A Stronger Bell Argument for (Some Kind of) Parameter Dependence

Paul M. Näger

Department of Philosophy, WWU Münster

Domplatz 6, D-48143 Münster, Germany

paul.naeger@wwu.de

2nd October 2015
Version 5.07

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 conclusion of the stronger Bell argument presented here provably is the strongest possible consequence from the violation of Bell inequalities on a qualitative probabilistic level (given usual background assumptions). Since among the excluded non-local theories are those whose only non-local probabilistic connection is a dependence between the space-like separated measurement outcomes of EPR/B experiments (a subset of outcome dependent theories), outcome dependence cannot be the crucial dependence for explaining a violation of Bell inequalities. Rather, the remaining non-local theories, which can violate Bell inequalities (among them quantum theory), are characterized by the fact that at least one of the measurement outcomes in some sense (which is made precise) probabilistically depends both on its local as well as on its distant measurement setting. While this is not to say that what is usually called parameter dependence has to hold, some kind of dependence on the distant parameter cannot be avoided. Against the received view, established by Jarrett and Shimony, that on a probabilistic level quantum non-locality amounts to outcome dependence, this result confirms and makes precise Maudlin's claim that some kind of parameter dependence is required.

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 fulfill 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. Formulating the stronger argument and deriving the new conclusion will make up a first part of this paper.

In a second part, we shall probabilistically analyze this new conclusion in a similar way as Jarrett (1984) famously analyzed the result of the standard Bell argument as outcome dependence and parameter dependence. The result of the new analysis will differ considerably from Jarrett's. Especially it will make explicit that some kind of parameter dependence cannot be avoided, while outcome dependence is irrelevant for the question whether Bell inequalities can be violated.

A third part is dedicated to comparing our result with existing positions and to drawing some immediate consequences. It will turn out that while literally correct Jarrett's classic analysis is misleading; the received view, holding that quantum non-locality is outcome dependence (Jarrett 1984; Shimony 1984, 1986), cannot be true; and Maudlin's information theoretic result (1994), that there must be some dependence between an outcome and its distant parameter, is confirmed and made precise in probabilistic terms; this will also resolve the ongoing tension between Jarrett's and Maudlin's opposing positions in favor of the latter.

Note that in this paper we restrict our investigations exclusively to the probabilistic level. More specifically, we mainly investigate which probabilistic dependences and independences there are in EPR/B experiments. Typically, from the formal result that there are certain non-local probabilistic dependences, far reaching conclusions about the existence of certain non-local physical or metaphysical connections are drawn, e.g. a non-separability or non-local causal relations. Since these latter inferences require further assumptions and are far from being trivial (especially they cannot reliably be made en passant), in this paper we shall constrain to establish a strengthening of Bell's core argument on the mathematical level and a probabilistic analysis of this result, while leav-

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

ing an appropriate physical and metaphysical interpretation of our findings for future work.

2 Bell arguments

2.1 EPR/B experiments and the standard Bell argument

We consider a usual EPR/B setup with space-like separated polarization measurements of an ensemble of photon pairs in an entangled quantum state $\boldsymbol{\psi} = \psi_0$ (Einstein, Podolsky, and Rosen 1935; Bohm 1951; Clauser and Horne 1974; see fig. 1). Possible hidden variables of the photon pairs are called $\boldsymbol{\lambda}$, so that the *complete* state of the particles at the source is $(\boldsymbol{\psi}, \boldsymbol{\lambda})$. Since in this setup the state $\boldsymbol{\psi}$ is the same in all runs, it will not explicitly be noted in the following (one may think of any probability being conditional on one fixed state $\boldsymbol{\psi} = \psi_0$). We denote Alice's and Bob's measurement setting as $\boldsymbol{a}$ and $\boldsymbol{b}$, respectively, and the corresponding (binary) measurement results as $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. 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 these five random variables.[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).

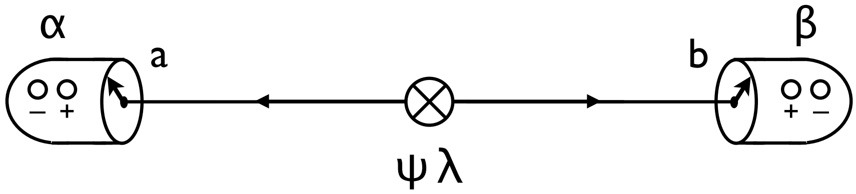

Figure 1: EPR/B setup

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 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'). A statistical evaluation of a series of many runs with similar pre-

[2] While the outcomes are discrete variables and the settings can be considered to be discrete (in typical EPR/B experiments there are two possible settings on each side), the hidden state may be continuous or discrete. In the following I assume $\boldsymbol{\lambda}$ to be discrete, but all considerations can be generalized to the continuous case.

paration procedures yields that the outcomes are strongly correlated given the settings and the quantum state.[3] For instance, in case the quantum state is the Bell state $\psi_0 = (|+\rangle|+\rangle + |-\rangle|-\rangle)/\sqrt{2}$ (and the settings are chosen with equal probability $\frac{1}{2}$) the correlations read:

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

These famous EPR/B correlations between space-like separated measurement outcomes have first been measured by Aspect et al. (1982) and have been confirmed under strict locality conditions (Weihs et al. 1998) as well as over large distances (Ursin et al. 2007). All these findings are correctly predicted by quantum mechanics: involving only empirically accessible variables, the quantum mechanical probability distribution essentially agrees with the empirical one.

Since according to (Corr), one outcome depends on both the other space-like separated outcome as well as on the distant (and local) setting, the observable part of the probability distribution (or the quantum mechanical distribution, respectively) clearly is non-local. Bell's idea (1964) was to 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, proceeds by showing that the empirically measured EPR/B correlations violate certain inequalities, the famous Bell inequalities. 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) employ two probabilistic assumptions, 'local factorisation'[4]

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

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 fact that there are perfect correlations (PCorr) between the outcomes for a certain relative angle of the measurement settings (e.g. for parallel settings given quantum state ψ_0).

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

[4] 'Local factorisation' is my term. Bell calls (ℓF) 'local causality', some call it 'Bell-locality', but most often it is simply called 'factorisation' or 'factorizability' (introduced by Fine). Bell's terminology already suggests a causal 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'.

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). While there are suggestions to explain the violation of the Bell inequalities by a failure of autonomy,[5] the main route in the debate has been to assume that it holds; and in order to focus on the factorisation condition (and possible modifications to it), this will also be one of our basic assumptions throughout this paper. If autonomy holds, the empirical violation of the Bell inequalities implies that local factorisation fails. 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'.

For my following critique of this standard Bell argument it is important to have a clear account of its logical structure. Here and in the following I shall presuppose 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)

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

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

(P5) Autonomy holds: (A)

(C1) Local factorisation fails: $\neg$(ℓF) (from P1–P5)

The conclusion of the argument is a probabilistic constraint that all theories of the quantum realm have to obey, and this constraint implies a non-locality, which is usually called quantum non-locality.

The core idea of my critique concerning this standard Bell argument for quantum non-locality is that its result is considerably weaker than it could be. I do not say that the argument is invalid (it is obviously not) 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 provide a tighter, more informative constraint for 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.

Specifically, I shall show that it is premise (P4) which can be made stronger. Stating that autonomy, perfect correlations and local factorisation imply the Bell inequalities, it

[5] A failure of autonomy can be realized by different kinds of models: conspiratorial models, simulation models (Fine 1982a), models with backwards causation (e.g. Price 1994; Corry 2015) or non-locality (San Pedro 2012).

is clear that one can make (P4) the stronger the weaker one can formulate the antecedent, i.e. the assumptions to derive the inequalities. Former improvements have concentrated on relaxing assumptions *except* the locality condition. In contrast, here I shall try to find *weaker alternatives to local factorisation*, which 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. This will make the overall argument stronger for it will allow for the conclusion that not only local theories but also those non-local ones that imply the inequalities are ruled out. I shall now first demonstrate this for one central class of non-local probability distributions, before in the subsequent section I consider the general case.

2.2 Bell inequalities from purely outcome dependent theories

Local factorisation is a specific product form of the *hidden joint probability* of the outcomes, as I shall call $P(\alpha\beta|ab\lambda)$.[6] A prominent non-local product form of this hidden joint probability is the following:

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda) \qquad (\mathrm{H}^{\alpha}_{16})$$

(For reasons that will become clear later the product form is tagged (H^{α}_{16}).) It differs from local factorisation in that it involves the distant outcome $\boldsymbol{\beta}$ in the first factor on the right hand side, which makes it a *non-local* product form (at least one of the factors involves at least one variable that is space-like separated to the respective outcome). Since product forms characterize probability distributions, which represent a whole class of theories, (H^{α}_{16}) represents a class of non-local theories. In the debate about Bell's theorem theories with such a non-local dependence between the outcomes in the product form are usually called *outcome dependent.* They represent physical theories according to which the outcomes are probabilistically or functionally dependent on another. The dependence between the space-like separated outcomes involved in these theories has emerged as the received view of what the violation of Bell inequalities amounts to: adequate theories of the quantum realm are widely believed to be correctly described as outcome dependent theories (Jarrett 1984; Shimony 1984).

In this section I shall prove that theories having the product form (H^{α}_{16}) are not consistent with the results of EPR/B experiments. In order to avoid misunderstandings, it is important to stress three central facts already at the outset of the argument. First, (H^{α}_{16}) is only one of several possible outcome dependent classes. So proving that (H^{α}_{16}) is impossible does not rule out all outcome dependent theories, but only very specific ones. For instance, the quantum mechanical distribution, which is well-known to be outcome

[6] 'Hidden' because the probability is conditional on the hidden state λ and thus is not empirically accessible.

dependent, is not correctly described by (H_{16}^{α}), but rather has the product form

$$P(\alpha\beta|ab) = P(\alpha|\beta ab)P(\beta), \tag{1}$$

i.e. according to quantum mechanics the outcome $\boldsymbol{\alpha}$ additionally depends on the distant setting $\boldsymbol{b}$ (and there is no dependence on a hidden variabel $\boldsymbol{\lambda}$). In order to distinguish (H_{16}^{α}) from such other outcome dependent classes I denote it as *purely* outcome dependent. So when in the following we show that purely outcome dependent theories are not consistent with results of EPR/B experiments, this does not mean that quantum mechanics is not consistent with these results.

Second, while not ruling out well established theories (there is, in fact, no serious theory that claims to have the product form (H_{16}^{α}) for EPR/B experiments), the result nevertheless has far reaching implications, because it informs us about the role that outcome dependence plays in the violation of Bell inequalities. This is because (H_{16}^{α}) contains outcome dependence in an isolated way, i.e. without being mixed with other non-local dependences. For this reason, the new result that purely outcome dependent theories *imply* the Bell inequalities shows that outcome dependence *per se* cannot explain the violation of the inequalities. This will be the central negative result of this paper. Emphatically, this is not to say that outcome dependent theories that are not purely outcome dependent, like quantum mechanics, cannot violate the Bell inequalities. Rather, it is to say that, contrary to what the reference to outcome dependent theories suggests, it is not the dependence between the outcomes that is central for the violation. Positively, we shall see in later sections that in order to violate the inequalities some kind of dependence on the distant setting is required (but not necessarily, as the quantum mechanical example shows, the kind of dependence that usually is called 'parameter dependence'). There we shall also explain in more detail, in which precise sense the fact that (H_{16}^{α}) implies Bell inequalities breaks with the received view.

Finally, we emphasize that the claims we shall be arguing for here are exclusively probabilistic ones: they are about *probabilistic* dependences (between the variables in the setup) and not about physical or metaphysical relations. It is important to stress the difference, because too often correlations are naively interpreted to indicate physical interactions, causal relations or the like. But correlation is not causation, and establishing the latter by the former involves non-trivial inferences, also invoking further assumptions. So saying that outcome dependence (which is by definition a probabilistic dependence) cannot explain the violation of the Bell inequalities is first of all not to say that a physical connection between the outcomes cannot explain the Bell inequalities. And likewise, saying that a violation of the inequalities requires probabilistic dependence of one outcome on both settings does not *per se* say that a physical connection between a setting and its distant outcome is implied. Since inferring (meta-)physical relations from probabilistic facts requires careful discussion, in this paper we shall not have the resources to treat that question as well. Rather we shall concentrate on deriving the probabilistic consequences of the violation of Bell inequalities, and therefore, if not explicitly stated otherwise, when I speak of 'dependence' in general or 'outcome dependence' in particular in the following I always mean probabilistic dependences (correlations), and not physical or metaphysical relations.[7]

After these preliminary remarks, I now turn to my argument against (H^{α}_{16}), which comes in two steps: I first show that if perfect correlations and perfect anti-correlations (and autonomy) hold, (H^{α}_{16}) is straightforwardly impossible. ('Straightforward' here means that the impossibility is not demonstrated via a Bell inequality, but in a more direct way.) This immediate inconsistency vanishes, when one relaxes perfect (anti-)correlations to nearly perfect (anti-)correlations (i.e. (anti-)correlations that show small deviations from perfectness). In this case, however, I demonstrate, second, that (H^{α}_{16}) and autonomy imply Bell inequalities. This is the genuine strengthening of the Bell argument that I have announced. Since the inequalities are empirically violated this also establishes an inconsistency (though a less direct one).

Let me start by putting my first claim in a precise form:

> **Lemma 1**: Autonomy, perfect correlations, perfect anti-correlations and (H^{α}_{16}) are an inconsistent set: (A) $\wedge$ (PCorr) $\wedge$ (PAcorr) $\rightarrow \neg$(H^{α}_{16})

(The proof of this lemma can be found in the appendix.)

Lemma 1 makes the surprising assertion that, given autonomy, purely outcome dependent theories are logically impossible if (for certain measurement settings) perfect correlations and (for certain other measurement settings) perfect anti-correlations hold. As can be seen in the proof of the lemma, the conflict does not require to formulate a Bell inequality: the inconsistency can be established in a much more direct way. Hence, in this case it does not even make sense to try to derive a Bell inequality, since the assumptions that would be needed for the derivation are already inconsistent. For this reason, lemma 1 is not in a literal sense a strengthening of the Bell argument. But since the aim of Bell's argument is to exclude certain theories of the micro-realm one might say that it is an *amendment* to the argument which strengthens its conclusion. The strengthening consists in the fact that lemma 1 precludes certain theories, namely purely outcome dependent ones, that usual Bell arguments do not rule out.

