally thought to be satisfied in a Bell's setup, Parameter Independence is true. To prove a theorem to this effect, let us first observe this fact:

Theorem 39 *Let A, B and B' be initial events and B have the same st-location as B'. Let also $A \rightarrowtail \mathbf{a}$ be a transition to scattered outcome $\mathbf{a}$ and $\mathbf{a}$ SLR B as well as $\mathbf{a}$ SLR B'. Then $cc_B(A \rightarrowtail \mathbf{a}) = cc_{B'}(A \rightarrowtail \mathbf{a})$.*

Proof: We need to show that for every $e \in pcl(A \rightarrowtail \mathbf{a})$: $e \not<_\exists B$ iff $e \not<_\exists B'$. But since $\mathbf{a}$ is SLR to B as well as to B', by No MFB there are histories $h_1 \in H_{\langle \mathbf{a} \rangle} \cap H_{[B]}$ and $h_2 \in H_{\langle \mathbf{a} \rangle} \cap H_{[B']}$. Since $H_{(e)} \subseteq H_{\langle \mathbf{a} \rangle}$, the two histories testify that e is consistent with B and e is consistent with B', respectively. The equivalence above follows then from same st-location. $\square$

From the theorem it follows that Parameter Independence is true. That is, if an outcome $\mathbf{a}$ of measurement A in one wing of the experiment is SLR to each alternative measurement selectable in the other wing, then this outcome is independent, propensity-wise, from measurements in the other wing. More precisely, we have: $pr(A \rightarrowtail \mathbf{a} \mid B) = pr(A \rightarrowtail \mathbf{a} \mid B')$.

As it stands, the proof of the theorem depends on the same st-location of B and B', which means of course that the two events are inconsistent. This suggests a possibility of proving the same conclusion from a different set of premises, in which same st-location of B and B' is replaced by the assumption that they are inconsistent, and the selection between them is affected by some "nicely" located selection event(s). For scarcity of space, however, we do not prove this theorem here, leaving the task to the reader.

7 Bell-Aspect Experiment

The experiment that Clauser, Shimony, Horne and Holt (1969) envisaged was carried out by Aspect and his team in 1980–1981. It seems to involve two levels of chanciness: The production of results, a_1, a_2 on the left and b_1 and b_2 on the right, seem to be chancy. And there is (apparently) freedom in the selection of settings on the apparatus to the left as well as on the apparatus to the right. We will denote these selection events by S_L and S_R, respectively. As a result of the selection on the left, there is either the measurement event Li or the measurement event Li', read as the measurement in the left wing with the setting put to i (i').[17] Similarly on the right, after selection event S_R there is either measurement Rj or measurement Rj'. To keep track of the wings and settings, we write Lia for result a of the measurement on the left, with setting i, and similarly for $Li'a, Rjb$ and $Rj'b$. As for spatiotemporal features, a selection event in one wing is space-like related to results occuring in the other wing, and, consequently, a measurement event performed in one wing is space-like related to results occurring in the other wing.

Let us now focus on a statistical aspect of the experiment described above. We assume the idealization that the experiment has infinitely many runs, and the results obtained on the runs permit one to estimate probabilities of "joint results," like $p^{ex}(Lia \wedge Rjb)$, and probabilities of "single" results, like $p^{ex}(Lia)$. (In total, there are 16 probabilities of the former kind, and 8 of the latter kind.) But how are the runs of this experiment differentiated? It is natural to say that the experiment is 'timed' by emissions for pairs of particles from the source. Yet, we will not lose on generality, and avoid a commitment to ontology of particles, if we postulate that runs of the experiment are differentiated by the occurrences of pairs of selection events.

To recall, run of the experiment is a modally thick notion, as it involves measurement events with all possible settings, and all possible results. As for types of runs, they are differentiated by propensity assignments to the transitions involved. This has a consequence that although some two occurrences of, for instance, S_L look perfectly similar, they may still belong to different types of runs, as transitions to some corresponding results occurring in two

[17] To add some physics, the measurement on the left of the electron's spin projection on the direction i (or i').

runs obtain different propensities. In a similar vein, two runs are subsumed to different types because some transition that has not occurred in either of these runs (this being said from some later point of view), receives different propensities on these runs.

