No-Forcing and No-Matching Theorems for Classical Probability Applied to Quantum Mechanics

Ehtibar N. Dzhafarov\(^1\) and Janne V. Kujala\(^2\)

\(^1\)Purdue University, ehtibar@purdue.edu
\(^2\)University of Jyväskylä, Department of Mathematical Information Technology, jvk@iki.fi

Abstract

Since Bell’s celebrated work, we know that correlations of spins measured along various axes in a system of entangled particles cannot be explained within the framework of classical probability theory, provided the identity of each spin is determined only by the choice of its axis, irrespective of those chosen for other particles. Here, we study classical probability models of a much more general kind, possessing two properties: (A) contextuality by default, implying that the spins for different combinations of spin axes across all particles are different random variables, and (B) constraining by marginals, meaning that the joint distributions of all spins are constrained by fixing distributions of some subsets thereof. We show that even at this level of generality the classical probability models have their limitations when applied to spin correlations in entangled particles. No-matching theorem says that no such model allows for all correlations quantum mechanics (QM) predicts and for no other: if it allows for all QM-compliant correlations, then it also allows for some correlations forbidden by QM; if it forbids all QM-forbidden correlations, then it also forbids some of the QM-compliant correlations. No-forcing theorem says that, in an important special class of these models (when the constraining subsets of spins are so-called connections), QM-forbidden correlations can be disallowed only together with all correlations violating classical mechanics (represented by Bell-type inequalities).

KEYWORDS: Bell/CH inequalities; cosphericity inequality; connections; couplings; EPR paradigm; Fine’s theorem; joint distribution; probability spaces.

Almost fifty years ago John Bell\(^1\) posed and answered in the negative the question of whether probability distributions of spins in entangled particles could be accounted for by a model written in the language of classical probability. These distributions being among the most basic predictions of QM, Bell’s theorem and its subsequent elaborations\(^2,3\) seem to mathematically isolate quantum determinism from the probabilistic forms of classical determinism, and establish the necessity for a quantum probability theory that is not reducible to the classical one. However, Bell-type theorems do not engage the full potential of classical probability theory, i.e., the theory adhering to Kolmogorov’s conceptual framework\(^4\). The use of probability theory in Bell-type theorems is
constrained by the following assumption:

\[ \text{(Bell)} \quad \text{A spin of a given particle is a random variable whose identity does not depend on measurement settings (axes) chosen for other particles.} \]

In a Kolmogorovian probability model for spin distributions this assumption can be replaced with a weaker one, known as \textit{marginal selectivity} \cite{5}:

\[ \text{(Gen-1)} \quad \text{The distribution (rather than identity) of a spin of a given particle does not depend on measurement settings chosen for other particles.} \]

Denial of statement (Bell) does not amount to admission of a “spooky action at a distance.” Rather it is based on our acknowledging, as a general principle, \textit{contextuality by default}:

\[ \text{(Gen-2)} \quad \text{No two spins recorded under different, mutually exclusive measurement settings (across all particles involved) ever co-occur, because of which they are defined on different probability spaces, i.e., possess no joint distribution (while the equality postulated in statement (Bell) implies a special case of joint distribution, see footnote 1)} \] \cite{6-8}.

The existence of a single probability space for all random variables imaginable (implying that they are all jointly distributed) is often mistakenly taken as one of the tenets of Kolmogorov’s theory. The untenability of this view is apparent by cardinality considerations alone. Also, it is easy to see that for any class of random variables of a given type (e.g., binary +1/-1 ones) one can construct a variable of the same type stochastically independent of each of them. The idea that all such variables can be defined on a single space therefore leads to a contradiction \cite{16}.