A defender of pure outcome dependence might try to avoid the conflict by asserting that perfect (anti-)correlations are not empirically confirmed. In fact, real experiments fail to yield the perfect (anti-)correlations that quantum theory predicts and that extrapolate the measured $\cos^2$-behavior (or $\sin^2$-behavior, respectively; cf. (Corr)) for non-parallel and non-perpendicular settings. Rather, the experiments show a certain deviation from perfect (anti-)correlations, such that perfect (anti-)correlations cannot be said to be empirically confirmed beyond doubt. Though it might seem reasonable to assume that they nevertheless do hold (because the experimental deviations from perfectness might be attributed to measurement errors and non-ideal detectors), it has become usual in the discussion about Bell's theorem to avoid the strong assumption of perfectness: either one does not make any reference to the correlations at parallel (or perpendicular) settings, or one assumes only *nearly* perfect correlations (nPCorr) (e.g. for parallel settings) and *nearly* perfect anti-correlations (nPACorr) (e.g. for perpendicular settings). Here we shall take the latter route and make the widely accepted

[7] Elsewhere I have derived what the probabilistic results derived here imply for the causal structure of the experiments (Näger 2013).

assumption of nearly perfect (anti-)correlations. Relaxing the perfect (anti-)correlations, a direct inconsistency similar to the one stated by lemma 1 does not follow any more (autonomy, nearly perfect correlations, nearly perfect anti-correlations and (H^{α}_{16}) are not an inconsistent set). Instead, in this case one can prove the following claim:

> **Lemma 2**: Given autonomy, nearly perfect correlations and nearly perfect anti-correlations, (H^{α}_{16}) implies Bell inequalities: (A)∧(nPCorr)∧(nPACorr)∧ (H^{α}_{16}) → (BI)

(For the proof of the lemma see the appendix.)

While this claim does not establish a straightforward inconsistency as the one given strictly perfect (anti-)correlations (cf. lemma 1), it is clear that, via the Bell argument, lemma 2 can be extended to argue for the inconsistency of (H^{α}_{16}) with autonomy, nearly perfect (anti-)correlations and the empirically confirmed EPR/B correlations. In this way, lemma 2 allows for a literal strengthening of Bell's theorem: it allows to modify premise (P4) of the Bell argument to say that both local as well as purely outcome dependent theories imply Bell inequalities. As local theories, purely outcome dependent theories do not produce correlations that are strong enough to violate Bell inequalities. Accordingly, the conclusion of the argument changes to preclude *more* theories than has been believed so far. Besides the local theories it also eliminates those non-local theories which assume an outcome to be dependent (functionally or probabilistically) not only on the local variables but also on the other, distant outcome.

There is a discrepancy between the original Bell argument and lemma 2, which hints to another aspect in which the latter helps to strengthen the former: while the original argument assumes *strictly* perfect correlations, lemma 2 only presumes *nearly* perfect correlations and anti-correlations. In fact, the proof of lemma 2 which shows how to derive Bell inequalities from outcome dependent theories and nearly perfect (anti-)correlations (and autonomy), can easily be adjusted to derive Bell inequalities from local theories and nearly perfect (anti-)correlations (and autonomy). Then, it is clear that one can relax premise (P3) to say that EPR/B correlations involve nearly perfect correlations as well as nearly perfect anti-correlations (instead of strictly perfect correlations). So the proof of lemma 2 also demonstrates the remarkable fact that *one can derive a Wigner-Bell inequality without strictly perfect correlations* (which were so far regarded to be a necessary assumption for deriving that type of Bell inequality).

The Bell inequality that follows by this new kind of prove is a generalized Wigner-Bell inequality,

$$P(\alpha_-\beta_+|a_1b_3) - 2\epsilon - \epsilon^2 \leq \frac{P(\alpha_-\beta_+|a_1b_2) + P(\alpha_-\beta_+|a_2b_3)}{(1-\epsilon^2)}, \tag{2}$$

that differs from a usual Wigner-Bell inequality

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

by certain correction terms involving a parameter $0 < \epsilon \ll 1$, which is a measure for the deviation from perfect correlations and perfect anti-correlations. (Precisely, ϵ^3 is the maximal fraction of photons deviating from perfect correlations or anti-correlations; see the proof of lemma 2.) It is easy to see that in the border case $\epsilon \to 0$ the generalized Wigner-Bell inequality agrees with the usual one. One can further show (see the proof of lemma 2) that the generalized inequality is violated by the usual statistics of EPR/B experiments, if at least 99.989% of the runs with parallel settings as well as those with perpendicular settings turn out to be perfectly correlated and perfectly anti-correlated, respectively. This defines the above condition of nearly perfect (anti-)correlations more precisely: only in worlds where the fraction of perfectly (anti-)correlated runs exceeds the indicated threshold, purely outcome dependent theories are ruled out.

This quantitative limit reveals a final resort for a defender of pure outcome dependence: she might hint to the fact that in actual experiments far less than 99.989% of the entangled objects show perfect (anti-)correlations. This indeed shows that the question whether purely outcome dependent theories can hold or not is not yet decided empirically beyond doubt. Let me stress, however, that the main aim in this paper is not to decide this empirical and quantitative question, but the conceptual and qualitative one, namely whether it is possible to amend Bell's argument for a stronger conclusion, ruling out even certain non-local theories.

That said, I can add that I think that there are good reasons not to take the mentioned empirical discrepancy to undermine the argument against purely outcome dependent theories. First, the derivation of the inequality (2) uses certain rather rough estimations, which contribute to the fact that the degree of perfectness that is required for a violation to take place is high. Improved future derivations, which include more precise (and expectedly more complicated) estimations, might lower that degree considerably. Second, the past has shown that experimental physicists have continuously been increasing the fraction of measured perfectly (anti-)correlated pairs of entangled objects, by using more and more sophisticated experimental techniques. So it is to be expected that the empirically confirmed degree of perfect correlation will increase in the future as well. Finally, quantum mechanics predicts perfect correlations and at present there is no further, independent evidence (besides the fact that experiments do not yield strictly perfect (anti-)correlations) to doubt that quantum mechanics is wrong; for this reason, it seems reasonable to assume that the deviation from perfectness in experiments is due to experimental imperfections.

Whether these arguments against the empirical discrepancy are conclusive or not: if my mathematical proofs are correct, the clear result of this section is that, given autonomy, purely outcome dependent theories cannot be adequate theories of the quantum realm if either *strictly* perfect (anti-)correlations or *nearly* perfect (anti-)correlations with a fraction of (dis-)agreement larger than 99.989% hold.

2.3 Generalization: A comprehensive scheme of possible theories

Strengthening an argument it is desirable to make it as strong as possible. We shall now generalize the stronger Bell argument that we have just presented so as to rule out *all* theories that can be ruled out by this type of argument. In order to capture all theories we shall proceed systematically and list a scheme of all logically possible theories, for each of which we check whether it is consistent under the given assumptions, and, if it is, whether it implies Bell inequalities or not. Note that this list will also contain theories that do not seem physically plausible. It is important, however, to include these theories into our investigation because in the end we aim to show that we have provided the strongest possible argument on a qualitative level (see section 2.6).

As we have said in the last section, local factorization and (H_{16}^{α}) are particular product forms of the hidden joint probability. In general, according to the product rule of probability theory, any hidden joint probability can equivalently be written as a product,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta ba\lambda)P(\beta|ab\lambda) \tag{4}$$

$$= P(\beta|\alpha ab\lambda)P(\alpha|ba\lambda). \tag{5}$$

Since there are two such general 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 (4), until in the next section we shall transfer the results to the other form (5).

We stress that the product form (4) 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 settings in both factors disappear, i.e. that so called parameter independence holds. Prima facie, any combination of variables in the two conditionals in (4) 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 (4) instead of (5)).

The specific product form of the hidden joint probability is *an 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 inequalities but also carries unambiguous information about which variables of the experiment are probabilistically independent of another. Therefore, 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

Table 1: Classes of probability distributions

	I	II	III	IV	V	VI	VII	VIII	IX
	(H_i^α):	$P(\alpha\beta \vert ab\lambda) = \ldots$					PCorr	nPCorr	
	i	$P(\alpha\vert\ \beta$	b	$a\ \ \lambda)$	$\cdot\ P(\beta\vert\ a$	$b\ \ \lambda)$	$\Box$(BI)	$\Box$(BI)	Notes
strong non-locality$^\alpha$	1	1	1	1	1	1	0	0	
	2	1	1	1	1	0	0	0	
	3	1	1	1	0	1	0	0	$\mathrm{QM_p}$
	4	1	1	0	1	1	—	0	
	5	1	0	1	1	1	—	0	
	6	0	1	1	1	1	0	0	$\mathrm{Bohm_s}$
	7	1	1	1	0	0	0	0	$\mathrm{QM_m}$
	8	0	1	1	1	0	0	0	
	9	0	1	1	0	1	0	0	$\mathrm{Bohm}_{\beta<a}$
	10	1	0	0	1	1	—	0	
	11	0	1	0	1	1	0	0	
	12	0	0	1	1	1	0	0	$\mathrm{Bohm}_{\alpha<b}$
	13	0	1	1	0	0	0	0	
	14	0	0	0	1	1	0	0	
weak non-locality$^\alpha$	15	1	1	0	1	0	—	1	
	16	1	0	1	0	1	—	1	pure outc. dep.
	17	1	0	1	1	0	—	—	
	18	1	1	0	0	1	—	—	
	19	1	1	0	0	0	—	—	
	20	1	0	1	0	0	—	—	
	21	1	0	0	1	0	—	—	
	22	0	1	0	1	0	1	1	
	23	0	0	1	1	0	—	—	
	24	1	0	0	0	1	—	—	
	25	0	1	0	0	1	—	—	
	26	1	0	0	0	0	—	—	
	27	0	1	0	0	0	—	—	
	28	0	0	0	1	0	—	—	
locality$^\alpha$	29	0	0	1	0	1	1	1	local factoriz.
	30	0	0	1	0	0	—	—	
	31	0	0	0	0	1	—	—	
	32	0	0	0	0	0	—	—	
Analysis:		$\neg(\mathrm{OI}_1)$	$\neg(\mathrm{PI}_1^\alpha)$	$\neg(\ell\mathrm{PI}_1^\alpha)$	$\neg(\mathrm{PI}_2^\beta)$	$\neg(\ell\mathrm{PI}_2^\beta)$			

experiment must belong to one of these 32 classes. In this systematic overview, the class constituted by local factorisation is (H_{29}^{α}), and it now also becomes clear why the product form of purely outcome dependent theories has been tagged (H_{16}^{α}) in the last section. Furthermore, if we allow that there might be no hidden states $\boldsymbol{\lambda}$, the quantum mechanical distribution as well as the empirical distribution (which as far as we know coincide, but see our discussion of perfect (anti-)correlations in the last section) belong to class (H_{7}^{α}) (if the photon state $\boldsymbol{\psi}$ is *maximally* entangled, noted by '$\mathrm{QM_m}$') or to (H_{3}^{α}), respectively (if $\boldsymbol{\psi}$ is *partially* entangled, noted by '$\mathrm{QM_p}$').[8] The de-Broglie-Bohm theory falls under class (H_{6}^{α}), when both settings are chosen before the detector at the respective other side has registered, i.e. $t(a) < t(\beta)$ and $t(b) < t(\alpha)$;[9] we label the corresponding probability distribution by '$\mathrm{Bohm_s}$' (the index standing for *s*ymmetrical time ordering). Otherwise, when the $\boldsymbol{\beta}$-measurement completes before $\boldsymbol{a}$ has been set to its final state (labelled by '$\mathrm{Bohm}_{\beta<a}$'), the theory falls in class (H_{9}^{α}); and when the $\boldsymbol{\alpha}$-measurement is over before $\boldsymbol{b}$ has been chosen (labelled by '$\mathrm{Bohm}_{\alpha<b}$'), we have class (H_{12}^{α}). Similarly, any other theory of the quantum realm has its unique place in one of the classes.

One crucial advantage of such an abstract 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) are consistent with the empirical probability distribution of EPR/B experiments. As in the previous section we discern two cases according to whether *strictly* perfect (anti-)correlations or *nearly* perfect (anti-)correlations hold.

2.3.1 Strictly perfect (anti-)correlations

For the case of strictly perfect (anti-)correlations, the following theorems hold:

> **Theorem 1.1**: Autonomy, perfect correlations, perfect anti-correlations and a class of probability distributions (H_{i}^{α}) form an inconsistent set if and only if (i) the product form of (H_{i}^{α}) involves at most one of the settings or (ii) the product form of (H_{i}^{α}) involves both settings but its first factor involves the distant outcome and at most one setting.

> **Corollary 1.1**: A class (H_{i}^{α}) is consistent with autonomy, perfect correlations and perfect anti-correlations if and only if (¬i) the product form of

[8] 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 an additional probabilistic dependence on the local setting in the second factor; hence, they fall in class (H_{3}^{α}). For an overview of the dependences and independences in the quantum mechanical probability distribution of maximally and partially entangled states see Näger (2015, table 1).

[9] Such temporal ordering between space-like separated events is, of course, only possible when there is a preferred frame of reference, which Bohm's theory presupposes.

(H$_i^\alpha$) involves both settings and (¬ii) in case the distant outcome appears in the first factor of (H$_i^\alpha$)'s product form, also both settings appear in that factor.

Theorem 1.2: Given autonomy, perfect correlations and perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i) and (¬ii)) implies Bell inequalities if and only if (iii) each factor of its product form involves at most one setting.

Corollary 1.2: Given autonomy, perfect correlations and perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i) and (¬ii)) does not imply Bell inequalities if and only if (¬iii) at least one factor of its product form involves both settings.

