Quantum Entanglement and the Issue of Selective Influences in Psychology: An Overview

Ehtibar N. Dzhafarov[1] and Janne V. Kujala[2]

[1]Purdue University
ehtibar@purdue.edu
[2]University of Jyväskylä
jvkujala@jyu.fi

Abstract. Similar formalisms have been independently developed in psychology, to deal with the issue of selective influences (deciding which of several experimental manipulations selectively influences each of several, generally non-independent, response variables), and in quantum mechanics (QM), to deal with the EPR entanglement phenomena (deciding whether an EPR experiment allows for a "classical" account). The parallels between these problems are established by observing that any two noncommuting measurements in QM are mutually exclusive and can therefore be treated as analogs of different values of one and the same input. Both problems reduce to that of the existence of a jointly distributed system of random variables, one variable for every value of every input (in psychology) or every measurement on every particle involved (in an EPR experiment). We overview three classes of necessary conditions (some of them also sufficient under additional constraints) for the existence of such joint distributions.

Keywords: Bell-CHSH-Fine inequalities, cosphericity test, EPR paradigm, joint distribution criterion, linear feasibility test, non-commuting measurements, pseudo-quasi-metrics on random variables, quantum entanglement, selective influences.

1 Introduction

Given a set of inputs into a system and a set of stochastically non-independent outputs, what is the precise meaning and means of ascertaining that a given output *is not influenced* by a given input? This paper reviews the developments related to this question.

The problem can be illustrated on the following *diagram of selective influences*:

$$\boxed{\begin{array}{ccc} \alpha^1 = \{w, x, y\} & \alpha^2 = \{x\} & \alpha^3 = \{w, z\} \\ \downarrow & \downarrow & \downarrow \\ A^1 & A^2 & A^3 \end{array}} \quad (1)$$

A^1, A^2, and A^3 here are *random outputs*, w, x, y, z are *inputs* (usually referred to as *external factors* in psychology and as *measurement settings* in QM), and arrows indicate the relation "may influence": thus, the diagram does not say that A^2 is necessarily influenced by x, but rather that A^2 is not influenced by w, y, z. The diagram is shown in the *canonical form*, i.e., the inputs are redefined, $\{w, x, y\}$ into α^1, $\{x\}$ into α^2, etc., so that each output A^i may only be influenced by a single input α^i that may not influence other outputs. We say then, for brevity, that (A^1, A^2, A^3) are *selectively influenced* by $(\alpha^1, \alpha^2, \alpha^3)$ and write this as

$$\left(A^1, A^2, A^3\right) \looparrowleft \left(\alpha^1, \alpha^2, \alpha^3\right). \tag{2}$$

Inputs $(\alpha^1, \alpha^2, \alpha^3)$ are treated as deterministic quantities, i.e., even if they are random variables, the joint distribution of the outputs is always conditioned on their specific values. Each input can have one of several values, and the joint distribution of (A^1, A^2, A^3) is known for each *allowable treatment*, a combination of input values. Thus, if w, x, y, z are all binary, then $\alpha^1, \alpha^2, \alpha^3$ may be viewed as inputs with 8, 2, and 4 values, respectively, but the number of allowable treatments cannot exceed $16 < 8 \times 2 \times 4$. It can be less than 16 because some of the combinations may be physically impossible or simply not used or observed.

As a motivating example, consider a double-detection experiment in which two stimuli, say brief flashes, are presented simultaneously (right-left) or in a succession (first-second), each on one of two levels of intensity. The observer is asked to state, for each of the two *observation areas* (i.e., locations or time intervals), whether it contains a flash (Yes/No). The results of such an experiment are statistical estimates of 16 probabilities

$$p\left(A^1, A^2 | \alpha^1, \alpha^2\right) = \Pr\left[A^1 : \begin{cases} Yes \\ No \end{cases}, A^2 : \begin{cases} Yes \\ No \end{cases} \middle| \, \alpha^1 : \begin{cases} \alpha_1^1 \\ \alpha_2^1 \end{cases}, \alpha^2 : \begin{cases} \alpha_1^2 \\ \alpha_2^2 \end{cases}\right], \tag{3}$$

where α^i ($i = 1, 2$) is the input representing the ith observation area (with values α_1^i, α_2^i), and A^i is the response (Yes or No) to the ith observation area. Assume that A^1 and A^2 for a given (α_i^1, α_j^2) are not independent (due to attention fluctuations, perceptual learning, fatigue, etc.) In what sense then can we say that $(A^1, A^2) \looparrowleft (\alpha^1, \alpha^2)$, and by what means can we find out if this is true?

