A Stronger Bell Argument for Quantum Non-Locality

Paul M. Näger

Department of Philosophy

University of Bremen, Enrique-Schmidt-Str. 7, D-28359 Bremen, Germany

paul.naeger@uni-bremen.de

15th August 2013
Version 3.21

It is widely accepted that the violation of Bell inequalities excludes local theories of the quantum realm. This paper presents a stronger Bell argument which even forbids certain non-local theories. The remaining non-local theories, which can violate Bell inequalities, are characterised by the fact that at least one of the outcomes in some sense probabilistically depends both on its distant as well as on its local setting. While this is not to say that parameter dependence in the usual sense necessarily holds, it shows that some kind of parameter dependence cannot be avoided, even if the outcomes depend on another. Our results are the strongest possible consequences from the violation of Bell inequalities on a qualitative probabilistic level.

Contents

1. Introduction

Bell's argument (1964; 1971; 1975) establishes a mathematical no-go theorem for theories of the micro-world. In its standard form, it derives that theories which are local (and fulfil certain auxiliary assumptions) cannot have correlations of arbitrary strength between events which are space-like separated. An upper bound for the correlations is given by the famous Bell inequalities. Since certain experiments with entangled quantum objects have results which violate these inequalities (EPR/B correlations), it concludes that the quantum realm cannot be described by a local theory. Any correct theory of the quantum realm must involve some kind of non-locality, a 'quantum non-locality'. This result is one of the central features of the quantum realm. It is the starting point for extensive debates concerning the nature of quantum objects and their relation to space and time.

Since Bell's first proof (1964) the theorem has evolved considerably towards stronger forms: there has been a sequence of improvements which derive the inequalities from weaker and weaker assumptions. The main focus has been on getting rid of those premises which are commonly regarded as auxiliary assumptions: Clauser et al. (1969) derived the theorem without assuming perfect correlations; Bell (1971) abandoned the assumption of determinism; Graßhoff et al. (2005) and Portmann and Wüthrich (2007) showed that possible latent common causes do not have to be *common* common causes.[1] What is common to all of these different derivations is that they assume one or another form of locality. Locality seems to be the central assumption in deriving the Bell inequalities—and hence it is the assumption that is assumed to fail when one finds that the inequalities are violated.

In this paper we are going to present another strengthening of Bell's theorem, which relaxes the central assumption: one does not have to assume locality in order to derive the Bell inequalities. Certain forms of non-locality, which we shall call 'weakly non-local' suffice: an outcome may depend on the other outcome or on the distant setting—as long as it does not depend on *both* settings, it still implies that the Bell inequalities hold. As a consequence, the violation of the Bell inequalities also excludes those weakly non-local theories. So it does not require *any* kind of non-locality, but a very *specific* one: at least one of the outcomes must depend probabilistically on both settings. While previous strengthenings of Bell's theorem secured that a certain auxiliary assumption is *not* the culprit, our derivation here for the first time strengthens the *conclusion* of the theorem.

We start by briefly reviewing the underlying experiments with entangled photons and introduce an appropriate notation (section 2). Second, we formulate the standard Bell argument in an explicit and clear form and state exactly which of its premisses we are going to strengthen (section 3). In section 4 we develop a comprehensive classification of probability distributions for EPR experiments and show which of these distributions implies Bell inequalities. Here we find the surprising fact that besides the local distributions also the weakly non-local ones imply the inequalities. Finally, we formulate the

[1]The debate about common common causes vs. separate common causes is to some degree still undecided (cf. Hofer-Szabó, 2008).

new stronger Bell argument (section 5) and discuss some of its immediate consequences (section 6).

2. EPR/B experiments and correlations

Many arguments for a quantum non-locality consider an EPR/B setup with polarisation measurements of photons (fig. 1; Einstein, Podolsky, and Rosen 1935; Bohm 1951; Clauser and Horne 1974). One run of the experiment goes as follows: a suitable source C (e.g. a calcium atom) is excited and emits a pair of photons whose quantum mechanical polarisation state $\boldsymbol{\psi}$ is entangled. Possible hidden variables of this state are called $\boldsymbol{\lambda}$, so that the *complete* state of the particle at the source is $(\boldsymbol{\psi}, \boldsymbol{\lambda})$. Since the preparation procedure is usually the same in all runs, the quantum mechanical state $\boldsymbol{\psi}$ is the same in all runs and will not explicitly be noted in the following. (One may think of any probability being conditional on one fixed state $\boldsymbol{\psi} = \psi_0$.) After the emission, the photons move in opposite direction towards two polarisation measurement devices A and B, whose measurement directions $\boldsymbol{a}$ and $\boldsymbol{b}$ are randomly chosen among two of three possible settings ($\boldsymbol{a} = 1, 2$; $\boldsymbol{b} = 2, 3$) while the photons are on their flight. A photon either passes the polariser (and is detected) or is absorbed by it (and is not detected), so that at each measuring device there are two possible measurement outcomes $\boldsymbol{\alpha} = \pm$ and $\boldsymbol{\beta} = \pm$.

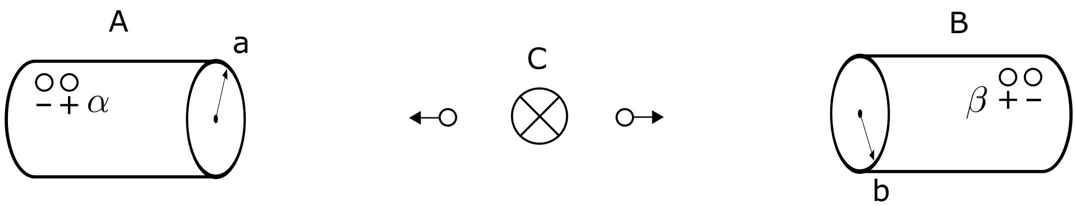

Fig. 1: EPR/B setup

On a probabilistic level, the experiment is described by the joint probability distribution $P(\alpha\beta ab\lambda) := P(\boldsymbol{\alpha} = \alpha, \boldsymbol{\beta} = \beta, \boldsymbol{a} = a, \boldsymbol{b} = b, \boldsymbol{\lambda} = \lambda)$ of the five random variables just defined.[2] We shall consistently use bold symbols $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{a}, \dots)$ for random variables and normal font symbols $(\alpha, \beta, a, \dots)$ for the corresponding values of these variables. We use indices to refer to *specific* values of variables, e.g. $\alpha_- = -$ or $a_1 = 1$, which provides useful shorthands, e.g. $P(\alpha_-\beta_+a_1b_2\lambda) := P(\boldsymbol{\alpha} = -, \boldsymbol{\beta} = +, \boldsymbol{a} = 1, \boldsymbol{b} = 2, \boldsymbol{\lambda} = \lambda)$. Expressions including probabilities with non-specific values of variables, e.g. $P(\alpha|a) = P(\alpha)$, are meant to hold for all values of these variables (if not otherwise stated).

Containing the hidden states $\boldsymbol{\lambda}$, which are by definition not measurable, the total distribution is empirically not accessible ('hidden level'), i.e. purely theoretical. Only the

[2] While the outcomes and settings are discrete variables, the hidden state may be continuous or discrete. In the following I assume λ to be discrete, but all considerations can be generalized to the continuous case.

marginal distribution which does not involve $\boldsymbol{\lambda}$, $P(\alpha\beta ab)$, *is* empirically accessible and is determined by the results of actual measurements in EPR/B experiments ('observable level').[3]

Although the EPR/B setup is constructed in order to weaken and minimize correlations between the involved variables,[4] a statistical evaluation of a series of many runs with similar preparation procedures yields that there *are* observable correlations: the outcomes are correlated given the parameters,[5]

$$P(\alpha\beta|ab) = P(\alpha|\beta ab)P(\beta) = \begin{cases} \cos^2\phi_{ab}\cdot\frac{1}{2} & \text{if } \alpha=\beta \\ \sin^2\phi_{ab}\cdot\frac{1}{2} & \text{if } \alpha\neq\beta \end{cases} \tag{Corr}$$

(where ϕ_{ab} is the angle between the measurement directions a and b). These famous EPR/B correlations between space-like separated measurement outcomes have first been measured by Aspect et al. (1982) and are correctly predicted by quantum mechanics.