(The proofs of these theorems can be found in the appendix.)

The consequences of these claims for the status of the different classes are represented in column VII of table 1. The heading of the column, '□(BI)', means *necessarily, Bell inequalities hold.* So the column indicates whether a certain product form implies Bell inequalities ('1') or does not imply them ('0') (according to theorem 1.2 or corollary 1.2, respectively); it further indicates when this question does not make sense ('—') because a product form is inconsistent with the background assumptions autonomy and perfect (anti-)correlations (according to theorem 1.1). It is understood that classes that are marked by either '0' or '1' are consistent with the background assumptions (cf. corollary 1.1). Clearly, all classes that are marked either by '—' or '1' are impossible if autonomy and perfect (anti-)correlations hold: the former yield a direct contradiction with the background assumptions, while the latter contradict the empirical probability distribution via the Bell argument.

The inconsistent classes ('—') divide into two subgroups, corresponding to which condition for inconsistency, (i) or (ii) (cf. theorem 1.1), is fulfilled:

Inconsistency due to condition (i): $\{(\mathrm{H}_{17}^\alpha), ..., (\mathrm{H}_{32}^\alpha)\}\backslash\{(\mathrm{H}_{22}^\alpha), (\mathrm{H}_{29}^\alpha)\}$

Inconsistency due to condition (ii): $\{(\mathrm{H}_{4}^\alpha), (\mathrm{H}_{5}^\alpha), (\mathrm{H}_{10}^\alpha), (\mathrm{H}_{15}^\alpha), (\mathrm{H}_{16}^\alpha)\}$

Positively, theorem 1.1 further says that all classes that are not inconsistent according to these criteria are consistent with the background assumptions: the criterion of consistency, (¬i) *and* (¬ii), as stated in corollary 1.1, is just the negation of the condition for inconsistency, (i) *or* (ii), in theorem 1.1.

We emphasize that the consistency and inconsistency claims of classes with the background assumptions have asymmetric consequences on the level of single probability distributions. On the one hand, a class being inconsistent with the background assumptions means that *every* probability distribution of that class forms an inconsistent set with the assumptions. It is the general product form defining the class which is in conflict with the assumptions, hence all members of the class are. The same, mutatis mutandis, however, is not true of the consistent classes. A class being consistent does *not* mean that *every* probability distribution of that class is consistent with the assumptions. Rather, by the laws of logic, it just means that *at least one* probability distribution of a class is consistent with the assumptions, showing that the general product form of that class is *not per se* in conflict with them. This is what consistency of a class means (when we define inconsistency in the natural way as just stated). This definition of consistency is perfectly compatible with the fact that there are distributions in a consistent class that are inconsistent with the assumptions due to their specific *numerical values*. For instance, one can easily imagine distributions falling under class (H_7^α) that, at parallel settings, involve correlations that are weaker than perfectness. These distributions are obviously not consistent with the background assumptions, although their general product form is. Hence, we have to keep in mind that being consistent with the background assumptions on the level of classes, which is the level the present analysis proceeds on, is just a necessary condition for the distributions in that class to be consistent.

Turning to theorem 1.2, all classes marked by '1', i.e. (H_{22}^α) and (H_{29}^α), can explicitly be shown to imply a Bell inequality. That (H_{29}^α), local factorization, implies the inequalities is well known, but that

$$P(\alpha\beta|ab\lambda) = P(\alpha|b\lambda)P(\beta|a\lambda), \tag{H_{22}^α}$$

a non-local class, does, has not been observed so far. That class is the symmetrical counterpart to local factorization, compared to which the settings are swapped, such that each outcome depends on its distant setting. For this reason the derivation of the Bell inequalities runs very similarly as for local factorization (just swap the settings in the original proof).

On the other hand, the theorem also says that any consistent class that violates (iii), can be shown *not* to imply the Bell inequalities. Here we have a similar asymmetry between the level of classes and that of distributions as in the case of (in-)consistency. Since a class implying Bell inequalities (given the background assumptions) means that *every* probability distribution having the product form in question obeys the inequalities, the claim that a class does *not* imply the inequalities (given the background assumptions) denotes the fact that *there is at least one probability distribution in that class that violates the inequalities.* Therefore, not implying Bell inequalities emphatically does *not* mean that *every* probability distribution in a class violates the inequalities. For this reason, given just the product form of one of the classes violating (iii) one cannot decide whether Bell inequalities hold; whether they do in these cases depends on the *numerical* features of the probability distribution in question. In this sense, one might reasonably say that probability distributions of these classes *can* violate Bell inequalities. So far for the meaning of implying and not implying the Bell inequalities. Let us now turn to the criterion which demarcates the two cases.

Condition (iii), that, in order to imply Bell inequalities, a consistent class may not involve more than one setting in each factor of its product form, is the essential characteristic (in terms of the product form) to tell apart classes marked by '1' from those marked by '0'. This criterion differs considerably from the usual message that local theories imply Bell inequalities (and non-local ones do not). In order to understand its content, let us partition the classes into three groups, depending on which variables appear in their constituting product forms:

Local$^{\alpha}$ classes: (H^{α}_{29})–(H^{α}_{32})
Each factor only contains time-like (or light-like) separated variables.

Weakly non-local$^{\alpha}$ classes: (H^{α}_{15})–(H^{α}_{28})
At least one of the factors involves space-like separated variables, but none of the factors involves both settings.

Strongly non-local$^{\alpha}$ classes: (H^{α}_{1})–(H^{α}_{14})
At least one of the factors involves both settings. (¬iii)

With these new concepts we can summarize theorems 1.1 and 1.2 as saying that *given autonomy, perfect correlations and perfect anti-correlations, every local*$^{\alpha}$ *and weakly non-local*$^{\alpha}$ *class either is inconsistent with autonomy and the perfect (anti-)correlations or (if it is consistent) obeys Bell inequalities. Certain strongly non-local*$^{\alpha}$ *classes are inconsistent with autonomy and the perfect (anti-)correlations as well; however, all consistent strongly non-local*$^{\alpha}$ *classes do not imply, i.e. can violate, the Bell inequalities.* (Strongly non-local$^{\alpha}$ classes are by definition just those classes that fulfill criterion (¬iii) not to imply the Bell inequalities.)

What does this result mean? On the one hand, it sounds familiar that local classes are impossible in the given situation. Local classes involve only time-like (or light-like) separated variables in the factors of their hidden joint probability, and local factorization, which is well-known to imply Bell inequalities, is the paradigm of product forms constituting these classes. Theorem 1.1 just adds the further claim that given autonomy and perfect (anti-)correlations all other local classes are directly inconsistent.

The surprising consequence of theorems 1.1 and 1.2 rather is that *even certain non-local classes are ruled out*. Every class in the group of weakly non-local$^{\alpha}$ classes is forbidden. Most of the classes in that group are directly inconsistent with the assumptions of autonomy and perfect (anti-)correlations, including (as we have shown in the previous section) the purely outcome dependent class (H^{α}_{16}). That purely outcome dependent theories are not even available when perfect (anti-)correlations (and autonomy) hold, is, as we have already remarked, a central result of this investigation, because it belongs to the group of classes that has evolved as the received view of what quantum non-locality amounts to on a probabilistic level. In fact, there is only one weakly non-local$^{\alpha}$ class, which is consistent with the background assumptions autonomy and perfect (anti-)correlations, viz. (H^{α}_{22}). Whether this class is physically plausible or not: the fact that it implies the inequalities proves the important insight that *Bell inequalities are not a locality condition* (because there is a class obeying Bell inequalities that is non-local).

Instead of locality, the hallmark of theories implying Bell inequalities rather is, as theorem 1.2 states, that they may not involve more than one setting in each factor of their product form. The negation of this condition, that at least one factor contains both settings, is exactly the defining feature of strongly non-local$^{\alpha}$ classes. That is why the above partition is so natural. However, this does not mean that all strongly non-local$^{\alpha}$ classes are allowed in the given situation; for some of them—(H_4^{α}), (H_5^{α}) and (H_{10}^{α})—are inconsistent with autonomy and perfect (anti-)correlations ('—'). This demonstrates that the criteria for being consistent with these background assumptions, (¬i) *and* (¬ii), and for not implying Bell inequalities, (¬iii), which is equivalent to being strongly non-local$^{\alpha}$, are not disjunct. A theory can be strongly non-local$^{\alpha}$ and still violate (¬ii) (all strongly non-local$^{\alpha}$ classes marked by '—'), e.g. (H_4^{α}); and there are theories fulfilling (¬i) and (¬ii) but fail to be strongly non-local$^{\alpha}$, e.g. (H_{22}^{α}). A successful theory must belong to one of the classes that takes both hurdles, and those are the ones marked by '0' in column VII of table 1.

2.3.2 Nearly perfect (anti-)correlations

Let us now relax the assumption of strictly perfect (anti-)correlations to nearly perfect (anti-)correlations and observe how that changes the situation. In this case, the following theorems can be proven:

> **Theorem 2.1**: Autonomy, nearly perfect correlations, nearly perfect anti-correlations, and a class of probability distributions (H_i^{α}) form an inconsistent set if and only if (i) the product form of (H_i^{α}) involves at most one of the settings.
>
> **Corollary 2.1**: A class (H_i^{α}) is consistent with autonomy, nearly perfect correlations and nearly perfect anti-correlations if and only if (¬i) the product form of (H_i^{α}) involves both settings.
>
> **Theorem 2.2**: Given autonomy, nearly perfect correlations and nearly perfect anti-correlations each consistent class (i.e. each class that fulfills (¬i)) implies Bell inequalities if and only if (iii) each factor of its product form involves at most one setting.
>
> **Corollary 2.2**: Given autonomy, nearly perfect correlations and nearly perfect anti-correlations each consistent class (i.e. each class that fulfills (¬i)) does not imply Bell inequalities if and only if (¬iii) at least one factor of its product form involves both settings.

(The proofs of the theorems can be found in the appendix.)

The consequences of these claims are represented in column VIII of table 1. Since nearly perfect (anti-)correlations are a considerably weaker requirement than that of strictly perfect ones, one essential change that occurs in these theorems compared to the former is that the conditions for consistency with the background assumptions (autonomy and nearly perfect (anti-)correlations in the new case) are considerably weaker as well: theorem 2.1 just requires condition (¬i) but not condition (¬ii). As a consequence, all classes that have been ruled out by condition (¬ii) (in theorem 1.1), viz. (H_4^α), (H_5^α), (H_{10}^α), (H_{15}^α) and (H_{16}^α), are now consistent with the new, less strict background assumptions.

Especially purely outcome dependent theories defined by

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda) \tag{H_{16}^α}$$

cease to be directly inconsistent with the background assumptions. However, since the criterion for implying Bell inequalities stays essentially unchanged (requirement (iii) still holds),[10] outcome dependent theories imply Bell inequalities (see our discussion in section 2.2), so they are still forbidden. It is just that the reason why they are forbidden changes. Similar facts are true for the symmetrical counterpart to purely outcome dependent theories,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta b\lambda)P(\beta|a\lambda), \tag{H_{15}^α}$$

which differs from (H_{16}^α) in that the settings are swapped between the factors, such that each outcome depends on the distant (instead of on the local) setting. In effect, also in the new situation it is still true that *all local*$^\alpha$ *and weakly non-local*$^\alpha$ *classes are forbidden.*

Concerning the strongly non-local$^\alpha$ classes, however, the situation changes. Formerly, certain strongly non-local$^\alpha$ classes, (H_4^α), (H_5^α) and (H_{10}^α), were forbidden because they were inconsistent with the background assumptions. Relaxing the background assumptions, we have already said that they become consistent. But unlike the weakly non-local$^\alpha$ classes that have become consistent, (H_{15}^α) and (H_{16}^α), these strongly non-local$^\alpha$ classes do not imply Bell inequalities, because they clearly do not fulfill condition (iii); by weakening the background assumptions, these classes cease to be ruled out by the theorems. As a consequence, all strongly non-local$^\alpha$ classes are now consistent with the background assumptions and do not imply Bell inequalities. The reason for this new situation is that abandoning criterion (¬ii) for consistency, the remaining criterion, (¬i), is entailed by the criterion for not implying Bell inequalities, viz. to be strongly non-local$^\alpha$ (so the two criteria are not logical independent any more).

In sum, the result is that *given autonomy, nearly perfect correlations and nearly perfect anti-correlations, every local*$^\alpha$ *and weakly non-local*$^\alpha$ *class either is inconsistent with autonomy and the perfect (anti-)correlations or (if it is consistent) obeys Bell inequalities. In contrast, all strongly non-local*$^\alpha$ *classes are consistent with autonomy and the perfect (anti-)correlations and do not imply Bell inequalities.* Unlike the case with strictly perfect correlations, there are no forbidden strongly non-local$^\alpha$ theories, which amounts to a slight modification of the set of precluded classes.

The main messages, however, have not changed: 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 are forbidden by the empirical statistics of

[10] There is just the slight difference that now both nearly perfect correlations *and* nearly perfect anti-correlations are required, whereas according to theorem 1.2 the anti-correlations were not needed for the derivation.

EPR/B experiments (if autonomy holds). Rather, we have found that 18 (21 in the case of strictly perfect (anti-)correlations) of the 32 logically possible classes are forbidden, among them 14 (17 in the case of strictly perfect (anti-)correlations) *non-local* classes. Some of these non-local classes are forbidden because they are directly inconsistent with the assumptions autonomy and (nearly) perfect (anti-)correlations. Others are forbidden because they imply Bell inequalities. This latter fact has two important consequences. First, it makes explicit that Bell inequalities are not a locality condition. Neither, second, is locality a necessary condition for deriving Bell inequalities. The criterion to imply the inequalities (if autonomy and (nearly) perfect anti-correlations hold) rather is a different one, which has not to do with the locality/non-locality divide: Bell inequalities are implied by each probability distribution whose product form involves at most one setting in each of its factors. So according to a probability distribution the outcomes might depend on their distant setting as well as on each other, (H^{α}_{15}), and still Bell inequalities follow. As a consequence, if one searches for theories which conform to the empirical fact that (nearly) perfect correlations hold and Bell inequalities are violated they can only be among the strongly non-local$^{\alpha}$ ones (which are defined to involve both settings in at least one factor). Contrary to the view suggested by Bell's original theorem it cannot be a weakly non-local$^{\alpha}$ class.