Many empirical situations have precisely the same formal structure. In QM, an example is provided by the Bohmian version of the EPR paradigm [3]: two subatomic particles are emitted from a common source in such a way that they retain highly correlated spins as they run away from each other. An experiment may consist, e.g., in measuring the spin of electron 1 along one of two axes, α_1^1 or α_2^1, and (in another location but simultaneously in some inertial frame of reference) measuring the spin of electron 2 along one of two axes, α_1^2 or α_2^2. The outcome of a measurement on electron 1, A^1, is a random variable with two possible values, "up" or "down," and the same holds for A^2, outcome of a measurement on electron 2. The question here is: for $i = 1, 2$, can we say that A^i may only depend on α^i, even though A^1 and A^2 are not independent? What makes this situation formally identical with the double-detection example is that the measurements along different axes, α_1^i and α_2^i, are *noncommuting*, i.e., they cannot be performed on the ith particle simultaneously. This makes it possible to consider them (measurements performed, not to be confused with their recorded outcomes) as mutually exclusive values of input α^i. The results of such an experiment are described by (3), with Yes/No interpreted as spin up/down. In the original EPR paradigm [14] the non-commuting measurements are those of momentum and location, each with a continuum of possible values. Our parallel with the issue of selective influences requires that the measurements of the momentum and of the location of a given particle be interpreted as mutually exclusive values of one and the same input, "(measurement of the) momentum-location of the particle." This may be less intuitive than the analogous interpretation of the spins along different axes.

The question of selective influences cannot generally be decided based on the marginal distributions of the outputs alone. The most important example here is the classical CHSH experiment [4] where the marginal distributions of A^1 and A^2 (in the case of two electrons) remain constant, with $\Pr[\textit{spin up}] = 1/2$. Examples from psychology are also readily available, especially if one adopts a copula view of the joint distributions. Thus, α^1 and α^2 may represent two stimuli presented in a succession (each having several values), and A^1, A^2 be *response times quantiles*. The marginal distributions then are always the same, unit-uniform.

2 A Historical Note

The issue of selective influences was introduced to psychology in Sternberg's influential paper [22], in the context of studying consecutive "stages" of information processing. Sternberg acknowledged that selective influences can hold even if the durations of the stages are not stochastically independent, but he lacked mathematical apparatus for dealing with this possibility. Townsend [24] proposed to formalize the notion of selectively influenced and stochastically interdependent random variables by the concept of "*indirect nonselectiveness*": the conditional distribution of the variable A^1 given any value a^2 of the variable A^2, depends on α^1 only, and, by symmetry, the conditional distribution of A^2 at any $A^1 = a^1$ depends on α^2 only. Under the name of "*conditionally selective influence*" this notion was mathematically characterized and generalized in [5]. Thus, if all combinations of values of inputs α^1, α^2 are allowable and random outputs A^1, A^2 are discrete, the diagram $\left(A^1, A^2\right) \overset{cond}{\leftarrow} \left(a^1, a^2\right)$, where $\overset{cond}{\leftarrow}$ means "is conditionally selectively influenced," holds if and only if $\Pr\left[A^1 = a^1, A^2 = a^2 \mid \alpha^1_x, \alpha^2_y\right]$ can be presented as

$$f_{12}\left(a^1, a^2\right) f_1\left(a^1, \alpha^1_x\right) f_2\left(a^2, \alpha^2_y\right) f\left(\alpha^1_x, \alpha^2_y\right), \tag{4}$$

for all values $\left(a^1, a^2\right)$ of $\left(A^1, A^2\right)$ at all treatments $\left(\alpha^1_x, \alpha^2_y\right)$. Conditional selectivity is a useful notion, but it is not a satisfactory formalization of the intuitive notion of selective influences. The reason is that $\left(A^1, A^2\right) \overset{cond}{\leftarrow} \left(a^1, a^2\right)$ can be shown [5] to violate the following obvious property of an acceptable definition:

the marginal distributions of A^1 and A^2 do not depend on, respectively, α^2 and α^1 ("*marginal selectivity*" [25]).

A different approach to selective influences, reviewed below, is based on [6,7,9,10,11,12,19]. As it turns out[1] this approach parallels the development in QM of the issue of whether an EPR experiment can have a "classical" explanation (in terms of non-contextual local variables). The Joint Distribution Criterion which is at the heart of this development (see below) was indirectly introduced in the celebrated work of Bell [2], and explicitly in [15,16,23].

3 Basic Notions

Aimed at providing a broad overview of concepts and results, the content of this paper partially overlaps with that of several previous publications, especially [11,12,19].

Random variables are understood in the broadest sense, as measurable functions $X : V_s \to V$, with no restrictions on the sample spaces (V_s, Σ_s, μ_s) and the induced probability spaces (*distributions*) (V, Σ, μ). In particular, any set X of jointly distributed random variables (functions on the same sample space) is a random variable, and its distribution (V, Σ, μ) is referred to as the *joint distribution* of its elements. We use symbol $\sim$ in the meaning of "has the same distribution as." A random variable in the narrow sense is a special case of a random entity, with V a finite product of countable sets and intervals of reals, and Σ the smallest sigma-algebra containing the corresponding product of power sets and Lebesgue sigma-algebras. Note that a vector of random variables in the narrow sense is a random variable in the narrow sense.