Statement (Gen-2) does not mean, of course, that stochastically unrelated (i.e., possessing no joint distribution) spins cannot be imposed a joint distribution on. In fact this can generally be done in a variety of ways, referred to as different \textit{couplings} \cite{9}. Each coupling is a way of designing a scheme of paring realizations of random variables, as if they co-occurred in an imaginary experiment.

\[ \text{(Gen-3)} \quad \text{The stochastically spins recorded under different, mutually exclusive measurement settings can be coupled arbitrarily, insofar as the joint distribution imposed on them (such as their being equal, opposite, independent, etc.) is consistent with the observable joint distributions of the spins recorded under the same measurement settings.} \]

Among the Kolmogorovian models this broadly understood one can easily find ones that allow for all QM-compliant spin distributions in entangled particles. Some of these models (e.g., the one in which all stochastically unrelated random variables are considered

\footnote{Clearly, random variables can be identically distributed without being equal, i.e., without possessing a joint distribution in which they are equal with probability 1.}
independent) are even compatible with all mathematically possible spin distributions. This is not particularly interesting.

An interesting question is whether a classical model can be constructed to “match” QM precisely, in the sense of allowing for all QM-compliant correlations and no other. At the end of this paper we answer this question in the negative for all Kolmogorovian models in which the couplings mentioned in statement (Gen-3) are constrained by fixing distributions of some subsets of all spins involved (no-matching theorem).

Prior to that, however, we consider a case when these constraining distributions are those of certain spin pairs, called connections. This special class of models is the most straightforward generalization of the models compatible with statement (Bell). For the models with connections we prove a stronger result (no-forcing theorem): such a model either allows for correlations forbidden by QM or it only allows for the correlations of classical mechanics, those satisfying the Bell/CH inequalities.

We consider the simplest Bohmian version of the EPR paradigm, depicted in Fig. [1]. For each of the four combined settings \((\alpha_i, \beta_j)\), the recorded spins form a random pair \((A_{ij}, B_{ij})\) whose distribution is

\[
\begin{array}{|c|c|c|}
\hline
\text{joint probability (jnt. prob.)} & B_{ij} = +1 & B_{ij} = -1 \\
\hline
A_{ij} = +1 & p_{ij} & \frac{1}{2} - p_{ij} \\
A_{ij} = -1 & \frac{1}{2} - p_{ij} & p_{ij} \\
\hline
\end{array}
\] (1)

The QM prediction is, for \(i, j \in \{1, 2\}\),

\[ p_{ij} = \frac{1}{4} - \frac{1}{4} \langle \alpha_i | \beta_j \rangle, \] (2)

where \(\langle \alpha_i | \beta_j \rangle\) is the cosine of the angle between axes \(\alpha_i\) and \(\beta_j\). The results of this experiment are uniquely described by the outcome vector

\[ p = (p_{11}, p_{12}, p_{21}, p_{22}). \] (3)

We say that \(p\) is QM-compliant if there exists some choice of the settings \(\alpha_1, \alpha_2, \beta_1, \beta_2\) under which \(p\) satisfies (2). The following inequality is known to be a necessary and sufficient condition for \(p\) being QM-compliant:

\[ |r_{11}r_{12} - r_{21}r_{22}| \leq \sqrt{1 - r_{11}^2} \sqrt{1 - r_{12}^2} + \sqrt{1 - r_{21}^2} \sqrt{1 - r_{22}^2}, \] (4)

where

\[ r_{ij} = 4p_{ij} - 1 \] (5)

is correlation between \(A_{ij}\) and \(B_{ij}\) \((i, j \in \{1, 2\})\). For geometric reasons obvious from Fig. [1] it is referred to as the cosphericity inequality.

In accordance with statement (Gen-2), in the Kolmogorovian probability theory the spin pairs recorded under mutually exclusive settings, say \((A_{11}, B_{11})\) and \((A_{12}, B_{12})\), should generally be treated as stochastically unrelated, possessing no joint distribution. Indeed, \(A_{11}\) and \(A_{12}\) cannot co-occur, because of which their realizations are not “naturally” coupled, and their joint distribution is not defined.
Figure 1: The experimental paradigm considered in this paper. Two spin-half particles created in a singlet state are running away from each other. Each particle has its spin measured along one of two axes (measurement settings): $\alpha_1$ or $\alpha_2$ for “Alice’s” particle (left), and $\beta_1$ or $\beta_2$ for “Bob’s” particle (right). Each measurement results in a random variable attaining one of two values, $+1$ (spin-up, shown by outward-pointing cones) or $-1$ (spin-down, inward-pointing cones). We confine the consideration to the case when these values are equiprobable (then the spins comply with statement (Gen-1) trivially).