2.4 Strengthening Bell's argument

It is clear that each set of theorems (1.1 and 1.2 as well as 2.1 and 2.2) can be used to strengthen Bell's argument. On the other hand, it is not clear which of these available new arguments should be considered to be the strongest. (The first set results in an argument that, compared to the argument resulting from the second set, requires the stronger assumption of strictly perfect correlations (weakening the argument), but allows for a stronger conclusion, because it rules out even some of the strongly non-local$^{\alpha}$ classes). Here we restrict our discussion to the argument resulting from the second set, because it avoids the controversial assumption of strictly perfect (anti-)correlations. (The argument from the first set can be formulated mutatis mutandis.)

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

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

(P3′) EPR/B correlations include nearly perfect correlations and nearly perfect anti-correlations: (Corr) $\rightarrow$ (nPCorr) $\wedge$ (nPACorr)

(P6) Those local$^{\alpha}$ and weakly non-local$^{\alpha}$ classes that involve at most one setting in their product form are inconsistent with autonomy, nearly perfect correlations and nearly perfect anti-correlations:

$$(\mathrm{A}) \wedge (\mathrm{nPCorr}) \wedge (\mathrm{nPACorr}) \rightarrow \bigwedge_{\substack{i=17..32\\ \setminus\{22,29\}}} \neg(\mathrm{H}^{\alpha}_{i})$$

(P4′) Bell inequalities can be derived from autonomy, nearly perfect correlations, nearly perfect anti-correlations and any local$^{\alpha}$ or weakly non-local$^{\alpha}$ class of probability distributions that involves both settings in its product form:

$$\Big[(\mathrm{A}) \wedge (\mathrm{nPCorr}) \wedge (\mathrm{nPACorr}) \wedge \Big(\bigvee_{\substack{i=15,16,\\ 22,29}} (\mathrm{H}_i^{\alpha}) \Big) \Big] \rightarrow (\mathrm{BI})$$

(P5) Autonomy holds: (A)

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

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

Compared to the original Bell argument (section 2.1) there are three substantial changes, which strengthen the argument. A first change concerns the fact that everywhere in the argument we have relaxed controversial strictly perfect correlations to uncontroversial nearly perfect correlations (in premisses (P3) and (P4) of the original argument). This is a strengthening in the sense that the argument makes weaker assumptions. At the same places in the argument where nearly perfect correlations occur we have additionally introduced nearly perfect (anti-)correlations. This might seem as a weakening of the argument; in fact, however, it is a neutral move, because it is uncontroversial that the nearly perfect anti-correlations follow from the EPR/B correlations (as the nearly perfect correlations do; see premise (P3′)), and these EPR/B correlations have already been assumed in the original argument (premise (P1)).

A second strengthening of the argument stems from introducing a completely new premise (P6), which states the content of theorem 2.1, that certain classes are not compatible with autonomy, nearly perfect correlations and perfect anti-correlations. Given that autonomy and perfect (anti-)correlations are assumed anyway (or derive from assumptions), it is clear that these classes will be ruled out by the overall argument. In this sense, (P6) provides a genuine strengthening of the conclusion of the theorem. Deriving a direct contradiction between the background assumptions and certain classes without involving a Bell inequality, premise (P6) has no counterpart in the original argument and rather has the status of an amendment—however, an amendment that naturally fits in. Note that assuming the additional premise (P6) does not weaken the argument because it can be proven mathematically (see the proof of theorem 2.1).

A third modification, indeed the central strengthening, consists in the adaption of premise (P4) to theorem 2.2, which says that one can derive Bell inequalities not only from local factorization but from all those local$^{\alpha}$ and weakly non-local$^{\alpha}$ classes that are consistent given autonomy and perfect (anti-)correlations. Accordingly, we have replaced local factorisation in the antecedent by the disjunction of these product forms. This makes the antecedent of (P4′) weaker than that in (P4) and, hence, the argument stronger. Since the overall Bell argument is a modus tollens argument to the negation of that premise, this modification also strengthens the conclusion of the theorem.

Making these changes has a considerable effect on the overall Bell argument. Instead of the standard conclusion (C1), that the violation implies the failure of local factorisation, by the modified argument we arrive at the essentially stronger conclusion (C1′). 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.

2.5 Further strengthening by a complementary partition

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 (4) and conceiving different possible product forms (table 1). However, we can as well write the hidden joint probability according to the second product rule (5), and similar arguments as above lead us to a similar table as table 1, whose classes, (H_1^β)–(H_{32}^β), differ to those in table 1 in that the outcomes and the settings are swapped. For instance, class (H_{16}^β) is defined by the product form $P(\alpha\beta|ab\lambda) = P(\beta|\alpha b\lambda)P(\alpha|a\lambda)$ in contrast to (H_{16}^α), 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, which reasonably might be called *complementary partition*. Any probability distribution must fall in exactly one of the classes (H_1^α)–(H_{32}^α) *and* in exactly one of the classes (H_1^β)–(H_{32}^β). Analogously to theorem 1 one can prove for the new partition that (given autonomy and nearly perfect (anti-) correlations) also each local$^\beta$ and weakly non-local$^\beta$ class either is inconsistent or implies Bell inequalities, so that we can reformulate (P6) and (P4′) as:

(P6′) Those local$^\alpha$, weakly non-local$^\alpha$, local$^\beta$ and weakly non-local$^\beta$ classes that involve at most one setting in their product form are inconsistent with autonomy, nearly perfect correlations and nearly perfect anti-correlations:

$$(\mathrm{A}) \wedge (\mathrm{nPCorr}) \wedge (\mathrm{nPACorr}) \rightarrow \bigwedge_{\substack{i=17..32\\ \setminus\{22,29\}}} \neg(\mathrm{H}_i^\alpha) \quad \wedge \bigwedge_{\substack{i=17..32\\ \setminus\{22,29\}}} \neg(\mathrm{H}_i^\beta)$$

(P4″) Bell inequalities can be derived from autonomy, nearly perfect correlations, nearly perfect anti-correlations and any local$^\alpha$, weakly non-local$^\alpha$, local$^\beta$ or weakly non-local$^\beta$ class of probability distributions that involves both settings in its product form:

$$\left[(\mathrm{A}) \wedge (\mathrm{nPCorr}) \wedge (\mathrm{nPACorr}) \wedge \Big(\bigvee_{\substack{i=15,16,\\ 22,29}} (\mathrm{H}_i^\alpha) \quad \vee \bigvee_{\substack{i=15,16,\\ 22,29}} (\mathrm{H}_i^\beta)\Big)\right] \rightarrow (\mathrm{BI})$$

With these new premises we can formulate an even stronger Bell argument from (P1), (P2), (P3′), (P6′), (P4″) and (P5) to

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

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

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 strongly non-local$^\alpha$ classes and one of the strongly non-local$^\beta$ classes 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 both product forms 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.

2.6 Immediate consequences

(1) On a rather general level, the fact that certain non-local theories imply Bell inequalities first of all illustrates that *Bell inequalities are not locality conditions* in the sense that, if a probability distribution obeys a Bell inequality, it must be local. In the discussion, Bell inequalities are so closely linked to locality that one could have this impression. Of 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 fact that up to now local factorisation has been the *only* product form which has been 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 some 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.*

Note that this result is not in conflict with Fine's insight (1982b) that an empirical probability distribution obeying a Bell inequality is equivalent with the existence of a hidden probability distribution that is local ('local stochastic hidden variable model'). My claim is that not every hidden probability distribution which obeys a Bell inequality is local; Fine's result, in contrast, implies that for every empirical distribution that is consistent with Bell inequalities one can find a local hidden distribution.

(2) The conclusions of the new Bell argument, which we have derived, are considerably stronger than those of previous versions. We have shown that the violation of Bell inequalities not only excludes local theories but also weakly non-local$^{\alpha}$ and weakly non-local$^{\beta}$ ones. In contrast, the conclusion of the standard Bell argument only forbids local theories and allows for *all* non-local ones, including the weakly non-local classes that we have shown to imply Bell inequalities. In this sense, the usual constraint following from the standard Bell argument, *is inappropriately weak.* While this is not to say that the standard argument is logically incorrect, it does mean that its conclusion is not as tight as it could be. We should keep in mind that any argument based on this standard conclusion, especially Jarrett's analysis, proceeds from a mixture of classes that can violate Bell inequalities with classes that imply them—and therefore might yield misleading results.

(3) The same is not true of our new result: all classes that it allows, all strongly non-local classes, can violate Bell inequalities. For this reason it is impossible to strengthen Bell's argument in such a way as to rule out more classes of probability distributions than we have ruled out here. In this sense, we can say that if our considerations have been correct and the typical background assumptions hold (autonomy and nearly perfect (anti-)correlations), by our systematic approach we can be sure that the conclusions from the new Bell argument are the *strongest possible consequences of the violation of Bell inequalities on a qualitative probabilistic level.* Note that this is not to say that further classes might not be ruled out due to other criteria, maybe due to their incompatibility with relativity or the like. The label 'qualitative probabilistic' indicates that we have only referred to classes of probability distributions defined by their probabilistic dependences and independences without referring to quantitative features.

It might be interesting to make explicit how we arrived at this strong conclusion. Especially, our considerations in this paper have two important features that preclude future strengthenings of the argument to rule out more classes. First, the central methodological procedure of our argument was to consider *all logically possible* classes of probability distributions. Hence, any probability distribution that conceivably might describe an EPR/B experiment must fall under one of the classes in our systematic overview (cf. table 1). For this reason, we can be sure that we have not overlooked any probability distribution for the EPR/B experiment. There simply are no probability distributions left that might bring in some surprise; we have captured them all.

A second important feature is that our argument provides sufficient *and* necessary conditions for classes to imply Bell inequalities. By stating that local classes imply Bell inequalities, former arguments typically have only provided sufficient criteria. This left open the possibility that there are further classes implying the inequalities—and, indeed, here we have found that many non-local classes, viz. the weakly non-local ones, do as well. On the other hand, by explicitly showing that the remaining classes, the strongly non-local classes, can violate the inequalities (see the proofs of theorems 1.2 and 2.2, where we have constructed explicit examples of distributions in those classes that violate the inequalities), we have precluded that future arguments might show one of the strongly non-local classes to imply the inequalities as well. And if this argument, that proceeds on the qualitative probabilistic level of the classes and their product forms, is correct, and the background assumptions we have presupposed hold, we cannot entail a stronger claim on that level than that local and weakly non-local classes imply Bell inequalities while strongly non-local classes can violate them.

(4) The latter claim also reveals a certain limitation of the argument presented here. It emphatically does not say that strongly non-local classes violate Bell inequalities; it only says that strongly non-local classes *can* violate Bell inequalities, meaning that some of the strongly non-local distributions do violate the inequalities while others do not. In fact, one can explicitly find examples for probability distributions in each of the strongly non-local$^{\alpha}$ classes (H_1^{α})–(H_{14}^{α}) (as well as in the strongly non-local$^{\beta}$ classes (H_1^{β})–(H_{14}^{β})) which *obey* Bell inequalities—and these distributions clearly could be ruled out by more precise arguments. However, belonging to the same class, discerning strongly non-local classes which violate the inequalities from those that obey them clearly can*not* be made on a *qualitative* probabilistic level. Any improvement of the argument must refer to the specific *numerical values* of the probability distribution in question, so there is no general claim that can be made on the basis of the mere product form; the product form of any strongly non-local class alone does not determine whether Bell inequalities hold or fail.

It follows that the consequence of my stronger Bell argument, that the quantum world can only be described correctly by a theory falling under a strongly non-local class, is only a *necessary* condition for violating Bell inequalities; it is not a sufficient one. (Note the difference between conditions for *violating* Bell inequalities and conditions for *not implying* them; we have provided necessary and sufficient conditions for the latter but only necessary ones for the former.) Sufficient criteria to violate Bell inequalities would have to involve conditions for the *strength* of the correlations. A common measure for how strong a correlation is, is mutual information, so 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 (cf. Maudlin 1994, ch. 6 and Pawlowski et al. 2010). These are important works, which can further sharpen the constraints for quantum non-locality following from EPR/B experiments. Such quantitative improvements, however, do not count against my claim here that the conclusion of my new stronger Bell argument captures the strongest possible consequences of the violation of Bell inequalities on a *qualitative* probabilistic level.

3 Analyzing the conclusions

Having strengthened Bell's argument to a more informative conclusion, we now have to make precise what this new, stronger constraint for quantum non-locality amounts to. Jarrett (1984) proved that the standard probabilistic constraint for quantum non-locality following from the usual Bell argument, the failure of local factorization, is equivalent to the disjunction of outcome dependence and parameter dependence. The *idea* of Jarrett's analysis is that a specific product form of the hidden joint probability (such as local factorisation), which is a *complex* independence condition, can be analysed by *pairwise* independences (such as outcome independence or parameter independence). Our new constraint for quantum non-locality, probabilistic Bell contextuality, is a conjunction of two disjunctions of several product forms and, hence, a complex independence condition as well. So we can apply Jarrett's idea to our new case and understand 'analysis' as providing an expression in terms of pairwise probabilistic independences which is equivalent to the new constraint.

Providing an analysis of the new stronger constraint will make explicit why the failure of purely outcome dependent theories to violate Bell inequalities (cf. section 2.2) rules out that outcome dependence can be responsible for violating Bell inequalities. In this way it will reveal Jarrett's distinction between outcome dependent and parameter dependent theories to be highly misleading. I first recall shortly Jarrett's analysis and introduce an appropriate set of independences, which will serve as analysantia. Then I shall develop an analysis for each of the classes (H_i^α) and subsequently of the new probabilistic constraint for quantum non-locality.