Consider an indexed set $\alpha = \left\{\alpha^\lambda : \lambda \in \Lambda\right\}$, with each α^λ being a set referred to as a (deterministic) *input*, with the elements of $\{\lambda\} \times \alpha^\lambda$ called *input points.* Input points therefore are pairs of the form $x = (\lambda, w)$, with $w \in \alpha^\lambda$, and should not be confused with *input values* w. A nonempty set $\Phi \subset \prod_{\lambda \in \Lambda} \alpha^\lambda$ is called a set of *(allowable) treatments.* A treatment therefore is a function $\phi : \Lambda \to \bigcup_{\lambda \in \Lambda} \alpha^\lambda$ such that $\phi(\lambda) \in \alpha^\lambda$ for any $\lambda \in \Lambda$.

Let there be a collection of sets of random variables A^λ_ϕ ($\lambda \in \Lambda$, $\phi \in \Phi$), referred to as (random) *outputs*, with distributions $\left(V^\lambda, \Sigma^\lambda, \mu^\lambda_\phi\right)$. Let

$$A_\phi = \left\{A^\lambda_\phi : \lambda \in \Lambda\right\}, \ \phi \in \Phi, \tag{5}$$

be a random variable with a known distribution (the joint distribution of all A^λ_ϕ in A_ϕ) for every treatment $\phi \in \Phi$. We define

$$A^\lambda = \left\{A^\lambda_\phi : \phi \in \Phi\right\}, \ \lambda \in \Lambda, \tag{6}$$

with the understanding that A^λ is not generally a random variable, i.e., A^λ_ϕ for different ϕ are not necessarily jointly distributed. The definition of the relation

$$\left\{A^\lambda : \lambda \in \Lambda\right\} \looparrowleft \left\{\alpha^\lambda : \lambda \in \Lambda\right\}, \tag{7}$$

[1] This was first pointed out to us by Jerome Busemeyer (personal communication, November 2010), for which we remain deeply grateful.

interpreted as "for each $\lambda \in \Lambda$, A^{λ} may be influenced by α^{λ} only," can be given in three equivalent forms:

(SI_1) there are independent random variables C, $\{S^{\lambda} : \lambda \in \Lambda\}$, and functions

$$\{R^{\lambda}(w, C, S^{\lambda}) : w \in \alpha^{\lambda}, \lambda \in \Lambda\}, \tag{8}$$

such that, for any treatment $\phi \in \Phi$,

$$\{R^{\lambda}(\phi(\lambda), C, S^{\lambda}) : \lambda \in \Lambda\} \sim A_{\phi}; \tag{9}$$

(SI_2) there is a random variable C and functions

$$\{P^{\lambda}(w, C) : w \in \alpha^{\lambda}, \lambda \in \Lambda\}, \tag{10}$$

such that, for any treatment $\phi \in \Phi$,

$$\{P^{\lambda}(\phi(\lambda), C) : \lambda \in \Lambda\} \sim A_{\phi}; \tag{11}$$

(JDC) there is a set of jointly distributed random variables

$$H = \{H_w^{\lambda} : w \in \alpha^{\lambda}, \lambda \in \Lambda\} \tag{12}$$

(one random variable for every value of every input), such that, for any treatment $\phi \in \Phi$,

$$\left\{H_{\phi(\lambda)}^{\lambda} : \lambda \in \Lambda\right\} \sim A_{\phi}. \tag{13}$$

The latter statement constitutes the *Joint Distribution Criterion* (JDC) for selective influences, and H is called the JDC *(indexed) set*. The proof of the equivalence [10] obtains essentially by the definition of a joint distribution, which seems to have been overlooked in the earlier derivations [15,16]. If $\Lambda = \{1, \ldots, n\}$ and all outputs A^{λ} are random variables in the narrow sense, then C in SI_2 and $C, S^1, \ldots, S^n$ in SI_1 can also be chosen to be random variables in the narrow sense; moreover, their distribution functions can be chosen arbitrarily, provided they are continuous and strictly increasing on their domains, e.g., unit uniform [11].

Two important consequences of (7) are as follows:

1. *(nestedness)* any subset Λ' of Λ, $\{A^{\lambda} : \lambda \in \Lambda'\} \looparrowleft \{\alpha^{\lambda} : \lambda \in \Lambda'\}$; in particular, $\{A^{\lambda} : \lambda \in \Lambda'\}$ may not depend on inputs outside Λ' (*complete marginal selectivity*);
2. (*invariance with respect to input-value-specific transformations*) for any set of measurable functions $\{F_w^{\lambda}(a) : w \in \alpha^{\lambda}, \lambda \in \Lambda, a \in V^{\lambda}\}$,

$$(B^{\lambda} : \lambda \in \Lambda) \looparrowleft \{\alpha^{\lambda} : \lambda \in \Lambda\} \tag{14}$$

where $B^{\lambda} = \left\{B_{\phi}^{\lambda} : \phi \in \Phi\right\}$, and $B_{\phi}^{\lambda} = F_{\phi(\lambda)}^{\lambda}\left(A_{\phi}^{\lambda}\right)$.

