

What Does a Hidden Variable Look Like?*

Adam Brandenburger[†]

H. Jerome Keisler[‡]

Version 03/03/11
Preliminary Version

Abstract

The hidden-variable question is whether or not correlations that are observed in the outcomes of an experiment can be explained via introduction of additional (“hidden”) variables which are unobserved by the experimenter. The question arises most famously in quantum mechanics (QM), but can also be asked in the classical realm. The nature of the experiment will tell us how to model the observable variables—i.e., the possible measurements and outcomes. But, by definition, we cannot know what structure to put on unobservable variables. Nevertheless, we show that, under one condition, the hidden-variable question can always be put into a canonical form. The condition is that the spaces of possible measurements and the spaces of possible outcomes, viewed as measurable spaces, are separable (i.e., the σ -algebras are countably generated). An argument based on Maharam’s Theorem ([10, 1942]) then shows that the hidden-variable space can always be taken to be the unit interval equipped with Lebesgue measure. As an application of our result, we give a hidden-variable characterization of the no-signaling property of QM.

1 Introduction

Hidden variables are extra variables added to the model of an experiment to explain correlations in the outcomes. Here is a simple example. Ann’s and Bob’s computers have been prepared with the same password. We know that the password is either p2s4w6r8 or 1a3s5o7d, but we do not know which it is. If Ann now types in p2s4w6r8 and this unlocks her computer, we immediately know what will happen when Bob types in one or other of the two passwords. The two outcomes—when Ann types a password and Bob types a password—are perfectly correlated. Clearly, it would be wrong to conclude that, when Ann types a password on her machine, this somehow causes Bob’s machine to acquire the same password. The correlation is purely informational: It is our state of knowledge that changes, not Bob’s computer. Formally, we can consider an r.v. (random variable) X for Ann’s password, an r.v. Y for Bob’s password, and an extra r.v. Z . The r.v. Z takes the value z_1 or z_2

*We are grateful to Samson Abramsky, Bob Coecke, Amanda Friedenberg, Barbara Rifkind, Gus Stuart, and Noson Yanofsky for valuable conversations, to John Asker, Axelle Ferriere, Andrei Savochnik, participants at the workshop on Semantics of Information, Dagstuhl, June 2010, and participants at the conference on Advances in Quantum Theory, Linnaeus University, Växjö, June 2010, for useful input, and to the Stern School of Business for financial support. wdhv-03-03-11

[†]Address: Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012, adam.brandenburger@stern.nyu.edu, www.stern.nyu.edu/~abranden

[‡]Address: Department of Mathematics, University of Wisconsin-Madison, Madison, WI 53706, keisler@math.wisc.edu, www.math.wisc.edu/~keisler

according as the two machines were prepared with the first or the second password. Then, even though X and Y will be perfectly correlated, they will also be independent (trivially so), conditional on the value of Z . In this sense, the extra r.v. Z explains the correlation.

Of course, even in the classical realm, there are much more complicated examples of hidden-variable analysis. But, the most famous context for hidden-variable analysis is quantum mechanics (QM). Starting with von Neumann [14, 1932], and including, most famously, Einstein, Podolsky, and Rosen [6, 1935], Bell [2, 1964], and Kochen and Specker [9, 1967], a vast literature has grown up around the question of whether a hidden-variable formulation of QM is possible. (The watershed no-go theorems of Bell and Kochen-Specker give conditions under which the answer is negative. The correlations that arise in QM—for example, in spin measurements—cannot be explained as reflecting the presence of hidden variables.)

Let us specify a little more what we mean by an experiment. We imagine that Ann can make one of several measurements on her part of a certain system and Bob can make one of several measurements on his part of the system. Each pair of measurements (one by Ann and one by Bob) leads to a pair of outcomes (one for Ann and one for Bob). We can build an **empirical model** of the experiment by choosing appropriate spaces for the sets of possible measurements and outcomes, and by specifying, for each pair of measurements, a probability measure over pairs of outcomes. An associated **hidden-variable** (henceforth **h.v.**) **model** is obtained by starting with the empirical model and then appending to it an extra r.v.. But, what structure should we put on the space on which this extra r.v., which constitutes our hidden variable(s), lives? After all, a hidden variable is a variable above and beyond those which are part of the actual experiment, and is therefore unobserved. There is no natural structure to impose.

Despite this apparent obstacle, we show that there is a canonical h.v. space. Fix an empirical model. Suppose there is an associated h.v. model which yields, for each pair of measurements, the same probability measure over pairs of outcomes. (We will say that the h.v. model **realizes** the empirical model.) *Then, under one condition, there is always an h.v. model, in which the h.v. space is the unit interval equipped with Lebesgue measure, which realizes the same empirical model. The unit interval with Lebesgue measure is, therefore, a canonical h.v. space.*

The condition for our theorem is that the spaces of possible measurements and the spaces of possible outcomes, viewed as measurable spaces, are separable (i.e., the σ -algebras are countably generated). Maharam's Theorem ([10, 1942]) on the classification of measure algebras is the key mathematical result that underlies the proof of our theorem.

We actually prove more than the statement above. The reason is that unrestricted h.v. models are not very interesting. Given any empirical model, it is trivial to build an h.v. model that realizes it, if no restrictions are placed on the h.v. model.¹ Hidden-variable analysis becomes interesting once we ask that the h.v. model satisfy various properties. Among such properties are: locality, parameter independence, outcome independence, λ -independence, and strong and weak determinism. (We define these properties presently.) *We show that the h.v. model we build preserves each such property satisfied by the original h.v. model.*

The final part of this paper offers an application of our construction of a canonical h.v. space. An important property of an empirical model is no signaling (Ghirardi, Rimini, and Weber [7, 1980]). No signaling says that, given a measurement by Ann, the marginal probabilities of her outcomes do not depend on what measurement Bob makes (and vice versa). QM satisfies no signaling, which ensures its compatibility with relativity. There are also hypothetical physical theories which are superquantum but still no-signaling (Popescu and Rohrlich [11, 1994]). We use our construction to

¹The idea of the construction is well known. We simply take the hidden-variable space to be a copy of the product of Ann's and Bob's outcome spaces, and build a probability measure on the diagonal of the product of these two products.

give an h.v. characterization of the no-signaling property: *An empirical model satisfies no signaling if and only if there is an h.v. model which realizes it and which satisfies parameter independence and λ -independence.* We interpret this result later.