3.1 Jarrett's analysis

Jarrett (1984) had the idea that one can be more explicit about the probabilistic nature of quantum non-locality by analyzing the probabilistic statement local factorisation (ℓF) in terms of pairwise conditional probabilistic independences. By a 'pairwise conditional probabilistic independence' I mean the fact that a random variable $\boldsymbol{x}$ is independent of another $\boldsymbol{y}$ given a conjunction of further variables $\boldsymbol{z}$. This is said to be true iff for all values of the variables the joint probability over the variables makes the following equation true:

$$P(x|yz) = P(x|z) \tag{6}$$

The independence is noted as $I(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})$. If, however, there is at least one set of values for which (6) does not hold, the variables $\boldsymbol{x}$ and $\boldsymbol{y}$ are called dependent given $\boldsymbol{z}$, and this probabilistic dependence is noted as $\neg I(\boldsymbol{x}, \boldsymbol{y}|\boldsymbol{z})$.

Jarrett uses three pairwise independences: 'outcome independence' is defined as

$I(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{ab\lambda})$ and 'parameter independence' as a conjunction of two independences, $I(\boldsymbol{\alpha},\boldsymbol{b}|\boldsymbol{a\lambda})\wedge I(\boldsymbol{\beta},\boldsymbol{a}|\boldsymbol{b\lambda})$. (Originally, Jarrett denotes these independences as 'completeness' and 'locality' respectively, but we shall use the now established names.) Jarrett proved mathematically that

(P7) Local factorisation is equivalent to the conjunction of outcome independence and parameter independence:

$$(\ell\mathrm{F}) \leftrightarrow I(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{ab\lambda}) \wedge I(\boldsymbol{\alpha},\boldsymbol{b}|\boldsymbol{a\lambda}) \wedge I(\boldsymbol{\beta},\boldsymbol{a}|\boldsymbol{b\lambda}) \tag{7}$$

From (C1), the conclusion of the standard Bell argument that local factorisation fails, and (P7) he concluded that

(C2) Outcome dependence or parameter dependence holds:

$$\neg I(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{ab\lambda}) \vee \neg I(\boldsymbol{\alpha},\boldsymbol{b}|\boldsymbol{a\lambda}) \vee \neg I(\boldsymbol{\beta},\boldsymbol{a}|\boldsymbol{b\lambda}) \tag{8}$$

which is the *analysis of the probabilistic constraint following from the standard Bell argument* ('Jarrett's analysis'). The analysis is correct, but since (as we have seen) the analysandum, the conclusion of the standard Bell argument, is inappropriately weak, its result is not as informative as it could be. In fact, we will shortly see by the analysis of the new stronger result that it is deceptive about the nature of quantum non-locality.

3.2 Pairwise independences

Aiming to analyze the new probabilistic constraint for quantum non-locality we first have to get an overview which concepts can play the role of the analysantia. In table 2 I introduce those nine pairwise independences which will be relevant. Among the relevant independences we find usual outcome independence, $I(\boldsymbol{\alpha},\boldsymbol{\beta}|\boldsymbol{ab\lambda})$, as well as $I(\boldsymbol{\alpha},\boldsymbol{b}|\boldsymbol{a\lambda})$, one independence of the conjunction which is usually called 'parameter independence'. Here we see a first problem with the standard names: how shall we call the latter if its conjunction with $I(\boldsymbol{\beta},\boldsymbol{a}|\boldsymbol{b\lambda})$ is called 'parameter independence'? I have tried to stay as close to the standard names as possible, but obviously further qualifications are needed. My suggestion is to continue to use the name 'parameter independence' for all independences between an outcome and its distant parameter, but to add the outcome in question, namely 'α-parameter independence' or 'β-parameter independence' respectively. Further differentiation in the nomenclature is required by the fact that there is another α-parameter independence in the table, $I(\boldsymbol{\alpha},\boldsymbol{b}|\boldsymbol{\beta a\lambda})$, which differs from the one already mentioned in the conditional variables (it additionally includes the outcome $\boldsymbol{\beta}$). Such independences of the same type but with different conditional variables are different independences and are in general logically independent of another: one can hold or not irrespective of whether the other does or does not. (One can show that only if one involves more than two independences logical restrictions appear.) I discern them by indices, e.g. the former is called 'α-parameter independence$_2$', the latter 'α-parameter independence$_1$'. Of course, there are further α-parameter independences (namely those conditional on $\boldsymbol{\beta\lambda}$ and $\boldsymbol{\lambda}$) which, however, do not play any role for the analysis here.

Similarly to the parameter independences I define local parameter independences (see table 2), which instead of the independence of an outcome on its distant parameter (e.g. $\boldsymbol{\alpha}, \boldsymbol{b}$) claim the independence of an outcome on its *local* parameter (e.g. $\boldsymbol{\alpha}, \boldsymbol{a}$). Besides these new names I have also introduced short labels for each independence, which we will mainly use in the following.

Table 2: Definition of conditional independences

independence	standard name	new name	label
$I(\boldsymbol{\alpha}, \boldsymbol{\beta}\|\boldsymbol{ab\lambda})$	outcome independence	outcome independence$_1$	(OI$_1$)
$I(\boldsymbol{\alpha}, \boldsymbol{b}\|\boldsymbol{\beta a\lambda})$	–	α-parameter independence$_1$	(PI$_1^\alpha$)
$I(\boldsymbol{\alpha}, \boldsymbol{b}\|\boldsymbol{a\lambda})$	[part of] parameter ind.	α-parameter independence$_2$	(PI$_2^\alpha$)
$I(\boldsymbol{\beta}, \boldsymbol{a}\|\boldsymbol{\alpha b\lambda})$	–	β-parameter independence$_1$	(PI$_1^\beta$)
$I(\boldsymbol{\beta}, \boldsymbol{a}\|\boldsymbol{b\lambda})$	[part of] parameter ind.	β-parameter independence$_2$	(PI$_2^\beta$)
$I(\boldsymbol{\alpha}, \boldsymbol{a}\|\boldsymbol{\beta b\lambda})$	–	α-local parameter independence$_1$	(ℓPI$_1^\alpha$)
$I(\boldsymbol{\alpha}, \boldsymbol{a}\|\boldsymbol{b\lambda})$	–	α-local parameter independence$_2$	(ℓPI$_2^\alpha$)
$I(\boldsymbol{\beta}, \boldsymbol{b}\|\boldsymbol{\alpha a\lambda})$	–	β-local parameter independence$_1$	(ℓPI$_1^\beta$)
$I(\boldsymbol{\beta}, \boldsymbol{b}\|\boldsymbol{a\lambda})$	–	β-local parameter independence$_2$	(ℓPI$_2^\beta$)

Having introduced these new concepts we are now in a position to clearly see one of the sources of confusion in the standard discussion. 'Outcome dependence or parameter dependence' does *not necessarily* mean that if you accept outcome dependence you can avoid parameter dependence in the sense of *any kind* of dependence of an outcome on its distant parameter (conditional on whatever variables). The slogan just says that in this case you can avoid parameter dependence in the usual sense of $\neg(\mathrm{PI}_2^\alpha) \vee \neg(\mathrm{PI}_2^\beta)$, while other kinds of parameter dependences like $\neg(\mathrm{PI}_1^\alpha)$ might still hold! And indeed the analysis of the new constraint will yield that at least one of the two parameter dependences $\neg(\mathrm{PI}_1^\alpha)$ and $\neg(\mathrm{PI}_2^\beta)$ *must* hold. Parameter dependence in this broader sense cannot be avoided but will turn out to be a necessary condition for violating the Bell inequalities.

3.3 Analyzing the classes

With these pairwise independences we can now attempt to analyse each class of probability distributions. For the analysis of the classes (H_i^α) in table 1 we shall need the first five independences in table 2 (the other four independences plus outcome independence$_1$ are only used for the analysis of the classes (H_j^β); see below). We have noted the corresponding dependences in the bottom line of table 1, i.e. each dependence is associated with one of the columns II–VI. The idea is that the dependence holds in a class if the column of that class contains '1'. Otherwise, i.e. if it contains '0', the corresponding independence holds. The result of this analysis is stated by the following theorem:

> **Theorem 3**: Each class in table 1 is equivalent to the conjunction of the specific pattern of independences (see the bottom line of the table) indicated by 0's in columns II–VI.

The proof of theorem 3 can be found in the mathematical appendix.

The theorem means that each pattern of independences corresponds to exactly one of the classes, e.g.

$$(\mathrm{H}^{\alpha}_{7}) \leftrightarrow (\mathrm{PI}^{\alpha}_{2}) \wedge (\ell\mathrm{PI}^{\alpha}_{2}). \tag{9}$$

One can see from the table that *each of the five independences corresponds to exactly one of the five variables in the conditionals of the factors*: if a certain independence *holds*, the corresponding variable does *not* appear (and vice versa), and if a certain independence *fails*, the corresponding variable does *appear* (and vice versa). Specifically, if (OI_1) holds, the first factor of the hidden joint probability does not involve the other outcome $\boldsymbol{\beta}$ (and vice versa), and if it does not, the first factor includes it (and vice versa). Similarly, $(\mathrm{PI}^{\alpha}_{1})$ and $(\ell\mathrm{PI}^{\alpha}_{1})$ correspond to the distant and the local parameter in the *first* factor respectively, while $(\mathrm{PI}^{\alpha}_{2})$ and $(\ell\mathrm{PI}^{\alpha}_{2})$ are linked to the distant and the local parameter in the *second* factor respectively. So the holding or failure of each of the five independences has a very well defined impact on the product form of the hidden joint probability (and vice versa), and the conjunction of *all* independences which hold according to a certain probability distribution determines its product form, i.e. its class (and vice versa).[11]

3.4 Analysis of the stronger conclusion

We can now use the analysis of the single classes to formulate the analysis of the new, stronger conclusion. This will provide us with sufficient and necessary conditions for a class being able to violate Bell inequalities. We had found that quantum non-locality is the failure of all local$^{\alpha}$, weakly non-local$^{\alpha}$, local$^{\beta}$ and weakly non-local$^{\beta}$ classes (C1$''$) and that these classes are characterized by the fact that their constituting product forms involve *at most one setting (parameter) in each of its factors.* Let us first give an analysis of the local$^{\alpha}$ and weakly non-local$^{\alpha}$ classes. Our analysis of the single classes $(\mathrm{H}^{\alpha}_{1})$–$(\mathrm{H}^{\alpha}_{32})$ has revealed that each variable in the conditionals of the factors corresponds to exactly one of the five independences in table 1. The distant parameter in the first factor corresponds to α-parameter independence$_1$, $(\mathrm{PI}^{\alpha}_{1})$, and the local parameter to α-local parameter independence$_1$, $(\ell\mathrm{PI}^{\alpha}_{1})$. So the first factor involves at most one parameter if and only if at least one of these independences holds, $(\mathrm{PI}^{\alpha}_{1}) \vee (\ell\mathrm{PI}^{\alpha}_{1})$. Similarly, at most one parameter appears in the second factor iff β-parameter independence$_2$ or β-local parameter indepence$_2$ hold, $(\mathrm{PI}^{\beta}) \vee (\ell\mathrm{PI}^{\beta})$. So we have found the following equivalence:

[11] Note that according to table 1 local factorisation is analysed as $(\mathrm{H}^{\alpha}_{29}) \leftrightarrow (\mathrm{OI}_1) \wedge (\mathrm{PI}^{\alpha}_{1}) \wedge (\mathrm{PI}^{\beta}_{2})$, while according to Jarrett it is $(\mathrm{H}^{\alpha}_{29}) \leftrightarrow (\mathrm{OI}_1) \wedge (\mathrm{PI}^{\alpha}_{2}) \wedge (\mathrm{PI}^{\beta}_{2})$, i.e. in Jarrett's claim $(\mathrm{PI}^{\alpha}_{1})$ is replaced by $(\mathrm{PI}^{\alpha}_{2})$. Given that (OI_1) holds, the replacement is logically correct, because one can show that $(\mathrm{OI}_1) \wedge (\mathrm{PI}^{\alpha}_{1}) \leftrightarrow (\mathrm{OI}_1) \wedge (\mathrm{PI}^{\alpha}_{2})$. So the two analyses of $(\mathrm{H}^{\alpha}_{29})$ are equivalent.