These properties should be viewed as desiderata for any reasonable definition of selective influences.

In QM, SI_1 corresponds to the existence of a *"classical" probabilistic* explanation. In psychology, statement SI_1 combined with auxiliary assumptions was used in [8] and [20] to analyze the representability of same-different pairwise discrimination probabilities by means of *Thurstonian-type models* in which two stimuli being compared are mapped into random entities (distributed in some hypothetical space of mental images) that in turn are mapped (deterministically or probabilistically) into a response, "same" or "different." Statement SI_1 was also used to analyze the response time distributions for *parallel-serial networks of mental operations* with selectively influenced components [13]. Note that the representation of the outputs A^λ as functions of the corresponding inputs α^i and unobservable sources of randomness, A^λ-specific (S^λ) and common (C), includes as special cases all conceivable generalizations and combinations of *regression* and *factor analyses*, with our term "input" corresponding to the traditional "regressor," and the term "source of randomness" to the factor-analytic "factor." This observation alone shows the potentially unlimited sphere of applicability of SI_1.

Statement SI_2 (corresponding in QM to *"classical" determinsitic* explanation) and JDC turn out to be more convenient in dealing with certain foundational probabilistic issues [9] and for the construction of the working *tests* (*necessary conditions*) for selective influences [10,11,12,19]. The tests are discussed below.

The following is a table of correspondences between the general terminology used in dealing with the issue of selective influences, and that of QM in dealing with EPR.

Selective Probabilistic Causality (general)	Quantum Entanglement Problem
observed random output	outcome of a given measurement on a given particle
input (factor)	set of noncommuting measurements on a given particle
input value	one of noncommuting measurements on a given particle
joint distribution criterion	joint distribution criterion
diagram of selective influences	"classical" explanation
representation in the form SI_1	probabilistic "classical" explanation
representation in the form SI_2	deterministic "classical" explanation

4 Tests for Selective Influences

Let $H = \left\{H_w^\lambda : w \in \alpha^\lambda, \lambda \in \mathit{\Lambda}\right\}$ be a *hypothetical* JDC-set, i.e., a set satisfying (13) but not necessarily jointly distributed. Denoting

$$\left\{H_{\phi(\lambda)}^\lambda : \lambda \in \mathit{\Lambda}\right\} = H_\phi,\ \phi \in \mathit{\Phi}, \tag{15}$$

let $\mathcal{H}$ be a *set of constraints* imposed on possible distributions of H_ϕ. For instance, $\mathcal{H}$ may be the requirement that all H_ϕ^λ be composed of Bernoulli variables, or multivariate-normally distributed.

A statement $S(H_{\phi_1}, \ldots, H_{\phi_s})$, with $\phi_1, \ldots, \phi_s \in \Phi$, is called a *test* for the relation (7) under constraints $\mathcal{H}$, if

1. (*observability*) its truth value only depends on the distributions of $H_{\phi_1}, \ldots, H_{\phi_s}$;
2. (*non-emptiness*) it is not true for all possible distributions of $H_{\phi_1}, \ldots, H_{\phi_s}$ satisfying $\mathcal{H}$,
3. (*necessity*) it is true if H is jointly distributed.

If $S(H_{\phi_1}, \ldots, H_{\phi_s})$ is false for all distributions of $H_{\phi_1}, \ldots, H_{\phi_s}$ satisfying $\mathcal{H}$ unless H is jointly distributed, the test is called a *criterion* for (7). In the following we assume that $\mathcal{H}$ always includes the requirement of complete marginal selectivity: for any $\Lambda' \subset \Lambda$, the joint distribution of $\left\{A^\lambda_{\phi(\Lambda')\cup\phi(\Lambda-\Lambda')} : \lambda \in \Lambda'\right\}$ does not depend on $\phi(\Lambda - \Lambda')$. If this condition is violated, (7) is ruled out trivially.

4.1 Pseudo-quasi-distance tests

A function $d : H \times H \to \mathbb{R}$ is a *pseudo-quasi-metric* (*p.q.-metric*) on H if, for any $H_1, H_2, H_3 \in H$,
(i) $d(H_1, H_2)$ only depends on the joint distribution of (H_1, H_2),
(ii) $d(H_1, H_2) \geq 0$,
(iii) $d(H_1, H_1) = 0$,
(iv) $d(H_1, H_3) \leq d(H_1, H_2) + d(H_2, H_3)$.

The conventional *pseudometrics* (also called *semimetrics*) obtain by adding the property $d(H_1, H_2) = d(H_2, H_1)$; the conventional *quasimetrics* are obtained by adding the property $\Pr[H_1 = H_2] < 1 \Rightarrow d(H_1, H_2) > 0$. A conventional *metric* is both a pseudometric and a quasimetric.