3. The standard Bell argument

Since according to (Corr), one outcome depends on both the other space-like separated outcome as well as on the distant (and local) parameter, the observable part of the probability distribution is clearly non-local. Bell (1964) could show that EPR/B correlations are so extraordinary that even if one allows for hidden states $\boldsymbol{\lambda}$ one cannot restore locality: given EPR/B correlations the *theoretical* probability distribution (including possible hidden states) must be non-local as well. Hence, *any possible* probability distribution which might correctly describe the experiment must be non-local.

This 'Bell argument for quantum non-locality', as I shall call it, runs as follows. Bell realised that EPR/B correlations have the remarkable feature to violate Bell inequalities. Since Bell then did not know that suitable measurements indeed yield the correlations, the violation was merely hypothetical, but today the violation of Bell inequalities is an empirically confirmed fact. It follows that at least one of the assumptions in the derivation of the inequalities must be false. Indeterministic generalizations (Bell, 1971; Clauser and Horne, 1974; Bell, 1975) of Bell's original deterministic derivation (1964)

[3] The (theoretical) transition from the total probability distribution to the observable marginal distribution is given by a marginalisation over λ, $P(\alpha\beta ab) = \sum_\lambda P(\alpha\beta ab\lambda)$. In order to be empirically adequate, any theoretical distribution must in this way yield the distribution which describes the statistics of EPR/B measurements.

[4] First, the settings are random and statistically independent. Second, the parameters are set *after* the emission, so that the setting may not influence the state of the particles at the emission. And, finally, but most importantly, the wings of the experiment are space-like separated, so that according to the First Signal Principle of relativity there cannot be any influence from one outcome to the other or from one setting to the outcome on the other wing.

[5] A correlation of the outcomes means that the joint probability $P(\alpha\beta|ab)$ is in general not equal to the product $P(\alpha|ab)P(\beta|ab) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$.

employ two probabilistic assumptions, 'local factorisation'[6]

$$P(\alpha\beta|ab\lambda) = P(\alpha|a\lambda)P(\beta|b\lambda) \tag{ℓF}$$

and 'autonomy'

$$P(\lambda|ab) = P(\lambda). \tag{A}$$

Another type of derivation (Wigner, 1970; van Fraassen, 1989; Graßhoff et al., 2005) additionally requires the empirical fact that there are perfect correlations (PCorr) between the outcomes if the measurement settings are equal. For both types of derivation we have the dilemma that any empirically correct probability distribution of the quantum realm must either violate autonomy or local factorisation (or both). Since giving up autonomy seems to be ad hoc and implausible ('cosmic conspiracy'), most philosophers conclude that the empirical violation of Bell inequalities implies that local factorisation fails.[7] And since local factorisation states the factorisation of the hidden joint probability distribution into *local* terms, the *failure* of local factorisation indicates a certain kind of *non*-locality, which is specific to the quantum realm—hence 'quantum non-locality'.

In order to make the logical structure clear let me note the Bell argument in an explicit logical form (where (I1), (I2), ... indicate intermediate conclusions). Here and in the following I shall use the Wigner-type derivation of Bell inequalities because, as we will see, it is the most powerful one allowing to derive Bell inequalities from the widest range of probability distributions:

(P1) There are EPR/B correlations: (Corr)

(P2) EPR/B correlations violate Bell inequalities: (Corr) $\rightarrow \neg$(BI)

(I1) Bell inequalities are violated: $\neg$(BI) (from P1 & P2, MP)

(P3) EPR/B correlations include perfect correlations: (Corr) $\rightarrow$ (PCorr)

(I2) There are perfect correlations: (PCorr) (from P1 & P3, MP)

(P4) Bell inequalities can be derived from autonomy, perfect correlations and local factorisation: (A) $\wedge$ (PCorr) $\wedge$ (ℓF) $\rightarrow$ (BI)

(I3) Autonomy or local factorisation has to fail: $\neg$(A) $\vee \neg$(ℓF) (from I1 & I2 & P4, MT)

[6] 'Local factorisation' is my term. Bell calls (ℓF) 'local causality', some call it 'Bell-locality', but most often it is simply called 'factorisation' or 'factorisability'. Bell's terminology already suggests a metaphysical interpretation, which I would like to avoid in this paper, and the latter two names are too general since, as I shall show, there are other forms of the hidden joint probability which can be said to factorise; hence '*local* factorisation'.

[7] Though not a majority view, there are suggestions to explain the violation of Bell inequalities by a violation of autonomy (e.g. Price, 1994; Szabó, 2000; San Pedro, 2012). Our analysis in this paper does not apply to these cases. We shall consistently assume that autonomy holds.

(P5) Autonomy holds: (A)

(C1) Local factorisation fails: $\neg(\ell\mathrm{F})$ (from I3 & P5)

(P6) Quantum non-locality is the failure of local factorisation:
$(\mathrm{QNL}) :\leftrightarrow \neg(\ell\mathrm{F})$ (definition)

It is obvious that the argument from (P1)–(P5) to (C1) is valid. It shows that if autonomy holds, EPR/B correlations mathematically imply a non-locality which is called quantum non-locality, (P6). (P6) is *not* a premise of the Bell argument but labels its result with an appropriate name; it determines what *quantum non-locality according to the standard view* amounts to an a probabilistic level.[8] It is clear that if the Bell argument could be modified to have a stronger conclusion, the definition (P6) would have to be adapted. What we call 'quantum non-locality' depends on the result of the Bell argument. Note that defining quantum non-locality as the conclusion of the Bell argument, the logical structure of the argument is such that quantum non-locality only provides *necessary conditions* for EPR/B correlations, i.e. for being empirically adequate. So we have to keep in mind that the analysis of quantum non-locality is *not* an analysis of EPR/B correlations but of a necessary condition for them.

The core idea of my critique concerning the standard view of quantum non-locality is that the result of the Bell argument is weaker as it could be. I do not say that the argument is invalid nor do I say that one of its premises is not sound, I just say that the argument can be made considerably stronger and that the stronger conclusion will allow us to define a tighter, more informative concept of quantum non-locality: one can be much more precise about what EPR/B correlations imply (if we assume that autonomy holds) than just saying that local factorisation has to fail. I shall show that besides the local classes EPR/B correlations also exclude certain non-local classes. Given this new result, the standard definition of quantum non-locality (P6) will turn out to be inappropriately weak, because it includes those non-local classes which can be shown to be forbidden.

Specifically, I shall show that it is premise (P4) which can be made stronger (while leaving the other premises basically at work). Being an implication from autonomy, perfect correlations and local factorisation to Bell inequalities, it is clear that we can make (P4) the stronger the weaker we can formulate the antecedens, i.e. the assumptions to derive the inequalities. This idea is not essentially a new one. Since Bell's original proof (1964) considerable efforts have been made to find derivations with weaker and weaker assumptions. For example, one of the milestones was to show that one can derive Bell

[8] We should note that the term 'quantum non-locality' is not unambiguous. According to the definition we use here, (P6), it refers to probabilistic facts—not to causal or metaphysical ones. Moreover, we do not mean the *specific* non-locality inherent in quantum mechanics (or the true theory of the quantum world, whatever it may be). It is clear that one could be more precise about this latter non-locality than saying that it violates local factorisation. Rather, by 'quantum non-locality' here and in the following we denote the general probabilistic fact, which follows from the violation of Bell inequalities and which characterises *any* viable theory of the quantum realm. Given the standard Bell argument this is just the failure of local factorisation.

inequalities without the original assumption of determinism. Currently, autonomy and local factorisation seem to constitute the weakest set of probabilistic assumptions which allow a derivation. What will be new about my approach is to try to find *alternatives to local factorisation*, which (given autonomy and perfect correlations) also imply that Bell inequalities hold. Since local factorization is the weakest possible form of local distributions, it is clear that such alternatives have to involve a kind of *non-locality*, i.e. what I am trying to show in the following is that we can derive Bell inequalities from certain non-local probability distributions.