2 Empirical and Hidden-Variable Models

Ann has a space of measurements, which is a measurable space (Y_a, \mathcal{Y}_a) , and a space of possible outcomes, which is a measurable space (X_a, \mathcal{X}_a) . Likewise, Bob has a space of measurements, which is a measurable space (Y_b, \mathcal{Y}_b) , and a space of possible outcomes, which is a measurable space (X_b, \mathcal{X}_b) . (Throughout, we will restrict attention to a system with just two parts. We comment later on the extension to more than two parts.) There is also an h.v. space, which is an unspecified measurable space (Λ, \mathcal{L}) . Write

$$(X, \mathcal{X}) = (X_a, \mathcal{X}_a) \otimes (X_b, \mathcal{X}_b)$$

$$(Y, \mathcal{Y}) = (Y_a, \mathcal{Y}_a) \otimes (Y_b, \mathcal{Y}_b)$$

$$\Psi = (X, \mathcal{X}) \otimes (Y, \mathcal{Y}).$$

$$\Omega = (X, \mathcal{X}) \otimes (Y, \mathcal{Y}) \otimes (\Lambda, \mathcal{L}).$$

Definition 2.1 *An **empirical model** is a probability measure e on Ψ .*

We see that an empirical model describes an experiment in which the pair of measurements $y = (y_a, y_b) \in Y$ is randomly chosen according to the probability measure $\text{marg}_Y e$, and y and the joint outcome $x = (x_a, x_b) \in X$ are distributed according to e .

Definition 2.2 *A **hidden-variable (h.v.) model** is a probability measure p on Ω .*

Definition 2.3 *We say that an h.v. model p **realizes** an empirical model e if $e = \text{marg}_\Psi p$. We say that two h.v. models are (**realization-**)**equivalent** if they realize the same empirical model.*

We see that an h.v. model is a model which has an extra component, viz., the h.v. space, and which reproduces a given empirical model when we average over the values of the h.v.. The interest in h.v. models is that we can ask them to satisfy properties that it would be unreasonable to demand of an empirical model. (In the example we began with, the property is conditional independence—which we would only expect once Z is introduced.) We review various properties of h.v. models in Section 4, after covering some preliminaries.

3 Preliminaries

Throughout the paper, we use the following two conventions. First, when p is a probability measure on a product space $(X, \mathcal{X}) \otimes (Y, \mathcal{Y})$ and $q = \text{marg}_X p$, then for each $J \in \mathcal{X}$ we write

$$p(J) = p(J \times Y) = q(J),$$

and for each q -integrable $f : X \rightarrow \mathbb{R}$ we write

$$\int_J f(x) dp = \int_{J \times Y} f(x) dp = \int_J f(x) dq.$$

Thus, in particular, a statement holds for p -almost all $x \in X$ if and only if it holds for q -almost all $x \in X$.

Second, when p is a probability measure on a product space $(X, \mathcal{X}) \otimes (Y, \mathcal{Y}) \otimes (Z, \mathcal{Z})$ and $J \in \mathcal{X}$, we let $p[J|\mathcal{Z}]$ be the function from Z into $[0, 1]$ such that

$$p[J|\mathcal{Z}]_z = p[J \times Y \times Z | \{X \times Y, \emptyset\} \otimes \mathcal{Z}]_{((x,y),z)} = \mathbb{E}[1_{J \times Y \times Z} | \{X \times Y, \emptyset\} \otimes \mathcal{Z}].$$

We use similar notation for (finite) products with factors to the left of (X, \mathcal{X}) or to the right of (Z, \mathcal{Z}) . Note that if $q = \text{marg}_{X \times Z} p$, then $q[J|\mathcal{Z}] = p[J|\mathcal{Z}]$. We also use the analogous notation for expected values of random variables: Given an integrable function $f : X \rightarrow \mathbb{R}$, we write $\mathbb{E}[f|\mathcal{Z}]$ for the conditional expectation $\mathbb{E}[f \circ \pi | \{X \times Y, \emptyset\} \otimes \mathcal{Z}]$ where π is the projection from $X \times Y \times Z$ to X .

Lemma 3.1 *The mapping $z \mapsto p[J|\mathcal{Z}]_z$ is the p -almost surely unique \mathcal{Z} -measurable function $f : Z \rightarrow [0, 1]$ such that for each set $L \in \mathcal{Z}$,*

$$\int_L f(z) dp = p(J \times L).$$

Proof. Let $f(z) = p[J|\mathcal{Z}]_z$. Using the definition of $p[J|\mathcal{Z}]$, we see that

$$\begin{aligned} \int_L f(z) dp &= \int_{X \times Y \times L} \mathbb{E}[1_{J \times Y \times Z} | \{X \times Y, \emptyset\} \otimes \mathcal{Z}] dp = \\ &= \int_{X \times Y \times L} 1_{J \times Y \times Z} dp = p((X \times Y \times L) \cap (J \times Y \times Z)) = p(J \times L), \end{aligned}$$

as required. ■

Corollary 3.2 *Let q be the marginal of p on $X \times Z$. Then, for each $J \in \mathcal{X}$, we have $p[J|\mathcal{Z}] = q[J|\mathcal{Z}]$ q -almost surely.*

Lemma 3.3 *If $p[J|\mathcal{Z}] \in \{0, 1\}$ p -almost surely, then $p[J|\mathcal{Y} \otimes \mathcal{Z}] = p[J|\mathcal{Z}]$ p -almost surely.*