We list now some properties, idealizations, and desiderata on a run of the Bell-CH experiment.

REQUIREMENTS:

1. Li is inconsistent with Li' and Rj is inconsistent with Rj';[18]
2. selection events S_L and S_R are consistent and $S_L <_\forall Li \cup Li'$, and $S_R <_\forall Rj \cup Rj'$;
3. there are exactly two setting at each wing: $\forall h : S_L \subset h \Rightarrow Li \subset h \vee Li' \subset h$ and $\forall h : S_R \subset h \Rightarrow Rj \subset h \vee Rj' \subset h$;
4. same st-location of alternative measurements: $s(Li) = s(Li')$, $s(Rj) = s(Rj')$, etc.
5. same st-location of alternative outcomes of each measurement: $s(Lia) = s(Lia')$, $s(Li'a) = s(Li'a')$, $s(Rjb) = s(Rjb')$, and $s(Rj'b) = s(Rj'b')$;
6. SLR relations: Lia SLR S_R, $Li'a$ SLR S_R, Rjb SLR S_L, and $Rj'b$ SLR S_L,
7. exclusiveness: if $h \in H_{[Li]}$, then there is a result a such that $h \in H_{\langle Lia \rangle}$, and analogously about Li', Rj and Rj'.
8. every measurement has three possible results: $+$, $-$, and f, where f is a "failed outcome;" that is: $a, b \in \{+, -, f\}$.
9. every outcome of every measurement on the left is consistent with every outcome of every measurement on the right.

We take it that selection events as well as measurement events are represented by BST initial events, whereas outcomes are interpreted as scattered outcome events. Then, as a result of the above conditions, in each run n of the experiment there are four sets of exclusive transitions of the same st-location (mind it that a run is a modally thick notion, and runs are differentiated by occurrences of pairs of selection events):

[18] Since $Lia \cup Rjb$ and the like are scattered outcomes, and $Li \cup R_j$ is below it, the latter is consistent.

$$\{Li^n \cup Rj^n \rightarrowtail Lia^n \cup Rjb^n \mid a, b \in \{+, -, f\}\} \tag{15}$$
$$\{Li'^n \cup Rj^n \rightarrowtail Li'a^n \cup Rjb^n \mid a, b \in \{+, -, f\}\} \tag{16}$$
$$\{Li^n \cup Rj'^n \rightarrowtail Lia^n \cup Rj'b^n \mid a, b \in \{+, -, f\}\} \tag{17}$$
$$\{Li'^n \cup Rj'^n \rightarrowtail Li'a^n \cup Rj'b^n \mid a, b \in \{+, -, f\}\}. \tag{18}$$

That is, above the selection $S_L^n \cup S_R^n$, there are the above four sets of transitions in the model. Yet, every history passing through this selection passes through exactly one of the initials $Li^n \cup Rj^n, Li'^n \cup Rj^n, Li^n \cup Rj'^n$ and $Li'^n \cup Rj'^n$. Hence, if a history σ in an MBS is to fully represent this experiment, we need to ensure that each of the four types of initials above occurs in σ sufficiently many times. This follows from conditions (3) and (4) of Definition (6). Note, however, that the first clauses of the above definition need to be modified, to accommodate *alternative* measurements, the gist of the Bell-CH set-up. The general definition of an MBS model fully representing a statistical experiment adequate for the Bell-CH setup is provided in the Appendix.

In the literature, the Bell-CH inequality (which is violated by quantum mechanics and most likely, by Nature as well) is derived from the three "independence" premises: Outcome Independence, Parameter Independence, and No Conspiracy. The first says that outcomes in two wings are probabilistically independent, conditional on the measurement settings and a value of the hidden variable. The second is the claim that outcome in one wing is probabilistically independent of the measurement setting selected in the other, conditional on a value of the hidden variable. The third amounts to assuming that the measurement settings and values of hidden variables are probabilistically independent.