However, a coupling scheme can always be imposed on such random variables $[9]$; one can create an infinity of eight-component random vectors

$$H = (A_{11}, B_{11}, A_{12}, B_{12}, A_{21}, B_{21}, A_{22}, B_{22})$$

(6)

(see Fig.2) whose two-component parts $(A_{ij}, B_{ij})$ have the distributions shown in the matrix $(7)$ In this conceptual framework, statement (Bell) is equivalent to the choice of the coupling scheme in which, for $i = 1, 2$ and $j = 1, 2$,

| jnt. prob. | $A_{12} = +1$ | $A_{12} = -1$ |
|------------|----------------|----------------|
| $A_{11} = +1$ | $\frac{1}{2}$ | $0$ |
| $A_{11} = -1$ | $0$ | $\frac{1}{2}$ |

(7)

| jnt. prob. | $B_{2j} = +1$ | $B_{2j} = -1$ |
|------------|----------------|----------------|
| $B_{1j} = +1$ | $\frac{1}{2}$ | $0$ |
| $B_{1j} = -1$ | $0$ | $\frac{1}{2}$ |

If this is assumed, then the spins can be single-indexed, $A_i$ and $B_j$, and the coupling vector $H$ in (6) can be written as

$$H = (A_1, B_1, A_2, B_2).$$

(8)

However, matrices $(7)$ have neither empirical nor theoretical justification, because de facto $A_{11}$ and $A_{22}$ (or $B_{1j}$ and $B_{2j}$) are not

---

2A rigorous formulation $[8, 17, 18]$ requires that $H$ be defined as $(A'_{ij}, B'_{ij} : i, j \in \{1, 2\})$ such that each pair $(A'_{ij}, B'_{ij})$ has the same distribution as (rather than being identical to) $(A_{ij}, B_{ij})$ for $i, j \in \{1, 2\}$. Our lax notation is common and unlikely to cause confusion.
No-Forcing Theorem

jointly distributed. There is therefore no prohibition against coupling them differently, so that generally,

| jnt. prob. | $A_{i2} = +1$ | $A_{i2} = -1$ |
|------------|----------------|----------------|
| $A_{i1} = +1$ | $1/2 - \epsilon_i^1$ | $\epsilon_i^1$ |
| $A_{i1} = -1$ | $\epsilon_i^1$ | $1/2 - \epsilon_i^1$ |

\[ (9) \]

In other words, a general coupling $H$ allows the spins $A_{i1}$ and $A_{i2}$ to be different with probability $2\epsilon_i$, where $0 \leq \epsilon_i \leq 1/2$ (and analogously for $B_{ij}$ and $B_{2j}$). We use the term connection to refer to the pairs $(A_{i1}, A_{i2})$ and $(B_{ij}, B_{2j})$. The four connections are uniquely characterized by the connection vector

$$\epsilon = (\epsilon_i^1, \epsilon_i^2, \epsilon_j^1, \epsilon_j^2),$$

or the corresponding correlations

$$r'_j = 1 - 4\epsilon_j'.$$

An outcome vector $p$ and a connection vector $\epsilon$ are mutually compatible if they can be embedded in one and the same coupling $H$, as shown in Fig. 2. In other words, $\epsilon$ and $p$ are mutually compatible if they can be computed as marginals from the probabilities assigned to the $2^8$ values of $H$: i.e., $p_{ij}$ is the sum of the probabilities for all values of $H$ with $A_{ij} = B_{ij} = 1$, and $\epsilon_i^1, \epsilon_i^2$ are the sums of the probabilities for all values of $H$ with, respectively, $A_{ij} = -A_{i2} = 1$ and $B_{1j} = -B_{2j} = 1$. The set of all compatible pairs $(p, \epsilon)$ forms an 8-dimensional polytope described by Lemma 1. But we need some notation first.