Proof. Let $L_0 = \{z \in Z : p[J|\mathcal{Z}]_z = 0\}$ and $L_1 = \{z \in Z : p[J|\mathcal{Z}]_z = 1\}$. Then $L_0, L_1 \in \mathcal{Z}$ and $p(L_0 \cup L_1) = 1$. By Lemma 3.1,

$$\begin{aligned} \int_{L_0} p[J|\mathcal{Z}]_z dp &= 0 = p(J \times L_0), \\ \int_{L_1} p[J|\mathcal{Z}]_z dp &= p(L_1) = p(J \times L_1). \end{aligned}$$

By Lemma 3.1 again,

$$\int_{Y \times L_0} p[J|\mathcal{Y} \otimes \mathcal{Z}]_{(y,z)} dp = p(J \times Y \times L_0) = p(J \times L_0) = 0,$$

so

$$p[J|\mathcal{Y} \otimes \mathcal{Z}]_{(y,z)} = 0 = p[J|\mathcal{Z}]_z \quad \forall (y, z) \in Y \times L_0.$$

Similarly,

$$\int_{Y \times L_1} p[J|\mathcal{Y} \otimes \mathcal{Z}]_{(y,z)} dp = p(J \times Y \times L_1) = p(J \times L_1) = p(L_1),$$

so

$$p[J|\mathcal{Y} \otimes \mathcal{Z}]_{(y,z)} = 1 = p[J|\mathcal{Z}]_z \quad \forall (y, z) \in Y \times L_1,$$

as required. ■

When $x \in X$, we write $p[x|\mathcal{Z}]_z = p[\{x\}|\mathcal{Z}]_z$. By the properties of probability measures, for X finite we have $\sum_{x \in X} p[x|\mathcal{Z}]_z = 1$ p -almost surely.

Given probability measures p on $(X, \mathcal{X}) \otimes (Y, \mathcal{Y})$ and r on (Y, \mathcal{Y}) , we say that p is an **extension** of r if $r = \text{marg}_Y p$. We say that two probability measures p and q on $(X, \mathcal{X}) \otimes (Y, \mathcal{Y})$ **agree on Y** if $\text{marg}_Y p = \text{marg}_Y q$.

In what follows, whenever we write an equation involving conditional probabilities, it will be understood to mean that the equation holds p -almost surely. By the term “measure” we will always mean “probability measure.”

4 Properties of Hidden-Variable Models

The definitions of properties of h.v. models in this section are due to Bell [2, 1964] (locality), Jarrett [8, 1984] (parameter independence and outcome independence), and Brandenburger and Yanofsky [4, 2008] (the distinction between strong and weak determinism). Some terminology is from Shimony [13, 1986] (parameter independence and outcome independence) and Dickson [5, 2005] (λ -independence).

The results in this paper hold when the spaces X_a, X_b, Y_a, Y_b are infinite, and the σ -algebras $\mathcal{X}_a, \mathcal{X}_b, \mathcal{Y}_a, \mathcal{Y}_b$ are countably generated. However, everything is conceptually simpler when X_a, X_b are finite, because in this case we can work with individual elements rather than with subsets. Furthermore, our results when X_a, X_b are finite can be used to prove similar results for the infinite case. This preliminary version of the paper treats the finite case; the next version will cover the infinite case.² Of course, there are cases where X_a, X_b are, in fact, finite: spin measurements in QM are one example.

Assumption: *From now on, except where noted otherwise, the outcome spaces X_a and X_b will be finite, and \mathcal{X}_a and \mathcal{X}_b will be the respective power sets.*

All expressions below which are given for Ann have counterparts for Bob, with a and b interchanged.

Definition 4.1 *The h.v. model p is λ -independent if for every event $L \in \mathcal{L}$,*

$$p[L|\mathcal{Y}]_y = p(L).$$

Note that the λ -independence property for p depends only on $r = \text{marg}_{Y \times \Lambda} p$. By well-known properties of product measures, we have:

Lemma 4.2 *The following are equivalent:*

- (i) *the measure p is λ -independent;*
- (ii) *the measure r is the product $r = \text{marg}_Y p \otimes \text{marg}_\Lambda p$;*

²For the proofs for the infinite case, see our online draft “Observable Implications of Unobservable Variables,” at www.stern.nyu.edu/~abranden/papers.

(iii) the σ -algebras \mathcal{Y} and \mathcal{L} are independent with respect to p , i.e.,

$$p(K \times L) = p(K) \times p(L)$$

for every $K \in \mathcal{Y}, L \in \mathcal{L}$.

Definition 4.3 The h.v. model p has **parameter independence** if for every $x_a \in X_a$ we have

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}].$$

Definition 4.4 The h.v. model p has **outcome independence** if for every $x = (x_a, x_b) \in X$ we have

$$p[x|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y} \otimes \mathcal{L}] \times p[x_b|\mathcal{Y} \otimes \mathcal{L}].$$

Definition 4.5 The h.v. model p is **local**, or has **locality**, if for every $x \in X$ we have

$$p[x|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] \times p[x_b|\mathcal{Y}_b \otimes \mathcal{L}].$$

The next proposition follows Jarrett [8, 1984, p.582].

Proposition 4.6 The h.v. model p is local if and only if p has parameter independence and outcome independence.

Proof. It is easily seen from the definitions that if p has parameter independence and outcome independence, then p is local.

Suppose that p is local. We have

$$\{x_a\} \times X_b = \bigcup_{x_b \in X_b} \{(x_a, x_b)\},$$

so

$$\begin{aligned} p[x_a|\mathcal{Y} \otimes \mathcal{L}] &= p[\{x_a\} \times X_b|\mathcal{Y} \otimes \mathcal{L}] = \sum_{x_b \in X_b} p[x_a, x_b|\mathcal{Y} \otimes \mathcal{L}] = \\ &= \sum_{x_b \in X_b} (p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] \times p[x_b|\mathcal{Y}_b \otimes \mathcal{L}]) = \\ p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] &\times \sum_{x_b \in X_b} p[x_b|\mathcal{Y}_b \otimes \mathcal{L}] = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] \times 1 = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}]. \end{aligned}$$

Similarly,

$$p[x_b|\mathcal{Y} \otimes \mathcal{L}] = p[x_b|\mathcal{Y}_b \otimes \mathcal{L}].$$

It follows that p has parameter independence.