As one might expect, a similar bunch of conditions will figure in our BST derivation of the Bell-CH inequality; there will, however, be some significant alterations. To draw a link between the two approaches, a value of a hidden variable is here understood as a type of runs of the experiment. Outcome Independence (recall, it is conditional on a hidden variable) amounts to No Probabilistic Funny Business: observe that the latter is supposed to hold in each type of runs. Parameter Independence presently takes on the following form:

$$pr(Li^{\Theta(\lambda)} \rightarrow Lia^{\Theta(\lambda)} \mid Rj^{\Theta(\lambda)}) = pr(Li^{\Theta(\lambda)} \rightarrow Lia^{\Theta(\lambda)} \mid Rj'^{\,\Theta(\lambda)}), \tag{19}$$

where $\Theta(\lambda) : \Lambda \rightarrow \mathcal{N}$ is a "selection function." Observe that the set-up considered satisfies the premises of Parameter Independence: points (6) and

(9) of the description of the set-up imply that every measurement in one wing is SLR to every outcome in the other wing. And, by clause (4), measurements in one wing have the same st-location. Hence Parameter Independence is provably true—cf. Theorem (39). This is the first surprise.

As for the condition of No Conspiracy, it is essential for our proof that in each run n of the experiment, each setting — Li^n, Li'^n, Rj^n and Rj'^n — is possible. It follows that in every run *of every type* each setting is possible. Note that this independence of selected settings and types of runs is modal and not probabilistic. This is the second surprise.[19]

The proof of Bell-CH inequalities takes advantage of the arithmetical fact which states that for any real numbers u, u', v, and v' from the $[0,1]$ interval,

$$-1 \leqslant uv + uv' + u'v' - u'v - u - v' \leqslant 0. \tag{20}$$

Make then these substitutions:

$$u = pr(Li^{\Theta(\lambda)} \to Li+^{\Theta(\lambda)} \mid Rj^{\Theta(\lambda)}) = pr(Li^{\Theta(\lambda)} \to Li+^{\Theta(\lambda)} \mid Rj'^{\,\Theta(\lambda)}) \tag{21}$$

$$v = pr(Rj^{\Theta(\lambda)} \to Rj+^{\Theta(\lambda)} \mid Li^{\Theta(\lambda)}) = pr(Rj^{\Theta(\lambda)} \to Rj+^{\Theta(\lambda)} \mid Li'^{\,\Theta(\lambda)}) \tag{22}$$

$$u' = pr(Li'^{\Theta(\lambda)} \to Li'+^{\Theta(\lambda)} \mid Rj^{\Theta(\lambda)}) = pr(Li'^{\Theta(\lambda)} \to Li'+^{\Theta(\lambda)} \mid Rj'^{\,\Theta(\lambda)}) \tag{23}$$

$$v' = pr(Rj'^{\,\Theta(\lambda)} \to Rj'+^{\Theta(\lambda)} \mid Li'^{\,\Theta(\lambda)}) = pr(Rj'^{\Theta(\lambda)} \to Rj'+^{\Theta(\lambda)} \mid Li^{\Theta(\lambda)}), \tag{24}$$

where the right-hand side equations rest on Parameter Independence. With the substitution made, multiply then the sides of Inequality (20) by w^λ, employ No PFB and sum over $\lambda \in \Lambda$. Finally, use Empirical Adequacy condition, that is, clause (4) of Definition (6). The result is the constraint on observable probabilities— the Bell-CH inequality:

$$\begin{aligned} & -1 \leqslant p^{ex}(Li + \wedge Rj+) + p^{ex}(Li + \wedge Rj'+) + p^{ex}(Li' + \wedge Rj'+) - \\ & p^{ex}(Li' + \wedge Rj+) - p^{ex}(Li+) - p^{ex}(Rj'+) \leqslant 0. \end{aligned} \tag{25}$$

The failure of this inequality means that there is no MBS fully representing the (idealized) statistical Bell-CH experiment, subject to the condition that outcomes are independent (in the sense that the pairs of corresponding transitions are not cases of PFB). More precisely, given the above condition, one cannot satisfy "probabilistic" conditions (1) and (3) of Definition (6).

[19] In recent papers of Hofer-Szabó 2008, Grasshoff, Portmann and Wüthrich (2005), and Portmann and Wüthrich (2007) Bell-type inequalities are derived from variously weakened versions of (probabilistic) No Conspiracy. One might thus have a premonition that probabilistic No Conspiracy is not needed for the proof.