A sequence of input points

$$x_1 = (\lambda_1, w_1), \ldots, x_l = (\lambda_l, w_l), \tag{16}$$

where $w_i \in \alpha^{\lambda_i}$ for $i = 1, \ldots, l \geq 3$, is called *treatment-realizable* if there are treatments $\phi^1, \ldots, \phi^l \in \Phi$ (not necessarily pairwise distinct), such that

$$\{x_1, x_l\} \subset \phi^1 \text{ and } \{x_{i-1}, x_i\} \subset \phi^i \text{ for } i = 2, \ldots, l. \tag{17}$$

If a JDC-set H exists, then for any p.q.-metric d on H we should have

$$d\left(H^{\lambda_1}_{w_1}, H^{\lambda_l}_{w_l}\right) = d\left(A^{\lambda_1}_{\phi^1}, A^{\lambda_l}_{\phi^1}\right) \tag{18}$$

and

$$d\left(H^{\lambda_{i-1}}_{w_{i-1}}, H^{\lambda_i}_{w_i}\right) = d\left(A^{\lambda_{i-1}}_{\phi^i}, A^{\lambda_i}_{\phi^i}\right),\ i = 2, \ldots, l, \tag{19}$$

whence

$$d\left(A^{\lambda_1}_{\phi^1}, A^{\lambda_l}_{\phi^1}\right) \leq \sum_{i=2}^{l} d\left(A^{\lambda_{i-1}}_{\phi^i}, A^{\lambda_i}_{\phi^i}\right). \tag{20}$$

This chain inequality constitutes a *p.q.-metric test* for selective influences. If this inequality is found not to hold for at least one treatment-realizable sequence of input points, selectivity (7) is ruled out [12].

It turns out that one needs to check the chain inequality only for *irreducible* treatment-realizable sequences $x_1, \ldots, x_l$, i.e., those with $x_1 \neq x_l$ and with the property that the only subsequences $\{x_{i_1}, \ldots, x_{i_k}\}$ with $k > 1$ that are subsets of treatments are pairs $\{x_1, x_l\}$ and $\{x_{i-1}, x_i\}$, for $i = 2, \ldots, l$. Inequality (20) is satisfied for all treatment-realizable sequences if and only if it holds for all irreducible sequences [12]. The situation is even simpler if $\Phi = \prod_{\lambda \in \Lambda} W^\lambda$ (all logically possible treatments are allowable). Then (20) is satisfied for all treatment-realizable sequences if and only if this inequality holds for all tetradic sequences of the form x, y, s, t, with $x, s \in \{\lambda_1\} \times \alpha^{\lambda_1}$, $y, t \in \{\lambda_2\} \times \alpha^{\lambda_2}$, $x \neq s$, $y \neq t$, $\lambda_1 \neq \lambda_2$ [10].

Order-distances constitute a special class of p.q.-metrics, defined as follows. Let the distribution of $H_w^\lambda \in H$ be $\left(V^\lambda, \Sigma^\lambda, \mu_w^\lambda\right)$. Let

$$R \subset \bigcup_{(\lambda_1,\lambda_2) \in \Lambda \times \Lambda} V^{\lambda_1} \times V^{\lambda_2}, \tag{21}$$

and let us write $a \preceq b$ for $(a, b) \in R$. Let R be a total order (transitive, reflexive, and connected). We assume that for any $(\lambda_1, \lambda_2) \in \Lambda \times \Lambda$, $\Pr\left[H_{w_1}^{\lambda_1} \preceq H_{w_2}^{\lambda_2}\right]$ is well-defined, i.e., $\left\{(a, b) : a \in V^{\lambda_1}, b \in V^{\lambda_2}, a \preceq b\right\}$ belongs to the product sigma-algebra over Σ^{λ_1} and Σ^{λ_2}. Then the function

$$\mathrm{D}\left(H_{w_1}^{\lambda_1}, H_{w_2}^{\lambda_2}\right) = \Pr\left[H_{w_1}^{\lambda_1} \prec H_{w_2}^{\lambda_2}\right], \tag{22}$$

where $\prec$ is the strict order induced by $\preceq$, is well-defined, and it is a p.q.-metric on H, called order-distance [12].