Again, supposing that p is local, we have

$$p[x_a, x_b|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] \times p[x_b|\mathcal{Y}_b \otimes \mathcal{L}],$$

and hence

$$p[x_a, x_b|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y} \otimes \mathcal{L}] \times p[x_b|\mathcal{Y} \otimes \mathcal{L}],$$

so p has outcome independence. ■

Definition 4.7 The h.v. model p has **strong determinism** if for each $x_a \in X_a$ we have

$$p[x_a|\mathcal{Y}_a \otimes \mathcal{L}]_{(y_a, \lambda)} \in \{0, 1\}.$$

This says that the set $Y_a \times \Lambda$ can be partitioned into sets $\{A_{x_a} : x_a \in X_a\}$ such that $p[x_a|A_{x_a}] = 1$ for each $x_a \in X_a$.

Definition 4.8 *The h.v. model p has **weak determinism** if for each $x \in X$ we have*

$$p[x|\mathcal{Y} \otimes \mathcal{L}]_{(y,\lambda)} \in \{0, 1\}.$$

This says that the set $Y \times \Lambda$ can be partitioned into sets $\{A_x : x \in X\}$ such that $p[x|A_x] = 1$ for each $x \in X$.

Lemma 4.9 *The following are equivalent:*

(i) *the measure p has weak determinism;*

(ii) *for each $x_a \in X_a$ we have*

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}]_{(y,\lambda)} \in \{0, 1\}.$$

Proof. It is clear that (ii) implies (i).

Assume (i). Then for p -almost all (y, λ) there is an $x \in X$ such that $p[x|\mathcal{Y} \otimes \mathcal{L}]_{(y,\lambda)} = 1$, and hence

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}]_{(y,\lambda)} = 1$$

for each $x_a \in X_a$. Therefore (ii) holds. ■

Proposition 4.10 *If p has strong determinism then p has weak determinism.*

Proof. Suppose p has strong determinism. By Lemma 3.3, we have

$$p[x_a|\mathcal{Y}_c \otimes \mathcal{L}] = p[x_a|\mathcal{Y} \otimes \mathcal{L}]$$

p -almost surely, and therefore

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}] \in \{0, 1\},$$

so p has weak determinism by Lemma 4.9(ii). ■

Proposition 4.11 *If p has weak determinism, then p has outcome independence.*

Proof. Suppose p has weak determinism. By Lemma 4.9, we have

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}] \in \{0, 1\}.$$

Therefore

$$p[x|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y} \otimes \mathcal{L}] \times p[x_b|\mathcal{Y} \otimes \mathcal{L}],$$

as required. ■

Proposition 4.12 *The h.v. model p has strong determinism if and only if p has weak determinism and parameter independence.*

Proof. Suppose p has strong determinism. By Lemma 3.3,

$$p[x_a|\mathcal{Y}_a \otimes \mathcal{L}] = p[x_a|\mathcal{Y} \otimes \mathcal{L}],$$

so p has parameter independence. By Proposition 4.10, p has weak determinism.

For the converse, suppose p has weak determinism and parameter independence. Fix $x_a \in X_a$. By weak determinism and Lemma 4.9,

$$p[x_a|\mathcal{Y} \otimes \mathcal{L}]_{(y,\lambda)} \in \{0, 1\}.$$

By parameter independence,

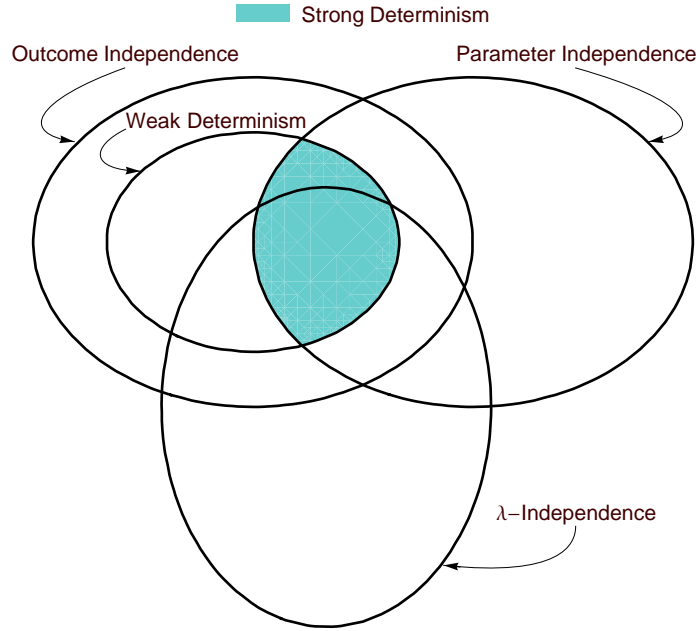
$$p[x_a|\mathcal{Y} \otimes \mathcal{L}] = p[x_a|\mathcal{Y}_a \otimes \mathcal{L}].$$

Therefore

$$p[x_a|\mathcal{Y}_a \otimes \mathcal{L}]_{(y,\lambda)} \in \{0, 1\},$$

so p has strong determinism. ■

We can summarize the properties we have considered and the relationships among them in the following Venn diagram.³



The definitions and results in this section extend immediately to systems with more than two parts, except that parameter independence must now be stated in terms of sets of parts instead of individual parts.

We need two more definitions for the next section. By the **Lebesgue unit interval** we mean the probability space $\mathbf{U} = ([0, 1], \mathcal{U}, u)$ where \mathcal{U} is the set of Borel subsets of $[0, 1]$ and u is Lebesgue measure on \mathcal{U} .