Given a connection vector $\epsilon = (\epsilon_i^1, \epsilon_i^2, \epsilon_j^1, \epsilon_j^2)$, consider the sums

$$1/4 (\pm r_1^1 \pm r_1^2 \pm r_2^1 \pm r_2^2)$$

where each $\pm$ is replaced with either $+$ or $–$. Let $s_0(\epsilon)$ denote the largest of the eight such sums with even numbers of plus signs. Let $s_1(\epsilon)$ denote the largest of the eight sums with odd numbers of plus signs. Since the components of $\epsilon$ belong to $[0, 1/2]$, the pairs $(s_0(\epsilon), s_1(\epsilon))$ fill in the triangular area connecting $(0, 0)$, $(1/2, 1)$, and $(1, 1/2)$. In particular,

$$s_0(\epsilon) + s_1(\epsilon) \leq 3/2,$$

$$0 \leq s_0(\epsilon) \leq 1,$$

$$0 \leq s_1(\epsilon) \leq 1.$$

We define $s_0(p)$ and $s_1(p)$ for any outcome vector $p = (p_{11}, p_{12}, p_{21}, p_{22})$ analogously (using $r_{ij}$ in place of $r'_j$). Since the components of $p$ also belong to $[0, 1/2]$, the pairs $s_0(p), s_1(p)$ have precisely the same properties as $s_0(\epsilon), s_1(\epsilon)$.
Figure 2: A coupling \( H = (A_{11}, B_{11}, A_{12}, B_{12}, A_{21}, B_{21}, A_{22}, B_{22}) \) for pairs \( (A_{ij}, B_{ij}) \) [??]. The number at a double-arrow connecting two random variables is their correlation. Horizontal and vertical arrows correspond to a connection vector \( \varepsilon = (\varepsilon_1^1, \varepsilon_1^2, \varepsilon_2^1, \varepsilon_2^2) \); diagonal arrows correspond to an outcome vector \( p = (p_{11}, p_{12}, p_{21}, p_{22}) \).

**Lemma 1.** \( p \) and \( \varepsilon \) are mutually compatible if and only if

\[
\begin{align*}
\mathsf{s}_0(\varepsilon) + \mathsf{s}_1(p) & \leq \frac{3}{2}, \\
\mathsf{s}_1(\varepsilon) + \mathsf{s}_0(p) & \leq \frac{3}{2}.
\end{align*}
\]

For the proof of this lemma see [8].

The following observation that we need later on is proved by simple algebra.

**Lemma 2.** The set \( E_0 \) of connection vectors \( \varepsilon \) with \( \mathsf{s}_0(\varepsilon) = 1 \) consists of the null vector \( \varepsilon_0 = (0, 0, 0, 0) \) and seven vectors obtained by replacing any two of or all four zeros in \( \varepsilon_0 \) with \( 1/2 \). For all \( \varepsilon \) in \( E_0 \), \( \mathsf{s}_1(\varepsilon) = 1/2 \).

The null (or identity) connection vector \( \varepsilon_0 \) plays a special role, as it corresponds to matrices (7) and premises all Bell-type theorems. It also plays a central role in the no-forcing theorem below. Note that according to Lemmas[1] and [2] an outcome vector \( p \) is compatible with \( \varepsilon_0 \) if and only if it is compatible with all connection vectors in \( E_0 \).

It is easy to see that a connection vector can be chosen so that it is compatible with all QM-compliant outcome vectors. The simplest example is the connection vector \( \varepsilon_{\text{ind}} = (1/4, 1/4, 1/4, 1/4) \) corresponding to the couplings (6) with all components pairwise independent, except, possibly, for pairs \( (A_{ij}, B_{ij}) \). Since \( \mathsf{s}_0(\varepsilon_{\text{ind}}) = \mathsf{s}_1(\varepsilon_{\text{ind}}) = 0 \), it follows from Lemma[1] that \( \varepsilon_{\text{ind}} \) is compatible with any \( p \), whether QM-compliant or not.

What is less obvious is what all the connection vectors are that are compatible only with QM-compliant outcome vectors. In other words, we are interested in the set \( \text{Force}_{\text{QM}} \) of connection vectors, defined as follows:

\( \text{Force}_{\text{QM}} \): all connection vectors \( \varepsilon \) such that if \( p \) is compatible with \( \varepsilon \), then \( p \) is QM-compliant, i.e., satisfies the cosphericity inequality (4).