4. Classes of probability distributions

We can find potential alternatives to local factorisation if we consider what it is: a particular feature of the *hidden joint probability*, as I shall call $P(\alpha\beta|ab\lambda)$. According to the product rule of probability theory, for *any* of the possible hidden probability distributions the joint probability of the outcomes (given the other variables) can be written as a product,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta ba\lambda)P(\beta|ab\lambda) \tag{1}$$
$$= P(\beta|\alpha ab\lambda)P(\alpha|ba\lambda). \tag{2}$$

Since there are two product forms, one whose first factor is a conditional probability of $\boldsymbol{\alpha}$ and one whose first factor is a conditional probability of $\boldsymbol{\beta}$, for the time being, let us restrict our considerations to the product form (1), until at the end of this section I shall generalize the results to the other form (2).

The product form (1) of the hidden joint probability holds in general, i.e. for all probability distributions. According to probability distributions with appropriate independences, however, the factors on the right-hand side of the equation reduce in that certain variables in the conditionals can be left out. If, for instance, outcome independence holds, $\boldsymbol{\beta}$ can disappear from the first factor, and the joint probability is said to 'factorise'. *Local* factorisation further requires that the distant parameters in both factors disappear, i.e. that parameter independence holds. Prima facie, any combination of variables in the two conditionals in (1) seems to constitute a distinct product form of the hidden joint probability. Restricting ourselves to *irreducibly hidden* joint probabilities, i.e. requiring $\boldsymbol{\lambda}$ to appear in both factors, there are $2^5 = 32$ *combinatorially* possible forms (for any of the three variables in the first conditional and any of the two variables in the second conditional *besides* $\boldsymbol{\lambda}$ can or cannot appear). Table 1 shows these conceivable forms which I label by (H_1^α) to (H_{32}^α) (the superscript α is due to the fact that we have used (1) instead of (2)).

The specific product form of the hidden joint probability is *the essential feature* of the probability distributions of EPR/B experiments. For, as we shall see, it not only determines whether a probability distribution can violate Bell inequalites but also carries unambiguous information about which variables of the experiment are probabilistically independent of another. Virtually any interesting philosophical question involving prob-

Table 1: Classes of probability distributions

	I		II	III	IV			V	VI		VII	VIII	IX
	(H_i^α):	$P(\alpha\beta\vert ab\lambda) = \dots$											
	i	$P(\alpha\vert$	β	b	a	$\lambda)$	$\cdot P(\beta\vert$	a	b	$\lambda)$	$\Box$(BI)	[Group]	Notes
strong non-locality$^\alpha$	1		1	1	1			1	1		0	iv	
	2		1	1	1			1	0		0	iv	
	3		1	1	1			0	1		0	iv	$\mathrm{QM}_{\mathrm{PE}}$
	4		1	1	0			1	1		0	iv	
	5		1	0	1			1	1		0	iv	
	6		0	1	1			1	1		0	iv	Bohm
	7		1	1	1			0	0		0	iv	$\mathrm{QM}_{\mathrm{ME}}$
	8		0	1	1			1	0		0	iv	
	9		0	1	1			0	1		0	iv	
	10		1	0	0			1	1		0	iv	
	11		0	1	0			1	1		0	iv	
	12		0	0	1			1	1		0	iv	
	13		0	1	1			0	0		0	iv	
	14		0	0	0			1	1		0	iv	
weak non-locality$^\alpha$	15		1	1	0			1	0		1	iii	
	16		1	0	1			0	1		1	iii	
	17		1	0	1			1	0		1	i	
	18		1	1	0			0	1		1	i	
	19		1	1	0			0	0		1	i	
	20		1	0	1			0	0		1	i	
	21		1	0	0			1	0		1	i	
	22		0	1	0			1	0		1	ii	
	23		0	0	1			1	0		1	i	
	24		1	0	0			0	1		1	i	
	25		0	1	0			0	1		1	i	
	26		1	0	0			0	0		1	i	
	27		0	1	0			0	0		1	i	
	28		0	0	0			1	0		1	i	
locality$^\alpha$	29		0	0	1			0	1		1	ii	local fact.
	30		0	0	1			0	0		1	i	
	31		0	0	0			0	1		1	i	
	32		0	0	0			0	0		1	i	

abilistic facts of EPR/B experiments depends on the specific product form of the hidden joint probability. Hence, it is natural to use the product form of the hidden joint probability in order to classify the probability distributions. We can say that each product form of the hidden joint probability constitutes a *class of probability distributions* in the sense that probability distributions with the same form (but different numerical weights of the factors) belong to the same class. In order to make the assignment of probability distributions to classes unambiguous let us require that each probability distribution belongs only to that class which corresponds to its *simplest* product form, i.e. to the form with the minimal number of variables appearing in the conditionals (according to the distribution in question).

This scheme of classes is comprehensive: Any probability distribution of the EPR/B experiment must belong to one of these 32 classes. In this systematic overview, the class constituted by local factorisation is (H^{α}_{29}) (see table 1, column IX), and if we allow that there might be no hidden states $\boldsymbol{\lambda}$, we can assign the quantum mechanical distribution to class (H^{α}_{7}) (for *maximally* entangled quantum states, noted by '$\mathrm{QM}_{\mathrm{ME}}$') or to ($\mathrm{H}^{\alpha}_{3}$), respectively (for *partially* entangled states, noted by '$\mathrm{QM}_{\mathrm{PE}}$').[9] The de-Broglie-Bohm theory falls under class (H^{α}_{6}), and similarly any other theory of the quantum realm has its unique place in one of the classes. The advantage of this classification is that it simplifies matters insofar we can now derive features of *classes* of probability distributions and can be sure that these features hold for all members of a class, i.e. for all theories whose probability distributions fall under the class in question.

The feature that we are most interested in is, of course, which of these classes (given autonomy) imply that Bell inequalities hold. We provide the answer by the following theorem:

> **Theorem 1**: Given autonomy, perfect correlations and perfect anti-correlations, a hidden joint probability implies Bell inequalities if each of the two factors in its product form involves at most one parameter.

Before we shall comment on the theorem, let us say that its proof can be found in the mathematical appendix. Although the proof is the core of our argument in this paper, nothing in what follows depends on understanding its details. We should remark that the perfect (anti-)correlations are required for letting a maximum of product forms imply Bell inequalities.

The result of theorem 1 is remarkable. So far it has been believed that *local* product forms imply Bell inequalities, but the theorem does not refer to this characteristic at all. It just requires that an outcome may not depend on both parameters. What does this mean? Which classes imply Bell inequalities according to the theorem? In order to make this clear let us partition the classes into three groups, depending on which

[9]The typical case for EPR/B experiments is to prepare a maximally entangled quantum state (e.g. $|\psi\rangle = \sqrt{p}|+\rangle|+\rangle + \sqrt{1-p}|-\rangle|-\rangle$ with $p = \frac{1}{2}$), because one wants to have a maximal violation of the Bell inequalities. The slightest deviation from maximal entanglement ($p \neq \frac{1}{2}$), however, breaks the symmetry of the state. The probability distribution of such *partially* entangled states shows a dependence on the local setting in the second factor; they fall in class (H^{α}_{3}).

variables appear in their constituting product forms:[10]

Local$^\alpha$ classes: (H^α_{29})–(H^α_{32}) (imply Bell inequalities)
each factor only contains time-like (or light-like) separated variables

Weakly non-local$^\alpha$ classes: (H^α_{15})–(H^α_{28}) (imply Bell inequalities)
at least one of the factors involves space-like separated variables, but none of the factors involves both parameters

Strongly non-local$^\alpha$ classes: (H^α_{1})–(H^α_{14}) (do not imply Bell inequalities)
at least one of the factors involves both parameters

(Note the superscript α, which indicates that we refer to classes deriving from (1) instead of from (2)).

Local classes involve only time-like (or light-like) separated variables in the factors of their hidden joint probability. When a product form is local, none of the outcomes depends on its distant parameter, so none of the outcomes depends on both parameters—hence, according to theorem 1, Bell inequalities follow. This is a well-known fact.