Definition 4.13 *The h.v. model p is **real-valued** if $(\Lambda, \mathcal{L}) = ([0, 1], \mathcal{U})$ and $\text{marg}_\Lambda p = u$.*

³By virtue of our Proposition 4.12, this diagram improves upon the Venn diagram in Brandenburger and Yanofsky [4, 2008], in showing that four of the regions in the latter diagram are empty.

5 A Canonical Hidden-Variable Space

We now state and prove our main result, which says that, given an h.v. model, there is an equivalent real-valued h.v. model which preserves properties.

Theorem 5.1 *Assume that the σ -algebras \mathcal{Y}_a and \mathcal{Y}_b are countably generated. Then every h.v. model p is equivalent to a real-valued h.v. model \bar{p} such that for each of the properties of λ -independence, parameter independence, outcome independence, strong determinism, and weak determinism, if p has the property then so does \bar{p} .*

Proof. We may assume without loss of generality that p has an atomless h.v. model, because the product $p \otimes u$ of p with the Lebesgue unit interval is equivalent to p and has the atomless h.v. model $(\Lambda, \mathcal{L}, \ell) \otimes \mathbf{U}$, and if p has any of the five properties above, then so does $p \otimes u$.

Let \mathcal{Y}_a^0 be a countable subset of \mathcal{Y}_a that generates the σ -algebra \mathcal{Y}_a , and similarly for \mathcal{Y}_b^0 . Let \mathcal{U}_0 be the family of all open subintervals of $[0, 1]$ with rational endpoints. Let

$$(x_{an}, x_{bn}, B_{an}, B_{bn}, C_n)$$

be an enumeration of the countable set $X_a \times X_b \times \mathcal{Y}_a^0 \times \mathcal{Y}_b^0 \times \mathcal{U}_0$. The Cartesian products $\{x_{an}\} \times \{x_{bn}\} \times B_{an} \times B_{bn} \times C_n$ generate the σ -algebra $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{U}$. For each n , let

$$D_n = \{\lambda \in \Lambda : p[\{(x_{an}, x_{bn})\} \times B_{an} \times B_{bn} | \mathcal{L}]_\lambda \in C_n\}.$$

To continue the proof, we need the following lemma. In what follows, we will write $\ell = \text{marg}_\Lambda p$.

Lemma 5.2 *There is a countably generated σ -algebra $\mathcal{D} \subseteq \mathcal{L}$ such that each of the sets D_n belongs to \mathcal{D} , and the restriction of ℓ to \mathcal{D} is atomless.*

Proof of lemma. Since ℓ is atomless, it follows from a result of Sierpinski [12, 1922] that for each set $L \in \mathcal{L}$ there is a set $L' \in \mathcal{L}$ such that $L' \subseteq L$ and $\ell(L') = \ell(L)/2$. Then, by the Axiom of Choice, there is a function $F : \mathcal{L} \rightarrow \mathcal{L}$ such that for each $L \in \mathcal{L}$, $F(L) \subseteq L$ and $\ell(F(L)) = \ell(L)/2$. Let \mathcal{E}_0 be the algebra of subsets of Λ generated by $\{D_n : n \in \mathbb{N}\}$. For each $m \in \mathbb{N}$, let \mathcal{E}_{m+1} be the algebra of subsets of Λ generated by $\mathcal{E}_m \cup \{F(L) : L \in \mathcal{E}_m\}$. Let $\mathcal{E} = \bigcup_m \mathcal{E}_m$, and let \mathcal{D} be the σ -algebra generated by \mathcal{E} . Clearly, each D_n belongs to \mathcal{D} , and \mathcal{E} is countable, so \mathcal{D} is countably generated.

We show that the restriction of ℓ to \mathcal{D} is atomless. Let \mathcal{D}' be the set of all $D \in \mathcal{D}$ that can be approximated by sets in \mathcal{E} with respect to ℓ , that is,

$$\mathcal{D}' = \{D \in \mathcal{D} : (\forall r > 0)(\exists E \in \mathcal{E})\ell(E \Delta D) < r\}.$$

It is clear that $\mathcal{E} \subseteq \mathcal{D}'$, and that \mathcal{D}' is closed under finite unions and intersections. The set \mathcal{D}' is also closed under unions of countable chains, because if $L_n \in \mathcal{D}'$ and $L_n \subseteq L_{n+1}$ for each n , and $L = \bigcup_n L_n$, then for each $r > 0$ there exists $n \in \mathbb{N}$ and $E \in \mathcal{E}$ such that $\ell(L \Delta L_n) < r/2$ and $\ell(E \Delta L_n) < r/2$. Therefore $\ell(L \Delta E) < r$, so $L \in \mathcal{D}'$. It follows that $\mathcal{D}' = \mathcal{D}$. Now suppose $D \in \mathcal{D}$, $\ell(D) > 0$, and $r > 0$. Then $D \in \mathcal{D}'$, so there is a set $G \in \mathcal{E}$ such that $\ell(D \Delta G) < r$. We have $F(G) \in \mathcal{E}$, $\ell(F(G)) = \ell(G)/2$, and $F(G) \subseteq G$. Then $D \cap F(G) \in \mathcal{D}$, and, by taking r small enough, we can guarantee that $\ell(D) > \ell(D \cap F(G)) > 0$. This shows that the restriction of ℓ to \mathcal{D} is atomless, and proves the lemma. ■

Proof of Theorem 5.1 contd. By Maharam's Theorem ([10, 1942]), the measure algebras of $(\Lambda, \mathcal{D}, \ell)$ and \mathbf{U} are isomorphic. This isomorphism maps the equivalence class (modulo null sets) of each $D \in \mathcal{D}$ to the equivalence class of a set $h(D) \in \mathcal{U}$ such that $u(h(D)) = \ell(D)$.