As a simple example, consider the results of a CHSH type experiment with two spin axes per each of two entangled $1/2$-spin particles. Enumerate the spin axes $1, 2$ for either particle, enumerate the two outcomes (up and down) of each measurement $1, 2$ for particle 1 and $1', 2'$ for particle 2, and denote

$$\Pr\left[H_i^1 = k, H_j^2 = l'\right] = \Pr\left[A_{(i,j)}^1 = k, A_{(i,j)}^2 = l'\right] = p_{kl|ij}, \tag{23}$$

where $i, j, k, l \in \{1, 2\}$. Define the order-distance D_1 by putting $1 \simeq 1' \prec 2 \simeq 2'$, where $\simeq$ is equivalence induced by $\preceq$. We have then the chain inequality

$$\begin{aligned} p_{12|12} &= \mathrm{D}_1(H_1^1, H_2^2) \\ &\leq \mathrm{D}_1(H_1^1, H_1^2) + \mathrm{D}_1(H_1^2, H_2^1) + \mathrm{D}_1(H_2^1, H_2^2) = p_{12|11} + p_{21|21} + p_{12|22}. \end{aligned} \tag{24}$$

Consider next a similar inequality for the order-distance D_2 defined by $1 \simeq 2' \prec 2 \simeq 1'$:

$$\begin{aligned} p_{11|12} &= \mathrm{D}_2(H_1^1, H_2^2) \\ &\leq \mathrm{D}_2(H_1^1, H_1^2) + \mathrm{D}_2(H_1^2, H_2^1) + \mathrm{D}_2(H_2^1, H_2^2) = p_{11|11} + p_{22|21} + p_{11|22}. \end{aligned} \tag{25}$$

By simple algebra, denoting

$$\Pr\left[H_i^1 = k\right] = p_{k\cdot|i\cdot},\ \Pr\left[H_j^2 = l'\right] = p_{\cdot l|\cdot j}, \tag{26}$$

the conjunction of (24) and (25) can be shown to be equivalent to

$$-1 \le p_{11|11} + p_{11|21} + p_{11|22} - p_{11|12} - p_{1\cdot|2\cdot} - p_{\cdot 1|\cdot 1} \le 0. \tag{27}$$

One derives analogously

$$\begin{array}{l} -1 \le p_{11|12} + p_{11|22} + p_{11|21} - p_{11|11} - p_{1\cdot|2\cdot} - p_{\cdot 1|\cdot 2} \le 0, \\ -1 \le p_{11|21} + p_{11|11} + p_{11|12} - p_{11|22} - p_{1\cdot|1\cdot} - p_{\cdot 1|\cdot 1} \le 0, \\ -1 \le p_{11|22} + p_{11|12} + p_{11|11} - p_{11|21} - p_{1\cdot|1\cdot} - p_{\cdot 1|\cdot 2} \le 0. \end{array} \tag{28}$$

The four double-inequalities (27)-(28) can be referred to as the *Bell-CHSH-Fine inequalities* [15,16], necessary and sufficient conditions for the CHSH type experiment to have a "classical" explanation.

4.2 Cosphericity Tests

Let the outputs A_ϕ^λ all be random variables in the narrow sense. Denote, for any distinct $\lambda_1, \lambda_2 \in \Lambda$ and any $\phi \in \Phi$ with $\phi(\lambda_1) = w_1$ and $\phi(\lambda_2) = w_2$,

$$\mathrm{Cor}\left[H_{w_1}^{\lambda_1}, H_{w_2}^{\lambda_2}\right] = \mathrm{Cor}\left[A_\phi^{\lambda_1}, A_\phi^{\lambda_2}\right] = \rho_{w_1 w_2}^{\lambda_1 \lambda_2}, \tag{29}$$

where Cor designates correlation. Let $\phi_1, \phi_2, \phi_3, \phi_4 \in \Phi$ be any treatments with

$$\begin{array}{l} \phi_1(\lambda_1) = \phi_2(\lambda_1) = w_1;\ \phi_1(\lambda_2) = \phi_3(\lambda_2) = w_2 \\ \phi_4(\lambda_1) = \phi_2(\lambda_1) = w_1';\ \phi_4(\lambda_2) = \phi_3(\lambda_2) = w_2'. \end{array} \tag{30}$$

Then, as shown in [19], if the components of H are jointly distributed,

$$\begin{array}{l} \left|\rho_{w_1 w_2}^{\lambda_1\lambda_2}\rho_{w_1 w_2'}^{\lambda_1\lambda_2} - \rho_{w_1' w_2}^{\lambda_1\lambda_2}\rho_{w_1' w_2'}^{\lambda_1\lambda_2}\right| \\ \quad \le \sqrt{1-\left(\rho_{w_1 w_2}^{\lambda_1\lambda_2}\right)^2}\sqrt{1-\left(\rho_{w_1 w_2'}^{\lambda_1\lambda_2}\right)^2} + \sqrt{1-\left(\rho_{w_1' w_2}^{\lambda_1\lambda_2}\right)^2}\sqrt{1-\left(\rho_{w_1' w_2'}^{\lambda_1\lambda_2}\right)^2}, \end{array} \tag{31}$$

This is the *cosphericity test* for (7), called so because geometrically (31) describes the possibility to place four points (w_1, w_2, w_1', w_2') on a unit sphere in 3D Euclidean space so that the angles between the corresponding radius-vectors have cosines equal to the correlations. Note that an outcome of this test does not allow to predict the outcome of the same test applied to nonlinearly input-value-specifically transformed random variables. Due to (14), this creates a multitude of cosphericity tests for one and the same initial set of outputs A_ϕ^λ .