The surprising consequence of theorem 1 is that *even certain non-local classes imply Bell inequalities*. We have called those non-local classes, which imply Bell inequalities, '*weakly* non-local$^\alpha$' (in contrast to *strongly* non-local$^\alpha$ ones, which do not). According to weakly non-local$^\alpha$ classes, at least one of the factors in the product form involves space-like separated variables, but none of the factors involves both parameters. So weakly non-local$^\alpha$ classes do fulfil the criterion of theorem 1, which means that they imply Bell inequalities, although they involve non-local dependences. For instance, the outcomes may depend on their distant parameters if they do not depend on their local one,

$$P(\alpha\beta|ab\lambda) = P(\alpha|b\lambda)P(\beta|a\lambda). \tag{H^α_{22}}$$

Or an outcome may depend on the other outcome,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda), \tag{H^α_{16}}$$

or both of these non-localities may occur,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta b\lambda)P(\beta|a\lambda). \tag{H^α_{15}}$$

Our proof of theorem 1 shows that all these product forms imply Bell inequalities, because none of its factors involves both parameters. From the perspective of the standard view this result is surprising because here we have cases where (as we shall show in the second part) parameter dependence or outcome dependence or both hold, and still Bell inequalities are implied. We emphasise that the proof of theorem 1, which shows that besides local theories also weakly non-local$^\alpha$ ones imply Bell inequalities, is the source of all new consequences we shall derive in this paper.

[10] Column VIII in table 1 introduces an even finer partition of the classes, which, however, is only relevant for the proof of theorem 1. We shall not comment on it in the main text.

As soon as there is a dependence on *both* parameters in at least one of the factors of the product form, which is how we have defined strongly non-local$^\alpha$ classes, one can*not* derive Bell inequalities from the product form any more. For each of these product forms one can easily find examples of probability distributions which do *violate* Bell inequalities. Thus, strongly non-local$^\alpha$ classes do *not* imply Bell inequalities, because *some* distributions in each of the classes violate them. This is, however, not to say that probability distributions in those classes necessarily violate Bell inequalities. On the contrary, one can as well find probability distributions of that form, which *obey* Bell inequalities. So depending on both parameters in one of the factors is only a *necessary* condition for violating Bell inequalities; it is not a sufficient one. Sufficient criteria to violate Bell inequalities would have to involve conditions for the *strength* of the correlations. As we have said in the introduction, mutual information between two variables is a measure for how strong the correlation between them is, so the information theoretic works which derive numerical values for how much mutual information has to be given in order to violate Bell inequalities, provide an answer to that question (see Maudlin 1994, ch. 6 and Pawlowski et al. 2010).

To sum it up: as opposed to what the standard discussion suggests, it is *not* true that local factorisation (and the other local product forms) are the *only* product forms which allow deriving Bell inequalities. Rather, we have found that 18 of the 32 logically possible classes imply Bell inequalities if autonomy (and perfect (anti-)correlations) hold, among them 14 *non-local* classes. Column VII of table 1 indicates which classes imply Bell inequalities and which do not ('$\Box$(BI)' means *necessarily, Bell inequalities hold*).

5. A stronger Bell argument

This consequence from theorem 1, that one can derive Bell inequalities also from certain non-local product forms, enables us to strengthen premise (P4) in the Bell argument. We can now write:

(P4′) Bell inequalities can be derived from autonomy, perfect correlations, perfect anti-correlations and any local$^\alpha$ or weakly non-local$^\alpha$ class of probability distributions:

$$\left[(\mathrm{A}) \wedge (\mathrm{PCorr}) \wedge (\mathrm{PACorr}) \wedge \left(\bigvee_{i=15}^{32} (\mathrm{H}_i^\alpha) \right) \right] \rightarrow (\mathrm{BI})$$

Compared to (P4), we have made two changes. First, we have replaced local factorisation in the antecedent by the *disjunction* of the local$^\alpha$ and weakly non-local$^\alpha$ classes (including local factorisation (H_{29}^α)). This makes the antecedent of (P4′) weaker than that in (P4) and, hence, the argument stronger. Second, we have added the condition that there are perfect anti-correlations (PACorr), since certain weakly non-local$^\alpha$ classes require them for the derivation. This additional assumption however does not weaken the argument since the perfect anti-correlations follow from the EPR/B correlations (as the perfect correlations do). We just have to modify premise (P3) to

(P3′) EPR/B correlations include perfect correlations and perfect anti-correlations: (Corr) $\rightarrow$ (PCorr) $\wedge$ (PACorr)

Changing these two premises has a considerable effect on the Bell argument. Instead of the standard conclusion (C1), that the violation implies the failure of local factorisation, by the modified argument from (P1), (P2), (P3′), (P4′) and (P5), we arrive at the essentially stronger conclusion:

(C1′) Both local$^{\alpha}$ and weakly non-local$^{\alpha}$ classes fail:

$$\Big(\bigwedge_{i=15}^{32} \neg(\mathrm{H}_i^{\alpha})\Big)$$

While the original result, the failure of local factorisation, implied that all local$^{\alpha}$ classes fail (because the other local classes are specializations of local factorisation), the new result additionally excludes all weakly non-local$^{\alpha}$ classes.

Our considerations leading to this new result of the Bell argument rest on the fact that we have found alternatives to local factorisation from writing the hidden joint probability according to the product rule (1) and conceiving different possible product forms (table 1). However, we can as well write the hidden joint probability according to the second product rule (2), and similar arguments as above lead us to a similar table as table 1, whose classes, (H_1^{β})–$(\mathrm{H}_{32}^{\beta})$, differ to those in table 1 in that the outcomes and the parameters are swapped. For instance, class $(\mathrm{H}_{16}^{\beta})$ is defined by the product form $P(\alpha\beta|ab\lambda) = P(\beta|\alpha b\lambda)P(\alpha|a\lambda)$ in contrast to $(\mathrm{H}_{16}^{\alpha})$, which is constituted by $P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda)$. Note that this new classification is a *different partition* of the possible probability distributions. Any probability distribution must fall in exactly one of the classes (H_1^{α})–$(\mathrm{H}_{32}^{\alpha})$ *and* in exactly one of the classes (H_1^{β})–$(\mathrm{H}_{32}^{\beta})$. Analogously to theorem 1 one can proof for the new partition that the local$^{\beta}$ and weakly non-local$^{\beta}$ classes imply Bell inequalities as well, so that we can reformulate (P4′) as:

(P4″) Bell inequalities can be derived from autonomy, perfect correlations, perfect anti-correlations and any local$^{\alpha}$, weakly non-local$^{\alpha}$, local$^{\beta}$ or weakly non-local$^{\beta}$ class of probability distributions:

$$\Big[(\mathrm{A}) \wedge (\mathrm{PCorr}) \wedge (\mathrm{PACorr}) \wedge \Big(\bigvee_{i=15}^{32} (\mathrm{H}_i^{\alpha}) \vee \bigvee_{i=15}^{32} (\mathrm{H}_i^{\beta})\Big)\Big] \rightarrow (\mathrm{BI})$$

Then we can formulate an even stronger Bell argument from (P1), (P2), (P3′), (P4″) and (P5) to

(C1″) All local$^{\alpha}$, weakly non-local$^{\alpha}$, local$^{\beta}$ and weakly non-local$^{\beta}$ classes fail:

$$\Big(\bigwedge_{i=15}^{32} \neg(\mathrm{H}_i^{\alpha}) \wedge \bigwedge_{i=15}^{32} \neg(\mathrm{H}_i^{\beta})\Big)$$

This is the conclusion of the new stronger Bell argument. It takes the usual result from *any* kind of non-locality (the mere failure of local factorisation) to a more *specific* one (namely exclusive the weakly non-local$^\alpha$ and weakly non-local$^\beta$ classes). Stating which classes are excluded, the result formulated here is a negative one. But it is easy to turn it into a positive formulation: since our scheme of logically possible classes is comprehensive, the failure of all local$^\alpha$ and weakly non-local$^\alpha$ classes is equivalent to the fact that one of the strongly non-local$^\alpha$ classes, (H_1^α)–(H_{14}^α), holds. Analogously, if a probability distribution is neither local$^\beta$ nor weakly non-local$^\beta$ it must be strongly non-local$^\beta$, i.e. belong to one of the classes (H_1^β)–(H_{14}^β). Therefore, equivalently to (C1″) we can say:

(C1‴) One of the classes strong non-locality$^\alpha$ and one of the classes strong non-locality$^\beta$ has to hold.

$$\left(\bigvee_{i=1}^{14}(\mathrm{H}_i^\alpha) \wedge \bigvee_{i=1}^{14}(\mathrm{H}_i^\beta)\right)$$

This is the positive conclusion of the stronger Bell argument in terms of classes.