Let \bar{p} be the probability measure on $\mathcal{X} \otimes \mathcal{Y} \otimes \mathcal{U}$ such that for each $Y \in \mathcal{X}$, $S \in \mathcal{Y}$, and $D \in \mathcal{D}$ we have $\bar{p}(Y \times S \times h(D)) = p(Y \times S \times D)$. It is clear that \bar{p} is an extension of $s \otimes \mu$. Moreover, \bar{p} has the same marginal as p on $X \times Y$, so \bar{p} is equivalent to p . The h.v. model \bar{p} is the Lebesgue unit interval, so \bar{p} is real-valued. For each $S \in \mathcal{Y}$ and $D \in \mathcal{D}$ we have

$$\bar{p}(S) = p(S), \bar{p}(h(D)) = p(D), \bar{p}(S \times h(D)) = p(S \times D).$$

Therefore, if p has λ -independence, then \bar{p} has λ -independence by Lemma 4.2.

The σ -algebra \mathcal{D} is large enough so that for each $K \in \mathcal{Y}$, the function $p[\{(x_a, x_b)\} \times K | \mathcal{L}]$ is \mathcal{D} -measurable. It follows that

$$p[x_a, x_b | \mathcal{Y} \otimes \mathcal{L}] = p[x_a, x_b | \mathcal{Y} \otimes \mathcal{D}],$$

$$p[x_a | \mathcal{Y}_a \otimes \mathcal{L}] = p[x_a | \mathcal{Y}_a \otimes \mathcal{D}],$$

and

$$p[x_b | \mathcal{Y}_b \otimes \mathcal{L}] = p[x_b | \mathcal{Y}_b \otimes \mathcal{D}].$$

From the definition of \bar{p} , one can see that joint distributions of the functions

$$p[x_a, x_b | \mathcal{Y} \otimes \mathcal{D}], p[x_a | \mathcal{Y}_a \otimes \mathcal{D}], p[x_b | \mathcal{Y}_b \otimes \mathcal{D}]$$

and

$$\bar{p}[x_a, x_b | \mathcal{Y} \otimes \mathcal{U}], \bar{p}[x_a | \mathcal{Y}_a \otimes \mathcal{U}], \bar{p}[x_b | \mathcal{Y}_b \otimes \mathcal{U}]$$

are the same. It follows that for each of the properties of parameter independence, outcome independence, strong determinism, and weak determinism, if p has the property then so does \bar{p} . ■

6 An Application to the No-Signaling Property

In this section we apply Theorem 5.1 to give an h.v.-characterization of the no-signaling property (Ghirardi, Rimini, and Weber [7, 1980]) of empirical models. (As before, all expressions below which are given for Ann have counterparts for Bob, with a and b interchanged.)

Definition 6.1 *The empirical model e is **no-signaling** if for every $x_a \in X_a$ we have*

$$e[x_a | \mathcal{Y}] = e[x_a | \mathcal{Y}_a].$$

This says that, the probability of a particular outcome for Ann, conditional on her measurement, is unaffected by also conditioning on Bob's measurement. QM satisfies no signaling ([7, 1980]). (For a proof that this is true for arbitrary sets of commuting observables, not just for multipartite situations, see Abramsky and Brandenburger [1, 2011, Section 4.2].) There are also no-signaling superquantum theories Popescu and Rohrlich ([11, 1994]). Satisfaction of no signaling is usually said to be necessary if a physical theory is to be compatible with relativity.

Here is our h.v.-characterization of no signaling.

Theorem 6.2 *Assume that the σ -algebras \mathcal{Y}_a and \mathcal{Y}_b are countably generated. Then, the empirical model e is no-signaling if and only if there is an h.v. model p which realizes e and which has parameter independence and λ -independence.*

We will prove the theorem and then discuss it. For the proof, we will need the notion of a fiber product of measures from Ben Yaacov and Keisler [3, 2009]. Let (X, \mathcal{X}) , (Y, \mathcal{Y}) , (Z, \mathcal{Z}) be measurable spaces. (In the next definition and lemma, the space X need not be finite.)

Definition 6.3 Let q, r be probability measures on $(X, \mathcal{X}) \otimes (Z, \mathcal{Z})$ and $(Y, \mathcal{Y}) \otimes (Z, \mathcal{Z})$ respectively. Assume that q and r have the same marginal s on (Z, \mathcal{Z}) . The **fiber product** $p = q \otimes_Z r$ is the probability measure p on $(X, \mathcal{X}) \otimes (Y, \mathcal{Y}) \otimes (Z, \mathcal{Z})$ such that

$$p(J \times K \times L) = \int_L q[J|\mathcal{Z}]_z \times r[K|\mathcal{Z}]_z ds$$

for all $J \in \mathcal{X}$, $K \in \mathcal{Y}$, and $L \in \mathcal{Z}$.

It follows from the Caratheodory Extension Theorem that when q, r are as above, the fiber product $p = q \otimes_Z r$ exists and is unique, and is a common extension of q and r . Next is a characterization of the fiber product in terms of conditional probabilities and extensions.

Lemma 6.4 Let q and r be as in Definition 6.3, and let p be a common extension of q, r on $(X, \mathcal{X}) \otimes (Y, \mathcal{Y}) \otimes (Z, \mathcal{Z})$. Then the following are equivalent.

(i) $p = q \otimes_Z r$.

(ii) $p[J \times K|\mathcal{Z}]_z = q[J|\mathcal{Z}]_z \times r[K|\mathcal{Z}]_z$ p -almost surely, for all $J \in \mathcal{X}$ and $K \in \mathcal{Y}$.

(iii) $p[J \times K|\mathcal{Z}]_z = p[J|\mathcal{Z}]_z \times p[K|\mathcal{Z}]_z$ p -almost surely, for all $J \in \mathcal{X}$ and $K \in \mathcal{Y}$.

(iv) $p[J|\mathcal{Y} \otimes \mathcal{Z}]_{(y,z)} = p[J|\mathcal{Z}]_z$ p -almost surely, for all $J \in \mathcal{X}$.