In the all-important for behavioral sciences 2×2 factorial design ($\Lambda = \{1, 2\}$, each input is binary, and Φ consists of all four possible treatments), the cosphericity test is a criterion for $\left(A^1, A^2\right) \looparrowleft \left(\alpha^1, \alpha^2\right)$ if (perhaps following some input-value-specific transformation) the outputs are bivariate normally distributed for all four treatments [19].

4.3 Linear Feasibility Test

The *Linear Feasibility Test* (LFT) is a criterion for selective influences in all situations involving finite sets of inputs/outputs, $\Lambda = \{1, \ldots, n\}$, with the ith input and ith output having finite sets of values, $\{1, \ldots, k_i\}$ and $\{1, \ldots, m_i\}$, respectively [11]. In other situations LFT can be used as a necessary condition because every set of possible values can be discretized. The distributions of $H_\phi = \left(H^1_{j_1}, \ldots, H^n_{j_n}\right)$ are represented by probabilities

$$\Pr\left[H^1_{j_1} = a_1, \ldots, H^n_{j_n} = a_n\right] = \Pr\left[A^1_\phi = a_1, \ldots, A^n_\phi = a_n\right], \quad (32)$$

with $\phi = (j_1, \ldots, j_n) \in \Phi$ and

$$(a_1, \ldots, a_n) \in \{1, \ldots, m_1\} \times \cdots \times \{1, \ldots, m_n\}. \quad (33)$$

We consider this probability the $[(a_1, \ldots, a_n), (j_1, \ldots, j_n)]$th component of the $m_1 \cdots m_n t$-vector P (with t denoting the number of treatments in Φ). The joint distribution of H in JDC, if it exists, is represented by probabilities

$$\Pr\left[H^1_1 = h^1_1 \ldots, H^1_{k_1} = h^1_{k_1}, \ldots, H^n_1 = h^n_1, \ldots, H^n_{k_n} = h^n_{k_n}\right], \quad (34)$$

with

$$\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right) \in \{1, \ldots, m_1\}^{k_1} \times \ldots \times \{1, \ldots, m_n\}^{k_n}. \quad (35)$$

We consider this probability the $\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right)$th component of the $(m_1)^{k_1} \cdots (m_n)^{k_n}$-vector Q. Consider now the Boolean matrix M with rows corresponding to components of P and columns to components of Q: let $M(r, c) = 1$ if and only if

1. row r corresponds to the $[(j_1, \ldots, j_n), (a_1, \ldots, a_n)]$th component of P,
2. column c to the $\left(h^1_1, \ldots, h^1_{k_1}, \ldots, h^n_1, \ldots, h^n_{k_n}\right)$th component of Q, and
3. $h^1_{j_1} = a_1, \ldots, h^n_{j_n} = a_n$.

Clearly, the vector Q exists if and only if the system

$$MQ = P, \ Q \geq 0 \quad (36)$$

has a solution (is *feasible*). This is a linear programming task in the standard form (with a constant objective function). Let $\mathcal{L}(P)$ be a Boolean function equal to 1 if and only if this system is feasible. $\mathcal{L}(P)$ is known to be computable, its time complexity being polynomial [18].

The potential of JDC to lead to LFT and provide an ultimate criterion for the Bohmian entanglement problem has not been utilized in quantum physics until relatively recently, when LFT was proposed in [26,27] and [1]. But the essence of the idea can be found in [21]. Given a set of numerical (experimentally estimated or theoretical) probabilities, computing $\mathcal{L}(P)$ is always preferable to dealing with explicit inequalities as their number becomes very large even for moderate-size vectors P. The classical Bell-CHSH-Fine inequalities (27)-(28) for

$n = 2$, $k_1 = k_2 = 2$, $m_1 = m_2 = 2$ (assuming that the marginal selectivity equalities hold) number just 8, but already for $n = 2$, $k_1 = k_2 = 2$ with $m_1 = m_2 = 3$ (describing, e.g., an EPR experiment with two spin-1 particles, or two spin-$1/2$ ones and inefficient detectors), our computations yield 1080 inequalitiies equivalent to $\mathcal{L}(P) = 1$. For $n = 3$, $k_1 = k_2 = k_3 = 2$ and $m_1 = m_2 = m_3 = 2$, corresponding to the GHZ paradigm [17] with three spin-$1/2$ particles, this number is 53792. Lists of such inequalities can be derived "mechanically" from the format of matrix M using well-known facet enumeration algorithms (see, e.g., program lrs at http://cgm.cs.mcgill.ca/~avis/C/lrs.html). Once such a system of inequalities S is derived, one can use it to prove necessity (or sufficiency) of any other system S' by showing, with the aid of a linear programming algorithm, that S' is redundant when added to S (respectively, S is redundant when added to S').