Finally, we can formulate the same result in terms of which features the hidden joint probability must have. Let us define the following concept:

> **Probabilistic Bell Contextuality (PBC)** holds if and only if according to each product form of the hidden joint probability $P(\alpha\beta|ab\lambda)$ at least one of the outcomes depends probabilistically on both settings.

Then, equivalently to (C1″) or (C1‴), we can say:

(C1⁗) Probabilistic Bell Contextuality holds.

(C1″), (C1‴) and (C1⁗) are equivalent conclusions of the stronger Bell argument.

Since in section 3 we have seen that the result of the original Bell argument defines what we call quantum non-locality, it is clear that given these new results we have to adapt the definition appropriately. It simply becomes implausible to stick to the old, looser definition *including* weakly non-local classes, given that we now know that we have to *exclude* them. Quantum non-locality cannot any more be regarded as the failure of local factorisation. Rather, with the new result of the Bell argument we can be much more precise about what it amounts to on a probabilistic level and re-define it as:

(P6′) Quantum non-locality is Probabilistic Bell Contextuality.

Of course, using (C1″) or (C1‴) instead of (C1⁗) one can as well have equivalent formulations in terms of classes.

According to the logic of the Bell argument, we have noted in section 3, quantum non-locality is a *necessary* condition for EPR/B correlations and their violation of Bell inequalities. Since the logical structure of the argument has not essentially changed, this is still true for the new result. Furthermore, we now know that the new concept of quantum non-locality is *not sufficient* for EPR/B correlations because we have seen that there are strongly non-local distributions which do *not* violate Bell inequalities. On the other hand, since we have also found that all strongly non-local classes include distributions which violate Bell inequalities and reproduce EPR/B correlations, we can say that on a qualitative level, which only considers the product forms of the hidden joint probability, i.e. probabilistic dependences and independences, we cannot improve the argument any more. It is impossible to reach a stronger conclusion than (C1″) by showing that we can derive Bell inequalities from still further classes. For if my argument is correct, there are no classes left which in general may imply Bell inequalities. Any future characterisation of quantum non-locality which is more detailed must involve reference to the *strengths* of the correlations in the strongly non-local classes.[11] In this sense, my new concept of quantum non-locality, although not being sufficient for EPR/B correlations, captures their *strongest possible consequences on a qualitative probabilistic level.*

6. Discussion

Why has this stronger consequence of the Bell argument, that we have derived, been overlooked so far? Obviously, it has wrongly been assumed that local factorisation is the *only* basis to derive Bell inequalities, and the main reason for neglecting other product forms of hidden joint probabilities might have been the fact that, originally, Bell inequalities were derived to capture consequences of a *local* worldview. The question that shaped Bell's original work clearly was Einstein's search for a local hidden variable theory and his main result was that it is impossible: locality has consequences which are in conflict with the quantum mechanical distribution—one cannot have a local hidden variable theory which yields the same predictions as quantum mechanics. Given this historical background, the idea to derive Bell inequalities from non-local assumptions maybe was beyond interest because the conflict with locality was considered to be the crucial point; or maybe it was neglected because Bell inequalities were so tightly associated with locality that a derivation from non-locality sounded totally implausible. Systematically, however, since it is now clear that the quantum mechanical distribution is empirically correct and Bell inequalities are violated, it is desirable to draw as strong consequences as possible, which requires to check without prejudice whether some non-local classes allow a derivation of Bell inequalities as well—and this is what we have done here.

My argument resolves two common misunderstandings in the debate about quantum non-locality. First, in the discussion Bell inequalities are so closely linked to locality that one could have the impression that Bell inequalities are locality conditions in the sense that, if a probability distribution obeys a Bell inequality, it must be local. Of

[11] Maudlin (1994, ch. 6) and Pawlowski et al. (2010) characterise quantum non-locality in terms of information. Since Shannon mutual information is a measure for the strength of correlations between variables, these approaches provide a more detailed analysis in the sense described.

course, Bell's argument never really justified that view, for the logic of the standard Bell argument is that local factorisation (given autonomy and perfect (anti-)correlations) is merely *sufficient* (and not necessary) for Bell inequalities. Maybe the association between Bell inequalities and locality might have arisen from the the fact that up to now local factorisation has been the *only* product form which was shown to imply Bell inequalities. Given only this information, it was at least possible (though unproven) that the holding of Bell inequalities implies locality. However, since we have shown that weakly non-local classes in general imply Bell inequalities and since the simulations show that even some strongly non-local distributions can conform to Bell inequalities, it has become explicit that this is not true. *Not all* probability distributions obeying Bell inequalities are *local*. So if a probability distribution obeys a Bell inequality it does not have to be local.

Second, sometimes it has been said that the violation of Bell inequalities by EPR/B correlations implies that there *cannot* be a screener-off for the correlations (e.g. van Fraassen, 1989). Table 1 shows that this is not true either. Among the strongly non-local classes there are classes according to which $\boldsymbol{\alpha}$ is screened-off from $\boldsymbol{\beta}$ (namely all classes with the value 0 in column II). Consider for example class (H_6^{α}) for which the product form reads $P(\alpha\beta|ab\lambda) = P(\alpha|ab\lambda)P(\beta|ab\lambda)$. Here we have it that the conjunction of $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{\lambda}$ screens the correlation off. Including the distant parameter for both outcomes, this is of course *not* a *local* screener-off. But that there can be a screener-off for the correlations, although non-local, shows that it is not true saying that EPR/B correlations disprove Reichenbach's Common Cause Principle (or the Causal Markov Condition). This claim would only be true if one excludes non-local screener-offs by adding the premise of locality, which, however, would be odd, because the violation of Bell inequalities shows that we cannot avoid a non-locality anyway. Of course, the true theory of the quantum realm might *in fact* have a distribution which does not screen-off, but to argue for that one needs further assumptions; this claim is not in general implied by the violation of Bell inequalities.

Most importantly, my argument points out that the discussion about the consequences of Bell's theorem so far has been based on results which—though not wrong—are weaker than they could be. Capturing all non-local classes the standard concept of quantum non-locality (the failure of local factorisation) includes classes which we have found to be compatible with Bell inequalities (weakly non-local classes). In this sense *the standard concept is inappropriately weak*. This is not to say that the discussion about Bell's theorem and its consequences has to be rewritten completely. The main result of Bell's theorem is that there must be a non-locality of some kind and this insight is well captured by the standard concept. All results which are based on the mere existence of a non-locality are save. However, if one explores deeper into the nature of this non-locality it might matter, whether one assumes the standard concept or starts from our stronger result. One example might be a probabilistic analysis of quantum non-locality in terms of pairwise independences. Jarrett (1984) analysed the standard concept of quantum non-locality, the failure of local factorisation, as what today usually is called 'outcome dependence or parameter dependence', which triggered a large discussion about which of the two dependences are realised, which of them is compatible with relativity, what each means in causal terms, etc. Given our new argument, however, we see that Jarrett's analysandum is too weak. It is to be expected that a similar analysis of our new stronger concept yields different and more informative results.

A precise analysis of the new concept of quantum non-locality will be work for future investigations. However, without providing a detailed analysis here, we can already say that Jarrett's result 'outcome dependence or parameter dependence' is at least highly misleading. According to our stronger result one cannot avoid that an outcome in some sense depends on its distant parameter. So contrary to what Jarrett's result suggests, one cannot avoid a dependence on the distant parameter even if the outcomes depend on another. Quantum non-locality necessarily involves a dependence on both parameters. We should stress, however, that this is not to say that parameter dependence in the usual sense has to hold. Probabilistic dependences between the same variables differ if they hold conditional on different sets of variables. Parameter dependence as it is usually defined in the debate is the dependence of an outcome on its distant parameter given a very specific set of other variables (namely given the local outcome and the hidden state). The parameter dependence which is required for quantum non-locality need not be conditional on that set of variables. In fact there are two ways in which a product form can depend on both parameters: either in the first or either in the second factor. A detailed analysis reveals that only a dependence on the distant parameter in the second factor implies parameter dependence in the usual sense. A dependence in the first factor implies a different kind of parameter dependence, which additionally involves the other outcome in the set of conditional variables—and this is a totally different kind of parameter dependence.