Proof. It is clear that (i), (ii), and (iii) are equivalent. Consider any $J \in \mathcal{X}, K \in \mathcal{Y}$, and $L \in \mathcal{Z}$. Assume (i). We have

$$\int_{K \times L} p[J|\mathcal{Z}] dp = \int_{Y \times L} p[J|\mathcal{Z}] \times 1_K dp.$$

By the rules of conditional expectations,

$$E[p[J|\mathcal{Z}] \times 1_K|\mathcal{Z}] = p[J|\mathcal{Z}] \times E[1_K|\mathcal{Z}] = p[J|\mathcal{Z}] \times p[K|\mathcal{Z}].$$

Therefore

$$\int_{Y \times L} p[J|\mathcal{Z}] \times 1_K dp = \int_L p[J|\mathcal{Z}] \times p[K|\mathcal{Z}] dp = \int_L q[J|\mathcal{Z}] \times r[K|\mathcal{Z}] dp.$$

By (i), this is equal to $p(J \times K \times L)$, so

$$\int_{K \times L} p[J|\mathcal{Z}] dp = p(J \times K \times L).$$

This shows that (i) implies (iv).

Now assume (iv). Then

$$p(J \times K \times L) = \int_{K \times L} p[J|\mathcal{Y} \otimes \mathcal{Z}] dp = \int_{K \times L} p[J|\mathcal{Z}] dp = \int_{Y \times L} p[J|\mathcal{Z}] \times 1_K dp.$$

As in the preceding paragraph,

$$\int_{Y \times L} p[J|\mathcal{Z}] \times 1_K dp = \int_L q[J|\mathcal{Z}] \times r[K|\mathcal{Z}] dp,$$

and condition (i) is proved. ■

Let e_a, s be the marginals of e on $X_a \times Y_a$ and Y respectively.

Proposition 6.5 *An empirical model e is no-signaling if and only if e is an extension of $e_a \otimes_{Y_a} s$.*

Proof. By Lemma 6.4. ■

Proof of Theorem 6.2. Suppose e is no-signaling. Then, the h.v. model where Λ is a singleton $\{\lambda\}$ realizes e and has parameter independence and λ -independence.

Now suppose that there is an h.v. model p which realizes e and which has parameter independence and λ -independence. By Proposition 6.5, we must show that e is an extension of the fiber product $e_a \otimes_{Y_a} s$.

By Theorem 5.1, e is realized by an h.v. model p which is real-valued and has parameter independence and λ -independence. Let $\Lambda = [0, 1]$, let \mathcal{L} be the σ -algebra of Borel subsets of $[0, 1]$, and let $\mathcal{L}^1, \mathcal{L}^2, \dots$ be an increasing chain of finite algebras of sets whose union generates \mathcal{L} . Let q_a, p_a, r be the marginals of p on $X_a \times Y_a \times \Lambda$, $X_a \times Y \times \Lambda$, and $Y \times \Lambda$ respectively. By parameter independence, p_a is the fiber product $p_a = q_a \otimes_{Y_a \times \Lambda} r$.

For each n , let q_a^n and r^n be the restrictions of q_a and r to $\mathcal{X}_a \otimes \mathcal{Y}_a \otimes \mathcal{L}^n$ and $\mathcal{Y} \otimes \mathcal{L}^n$ respectively. In general, p will not be an extension of the fiber product $q_a^n \otimes_{Y_a \times \Lambda} r^n$. Our plan is to show that $q_a^n \otimes_{Y_a \times \Lambda} r^n$ is an extension of $e_a \otimes_{Y_a} s$, and converges to p_a as $n \rightarrow \infty$.

We first prove convergence. Fix an integer $k > 0$, and element $x_a \in X_a$, and sets $U \in \mathcal{Y}_a \otimes \mathcal{L}^k$ and $K_b \in \mathcal{Y}_b$. Then $q_a^n[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n]$ is a uniformly bounded martingale with respect to the sequence of σ -algebras $\mathcal{Y}_a \otimes \mathcal{L}^n$, for $n \geq k$. By the Martingale Convergence Theorem, $q_a^n[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n]$ converges to $q_a[x_a || \mathcal{Y}_a \otimes \mathcal{L}]$ p -almost almost surely. Similarly, for each $K_b \in \mathcal{Y}_b$, $r^n[K_b || \mathcal{Y}_a \otimes \mathcal{L}^n]$ converges to $r[K_b || \mathcal{Y}_a \otimes \mathcal{L}]$ p -almost surely. We have

$$(q_a^n \otimes_{Y_a \times \Lambda} r^n)(\{x_a\} \times U \times K_b) = \int_U q_a^n[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n] \times r^n[K_b || \mathcal{Y}_a \otimes \mathcal{L}^n] dp$$

and

$$p_a(\{x_a\} \times U \times K_b) = \int_U q_a[x_a || \mathcal{Y}_a \otimes \mathcal{L}] \times r[K_b || \mathcal{Y}_a \otimes \mathcal{L}] dp.$$

Moreover, as $n \rightarrow \infty$,

$$q_a^n[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n] \times r^n[K_b || \mathcal{Y}_a \otimes \mathcal{L}^n] \rightarrow q_a[x_a || \mathcal{Y}_a \otimes \mathcal{L}] \times r[K_b || \mathcal{Y}_a \otimes \mathcal{L}]$$

p -almost surely. By Fatou's Lemma,

$$\int_U q_a^n[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n] \times r^n[K_b || \mathcal{Y}_a \otimes \mathcal{L}^n] dp \rightarrow \int_U q_a[x_a || \mathcal{Y}_a \otimes \mathcal{L}] \times r[K_b || \mathcal{Y}_a \otimes \mathcal{L}] dp.$$

Therefore

$$(q_a^n \otimes_{Y_a \times \Lambda} r^n)(\{x_a\} \times U \times K_b) \rightarrow p_a(\{x_a\} \times U \times K_b).$$