8 Discussion

The main result of this paper is this: given (1) the assumptions needed to construct causal probability spaces and (2) the idealizations made in the construction of propensity-frequentist models for Bell-CH experiment, a frequentist-propenisty model for Bell-CH that satisfies (i) (non-probabilistic) No Conspiracy and (ii) No Probabilistic Funny Business (PFB), is not possible. No Conspiracy requires that in each run n of the experiment, it is possible to select each Li and Li' on the left and each Rj and Rj' on the right. No PFB is a BST rendition of Outcome Independence. Notably, Parameter Independence, which is another premise of Bell's theorem, is unnecessary, as it is derivable from the assumptions (1) and the idealizations (2), mentioned above. Thus, given that our world satisfies the assumptions (1) and the idealizations are "reasonably close to the reality," there is conspiracy in our world, or probabilistic funny business, or both. By Theorem 37, probabilistic funny business implies that the corresponding μ measure is correlated, which means that some *basic* transitions, with SLR initials, are correlated.

Since the main result is conditional on the assumptions (1) and the idealizations (2), let us examine them, starting with the former. First, there are finitistic assumptions in Müller's (2005) construction of causal probability spaces: the underlying set of choice events must be finite, and each element thereof — finitely splitting. As Müller himself indicates, there are ways to remove these limitations. Also, an important Fact 21 has an infinitistic version, i.e., with no finitistic assumptions in the premises. Another important assumption is No MFB. This means that although a joint occurrence of some SLR results is not observed in an experiment, a BST model of the phenomenon nevertheless has histories with the joint occurrence of the two results, but it has probability zero.

A similar move occurs in the idealizations (2). It is assumed that if some type-level measurement result $\tilde{a}$ of measurement event $\tilde{A}$ is observed in a statistical experiment, then in a model, if in the n-th run a corresponding (token-level) measurement event A^n occurs, then there is a history in which both A^n and $\mathbf{a}^n$ occur. This sounds like a proliferation of possibilities: if a result is possible, then it is possible in every run in which the corresponding measurement event occurs. The proliferation is mitigated by probabilities, however, as the results might have probability zero assigned. Other idealizations concern spatiotemporal locations of events, yet as we indicated above, they can be weakened, if one wants, but in my opinion they are still quite reasonable as they stand.

Thus, as the assumptions are rather weak, and the idealizations either realistic or cable of being appropriately weakened, we are facing a choice:

there is conspiracy in our world, or some *basic* transitions, with SLR initials, are correlated.

Long live experimental metaphysics!

9 Appendix: General Propensity Models

By comparing a simplified EPR experiment of Section (3) with the Bell-CH experiment, we have seen that runs of experiments might be counted differently. Also, we have seen that, since the measurements considered might be incompatible, it should not be required that all measurement events occur in a run. This suggests a simplified and abstract approach to be sketched here.

To begin with, we assume that an experimenter somehow knows how to count runs of the experiment; suffice it to say that she counts the runs by natural numbers. Suppose next that our experimenter believes that there is a finite set $\mathfrak{X}$ of measurements (type-level) that can be carried out (instantiated), separately or jointly, at each run of the experiment. Let her further estimate that every $\tilde{X} \in \mathfrak{X}$ has a finite set $\{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k\}$ of possible outcomes. Now, to capture this structure of possible measurements and their possible outcomes, we postulate that for every $n \in \mathcal{N}$ there is a set T^n consisting of transitions from possible measurements (token-level) to their possible outcomes (also token-level) that are available at the nth run. In the context of Bell-CH, T^n should contain usual "measurement" transitions, like $(Li \cup Rj)^n \rightarrowtail (Lia \cup Rjb)^n$ as well as less standard "selection" transitions, like $S_L^n \rightarrowtail Li^n$. First we define what it means when we say that a history in an MBS weakly (i.e., modally and spatiotemporally) represents a statistical experiment:

Definition 40 (MBS weakly represents) *Consider a statistical experiment in which there is a set $\mathfrak{X}$ of measurements such that each is possible at every run and for every $\tilde{X} \in \mathfrak{X}$ there is a finite set $\mathfrak{P}_X := \{\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_k\}$ of possible outcomes. An MBS $\mathfrak{M}$ weakly represents this statistical experiment in terms of mapping Y, history σ and sequence $T^1, T^2, T^3, \ldots$ of sets of transitions iff*

1. *Y is an injective function from the set of triples $\langle n, \tilde{X}, \tilde{x}_i \rangle$, where $n \in \mathcal{N}$, $\tilde{X} \in \mathfrak{X}$ and $\tilde{x}_i \in \mathfrak{P}_X$, to the set of transitions from initials to scattered outcomes in $\mathfrak{M}$;*
2. *$\mathfrak{I}_X^n := \{Y(n, \tilde{X}, \tilde{x}_i) \mid \tilde{x}_i \in \mathfrak{P}_X\}$ is the set of alternative exclusive outcomes of the same st-location;*

3. *the sequence* $T^1, T^2, \ldots$ *is shifted in time, where* $T^n = \bigcup_{\tilde{X} \in \mathfrak{X}} \mathfrak{I}^n_X$*;*

4. *for every* $n \in \mathcal{N}$*, there is at least one* $t \in T^n$*, whose outcome occurs in* σ*.*

We say $t \in T^m$ *and* $t' \in T^n$ *are corresponding transitions iff there is* $\tilde{X} \in \mathfrak{X}$ *and* $\tilde{x}_i \in \mathfrak{P}_X$ *such that* $t = Y(\langle m, \tilde{X}, \tilde{x}_i\rangle)$ *and* $t' = Y(\langle n, \tilde{X}, \tilde{x}_i\rangle)$*.*

Then we define types of runs.

Definition 41 *The* m*-th run and the* n*-th run belong to the same type of runs iff for every* $t \in T^m$ *and* $t' \in T^n$*, if* t *and* t' *are corresponding transitions, then* $pr(t) = pr(t')$*.*

Definition 42 (MBS model fully represents) *Let* $\mathfrak{M}$ *be an MBS that weakly represents, in terms of mapping* Y*, history* σ *and sequence* $T^1, T^2, T^3, \ldots$*, a statistical experiment characterized by sets* $\mathfrak{X}$ *and* $\{\mathfrak{P}_X\}_{\tilde{X} \in \mathfrak{X}}$*. We say that* $\mathfrak{M}$ *represents fully the experiment in terms of* Y*,* σ *and propensity assignment* pr *iff there is a partition* Λ *of the set* $\mathcal{N}$ *of natural numbers such that*

1. *for every* $l, m \in \mathcal{N}$*:* $l, m \in \lambda$ *for some* $\lambda \in \Lambda$ *iff for every* $t \in T^l$ *and* $t' \in T^m$*, if* t *and* t' *are corresponding transitions, then* $pr(t) = pr(t')$*;*

2. $p^{ex}(\tilde{\mathbf{x}}) = \sum_\lambda w_\lambda pr(Y(\langle \Theta(\lambda), \tilde{X}, \tilde{x}\rangle))$*,*
   where $\Theta : \Lambda \to \mathcal{N}$ *is a "selection function" such that* $\Theta(\lambda) \in \lambda$*, and*
   $$w_\lambda = lim_{n\to\infty} \frac{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda\}}{n},$$

3. *for every* $\lambda \in \Lambda$*:*
   $$pr(Y(\langle \Theta(\lambda), \tilde{X}, \tilde{x}\rangle)) = \lim_{n\to\infty} \frac{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda \wedge \sigma \in H_{\langle \mathbf{x}\rangle}\}}{\#\{m \in \mathcal{N} \mid m \leqslant n \wedge m \in \lambda\}}.$$

Let us write $X^n \rightarrowtail \mathbf{x}^n$ for $Y(\langle n, \tilde{X}, \tilde{x}\rangle)$. The transitions to consider in constructing a model of the Bell-CH experiment are the transitions from the sets of Equations (15)–(18), as well as "selection" transitions: $(S_L \cup S_R)^n \rightarrowtail (Li \cup Rj)^n$, $(S_L \cup S_R)^n \rightarrowtail (Li' \cup Rj)^n$, $(S_L \cup S_R)^n \rightarrowtail (Li \cup Rj')^n$, and $(S_L \cup S_R)^n \rightarrowtail (Li' \cup Rj')^n$. Because selection event $(S_L \cup S_R)^n$ is below any other event in any transition from T^n, by clause (4) of Definition (40) we get that $(S_L \cup S_R)^n$ occurs in history σ. As the experimenter strives to achieve a non-zero experimental probability of any combination of settings, in history σ there must be an infinite number of runs with each combination of settings.