Finally, we should note that there is another approach to quantum non-locality whose results seem to converge with ours. Maudlin (1994, ch. 6) examines the quantum non-locality not via Bell's theorem, but directly investigates the EPR/B correlations by information theoretic methods. He proves that at least one of the outcomes must depend on *information* about both settings. Since (Shannon mutual) information implies correlation, it seems that Maudlin's claim is in accordance with our results. On the one hand this seems to confirm our findings. Two different methods yielding the same results are good evidence for the stability of a claim. On the other hand, at this point of the analysis we can only say that the two approaches agree in the rough sense we have stated. Whether they do agree in their details would require further investigation. The crucial point is the one just mentioned that a precise claim about a dependence of an outcome on its distant parameter would have to state conditional on which other variables the dependence holds. Since the product forms of the hidden joint probability are explicit probabilistic statements one can make these facts precise from my results. It is not clear, however, if one can make them precise from Maudlin's. Having the exact conditional dependences might, for instance, be important for causal inference, which is very sensitive to the exact probabilistic dependences that are present in a given situation.

To sum it up: in this paper we have considerably strengthened Bell's theorem. We have given a comprehensive overview of the possible types of probability distributions, which can describe EPR/B experiments (table 1). The overview has revealed that one can derive Bell inequalities not only from local theories but also from a large range of non-local ones, which we have called weakly non-local. This has enabled us to formulate a stronger Bell argument than usual to exclude local *and* weakly non-local theories of the quantum world. The argument has three immediate consequences. First, the argument makes explicit that Bell inequalities are *not* locality conditions because it shows that weakly non-local theories imply the inequalities. Second, the overview of the possible probability distributions makes explicit that the violation of the inequalities does *not* necessarily imply that there is no screener-off for the correlations. There are distributions according to which there are non-local screener-offs. Third and most importantly, since the result of the Bell argument defines what we appropriately call quantum non-locality, this new result yielded a tighter, more informative concept of quantum non-locality, which describes more precisely what the violation of Bell inequalities implies on a probabilistic level. In fact, my new concept of quantum non-locality—although not being sufficient for EPR/B correlations—captures the strongest possible consequences of EPR/B correlations on a *qualitative* probabilistic level. In this sense, the standard concept of quantum non-locality, the failure of local factorisation, is too weak, while the new concept is an appropriate basis for future explorations of the detailed consequences of Bell's theorem.

A. Appendix: Proof of theorem 1

In order to prove theorem 1 it will provide useful to partition the classes into four groups depending on which variables appear in their constituting product forms (see table 1, column VIII):

(i) At least one of the parameters does not appear at all: (H^α_{17})–(H^α_{21}), (H^α_{23})–(H^α_{28}), (H^α_{30})–(H^α_{32})

(ii) Both parameters appear separately, one in each factor: (H^α_{22}), (H^α_{29})

(iii) As (ii) but the first factor additionally involves the outcome $\boldsymbol{\beta}$: (H^α_{15}), (H^α_{16})

(iv) Both parameters appear together in at least one of the factors: (H^α_{1})–(H^α_{14})

While group (iv) is identical to the group of strongly non-local$^\alpha$ classes, the groups (i)–(iii) provide a finer partition of the union of the local$^\alpha$ and weakly non-local$^\alpha$ classes. In terms of the partition (i)–(iv), confirming theorem 1 means to prove that *given autonomy and perfect (anti-)correlations, the classes belonging to groups (i), (ii) and (iii) imply Bell inequalities.* Additionally, we shall show that classes in group (iv) in general do *not* imply Bell inequalities.

A.1. Group (i)

Consider one version of the Wigner-Bell inequality (Wigner, 1970; van Fraassen, 1989),

$$P(\alpha_-\beta_+|a_1b_3) \leq P(\alpha_-\beta_+|a_1b_2) + P(\alpha_-\beta_+|a_2b_3). \tag{3}$$

We can write the probabilities in terms of the hidden probability distribution if we sum over λ,

$$P(\alpha_-\beta_+|ab) = \sum_\lambda P(\alpha_-\beta_+|ab\lambda)P(\lambda|ab), \tag{4}$$

and assuming autonomy (A), we can further rewrite it as

$$P(\alpha_-\beta_+|ab) = \sum_\lambda P(\alpha_-\beta_+|ab\lambda)P(\lambda). \tag{5}$$

It is obvious that in this form the empirical joint probability $P(\alpha_-\beta_+|ab)$ depends on the parameters only via the hidden joint probability $P(\alpha_-\beta_+|ab\lambda)$. Hence, if a certain parameter does not appear in a specific product form of the hidden joint probability (group (i)), the empirical joint probability becomes independent of this parameter. Consider, for instance, how class (H^α_{17}), the product form of which does not involve the parameter $\boldsymbol{b}$,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|a\lambda), \tag{6}$$

makes the empirical joint probability independent of $\boldsymbol{b}$:

$$P(\alpha_-\beta_+|ab) = \sum_\lambda P(\alpha_-|\beta_+a\lambda)P(\beta_+|a\lambda)P(\lambda) = P(\alpha_-\beta_+|a). \tag{7}$$

Inserting this empirical joint probability, which does not depend on $\boldsymbol{b}$, into the Bell-Wigner inequality, reveals that in this case the inequality holds trivially, just because it has lost its functional dependence on $\boldsymbol{b}$:[12]

$$P(\alpha_-\beta_+|a_1) \leq P(\alpha_-\beta_+|a_1) + P(\alpha_-\beta_+|a_2) \tag{8}$$

(H^α_{17}) implying that Bell inequalities hold is surprising because its constituting product form is both *non-local* and *non-factorising*: $\boldsymbol{\beta}$ depends on the distant parameter $\boldsymbol{a}$ in the second factor, $P(\beta|a\lambda)$, and $\boldsymbol{\alpha}$ depends on $\boldsymbol{\beta}$ in the first, $P(\alpha|\beta a\lambda)$, i.e. $\boldsymbol{\lambda}$ and $\boldsymbol{a}$ do *not* screen-off the outcomes from another. However, very similarly, we can show that all other classes in group (i) meet the requirements of Bell inequalities: no matter what kind of non-localities they involve, if at least one of the parameters does not appear in the product form, Bell inequalities hold trivially. Hence, we can conclude that if autonomy holds (which we have used to simplify the expectation value in (5)) distributions in group (i) imply that Bell inequalities hold:[13]

$$\left[(\mathrm{A}) \wedge \Big(\bigvee_{\substack{i=17-21\\23-28\\30-32}} (\mathrm{H}^\alpha_i)\Big)\right] \rightarrow (\mathrm{BI}) \tag{9}$$

A.2. Group (ii)

Let us now turn to distributions in group (ii). Since according to this group both parameters appear in the product form (one in each factor), it is clear that, contrary to group (i), here Bell inequalities do not hold just because of the functional dependences. However, local factorisation (H^α_{29}) belongs to this group and we know how we can derive Bell inequalities with this product form of the hidden joint probability. Since the derivation from the other class in this group, (H^α_{22}), is very similar, let me first sketch a derivation with local factorisation, which is based on the ideas of Wigner (1970) and van Fraassen (1989).

We proceed from the empirical fact that there are perfect correlations between the measurement outcomes if the settings equal another:

$$P(\alpha_\pm\beta_\mp|a_ib_i) = 0 \tag{10}$$

Similarly to (5), using autonomy and local factorisation we can rewrite the empirical

[12] Note that (7) even *directly* contradicts the empirical distribution (not only indirectly by making Bell inequalities true), because it states that the empirical joint probability does not depend on one of the parameters, which is wrong.