(P7′a) The disjunction of local$^\alpha$ and weakly non-local$^\alpha$ classes is equivalent to the fact that $\boldsymbol{\alpha}$ is independent$_1$ of at least one parameter and $\boldsymbol{\beta}$ is independent$_2$ of at least one parameter:

$$\Big(\bigvee_{i=15}^{32}(\mathrm{H}_i^\alpha)\Big) \leftrightarrow \Big[\Big((\mathrm{PI}_1^\alpha) \vee (\ell\mathrm{PI}_1^\alpha)\Big) \wedge \Big((\mathrm{PI}_2^\beta) \vee (\ell\mathrm{PI}_2^\beta)\Big)\Big]$$

In a very similar way as we have proceeded for the classes (H_1^α)–(H_{32}^α) one can find an analysis for the classes (H_1^β)–(H_{32}^β) (remember the table which is symmetric to table 1 in swapping the outcomes and the parameters and apply all considerations mutatis mutandis):

(P7′b) The disjunction of local$^\beta$ and weakly non-local$^\beta$ classes is equivalent to the fact that $\boldsymbol{\beta}$ is independent$_1$ of at least one parameter and $\boldsymbol{\alpha}$ is independent$_2$ of at least one parameter:

$$\Big(\bigvee_{i=15}^{32}(\mathrm{H}_i^\beta)\Big) \leftrightarrow \Big[\Big((\mathrm{PI}_1^\beta) \vee (\ell\mathrm{PI}_1^\beta)\Big) \wedge \Big((\mathrm{PI}_2^\alpha) \vee (\ell\mathrm{PI}_2^\alpha)\Big)\Big]$$

Since according to the conclusion of the stronger Bell argument (C1″) the disjunction of all local$^\alpha$, weakly non-local$^\alpha$, local$^\beta$ and weakly non-local$^\beta$ classes *fails*, the negation of the disjunction of (P7'a) and (P7'b) finally yields the analysis of (C1″):

(C2′) $\boldsymbol{\alpha}$ depends$_1$ on both parameters or $\boldsymbol{\beta}$ depends$_2$ on both parameters and $\boldsymbol{\beta}$ depends$_1$ on both parameters or $\boldsymbol{\alpha}$ depends$_2$ on both parameters:

$$\Big[\Big(\neg(\mathrm{PI}_1^\alpha) \wedge \neg(\ell\mathrm{PI}_1^\alpha)\Big) \vee \Big(\neg(\mathrm{PI}_2^\beta) \wedge \neg(\ell\mathrm{PI}_2^\beta)\Big)\Big] \wedge \Big[\Big(\neg(\mathrm{PI}_1^\beta) \wedge \neg(\ell\mathrm{PI}_1^\beta)\Big) \vee \Big(\neg(\mathrm{PI}_2^\alpha) \wedge \neg(\ell\mathrm{PI}_2^\alpha)\Big)\Big]$$

While the conclusion (C1″) of the stronger Bell argument was in terms of classes, here we have the equivalent expression, the analysis, in terms of pairwise independences. It is a rather complex logical expression whose meaning and implications are not easy to grasp. A first understanding might be attained by making explicit how this analysis of the conclusion (C1″) is also an analysis of the equivalent conclusion (C1‴), which says that the conjunction of strongly non-local$^\alpha$ and strongly non-local$^\beta$ classes holds. These classes were characterized by the fact that at least one of the factors in each product form must involve both parameters and this is exactly what (C2′) says: The first term in the first disjunction, $\neg(\mathrm{PI}_1^\alpha) \wedge \neg(\ell\mathrm{PI}_1^\alpha)$ ('α-double parameter dependence$_1$'), guarantees a dependence on both parameters in the first factor of the product forms (H_i^α), the second term in the first disjunction, $\neg(\mathrm{PI}_2^\beta) \wedge \neg(\ell\mathrm{PI}_2^\beta)$ ('β-double parameter dependence$_2$'), implies a similar fact for the second factor of these forms, and analogously, the second disjunction entails a dependence on both parameters in at least one of the factors of the product forms (H_i^β) (and vice versa).

So the analysis involves double parameter dependences for each outcome in two different forms, either conditional on all other variables (double parameter dependence$_1$)

or conditional on all other variables excluding the other outcome (double parameter dependence$_2$). The logic of the expression has it that these can hold in different combinations, but whichever combination does, there is one thing that necessarily follows if (C2′) is true:

(C3) Double parameter dependence: at least one of the outcomes depends probabilistically on both parameters (in at least one of the forms double parameter dependence$_1$ or double parameter dependence$_2$).

For one *can* avoid that *one* of the outcomes is double parameter dependent$_1$ and double parameter dependent$_2$, but then it follows that the respective other outcome must be double parameter dependent$_1$ as well as double parameter dependent$_2$. Of course, you can also have mixed cases in which both outcomes are double parameter dependent (in one or both of the two forms), but in any case you have double parameter dependence of at least one of the outcomes.

So we have found two results: the precise probabilistic analysis of the new stronger conclusion (C2′) and a general feature of and deriving from that analysis (C3), that at least one of the outcomes must be double parameter dependent. Since the conclusion is a necessary condition for EPR/B correlations (if autonomy and nearly perfect (anti-)correlations hold), double parameter dependence of at least one of the outcomes, which is implied by quantum non-locality, is a *necessary* condition for EPR/B correlations as well: whenever we find that EPR/B correlations hold, double parameter dependence (C3) must hold as well. So given that measurement results in our world yield EPR/B correlations (and assuming autonomy), we can be sure that at least one of the outcomes depends both on the local as well as on the distant parameter.

On the other hand, since here we have derived an analysis of a conclusion following from the violation of Bell inequalities, neither the analysis (C2′) nor its consequence double parameter dependence (C3), is *sufficient* for the violation of Bell inequalities. If, according to a certain probability distribution, an outcome depends on both parameters in the sense of (C2′) the correlations between the two wings *might* be strong enough to violate Bell inequalities—but they need not be (see section 2.6). However, we also know (from that section) that the conclusion of the argument is *sufficient for a class to be able to violate Bell inequalities*, in the sense that if a class fulfills the conditions mentioned in the conclusion, there is at least one probability distribution in that class which violates the inequalities. Hence, the analysis (C2′) is also a sufficient condition for a class to be able to violate Bell inequalities (but not its implication C3).

It is obvious that the result of this new analysis differs from Jarrett's. There are two main differences: first, while Jarrett's analysis suggests that either the outcome or an outcome and its distant parameter depend on another, according to the new analysis it seems that one cannot avoid some kind of parameter dependence. Second, the concept of outcome dependence does not appear in the new analysis at all. What do these discrepancies mean? We now have to compare our result in detail to Jarrett's and the received view.

4 Consequences

4.1 Shortcomings of Jarrett's analysis

(1) The main message of our new result is that given EPR/B correlations and autonomy *one cannot avoid some kind of dependence between at least one of the outcomes and both parameters (C3).* This is a necessary condition for the violation of Bell inequalities according to my new analysis. Jarrett's analysis, in contrast, does not bring out this essential requirement: from his result 'outcome dependence or parameter dependence' one just cannot see that, necessarily, there must be some kind of double parameter dependence. This is a first shortcoming of Jarrett's analysis.

(2) A second problematic feature of Jarrett's result consists in the fact that unlike our new result it seems to suggest that one can avoid *any* dependence of the outcomes on their distant settings if the outcomes depend on another. If this suggestive interpretation were the correct reading of Jarrett's result, it would be plainly wrong. However, this is not what it literally says. 'Parameter dependence' here does not mean *any* kind of parameter dependence but a very specific kind, namely parameter dependence$_2$. Saying that one can avoid this specific kind does not mean that there is no dependence of the outcomes on their distant parameters at all. Our presentation of different kinds of parameter dependences (see table 2) has made explicit that parameter dependence$_2$ is only one among several kinds, all of which might hold if parameter dependence$_2$ fails. So in a careful literal reading Jarrett's result does not contradict our result (C2′), that one can avoid parameter dependence$_2$ only if parameter dependence$_1$ holds.

While there is no logical inconsistency between the two analyses, this reasoning shows that Jarrett's result is *liable to be misunderstood to its non-literal false sense*, that one can avoid *any* kind of parameter dependence if outcome dependence holds. In this sense, Jarrett's result is misleading. In fact, it seems that Jarrett's result has to a large extent received this unfortunate interpretation. There is a bunch of literature about quantum non-locality (on any level, whether causal, spatio-temporal or metaphysical) which is based on Jarrett's distinction, and which discusses in detail what outcome dependence or parameter dependence would amount to, the preferred solution being outcome dependence without parameter dependence. But in most cases this makes only sense, if one believes that by neglecting parameter dependence one can avoid *any* kind of parameter dependence! If the authors in that debate would have known that one cannot avoid some kind of parameter dependence anyway, they would surely not have spent so much time on finding arguments why outcome dependence rather than parameter dependence holds. Much of the debate based on Jarrett's analysis seems to adhere to the wrong non-literal reading of Jarrett's result.

(3) Thirdly, even in a correct literal reading, Jarrett's result is problematic: understood as providing insights about quantum non-locality (on a probabilistic level), *it is highly deceptive because it rests on inappropriate categories.* This point becomes clear, if one investigates how Jarrett's analysantia outcome dependence$_1$, α-parameter dependence$_2$ and β-parameter dependence$_2$ relate to the new concept of quantum non-locality. Here is, first, how they do *not* relate: Jarrett's analysis of the weaker concept

is a disjunction of these three dependences and it could have been that the analysis of the stronger concept just cancels one or two of the elements in the disjunction, revealing them as options which are not really available. For instance it might have been that the new analysis yields just $\neg(\mathrm{PI}_2^\alpha) \vee \neg(\mathrm{PI}_2^\beta)$, cancelling $\neg(\mathrm{OI}_1)$. However, it turns out that this is not the case. The logical structure of the new analysis is not just a simplification of the former, but, in fact, is much more complicated involving new concepts (parameter dependence$_1$, local parameter dependence$_1$ and local parameter dependence$_2$) and not involving others (outcome dependence$_1$). This suggests that *Jarrett's categories outcome dependence$_1$ and parameter dependence$_2$ cannot capture the conclusion of the stronger Bell argument.*

Table 3: Jarrett's classes of possible probability distributions

Label	$\neg(\mathrm{OI}_1)$	$\neg(\mathrm{PI}_2^\alpha) \vee \neg(\mathrm{PI}_2^\beta)$	Notes	
(J_1)	1	1		quantum non-locality
(J_2)	0	1	Bohm	
(J_3)	1	0	QM	
(J_4)	0	0		locality

To make this explicit, consider the partition of the probability distributions according to the dependences in Jarrett's analysis (table 3). There are four classes, which I call 'Jarrett's classes' and label as (J_1)–(J_4). Any of the 32 possible classes from table 1 must fall into one of Jarrett's coarse-grained classes. While the local classes belong to (J4), any of the classes (J1)–(J3) includes both weakly and strongly non-local classes. So Jarrett's non-local classes, which are assumed to be able to violate Bell inequalities, *mix probability distributions which can with such which cannot* (see fig. 2). *They do not cut the probability distributions at their natural joints!*

This means that neither outcome dependence$_1$ nor parameter dependence$_2$ are necessary or (contrary to Jarrett's analysis) sufficient for characterising quantum non-locality according to the new conclusion of the Bell argument. Providing, for instance, the information that a certain probability distribution is outcome dependent$_1$ does not tell you whether it can violate Bell inequalities or not. The crucial fact is whether double parameter dependence of a certain kind holds. α-parameter dependence$_2$ and β-parameter dependence$_2$ at least play a *certain* role in this complex condition; outcome dependence$_1$, however, does *not* play any role in the formulation of the new conclusion. Not being able to capture the new conclusion, we conclude that the partition according to Jarrett's categories outcome dependence$_1$ and parameter dependence$_2$ is inappropriate or unnatural for the analysis of quantum non-locality.

So it seems that a significant amount of the debate after Jarrett's paper which has focused on the question of the formal, physical and metaphysical differences between outcome dependence$_1$ and parameter dependence$_2$, in order to decide which of the two does hold, is misguided. *'Outcome dependence or parameter dependence?' is just the wrong question if one wants to explore deeper into the nature of quantum non-locality,*

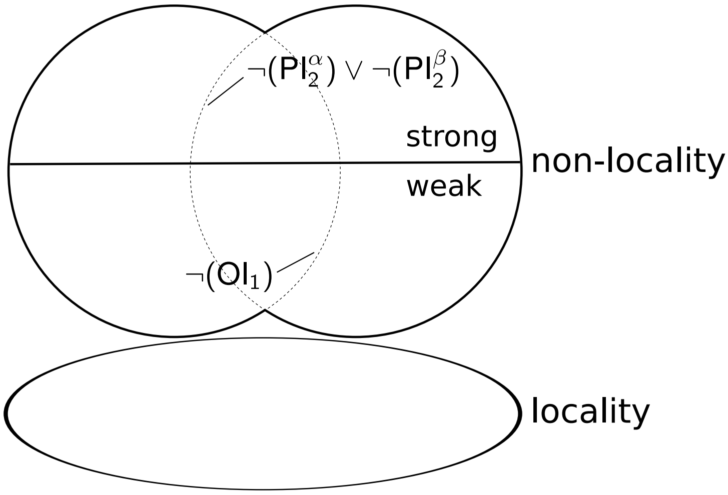

Figure 2: Outcome dependence$_1$ and parameter dependence$_2$ vs. weak and strong non-locality. 'Strong non-locality' means *strong$^\alpha$* <u>*and*</u> *strong$^\beta$ non-locality* (i.e. those distributions which can violate Bell inequalities), while 'weak non-locality' means *weak$^\alpha$* <u>*or*</u> *weak$^\beta$ non-locality* (i.e. those non-local distributions which imply Bell inequalities).

because each of the two options subsumes probability distributions which can and such which cannot violate Bell inequalities. Making this question a guide to quantum non-locality is like asking whether those humans which can get pregnant have dark or fair hair. Rather, the natural question, the new analysis shows, is which of the outcomes is double parameter dependent and whether it is double parameter dependent$_1$ or double parameter dependent$_2$.

In sum, Jarrett's analysis is not false if one understands it literally; it is however, strongly liable to be misunderstood (and has widely been misunderstood). Being based on the standard Bell argument, it is not as informative as one can get with Bell's theorem. Moreover, the categories it uses do not cut the problem at its natural joints.

4.2 Failure of the received view: outcome dependence cannot explain the violation of Bell inequalities

We have seen that Jarrett's analysis, though liable to be misunderstood, is nevertheless true in a literal sense. Based on Jarrett's analysis there has emerged the received view that quantum non-locality on a probabilistic levels is outcome dependence (and not parameter dependence). This view, however, is provably wrong. Its failure is not obvious; given just Jarrett's analysis it might have been true. But given our stronger result here, one can show that it cannot be true.

(1) How can it be that the received view, whose central claim, that quantum non-locality is outcome dependence, is based on a true analysis, turns out to be wrong? The reason is that this claim is stronger than Jarrett's analysis licenses. Jarrett's analysis

just says that *outcome dependence or parameter dependence* is a *necessary* condition for violating Bell inequalities. But claiming that outcome dependence (and parameter independence) holds commits one to the stronger claim that outcome dependence (without parameter dependence) can *explain* the violation of the Bell inequalities.

What would it mean that outcome dependence can explain a violation? By our previous considerations it is clear that it would be too strong to require that every theory (or probability distribution) which is outcome dependent violates the inequalities. For we have seen that a dependence in itself never suffices to guarantee a violation; only if dependences have a certain numerical strength, a violation can be inferred (see section 2.6). What we should require, however, when someone claims that outcome dependence can explain the violation of the Bell inequalities, is that if a class of theories is outcome dependent, there are at least *some* probability distributions belonging to that class that violate the inequalities. In the terminology we have used in this paper that claim is to say:

(i) Outcome dependence is sufficient for a class being able to violate Bell inequalities.