The name of the set is to indicate that \( \varepsilon \in \text{Force}_{\text{QM}} \) forces every \( p \) compatible with it to be QM-compliant. The set is not empty, because, as the next lemma shows, it includes \( E_0 \) of Lemma[2].
Lemma 3. $E_0 \subset \text{Force}_{QM}$.

Proof. By Fine’s theorem [4], $p$ is compatible with $\varepsilon_0$ (hence also with other members of $E_0$) if and only if it satisfies the Bell/CH inequalities,

$$0 \leq p_{11} + p_{12} + p_{21} + p_{22} - 2p_{ij} \leq 1 \text{ for all } i, j \in \{1, 2\}.$$  \hspace{1cm} (15)

Since the cosphericity inequality is a necessary condition for the compatibility of $p$ with $\varepsilon_0$ [13], QM-compliance follows from the Bell/CH inequalities. \hfill \square

We are thus led to the following questions:

(Q1) What is the entire set $\text{Force}_{QM}$ (what connection vectors it contains beside $E_0$)?

(Q2) What is the set $P_{QM}$ of the outcome vectors $p$ each of which is compatible with at least one of the connection vectors in $\text{Force}_{QM}$?

The questions are significant for the following reason. If $P_{QM}$ turned out to coincide with the set of all QM-compliant $p$, we would have a hope of constructing a variant of classical probability theory that would match QM in the sense of allowing all those $p$ that are possible in QM and forbidding all those $p$ that QM forbids. Quantum determinism could then be “explained” by pointing out that the multiple spaces for different measurement settings can be coupled by using appropriately chosen connection vectors for different settings. However, this hope should be abandoned, because $\text{Force}_{QM}$ in fact coincides with $E_0$, whence $P_{QM}$ includes only those $p$ that satisfy the Bell/CH inequalities. It is well known that the Bell/CH inequalities do not describe all QM-compliant vectors: e.g., they are violated if we use in (2) coplanar vectors at the angles $\alpha_1 = 0, \alpha_2 = \pi/2, \beta_1 = \pi/4, \beta_2 = -\pi/4$.

The proof makes use of the following observation.

Lemma 4. If $p$ belong to the set $P_0$ described by

$$s_0 (p) + s_1 (p) = 3/2,$$

$$s_0 (p) < 1,$$

then $p$ is not QM-compliant (violates the cosphericity inequality).

Proof. One easily checks that

$$s_i (p) = \frac{1}{4} (|r_{11}| + |r_{12}| + |r_{21}| + |r_{22}|)$$

$$s_{1-i} (p) = s_i (p) - \frac{1}{2} \min (|r_{11}|, |r_{12}|, |r_{21}|, |r_{22}|),$$

where $i$ is 0 or 1 according as the number of positive correlations $r_{ij}$ is even or odd. Without loss of generality, let the minimum in the second expression equal $|r_{22}|$. Then

$$s_0 (p) + s_1 (p) = \frac{1}{2} (|r_{11}| + |r_{12}| + |r_{21}|),$$
and this can only equal $\frac{3}{2}$ if each of the three correlations equals $\pm 1$. The cosphericity inequality \[4\] can only be satisfied then if
\[r_{22} = \pm 1 \text{ and } r_{11}r_{12} = r_{21}r_{22}.
\]

It is easy to see that the latter is possible only if the number of $+1$’s among the four $\pm 1$ correlations is even. It follows that $s_0(p) = 1$. \hfill \square

**Theorem 5** (no-forcing). The answer to Q1 is: $\text{Force}_{\text{QM}} = E_0$ (whence the answer to Q2 is: $P_{\text{QM}}$ is the set of all $p$ satisfying Bell/CH inequalities).