[13] The sign '$\bigvee$' denotes a multiple disjunction, e.g. $\bigvee_{i=1..n}(\mathrm{H}^\alpha_i) := (\mathrm{H}^\alpha_1) \vee (\mathrm{H}^\alpha_2) \vee \ldots \vee (\mathrm{H}^\alpha_n)$.

joint probability in terms of the hidden joint probability,

$$P(\alpha_\pm \beta_\mp | a_i b_i) = 0 = \sum_\lambda P(\lambda) P(\alpha_\pm | a_i \lambda) P(\beta_\mp | b_i \lambda). \tag{11}$$

Since probabilities are non-negative (and we assume $P(\lambda) > 0$ for all λ), at least one of the two remaining factors in each summand must be zero, i.e. for all values of i and λ we must have:

$$\begin{aligned} &\big[P(\alpha_+|a_i\lambda) = 0 \quad \vee \quad P(\beta_-|b_i\lambda) = 0\big] &(12)\\ \wedge \quad &\big[P(\alpha_-|a_i\lambda) = 0 \quad \vee \quad P(\beta_+|b_i\lambda) = 0\big] &(13)\end{aligned}$$

There are two cases. Suppose first that $P(\alpha_+|a_i\lambda) = 0$. From there all other probabilities follow as either 0 or 1:

$$\overset{\text{(CE)}}{\Rightarrow} P(\alpha_-|a_i\lambda) = 1 \qquad \overset{(13)}{\Rightarrow} P(\beta_+|b_i\lambda) = 0 \qquad \overset{\text{(CE)}}{\Rightarrow} P(\beta_-|b_i\lambda) = 1$$

Here, '(CE)' stands for 'complementary event' and refers to a theorem of probability theory that the sum of the probability of an event A and of its complementary event $\overline{A}$ is 1, e.g. $P(\alpha_+|a_i\lambda) + P(\alpha_-|a_i\lambda) = 1$.

Assume, second, that $P(\beta_-|b_i,\lambda) = 0$. Again all other probabilities are determined to be either 0 or 1:

$$\overset{\text{(CE)}}{\Rightarrow} P(\beta_+|b_i\lambda) = 1 \qquad \overset{(13)}{\Rightarrow} P(\alpha_-|a_i\lambda) = 0 \qquad \overset{\text{(CE)}}{\Rightarrow} P(\alpha_+|a_i\lambda) = 1$$

In order to avoid contradiction the two cases have to be disjunct. So given a certain measurement direction i, the two cases define a partition of the values of $\boldsymbol{\lambda}$: all values of $\boldsymbol{\lambda}$ for which $P(\alpha_+|a_i\lambda) = 0$ belong to the set $\Lambda(i)$, while all other values, for which $P(\alpha_-|a_i\lambda) = 0$, belong to $\overline{\Lambda(i)}$. Note that each value of i defines a different partition.

We can use the fact that the $\boldsymbol{\lambda}$-partitions depend on just *one* parameter i to calculate the hidden joint probability $P(\alpha\beta|ab\lambda)$ for *any* choice of measurement directions $a_i b_j$ by forming intersections of partitions for different parameters (see table 2). Since all values are either 0 or 1 we have shown that determinism holds on the hidden level.

Given table 2, i.e. determinism and the composability of the $\boldsymbol{\lambda}$-partitions (each of which depends on just one parameter), it is easy to show that Wigner-Bell inequalities must hold. Consider the inequality

$$P(X \cap \overline{Z}) \leq P(X \cap \overline{Y}) + P(Y \cap \overline{Z}), \tag{14}$$

which in general holds for any events X, Y, Z of a measurable space (the validity of the inequality is obvious if one draws a Venn diagram, see (Neapolitan and Jiang, 2006)). Assuming $X = \Lambda(1)$, $Y = \Lambda(2)$ and $Z = \Lambda(3)$ gives the inequality

$$P(\Lambda(1) \cap \overline{\Lambda(3)}) \leq P(\Lambda(1) \cap \overline{\Lambda(2)}) + P(\Lambda(2) \cap \overline{\Lambda(3)}). \tag{15}$$

We can now express the probabilities in the inequality by the empirical probability

Table 2: Values of the hidden joint probability

	$\lambda \in$ $\Lambda(i) \cap \Lambda(j)$	$\Lambda(i) \cap \overline{\Lambda(j)}$	$\overline{\Lambda(i)} \cap \Lambda(j)$	$\overline{\Lambda(i)} \cap \overline{\Lambda(j)}$
$P(\alpha_+\beta_+\|a_ib_j\lambda) =$	0	0	0	1
$P(\alpha_+\beta_-\|a_ib_j\lambda) =$	0	0	1	0
$P(\alpha_-\beta_+\|a_ib_j\lambda) =$	0	1	0	0
$P(\alpha_-\beta_-\|a_ib_j\lambda) =$	1	0	0	0

distribution if we use the hidden joint probability from table 2, e.g.:

$$\begin{aligned} P(\Lambda(1) \cap \overline{\Lambda(2)}) &\overset{(\sigma\text{-additivity})}{=} \sum_{\lambda \in \Lambda(1) \cap \overline{\Lambda(2)}} P(\lambda) = \\ &\overset{(\text{table } 2)}{=} \sum_{\lambda} P(\lambda) P(\alpha_-\beta_+|a_1b_2\lambda) = \\ &\overset{(\text{A})}{=} P(\alpha_-\beta_+|a_1b_2) \end{aligned} \tag{16}$$

The resulting inequality is the Wigner-Bell inequality (3).

This derivation reminds us how local factorisation together with autonomy and perfect correlations implies Bell inequalities. The other class in group (ii), (H_{22}^{α}), differs from local factorisation in that the parameters are swapped: instead of a dependence of each outcome on the local parameters it involves a dependence on the *distant* ones. Regardless of the implicit non-locality it can be used to derive a Bell inequality in a very similar way: given (H_{22}^{α}), instead of (11) we have

$$P(\alpha_\pm\beta_\mp|a_ib_i) = 0 = \sum_{\lambda} P(\lambda)P(\alpha_\pm|b_i\lambda)P(\beta_\mp|a_i\lambda). \tag{17}$$

and by very similar arguments we arrive at a similar partition of the values of $\boldsymbol{\lambda}$: $\Lambda(i)$ denotes all values of lambda for which $P(\alpha_+|b_i\lambda) = 0$, while the complementary set $\overline{\Lambda(i)}$ is defined by $P(\alpha_-|b_i\lambda) = 0$. If we then calculate the values of the hidden joint probability we arrive at the very same result as in table 2—and the rest of the derivation runs identically up to the Bell-Wigner inequality (3). Hereby we have found another *non-local* hidden joint probability which implies Bell inequalities and the result for group (ii) reads

$$\left[(\text{A}) \wedge (\text{PCorr}) \wedge \Big(\bigvee_{i=22,29} (\mathrm{H}_i^{\alpha}) \Big)\right] \to (\text{BI}). \tag{18}$$

A.3. Group (iii)

Up to this point one might have been surprised about how easy one can derive Bell inequalities from product forms other than local factorisation, but that one can do it even from classes in group (iii) is my strongest claim. These classes include both parameters, one in each factor, so they do not trivially imply Bell inequalities as classes in group (i). Neither, it seems, can they imply Bell inequalities in the same way as classes in group (ii) because they additionally involve $\boldsymbol{\beta}$ in the first factor. However, they do imply Bell inequalities, and they do it in a very similar (yet slightly more complicated) way than classes in group (ii), if besides perfect correlations for equal settings (10) we also assume perfect anti-correlations (PACorr) for perpendicular settings $(a_i b_{i_\perp})$:

$$P(\alpha_\pm, \beta_\pm | a_i, b_{i_\perp}) = 0 \tag{19}$$

Let me sketch the proof for class $(\mathrm{H}_{16}^{\alpha})$ which follows along the lines of that for local factorisation. By autonomy and the product form of $(\mathrm{H}_{16}^{\alpha})$ we rewrite (10) and (19) as

$$P(\alpha_\pm \beta_\mp | a_i b_i) = 0 = \sum_\lambda P(\lambda) P(\alpha_\pm | \beta_\mp a_i \lambda) P(\beta_\mp | b_i \lambda) \tag{20}$$

$$P(\alpha_\pm \beta_\pm | a_i b_{i_\perp}) = 0 = \sum_\lambda P(\lambda) P(\alpha_\pm | \beta_\pm a_i \lambda) P(\beta_\pm | b_{i_\perp} \lambda), \tag{21}$$

and again, at least one of the factors in each summand must vanish, i.e. for all values of i and $\boldsymbol{\lambda}$ (assuming $P(\lambda) > 0$) we must have:

$$\begin{aligned}
&\big[P(\alpha_+ | \beta_- a_i \lambda) = 0 \quad \vee \quad P(\beta_- | b_i \lambda) = 0\big] && (22)\\
\wedge\ &\big[P(\alpha_- | \beta_+ a_i \lambda) = 0 \quad \vee \quad P(\beta_+ | b_i \lambda) = 0\big] && (23)\\
\wedge\ &\big[P(\alpha_+ | \beta_+ a_i \lambda) = 0 \quad \vee \quad P(\beta_+ | b_{i_\perp} \lambda) = 0\big] && (24)\\
\wedge\ &\big[P(\alpha_- | \beta_- a_i \lambda) = 0 \quad \vee \quad P(\beta_- | b_{i_\perp} \lambda) = 0\big] && (25)
\end{aligned}$$

As above, from these conditions we can infer that all involved probabilities must be 0 or 1, depending on which of the following two cases holds.