References

Aspect, A., Dalibard, J., and Roger, G. (1982). Experimental test of Bell's inequalities using time-varying analyzers. *Physical Review Letters*, 49:1804–1807.

Bell, J. S. (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, Cambridge.

Belnap, N. (1992). Branching space-time. *Synthese*, 92:385–434. 'Postprint' archived at PhilSci Archive, http://philsci-archive.pitt.edu/archive/00001003.

Belnap, N. (2002). EPR-like "funny business" in the theory of branching space-times. In Placek, T. and Butterfield, J., editors, *Nonlocality and Modality*, NATO Science Series, pages 293–315, Dordrecht. Kluwer Academic Publisher.

Belnap, N. (2005). A theory of causation: causae causantes (originating causes) as inus conditions in branching space-times. *British Journal for the Philosophy of Science*, 56:221–253.

Belnap, N. and Szabó, L. (1996). Branching space-time analysis of the GHZ theorem. *Foundations of Physics*, 26(8):982–1002.

Clauser, J., Holt, R., Shimony, A., and Horne, M. (1969). Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23:880–884.

Clauser, J. and Horne, M. (1974). Experimental consequences of objective local theories. *Physical Review D*, 10:526–535.

Grasshoff, G., Portmann, S., and Wüthrich, A. (2005). Minimal assumption derivation of a bell-type inequality. *British Journal for the Philosophy of Science*, 56:663–680. preprint: quant-ph/0312176.

Greenberger, D., Horne, M., Shimony, A., and Zeilinger, A. (1990). Bell's theorem without inequalities. *American Journal of Physics*, 58(12):69–72.

Hofer-Szabó, G. (2008). Separate- versus common -common-cause-type derivations of the bell inequalities. *Synthese*, 163(2):199–215.

Hofer-Szabó, G., Rédei, M., and Szabó, L. (1999). On Reichenbach's common cause principle and Reichenbach's notion of common cause. *British Journal for the Philosophy of Science*, 50:377–399.

Kowalski, T. and Placek, T. (1999). Outcomes in branching space-time and GHZ-Bell theorems. *British Journal for the Philosophy of Science*, 50:349–375.

Müller, T. (2002). Branching space-time, modal logic and the counterfactual conditional. In Placek, T. and Butterfield, J., editors, *Nonlocality and Modality*, NATO Science Series, pages 273–291, Dordrecht. Kluwer Academic Publisher.

Müller, T. (2005). Probability theory and causation: a Branching Space-Times analysis. *British Journal for the Philosophy of Science*, 56(3):487–520.

Müller, T., Belnap, N., and Kishida, K. (2008). Funny business in branching space-times: Infinite modal correlations. *Synthese*, 164(1):141–159.

Placek, T. (2000a). *Is Nature Deterministic? A Branching Perspective on EPR Phenomena*. Jagiellonian University Press, Cracow.

Placek, T. (2000b). Stochastic outcomes in branching space-time. An analysis of the Bell theorems. *British Journal for the Philosophy of Science*, 51(3):445–475.

Placek, T. and Wroński, L. (2009). On infinite epr-like correlations. *Synthese*, 167(1):1–32.

Portmann, S. and Wüthrich, A. (2007). Minimal assumption derivation of a weak Clauser-Horne inequality. *Studies in History and Philosophy of Modern Physics*, 38(4):844–862.

Weiner, M. and Belnap, N. (2006). How causal probabilities might fit into our objectively indeterministic world. *Synthese*, 149(1):1–36.

Wroński, L. and Placek, T. (2009). On Minkowskian branching structures. *Studies in History and Philosophy of Modern Physics*, 40:251–258.