**Proof.** We know from \[13\] that $s_0(\varepsilon) \leq 1$, and from Lemmas 2 and 3 that $s_0(\varepsilon) = 1$ describes the set $E_0 \subset \text{Force}_{\text{QM}}$. The theorem is proved by showing that $\text{Force}_{\text{QM}}$ does not contain any $\varepsilon$ with $s_0(\varepsilon) < 1$. From the definition of $\text{Force}_{\text{QM}}$, if there is a $p$ with which $\varepsilon$ is compatible but which does not satisfy the cosphericity inequality \[4\], then $\varepsilon \notin \text{Force}_{\text{QM}}$. By Lemma 1 if for a given $\varepsilon$ one chooses a $p$ such that
\[s_0(\varepsilon) + s_1(p) \leq s_0(p) + s_1(p) = \frac{3}{2},
\]
\[s_1(\varepsilon) + s_0(p) \leq s_0(p) + s_1(p) = \frac{3}{2},
\]
then $p$ and $\varepsilon$ are compatible. If $s_0(\varepsilon) \leq \frac{1}{2}$, then choose $p$ with $s_1(p) = 1$ and $s_0(p) = \frac{1}{2}$ to satisfy this system. If $\frac{1}{2} < s_0(\varepsilon) < 1$, then choose $p$ with $s_1(p) = s_1(\varepsilon)$ and $s_0(p) = \frac{3}{2} - s_1(\varepsilon)$ to satisfy this system. By Lemma 3 all these choices of $p$ belong to $P_0$ and therefore violate the cosphericity inequality. It follows that all $\varepsilon$ with $s_0(\varepsilon) < 1$ do not belong to $\text{Force}_{\text{QM}}$. \hfill \square

One consequence of this theorem is that classical probability theory in which couplings are constrained by connections cannot match QM precisely: if it allows only for QM-compliant $p$ (as does the choice of $\varepsilon = \varepsilon_0$), then it only allows for a proper subset thereof; and if it allows for all QM-compliant $p$, then it also allows for some $p$ that are QM-contravening (as does the connection vector $\varepsilon_{\text{ind}} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ that is compatible with all possible outcome vectors $p$). We now generalize this no-matching statement for connection vectors to arbitrary marginal distributions imposed on couplings $H$ in \[8\].

A subset of any $k \leq 8$ components of $H$ is referred to as its $k$-marginal. So far we only considered 1-marginals, that we posited to have equiprobable $+1/-1$ values, and certain 2-marginals. Of the latter, the distributions of $(A_{ij}, B_{ij})$ are the mandatory constraints imposed on $H$, in fact underlying the definition of $H$. The empirically inaccessible 2-marginals $(A_{11}, A_{12})$ and $(B_{11}, B_{22})$ constrain $H$ by specifying a connection vector $\varepsilon$. There are, however, other empirically inaccessible marginals, such as $(A_{11}, A_{22}, A_{12}, A_{21}, A_{22})$, $(A_{11}, B_{22}, B_{12})$, etc. Each of them, once its distribution is specified, constrains the possible distributions of $H$. The connection vectors do occupy a privileged place among these constraints, because they are the only constraints that can specialize $H$ in \[6\], by choosing $\varepsilon \in E_0$, to the unconstrained reduced coupling \[8\]. This privileged status, however, is immaterial for the next theorem, because it applies to any set of marginal distributions, even if only some of them (e.g., only connection vectors) can be well justified.

**Theorem 6** (no-matching). There is no set of marginal distributions imposed on $H$ such that outcome vectors $p$ are compatible with this set if and only if they are QM-compliant.
Figure 3: Areas of outcome vectors $p = (p_{11}, p_{12}, p_{21}, p_{22})$, two components of which (no matter which) define panels on the left, and the remaining two form the axes of each panel, as shown on the right. The pink area contains all $p$ satisfying the Bell/CH inequalities (15). One can constrain the couplings $H$ by marginal distributions (notably, by putting $\epsilon = (0, 0, 0, 0)$), so that the set of all $p$ compatible with them coincides with the pink area \([15]\). The gray area contains all $p$ satisfying the Tsirelson inequalities (18). One can constrain the couplings $H$ by marginal distributions (e.g., by putting $\epsilon = ((\sqrt{2} - 1)/8, (\sqrt{2} - 1)/8, (\sqrt{2} - 1)/8)$), so that the set of all $p$ compatible with them coincides with the gray area \([18]\). The blue area (that includes the pink area and is included in the gray one) contains all $p$ satisfying the cosphericity inequality (4), i.e., all QM-compliant $p$. The no-matching theorem says that, for any set of marginal distributions, the set of all $p$ compatible with them never coincides with the blue area precisely.