If $P(\alpha_+ | \beta_- a_i \lambda) = 0$:

$$\begin{aligned}
&\overset{\text{(CE)}}{\Rightarrow} P(\alpha_- | \beta_- a_i \lambda) = 1 \qquad && \overset{(25)}{\Rightarrow} P(\beta_- | b_{i_\perp} \lambda) = 0\\
&\overset{\text{(CE)}}{\Rightarrow} P(\beta_+ | b_{i_\perp} \lambda) = 1 && \overset{(24)}{\Rightarrow} P(\alpha_+ | \beta_+ a_i \lambda) = 0\\
&\overset{\text{(CE)}}{\Rightarrow} P(\alpha_- | \beta_+ a_i \lambda) = 1 && \overset{(23)}{\Rightarrow} P(\beta_+ | b_i \lambda) = 0\\
&\overset{\text{(CE)}}{\Rightarrow} P(\beta_- | b_i \lambda) = 1 &&
\end{aligned}$$

If $P(\beta_-|b_i\lambda) = 0$:

$$
\begin{aligned}
&\overset{\text{(CE)}}{\Rightarrow} P(\beta_+|b_i\lambda) = 1 && \overset{(23)}{\Rightarrow} P(\alpha_-|\beta_+a_i\lambda) = 0 \\
&\overset{\text{(CE)}}{\Rightarrow} P(\alpha_+|\beta_+a_i\lambda) = 1 && \overset{(24)}{\Rightarrow} P(\beta_+|b_{i_\perp}\lambda) = 0 \\
&\overset{\text{(CE)}}{\Rightarrow} P(\beta_-|b_{i_\perp}\lambda) = 1 && \overset{(25)}{\Rightarrow} P(\alpha_-|\beta_-a_i\lambda) = 0 \\
&\overset{\text{(CE)}}{\Rightarrow} P(\alpha_+|\beta_-a_i\lambda) = 1
\end{aligned}
$$

The cases are disjunct and, hence, define a partition for the values of $\boldsymbol{\lambda}$ for each measurement direction i: $\Lambda(i)$ includes all values of $\boldsymbol{\lambda}$ for which $P(\beta_+|b_i\lambda) = 0$, while $\overline{\Lambda(i)}$ includes the complementary values, which make $P(\beta_-|b_i\lambda) = 0$. Calculating the hidden joint probability $P(\alpha\beta|ab\lambda)$ for an arbitrary choice of measurement directions a_ib_j gives us the very same result as in table 2—and again a Wigner Bell inequality follows.

Since the derivation for class (H_{15}) runs mutatis mutandis, we have shown:

$$\left[(\mathrm{A}) \wedge (\mathrm{PCorr}) \wedge (\mathrm{PACorr}) \wedge \left(\bigvee_{i=15,16} (\mathrm{H}_i^\alpha)\right)\right] \rightarrow (\mathrm{BI}) \tag{26}$$

A.4. Group (iv)

Finally, classes of group (iv) do *not* imply Bell inequalities. Involving both parameters in at least one of the factors, they neither fulfill Bell inequalities by their functional dependences nor do they admit of deriving a Bell inequality in the manner of classes in group (ii) or (iii). In order to rule out that there are other kinds of derivations one has to find explicit examples of probability distributions for each class in the group which *violate* Bell inequalities. Requiring just *any* example we can assume a toy model with only two possible hidden states ($\boldsymbol{\lambda} = 1, 2$). Then the probability distribution $P(\alpha\beta ab\lambda)$ is determined by assigning a value to each of the $2^5 = 32$ probabilities which conform to the laws of probability theory (each value lies in the interval $[0, 1]$ and all values sum to 1). Furthermore, the values have to be chosen such that autonomy and the specific product form of the class in question hold and that Bell inequalities are violated. I have found appropriate distributions for each class (H_1^α)–(H_{14}^α) by solving numerically a corresponding set of equations. This fact, that some probability distributions of these classes violate Bell inequalities, means that none of these classes implies Bell inequalities in general, i.e. by its constituting product form. Of course, this does *not* mean that *all* probability distributions in these classes *violate* Bell inequalities: in fact one can as well find examples of probability distributions in each class (H_1^α)–(H_{14}^α) which *fulfill* Bell inequalities. This means that for these classes the product form alone does not determine whether Bell inequalities hold or fail; whether they do depends on the numerical values of the specific distribution. On the general level of the classes we can only say that classes in group (iv) neither imply Bell inequalities nor do they imply their failure. Probability distributions in those classes *can* violate Bell inequalities.

References

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

Bell, J. S. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics 1*(3), 195–200.

Bell, J. S. (1971). Introduction to the hidden-variable question. In B. d'Espagnat (Ed.), *Foundations of quantum mechanics: Proceedings of the International School of Physics 'Enrico Fermi', Course XLIX*, pp. 171–181. New York: Academic.

Bell, J. S. (1975). The theory of local beables. *TH-2053-CERN*.

Bohm, D. (1951). *Quantum Theory*. Englewood Cliffs: Prentice-Hall.

Clauser, J. F. and M. A. Horne (1974). Experimental consequences of objective local theories. *Physical Review D 10*(2), 526–535.

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

Einstein, A., B. Podolsky, and N. Rosen (1935). Can quantum mechanical description of physical reality be considered complete? *Physical Review 47*, 777–780.

Graßhoff, G., S. Portmann, and A. Wüthrich (2005). Minimal assumption derivation of a Bell-type inequality. *British Journal for the Philosophy of Science 56*(4), 663–680.

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

Jarrett, J. P. (1984). On the physical signifance of the locality conditions in the Bell arguments. *Nous 18*, 569–590.

Maudlin, T. (1994). *Quantum Non-locality and Relativity : Metaphysical Intimations of Modern Physics*. Oxford: Blackwell.

Neapolitan, R. E. and X. Jiang (2006). A tutorial on learning causal influence. In D. E. Holmes and L. C. Jain (Eds.), *Innovations in machine learning : theory and applications*, pp. 29–72. Berlin ; New York: Springer.

Pawlowski, M., J. Kofler, T. Paterek, M. Seevinck, and C. Brukner (2010). Non-local setting and outcome information for violation of Bell's inequality. *New Journal of Physics 12*(8), 083051.

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

Price, H. (1994). A neglected route to realism about quantum mechanics. *Mind 103*, 303–336.

San Pedro, I. (2012). Causation, measurement relevance and no-conspiracy in EPR. *European Journal for Philosophy of Science 2*, 137–156.

Szabó, L. E. (2000). Attempt to resolve the EPR-Bell paradox via Reichenbach's concept of common cause. *International Journal of Theoretcial Physics 39*, 901–911.

van Fraassen, B. C. (1989). The charybdis of realism: Epistemological implications of Bell's inequality. In J. T. Cushing and E. McMullin (Eds.), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem*, pp. 97–113. Notre Dame: University of Notre Dame Press.

Wigner, E. P. (1970). On hidden variables and quantum mechanical probabilities. *American Journal of Physics 38*, 1005–1009.