We shall now show that in the light of our new results this claim truns out to be wrong. It is both wrong in a non-intended literal as well as in a reasonable non-literal reading; in a very liberal third reading it is true, but the claim ceases to be about outcome dependence.

Taken literally, (i) would mean that class (H_{26}^{α}),

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta\lambda)P(\beta|\lambda), \tag{10}$$

whose *only* dependence (among the five relevant dependences) is outcome dependence, is sufficient for being able to violate Bell inequalities. This assertion is obviously wrong because the class trivially implies Bell inequalities (due to the non-appearance of the settings in the product form). However, though (as far as I can see) explicit precise statements are missing, it is also very likely that defenders of the received view never intended their claim to be read in this way: the class does not seem physically reasonable due to a lack of dependence on the local settings.

Rather, what proponents of the received view more likely had in mind is the following:

(ii) Outcome dependence is sufficient for a class being able to violate Bell inequalities *if appropriate local parameter dependences hold* ($\neg(\ell\mathrm{PI}_1^{\alpha})$ and $\neg(\ell\mathrm{PI}_2^{\alpha})$ in the case of α-classes).

That would amount to the claim that class (H_{16}^{α}), defined by

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda), \tag{11}$$

is able to violate Bell inequalities (the class of purely outcome dependent theories). However, we have shown in section 2.2 that this class (non-trivially) implies Bell inequalities as well. So if that is what proponents of outcome dependence had in mind, it is not correct either.

A proponent of outcome dependence might wish to defend her position by claiming that the reading we have just given is still too literal. In fact, she might claim, what proponents of outcome dependence really mean is:

(iii) Outcome dependence is sufficient for being able to violate Bell inequalities if appropriate local parameter dependences hold ($\neg(\ell\mathrm{PI}_1^\alpha)$ and $\neg(\ell\mathrm{PI}_2^\alpha)$ in the case of α-classes) *and an appropriate non-local parameter dependence holds* ($\neg(\mathrm{PI}_1^\alpha)$ in the case of α-classes).

These dependences are realized in class (H_3^α),

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta ab\lambda)P(\beta|b\lambda), \tag{12}$$

which essentially describes the quantum mechanical distribution (if one neglects the dependence on the hidden variable), and in class (H_1^α). The non-local parameter dependence in the first factor, (PI_1^α), it might be said, is required, because from a physical perspective the probability distribution has to reflect which measurement has been carried out at the other wing. That dependence is unproblematic because—unlike usual parameter dependence (PI_2^α)—it conditions on the distant outcome and hence is not in tension with relativity, because it cannot be used to send signals.

In this strongly non-literal reading claim (i), that outcome dependence is sufficient for a class being able to violate Bell inequalities, finally becomes true: both classes are among those that can violate the Bell inequalities. However, this reading is highly problematic. The main reason is not that it makes (i) a highly misleading claim (because it smuggles in a further non-local dependence via a suppressed background assumption, hiding the essential point that another non-local dependence is included as well); while this clearly is a weird way of understanding a claim, it still does not make (i) in reading (iii) wrong.

Rather, the main problem about this reading concerns a different, stronger point, which can be seen as follows. In an explicit form, the most promising claim the defender of outcome dependence can advance is:

(iii′) The conjunction of outcome dependence and parameter dependence$_1$ ($\neg(\mathrm{PI}_1^\alpha)$ in the case of α-classes) is sufficient for being able to violate Bell inequalities if appropriate local parameter dependences hold ($\neg(\ell\mathrm{PI}_1^\alpha)$ and $\neg(\ell\mathrm{PI}_2^\alpha)$ in the case of α-classes).

While literally true, the interesting point about this claim is that the condition of outcome dependence is not needed at all, in order for it to be true. As can be seen from table 1, outcome dependence can be dropped without making the claim false: (H_9^α) which differs from (H_3^α) just in that it is outcome independent still is among the classes that are able to violate Bell inequalities (and the same is true of (H_6^α) and (H_1^α)). In contrast, parameter dependence$_1$ is not a redundant constraint of the present sufficiency claim: (H_{16}^α) which differs from (H_3^α) just in that it is α-parameter independent$_1$ is emphatically *not* among the classes that are able to violate Bell inequalities. So the true, non-redundant, most general version of the present sufficiency claim is:

(iii″) α-parameter dependence$_1$ is sufficient for being able to violate Bell inequalities if appropriate local parameter dependences hold ($\neg(\ell\mathrm{PI}^\alpha_1)$ and $\neg(\ell\mathrm{PI}^\alpha_2)$ in the case of α-classes).

In this last formulation, however, outcome dependence has disappeared from the claim altogether: it has turned out to be a *redundant* part of the original sufficiency claim (iii′). This result is problematic for the defender of outcome dependence for two reasons. First, being redundant, outcome dependence does not play any role for characterizing classes that are able to violate Bell inequalities, and hence cannot explain the violation of the inequalities. Second, not involving outcome dependence, the new sufficiency claim without redundancy (3″) cannot be viewed as an interpretation of (i) any more (it is not connected to its original formulation (i) besides that it is logically consistent with it)—which reduces this attempt of interpretation ad absurdum.

In sum, claim (i) which is at the heart of the received view, is not substantiated in any reading. It is either plainly false (as in the first two readings) or, if it is true, involves outcome dependence only in a redundant way. In any case, it turns out that outcome dependence is not sufficient for a class being able to violate Bell inequalities in any reasonable sense. Note that this does not mean that outcome dependence in fact does not hold—it might. But even if it does it cannot be regarded as being the crucial probabilistic dependence between the wings. Probabilistic quantum non-locality is not essentially outcome dependence. In this sense, the received view misjudges the role of outcome dependence for the violation of Bell inequalities. This is another main result of the present paper.

(2) We might note that claims about usual parameter dependence$_2$ do not lead into similar difficulties. Of course, a literal reading

(iv) Parameter dependence$_2$ is sufficient for a class being able to violate Bell inequalities.

is not true either (because it would wrongly imply that class (H^α_{28}) can violate Bell inequalities). But a plausible interpretation of this claim as

(v) Parameter dependence$_2$ is sufficient for a class being able to violate Bell inequalities *if appropriate local parameter dependences hold* ($\neg(\ell\mathrm{PI}^\alpha_1)$ and $\neg(\ell\mathrm{PI}^\alpha_2)$ in the case of α-classes).

is true (because the classes (H^α_1), (H^α_5), (H^α_6) and (H^α_{12}) can violate Bell inequalities). So we have found an asymmetry in the possible positions proposed by Jarrett's analysis: while it is reasonable to say that parameter dependence$_2$ (given appropriate local parameter dependences) is sufficient for a class being able to violate Bell inequalities, the same is not true for outcome dependence.

(3) Finally, we should explain how quantum mechanics fits into this picture. According to the standard view, quantum mechanics is regarded as the paradigm of an outcome dependent (and parameter independent) theory violating Bell inequalities. So, is it not a counterexample to my result that theories necessarily have to depend on the

distant parameter in order to violate the inequalities? And to the claim that outcome dependence cannot explain the violation of the inequalities?

It is true, quantum mechanics is well known to be 'outcome dependent and parameter independent', but again this is not to be understood that according to quantum mechanics there is no probabilistic dependence of an outcome on the distant parameter at all. In fact, it is easy to check, which independences hold according to quantum mechanics: one can calculate all relevant conditional probabilities from the quantum mechanical probability distribution for the EPR/B experiment(Corr). A simple comparison of these probabilities then shows which of the independences hold and which do not, and it turns out that besides being outcome dependent, quantum mechanics is parameter dependent$_1$, $\neg(\mathrm{PI}_1^{\alpha})$ and $\neg(\mathrm{PI}_1^{\beta})$, so according to quantum mechanics each outcome *does* depend on its distant parameter!

This parameter dependence in quantum mechanics is not as surprising as it may seem since, according to the formalism, the measurement direction at A determines the possible collapsed states at B and the actual outcome at A only determines in which of the (two) possible states the photon state at B collapses. So contrary to what the standard talk suggests, quantum mechanics is parameter dependent (in some sense), and it is important to see that it is as well local parameter dependent$_1$, $\neg(\ell\mathrm{PI}_1^{\alpha})$ and $\neg(\ell\mathrm{PI}_1^{\beta})$ (while it is local parameter independent$_2$, $(\ell\mathrm{PI}_2^{\alpha})$ and $(\ell\mathrm{PI}_2^{\beta})$), because then, the quantum mechanical distribution fulfills the requirement of the stronger Bell argument by rendering the first terms of the two disjunctions in (C2′) true. If my argument in this paper is true, it cannot be otherwise. For if quantum mechanics were not parameter dependent in this double sense, it could not (as it does) violate Bell inequalities.

The example of quantum mechanics also illustrates that my results do not mean that all outcome dependent theories imply the inequalities; it just says that a probabilistic dependence between the outcomes *per se* does not suffice to explain a violation. In conjunction with parameter dependence, outcome dependence might even contribute to a violation of the inequality; but in contrast to the former it cannot per se explain a violation. Assuming that outcome dependence can do this explanatory job is the main error of the received view.

4.3 Resolving the Jarrett-Maudlin debate

(1) Opposed to the received view, there is another position concerning quantum nonlocality, whose result seems to agree with ours. Maudlin (1994, ch. 6; cf. also a recent refinement by Pawlowski et al. 2010) proves that, in order to reproduce the EPR/B correlations, at least one of the outcomes must depend on *information* about both settings. Since (Shannon mutual) information implies correlation,[12] one can infer that at least one of the outcomes must depend probabilistically on both settings—and this is exactly my result (C3).

(2) This convergence is good news, both for Maudlin's as well as my argument here, because the two investigations use very different different methods, and two different

[12] Shannon mutual information, which is the concept that Maudlin's and Pawlowski et al.'s considerations essentially are based on, is a measure for the strength of a correlation.

methods yielding the same result are evidence for the stability of a claim. On the one hand, Maudlin's approach is an information theoretic investigation proceeding from the EPR/B correlations without invoking Bell's theorem. In contrast, my argument in this paper approaches quantum non-locality via Bell inequalities and probabilistic analysis, i.e. it stands methodologically in the Bell-Jarrett tradition, which has started and shaped the debate. In this sense, the two arguments are methodologically quite different, and it is fair to say that my approach here confirms Maudlin's result by a different method.

(3) We should note that our result is in one sense weaker and in one sense stronger than the information theoretic one. It is weaker because it is purely qualitative: it just says which probabilistic dependences are required, but it is tacit about how strong the correlations have to be in order to violate Bell inequalities. In section 2.6 I have argued that such qualitative results can only be necessary conditions for a violation, because having the right dependences for violating Bell inequalities does not mean that the inequalities are in fact violated. In contrast, sufficient criteria must involve conditions on the strength of the correlations, and the information theoretic approach derives such criteria by calculating the amount of information (the quantitative strength of the correlations) that is required in order to reproduce EPR/B correlations; these are important results.

In another sense, however, our result is also stronger than the information theoretic one. Maudlin's analysis just implies *some kind of* dependence of an outcome on its distant parameter. But which precisely? We have seen that there are different kinds of parameter dependences, which differ in the conditional variables. Especially it cannot be an unconditional parameter dependence because that would contradict the empirical distribution. So which are the ones that are required? My detailed result (C2′) makes precise which kind of parameter dependences are required. Present information theoretic results do not provide a similar detailed characterization. However, a complete list of which dependences exactly hold or fail might be important for the discussion of quantum non-locality on other levels: causal inference (cf. Spirtes 1993; Pearl 2000), for instance, is very sensitive to the exact pattern of dependences and independences.

(4) My results also allow to resolve the tension between Maudlin's approach on the one hand and Jarrett's analysis and the received view on the other hand. While Jarrett's result seems to suggest that there is a choice to make between outcome dependence or parameter dependence and the received view holds that it is the dependence between the outcomes which is realized, Maudlin's informational approach opposes to these positions by saying that one of the outcomes must depend on information about the distant setting. This tension has been stood unresolved for over 20 years now.

During that time the two analyses have coexisted; Maudlin's critique of Jarrett's analysis and the resulting standard position did not succeed in convincing the adherents of outcome dependence—though Maudlin did have good arguments: he realized that Jarrett's analysis can be misleading because there are different kinds of parameter dependences and outcome dependences, according to which variables appear in the conditional (Maudlin 1994, ch. 4); he argued that Jarrett's analysis is also misleading because his informational approach unveiled that some kind of parameter dependence is unavoidable; and, hence, it is wrong to assume that outcome dependence per se can explain EPR/B correlations (Maudlin 1994, ch. 6). To me it is not exactly clear, why Jarrett's analysis and the received view based on it could keep on for such a long time, given Maudlin's critique. One reason might have been that Maudlin's arguments do not connect to the Bell-Jarrett methodology, such that it was hard to compare the two approaches and to see which in fact is right.

In this paper, however, we have provided that connection. We have strengthened the Bell-Jarrett approach to our new results and these (i) confirm Maudlin's claim that there must be a dependence of at least one outcome on the distant parameter (we have furthermore derived, which exact combinations of dependences are required, section 3.4), (ii) show that Jarrett's analysis is indeed misleading (we have made precise how exactly, section 4.1) and (iii) prove that the received view is wrong (because outcome dependence is not sufficient for a class to be able to violate Bell inequalities; section 4.2). This clearly resolves the Jarrett-Maudlin controversy in favour of the latter.

5 Discussion

In this paper we have presented a stronger version of Bell's theorem and spelt out its consequences on a probabilistic level. Here we shall summarize and discuss the main results.