Proof. Observe first that a distribution of any $k$-marginal $(X_1, \ldots, X_k)$ of $H$ can be presented by $2^k$ probabilities

$$
Pr [X_{i_1} = 1, \ldots, X_{i_k} = 1],
$$

with all values equated to 1, for all $k'$-(sub)marginals of $X_1, \ldots, X_k$ ($0 \leq k' \leq k$). This includes the empty subset, for which we put $Pr [\emptyset] = 1$. If we fix distributions of several marginals, with the numbers of components $k_1, \ldots, k_m$, then the total number of different probabilities is $N < 2^{k_1} + \ldots + 2^{k_m}$. This set of probabilities constrains the set of possible outcome vectors $p$ to those for which one can find a $2^8$-component vector $Q$ with the following properties:

$$
M_1 Q = p
$$

subject to

$$
M_2 Q = P, Q \geq 0.
$$

Here, $Q$ is the vector of probabilities assigned to all possible values of $H$ (and $Q \geq 0$ is understood componentwise), $M_1, M_2$ are Boolean (0/1) matrices with dimensions $4 \times 2^8$ and $N \times 2^8$, respectively, and $P$ is the vector of all probabilities of the form \([16]\) that define the distributions of the marginals chosen. The entries of the matrices are defined by the following rule: (1) choose the row of the matrix corresponding to the probability $Pr [X_1 = 1, \ldots, X_k = 1]$; (2) choose the column of $M$ corresponding to values

$$(A_{ij} = a_{ij}, B_{ij} = b_{ij} : i, j \in \{1, 2\})$$
of $H$ ($a_{ij}, b_{ij} \in \{-1, 1\}$); (3) put 1 in the intersection of this row and this column if and only if $a_{ij}$ and $b_{ij}$ equal 1 for all $A_{ij}$ and $B_{ij}$ that belong $(X_1, \ldots, X_k)$; (4) for the 0-marginal (empty set), all entries are 1. In matrix $M_1$ the marginals for its four rows are $(A_{ij}, B_{ij}), i, j \in \{1, 2\}$, and the row corresponding to, e.g., $(A_{11}, B_{11})$ contains 1 in each cell whose column corresponds to $H$-values with

$$
\begin{pmatrix}
A_{11} & B_{11} & A_{12} & \cdots & B_{22} \\
1 & 1 & \text{any} & \cdots & \text{any}
\end{pmatrix}.
$$

To illustrate the structure of matrix $M_2$ and vector $P$, assume that one of the marginals chosen is $(A_{11}, A_{12}, A_{21}, A_{22})$. Then $P$ includes the 16 probabilities

$$
\begin{align*}
\Pr[A_{11} = 1, A_{12} = 1, A_{21} = 1, A_{22} = 1] \\
\Pr[A_{11} = 1, A_{12} = 1, A_{21} = 1, \ldots, \Pr[A_{12} = 1, A_{21} = 1, A_{22} = 1] \\
\Pr[A_{11} = 1, A_{12} = 1, \ldots, \Pr[A_{21} = 1, A_{22} = 1] \\
\Pr[A_{11} = 1, \ldots, \Pr[A_{22} = 1] = 1/2 \\
\Pr[] = 1.
\end{align*}
$$

The row of $M_2$ corresponding to, say, $(A_{11}, A_{12}, A_{21})$, contains 1 for all columns with $H$-values

$$
\begin{pmatrix}
A_{11} & B_{11} & A_{12} & A_{21} & B_{21} & A_{22} & B_{22} \\
1 & \text{any} & 1 & \text{any} & 1 & \text{any} & \text{any}
\end{pmatrix}.
$$