Acknowledgments. This research has been supported by the NSF grant SES-1155956 to Purdue University and the Academy of Finland grant 121855 to University of Jyväskylä.

References

1. Basoalto, R.M., Percival, I.C.: BellTest and CHSH experiments with more than two settings. Journal of Physics A: Mathematical & General 36, 7411–7423 (2003)
2. Bell, J.: On the Einstein-Podolsky-Rosen paradox. Physics 1, 195-200 (1964)
3. Bohm, D., Aharonov, Y.: Discussion of Experimental Proof for the Paradox of Einstein, Rosen and Podolski. *Physical Review* 108, 1070-1076 (1957)
4. Clauser, J.F., Horne, M.A., Shimony, A., Holt, R.A.: Proposed experiment to test local hidden-variable theories. Physical Review Letters 23, 880-884 (1969)
5. Dzhafarov, E.N.: Conditionally selective dependence of random variables on external factors. Journal of Mathematical Psychology 43, 123-157 (1999)
6. Dzhafarov, E.N.: Unconditionally selective dependence of random variables on external factors. Journal of Mathematical Psychology 45, 421-451 (2001)
7. Dzhafarov, E.N.: Selective influence through conditional independence. Psychometrika 68, 7-26 (2003)
8. Dzhafarov, E.N.: Thurstonian-type representations for "same-different" discriminations: Probabilistic decisions and interdependent images. Journal of Mathematical Psychology 47, 229-243 (2003) [see Dzhafarov, E.N.: Corrigendum to "Thurstonian-type representations for 'same–different' discriminations: Probabilistic decisions and interdependent images." Journal of Mathematical Psychology 50, 511 (2006)]
9. Dzhafarov, E.N., Gluhovsky, I.: Notes on selective influence, probabilistic causality, and probabilistic dimensionality. Journal of Mathematical Psychology 50, 390–401 (2006)
10. Dzhafarov, E.N., Kujala, J.V.: The Joint Distribution Criterion and the Distance Tests for Selective Probabilistic Causality. Frontiers in Quantitative Psychology and Measurement 1:151 doi: 10.3389/fpsyg.2010.0015 (2010)
11. Dzhafarov, E.N., Kujala, J.V.: Selectivity in probabilistic causality: Where psychology runs into quantum physics. Journal of Mathematical Psychology 56, 54-63 (2012)

12. Dzhafarov, E.N., Kujala, J.V.: Order-distance and other metric-like functions on jointly distributed random variables. Proceedings of the American Mathematical Society (in press as of 2011)
13. Dzhafarov, E.N., Schweickert, R., Sung, K.: Mental architectures with selectively influenced but stochastically interdependent components. Journal of Mathematical Psychology 48, 51-64 (2004)
14. Einstein, A., Podolsky, B., Rosen N.: Can Quantum-Mechanical Description of Physical Reality be Considered Complete? Physical Review 47, 777–780 (1935)
15. Fine, A.: Joint distributions, quantum correlations, and commuting observables. Journal of Mathematical Physics 23, 1306-1310 (1982)
16. Fine, A.: Hidden variables, joint probability, and the Bell inequalities. Physical Review Letters 48, 291-295 (1982)
17. Greenberger, D.M., Horne, M.A., Zeilinger, A.: Going beyond Bell's theorem. In: M. Kafatos (ed.) Bell's Theorem, Quantum Theory and Conceptions of the Universe, pp. 69–72. Kluwer, Dordrecht (1989)
18. Karmarkar, N.: A new polynomial-time algorithm for linear programming. Combinatorica 4, 373-395 (1984)
19. Kujala, J.V., Dzhafarov, E.N.: Testing for selectivity in the dependence of random variables on external factors. Journal of Mathematical Psychology 52, 128–144 (2008)
20. Kujala, J.V., Dzhafarov, E.N.: Regular Minimality and Thurstonian-type modeling. Journal of Mathematical Psychology 53, 486–501 (2009)
21. Pitowski, I.: Quantum Probability – Quantum Logic. Springer, Berlin (1989)
22. Sternberg, S.: The discovery of processing stages: Extensions of donders' method. Acta Psychologica 30, 276–315 (1969)
23. Suppes, P., Zanotti, M.: When are probabilistic explanations possible? Synthese 48, 191-199 (1981)
24. Townsend, J. T.: Uncovering mental processes with factorial experiments. Journal of Mathematical Psychology 28, 363–400 (1984)
25. Townsend, J.T., Schweickert, R.: Toward the trichotomy method of reaction times: Laying the foundation of stochastic mental networks. Journal of Mathematical Psychology 33, 309–327 (1989)
26. Werner, R.F., Wolf, M.M.: All multipartite Bell correlation inequalities for two dichotomic observables per site. arXiv:quant-ph/0102024v1 (2001)
27. Werner, R.F., Wolf, M.M.:. Bell inequalities and entanglement. arXiv:quant-ph/0107093 v2 (2001)