It follows that for each $x^a \in X_a$, $K_a \in \mathcal{Y}_a$, and $K_b \in \mathcal{Y}_b$,

$$(q_a^n \otimes_{Y_a \times \Lambda} r^n)(\{x_a\} \times K_a \times K_b) \rightarrow p_a(\{x_a\} \times K_a \times K_b) = e(\{x_a\} \times K_a \times K_b).$$

We next prove that for each n , $q_a^n \otimes_{Y_a \times \Lambda} r^n$ is an extension of $e_a \otimes_{Y_a} s$. Let \mathcal{X}^n be the set of all atoms of \mathcal{L}^n of positive Lebesgue measure. Then \mathcal{X}^n is a finite collection of pairwise disjoint subsets of Λ whose union has Lebesgue measure 1. Let $u = q_a^n \otimes_{Y_a \times \Lambda} r^n$. By Lemma 6.4,

$$u[x_a || \mathcal{Y} \otimes \mathcal{L}^n] = u[x_a || \mathcal{Y}_a \otimes \mathcal{L}^n].$$

The conditional probability $u[x_a|\mathcal{Y} \otimes \mathcal{L}^n]_{(y,\lambda)}$ depends only on y and the atom $A \in \mathcal{X}^n$ that contains λ , so we may write

$$u[x_a|\mathcal{Y} \otimes \mathcal{L}^n]_{(y,\lambda)} = u[x_a|\mathcal{Y} \otimes \mathcal{L}^n]_{(y,A)}$$

whenever $\lambda \in A \in \mathcal{X}^n$. We have

$$u[x_a|\mathcal{Y}]_y = \sum_{A \in \mathcal{X}^n} u[x_a|\mathcal{K} \otimes \mathcal{L}^n]_{(y,A)} \times p[A|\mathcal{Y}]_y.$$

A similar computation holds with \mathcal{Y}_a in place of \mathcal{Y} . Since p has λ -independence,

$$p[A|\mathcal{Y}]_y = p(A) = p[A|\mathcal{Y}_a]_y$$

for each $A \in \mathcal{X}^n$ and $y \in Y$. Therefore

$$u[x_a|\mathcal{Y}] = u[x_a|\mathcal{Y}_a].$$

Since q_a^n is an extension of e_a , and r^n is an extension of s , we see from Lemma 6.4 that $u = q_a^n \otimes_{Y_a \times \Lambda} r^n$ is an extension of $e_a \otimes_{Y_a} s$. Thus

$$(e_a \otimes_{Y_a} s)(\{x_a\} \times K_a \times K_b)$$

is a constant sequence that converges to $e(\{x_a\} \times K_a \times K_b)$, and hence

$$(e_a \otimes_{Y_a} s)(\{x_a\} \times K_a \times K_b) = e(\{x_a\} \times K_a \times K_b)$$

for all $x_a \in X_a$, $K_a \in \mathcal{Y}_a$, and $K_b \in \mathcal{Y}_b$. This shows that e is an extension of $e_a \otimes_{Y_a} s$. A similar argument holds for b in place of a , so e is no-signaling by Proposition 6.5. ■

By definition, empirical models that arise in the classical world can be realized by h.v. models that have λ -independence and locality. Remember that locality is equivalent to the conjunction of parameter independence and outcome independence (Proposition 4.6). Thus, Theorem 6.2 says that dropping just outcome independence takes us from the classical world all the way to the superquantum—i.e., no-signaling—world. It is well known that locating the QM world inside the no-signaling world is a hard problem. Still, Theorem 6.2 suggests that there may be weakenings of outcome independence that will help shed light on this problem.

References

- [1] Abramsky, S., and A. Brandenburger, “A Unified Sheaf-Theoretic Treatment of Non-Locality and Contextuality,” February 2011, available at <http://arxiv.org/abs/1102.0264v2>.
- [2] Bell, J., “On the Einstein-Podolsky-Rosen Paradox,” *Physics*, 1, 1964, 195-200.
- [3] Ben Yaacov, I., and H.J. Keisler, “Randomizations of Models as Metric Structures,” *Confluentes Mathematici*, 1, 2009, 197-223. Available at www.math.wisc.edu/~keisler.
- [4] Brandenburger, A., and N. Yanofsky, “A Classification of Hidden-Variable Properties,” *Journal of Physics A: Mathematical and Theoretical*, 41, 2008, 425302.
- [5] Dickson, W.M., *Quantum Chance and Non-Locality: Probability and Non-Locality in the Interpretations of Quantum Mechanics*, Cambridge University Press, 2005.

- [6] Einstein, A., B. Podolsky, and N. Rosen, "Can Quantum-Mechanical Description of Physical Reality be Considered Complete?" *Physical Review*, 47, 1935, 777-780.
- [7] Ghirardi, G.C., A. Rimini, and T. Weber, "A General Argument Against Superluminal Transmission Through the Quantum Mechanical Measurement Process," *Lettere Al Nuovo Cimento (1971-1985)*, 27, 1980, 293-298.
- [8] Jarrett, J., "On the Physical Significance of the Locality Conditions in the Bell Arguments," *Noûs*, 18, 1984, 569-589.
- [9] Kochen, S., and E. Specker, "The Problem of Hidden Variables in Quantum Mechanics," *Journal of Mathematics and Mechanics*, 17, 1967, 59-87.
- [10] Maharam, D., "On Homogeneous Measure Algebras," *Proc. Nat. Acad. Sci. U.S.A.*, 28, 1942, 108-111.
- [11] Popescu, S., and D. Rohrlich, "Quantum Nonlocality as a Resource," *Foundations of Physics*, 24, 1994, 379-385.
- [12] Sierpinski, W., "Sur les Fonctions d'Ensemble Additives et Continues," *Fundamenta Mathematicae*, 3, 1922, 240-246.
- [13] Shimony, A., "Events and Processes in the Quantum World," in Penrose, R., and C. Isham (eds.), *Quantum Concepts in Space and Time*, Oxford University Press, 1986, 182-203.
- [14] Von Neumann, J., *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag, 1932. (Translated as *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.)