Now, the set of all vectors $p$ for which a $Q$ exists satisfying (17) forms a polytope confined within a $[0, 1/2]^4$ cube. This polytope can be empty (if the distributions for the marginals chosen are not compatible), consist of a single point (e.g., if the marginals chosen include $H$ itself), or have any dimensionality between 1 and 4. The statement of the theorem follows from the fact that the set of QM-compliant $p$, those satisfying (4), is not a polytope. Fig. 3 makes this fact obvious, by showing the curvilinear shape of the two-dimensional cross-sections of the set of QM-compliant $p$. □

Fig. 3 serves an additional purpose of demonstrating that the failure of classical probability theory to match QM is not due to its general inability to match theories outside the scope of classical mechanics. It is the nonlinearity of the area of QM-compliant outcome vectors that is responsible for the no-matching theorem. Thus, the Tsirelson inequalities

$$
\frac{1 - \sqrt{2}}{2} \leq p_{11} + p_{12} + p_{21} + p_{22} - 2p_{ij} \leq \frac{1 + \sqrt{2}}{2} \quad (i, j \in \{1, 2\}).
$$

(18)

are known to be necessary (but not sufficient) for QM-compliance of $p$ [14,15]. It has been shown [8] that by choosing, e.g., all components of $\epsilon$ equal to $(\sqrt{2}-1)/8$, one obtains the class of couplings that agree with $p$ if and only if $p$ satisfies the Tsirelson inequalities.

This work was supported by NSF grant SES-1155956.
References and Notes

1. J. Bell: On the Einstein-Podolsky-Rosen paradox. *Physics* 1, 195-200 (1964).

2. J.F. Clauser, M.A. Horne: Experimental consequences of objective local theories. *Phys. Rev. D*. 10, 526-535 (1974).

3. A. Fine: Hidden variables, joint probability, and the Bell inequalities. *Phys. Rev. Lett.* 48, 291-295 (1982).

4. A. Kolmogorov: *Foundations of the Theory of Probability* (New York, Chelsea, 1956).

5. E.N. Dzhafarov, J.V. Kujala: On selective influences, marginal selectivity, and Bell/CHSH inequalities. *Topics Cog. Sci.* (in press).

6. A.Yu. Khrennikov: Bell-Boole inequality: Nonlocality or probabilistic incompatibility of random variables? *Entropy* 10, 19-32 (2008).

7. A.Yu. Khrennikov: EPR–Bohm experiment and Bell’s inequality: Quantum physics meets probability theory. *Theor. Math. Phys.* 157, 1448–1460 (2008).

8. E.N. Dzhafarov, J.V. Kujala: All-possible-couplings approach to measuring probabilistic context. *PLoS ONE* 8(5): e61712. doi:10.1371/journal.pone.0061712 (2013).

9. H. Thorisson: *Coupling, Stationarity, and Regeneration* (New York, Springer, 2000).

10. D. Bohm, Y. Aharonov: Discussion of experimental proof for the paradox of Einstein, Rosen and Podolski. *Phys. Rev.*, 108: 1070-1076 (1957).

11. L.J. Landau: Empirical two-point correlation functions. *Found. Phys.* 18: 449–460 (1988).

12. A. Cabello: How much larger quantum correlations are than classical ones. *Phys. Rev. A* 72: 12113 1-5 (2005).

13. J.V. Kujala, E.N. Dzhafarov: Testing for selectivity in the dependence of random variables on external factors. *J. Math. Psych.* 52, 128-144 (2008).

14. B.S. Tsirelson: Quantum generalizations of Bell’s inequality. *Lett. Math. Phys.* 4: 93–100 (1980).

15. L.J. Landau: On the violation of Bell’s inequality in quantum theory. *Phys. Lett. A* 120: 54–56 (1987).

16. E.N. Dzhafarov, J.V. Kujala: A qualified Kolmogorovian account of probabilistic contextuality. *Lect. Notes in Comp. Sci.* (in press).

17. E.N. Dzhafarov, J.V. Kujala: Selectivity in probabilistic causality: Where psychology runs into quantum physics. *J. Math. Psych.* 56: 54-63 (2012).

18. E.N. Dzhafarov, J.V. Kujala: Order-distance and other metric-like functions on jointly distributed random variables. *Proc. Amer. Math. Soc.* (in press).