(1) The strengthening of the Bell argument rests on the insight that the members of a range of non-local theories, which we have called *weakly non-local*, either are inconsistent with autonomy and nearly perfect correlations or imply Bell inequalities (as do local theories). For instance, it is impossible to violate Bell inequalities even if a dependence on the distant outcome holds as in the product form

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

Consequently, the empirical violation of the inequalities does rule out local theories (which is well known from the original argument) *and* these weakly non-local ones (which is one central result of this paper). Showing that the violation of Bell inequalities excludes more theories than the standard Bell argument suggests, *the new argument has a stronger conclusion than the original one.*

The remaining theories, which are compatible with a violation of Bell inequalities, are called strongly non-local; a list of their product forms can be found in table 1 (H^α_1–H^α_{14}). They are characterized by the fact that *at least one of the factors in the product form involves both settings in its conditionals*, i.e. at least one of the outcomes must depend probabilistically (or functionally, respectively) on both settings (*probabilistic Bell contextuality*). Without such a dependence between an outcome and both settings Bell inequalities cannot be violated.

(2) On a rather general level, the fact that certain non-local probability distributions imply Bell inequalities, first of all makes explicit that *Bell inequalities are not locality conditions* in the sense that, if a probability distribution obeys a Bell inequality, it must be local. Local theories obey Bell inequalities but not vice versa.

(3) More importantly, the new result reveals that *the usual constraint for quantum*

non-locality, which follows from the standard Bell argument, *is inappropriately weak*. For the latter states a failure of the local factorisation condition, suggesting that just *any* non-local dependence is required. Allowing for all non-local classes, however, this includes classes which we have found to be compatible with Bell inequalities (weakly non-local classes). For this reason, the standard constraint, to require just a failure of local factorisation, mixes classes which can violate Bell inequalities with classes that cannot. While this is not to say that the standard argument is logically incorrect, it means that its conclusion is not an appropriate, tight characterization of quantum non-locality. This explains, why arguments proceeding from this result, like Jarrett's analysis, can yield misleading conclusions.

(4) It is a crucial feature of our new argument that its conclusion only contains classes of theories which can violate Bell inequalities. This, first, qualifies the new constraint as an appropriate basis for further explorations. Second, it precludes speculations whether the argument could be made even stronger: the argument we have presented has the *strongest possible conclusion from the violation of Bell inequalities on a qualitative probabilistic level* (which only takes into account dependences and independences rather than numerical strengths of correlations). It has been essential for arriving at this result to have the complete list of logically possible classes (instead of just the physically plausible ones, see table 1), because in this way we could be sure not to have neglected possible classes of theories.

(5) Jarrett has analysed the conclusion of the standard Bell argument, the failure of local factorization as *outcome dependence or parameter dependence*. In a similar way we have analyzed the conclusion of the stronger Bell argument, the failure of local and weakly non-local classes, and the precise result is:

$$\begin{gathered}\Big[\Big(\neg I(\alpha,b|\beta,a,\lambda)\wedge\neg I(\alpha,a|\beta,b,\lambda)\Big)\vee\Big(\neg I(\beta,a|b,\lambda)\wedge\neg I(\beta,b|a,\lambda)\Big)\Big]\wedge\\ \Big[\Big(\neg I(\beta,a|\alpha,b,\lambda)\wedge\neg I(\beta,b|\alpha,a,\lambda)\Big)\vee\Big(I(\alpha,b|a,\lambda)\wedge\neg I(\alpha,a|b,\lambda)\Big)\Big]\end{gathered} \tag{13}$$

Since the disjunction of $\neg I(\alpha,a|b,\lambda)$ and $\neg I(\beta,b|a,\lambda)$ is what is usually called parameter dependence, the result shows that one can avoid usual parameter dependence, but only if one accepts other kinds of parameter dependences, which additionally condition on the distant outcome, viz. $\neg I(\alpha,a|\beta,b,\lambda)$ and $\neg I(\beta,b|\alpha,a,\lambda)$. So in any case there must be some kind of parameter dependence. More precisely, the essential requirement is that at least one of the outcomes depends (in one of two senses) on *both* parameters, the distant and the local one (double parameter dependence).

Outcome dependence, in contrast, does not play any role in this new analysis: if a class fulfills the indicated condition, it can violate Bell inequalities, and if it does not, it cannot; whether *additionally* outcome dependence holds is irrelevant for the question whether the class can violate the inequalities.

(6) The result of the new analysis obviously is in tension with Jarrett's result. While Jarrett's result is not wrong, if one understands it in its correct literal sense, it is liable to be misunderstood (and has been misunderstood) as that there is a choice to make

between a dependence between the outcomes and a dependence between an outcome and its distant parameter. In fact, as our result shows there is only a choice to make between different types of parameter dependences.

We have also argued that Jarrett's categories (outcome dependence and parameter dependence) do not cut the classes at their natural joints. Outcome dependent theories as well as parameter dependent theories both mix classes which can violate Bell inequalities (strongly non-local classes) and those which cannot (weakly non-local classes). In the light of our present analysis, this is not surprising, because Jarrett already starts from a constraint that mixes classes.

(7) The received view about quantum non-locality, which is based on Jarrett's analysis, holds that, in fact, the quantum non-locality realized by nature on a probabilistic level amounts to outcome dependence (and not parameter dependence) (cf. Jarrett (1984) and Shimony (1984, 1986), who have introduced the view). We have discussed different readings of this claim but it has turned out that it is either plainly wrong (because classes which are outcome dependent without being double parameter dependent are weakly non-local and hence cannot violate Bell inequalities) or the claim reduces to a claim without outcome dependence (because for classes including a double parameter dependence outcome dependence becomes redundant). In sum, whether outcome dependence holds does not change whether a class can violate Bell inequalities or not. Adding outcome dependence does not make a class which implies Bell inequalities able to violate them; and vice versa subtracting outcome dependence from a class which can violate Bell inequalities does not make that class imply the inequalities. One example for the former is the fact that adding outcome dependence to a local theory, which means to have the product form of class (H_{16}^{α}) instead of local factorization, does not allow for a violation, because theories from class (H_{16}^{α}), as we have shown, still imply Bell inequalities. *A probabilistic dependence between the outcomes is just too weak to make the difference for a violation of Bell inequalities.* It is for this reason that a dependence between the outcomes cannot explain the violation of Bell inequalities.

This is emphatically not to say that outcome dependence in fact does not hold—it might or it might not. But even if it does, it cannot be the only non-local dependence. Neither is this to say that outcome dependence, if it holds, does not contribute to a violation of the Bell inequalities; to the contrary: if outcome dependence holds it will make such a contribution.[13] But what this result does deny against the received view is that outcome dependence can be the crucial dependence, i.e. the dependence that is *per se* responsible for the violation of Bell inequalities. And since this is the central claim of the received view, that position has to be considered wrong.

If outcome dependence is not the correct view on a probabilistic level, this might also affect the physical and metaphysical claims of the received view. Having focused on the probabilistic level, these latter levels have not been the subject of the present paper.

[13] A dependence on the distant outcome does matter when one considers not only the violation of Bell inequalities but the exact quantitative reproduction of EPR/B correlations. Pawlowski et al. (2010) have shown that there must be information about the distant outcome and that information can either be available by a direct correlation (as in the case of quantum mechanics) or be revealed by a hidden variable (which, however, is not available in the case of quantum mechanics).

The standard view concerning these matters so far has been that there is a non-causal influence between the outcomes, which is metaphysically realized by a non-local connection between the outcomes (a so called non-separability, according to most authors). However, given that the probabilistic picture has changed considerably, it remains to be investigated on the basis of the new analysis whether this view can still be maintained. Since probabilistic outcome dependence cannot account for a violation of the inequalities, it might seem tempting to immediately conclude that also the metaphysical picture must be wrong. But inferring physical or metaphysical relations from probabilistic facts requires careful analysis, since the transition is well known to be vulnerable to fallacies ('correlation is not causation'). For this reason, establishing the right kind of (meta-)physical connection would have required a further lengthy analysis—so here we have to remain tacit on this question. Having said this, it might be interesting to remark that there seem to be good arguments that the current result, that a probabilistic dependence between the outcomes is too weak to explain a violation of the Bell inequalities, most plausibly entails that also an influence between the outcomes is not strong enough to account for a violation (Näger 2013).

(8) What about quantum mechanics in this new picture? Quantum mechanics is well-known to be outcome dependent *and* to correctly reproduce the EPR/B correlations—so how does this fit with the present results? An answer can be seen by realizing that my result does not mean that all outcome dependent theories (in the probabilistic sense) imply the inequalities; it just says that a probabilistic dependence between the outcomes *per se* does not suffice to explain a violation. It follows that quantum mechanics cannot only involve a probabilistic dependence between the outcomes, and especially it cannot belong to class (H_{16}^{α}). Rather, the quantum mechanical product form (here: for maximally entangled states),

$$P(\alpha\beta|ab) = P(\alpha|\beta ab)P(\beta) = P(\beta|\alpha ab)P(\alpha) \tag{14}$$

additionally involves a dependence on the distant setting in each first factor—and it is the dependence on both settings in these factors (rather than the dependence between the outcomes), which is crucial for violating the Bell inequalities.

(9) We have noted that there is another approach to quantum non-locality whose result seems to converge with ours. Maudlin (1994, ch. 6; cf. also a recent refinement by Pawlowski et al. 2010) 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, Maudlin's claim is—at least roughly—in accordance with our results. On the one hand, this is good news because two different methods yielding the same results are evidence for the stability of a claim.

On the other hand, I stress that there are at least three non-trivial differences between Maudlin's approach and the one we have presented in this paper. First, each proposal has its own, very different methodology. Maudlin analyses the correlations information theoretically and does not connect his considerations to Bell's argument. In contrast, our approach here is in continuity with Bell's thoughts, which have started and shaped the discussion; it develops and strengthens the method that is most common in the debate, viz. the access via Bell inequalities. Second, the information theoretic approaches are stronger in that they indicate the amount of information (the quantitative strength of the correlations) that is required in order to reproduce EPR/B correlations. This is an important result providing sufficient criteria for reproducing the correlations. In another sense, however, third, the information theoretic considerations are also weaker than our results in this paper. Maudlin only claims a dependence between an outcome and both settings, whereas we have presented a precise condition of required dependences The precise result might relevant for further analysis such as causal inference, which is very sensitive to which (in-)dependences exactly hold.

(10) Finally, we may ask, *why these stronger consequences of the Bell argument*, that we have derived in this paper, *have 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 such a theory 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 today it is clear that the quantum mechanical distribution is empirically correct and Bell inequalities are violated, it is desirable to draw as strong consequences as possible from the argument, which requires to check without prejudice whether some non-local classes allow a derivation of Bell inequalities as well. That this is indeed the case is the result of this paper.

Acknowledgements

I would like to thank Frank Arntzenius, Tjorven Hetzger, Gábor Hofer-Szabó, Meinard Kuhlmann, Wayne Myrvold, Thorben Petersen, Manfred Stöckler, Nicola Vona, Adrian Wüthrich and audiences at the conference 'Philosophy of Physics in Germany' (Hanover, 2010) and at the '14th Congress of Logic, Methodology and Philosophy of Science' (Nancy, 2011) for helpful comments and discussion. I am also grateful to two referees for valuable comments. This paper is a partial result of a research project funded by the Deutsche Forschungsgemeinschaft (DFG).

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. (Reprinted in Bell 1987b). New York: Academic Press.

Bell, J. S. (1975). The theory of local beables. *TH-2053-CERN* (Reprinted in Bell 1987b).

Bohm, D. (1951). *Quantum Theory.* Englewood Cliffs, NJ: 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.

Corry, R. (2015). Retrocausal models for EPR. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 49*, 1–9.

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

Fine, A. (1982a). Some Local Models for Correlation Experiments *Synthese 50*, 279–294.

Fine, A. (1982b). Hidden Variables, Joint Probability, and the Bell Inequalities. *Physical Review Letters 48*, 291–295.

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. *Noûs 18*, 569–590.

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

Näger, P. M. (2013). Causal graphs for EPR experiments. *Preprint*. http://philsci-archive.pitt.edu/9915/

Näger, P. M. (2015). The causal problem of entanglement. *Synthese*. Advance online publication. doi: 10.1007/s11229-015-0668-6.

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

Price, H. (1994). A Neglected Route to Realism About Quantum Mechanics. *Mind 103*, 303–336.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference.* Cambridge: Cambridge University Press. (2nd ed. 2009)

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.

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

Shimony, A. (1984). Controllable and uncontrollable non-locality. In S. Kamefuchi (Ed.), *Foundations of Quantum Mechanics in the Light of New Technology*, pp. 225–230. Tokyo: The Physical Society of Japan.

Shimony, A. (1986). Events and processes in the quantum world. In R. Penrose and C. J. Isham (Eds.), *Quantum Concepts in Space and Time*, pp. 182–203. Oxford: Clarendon Press.

Spirtes, P., Glymour, C., Scheines, R. (1993). *Causation, Prediction, and Search.* Springer: New York. (2nd ed. Cambridge, MA: MIT-Press, 2000)

Sutherland, R. I. (1983). Bell's theorem and backwards-in-time causality. *International Journal of Theoretcial Physics 22*(4), 377–384.

Ursin, R., Tiefenbacher, F., Schmitt-Manderbach, T., Weier, H., Scheidl, T., Lindenthal, M., Blauensteiner, B., Jennewein, T., Perdigues, J., Trojek, P., Omer, B., Furst, M., Meyenburg, M., Rarity, J., Sodnik, Z., Barbieri, C., Weinfurter, H., and Zeilinger, Anton (2007). Entanglement-based quantum communication over 144 km. *Nature Physics 3*(7), 481–486.

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. (Reprinted with additional comments from Synthese 52, 25–38). Notre Dame: University of Notre Dame Press.

Weihs, G., Jennewein, T., Simon, C., Weinfurter, H., and Zeilinger, A. (1998). Violation of Bell's inequality under strict Einstein locality conditions. *Physical Review Letters 81*(23), 5039–5043.

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

Appendix

Proof of lemma 1

We proceed by reductio. By autonomy and (H^{α}_{16}) we rewrite the conditions for perfect correlations and for perfect anti-correlations:

$$P(\alpha_\pm\beta_\mp|a_ib_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\mp a_i\lambda)P(\beta_\mp|b_i\lambda) \tag{15}$$

$$P(\alpha_\pm\beta_\mp|a_{i_\perp}b_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda)P(\beta_\mp|b_{i_\perp}\lambda) \tag{16}$$

$$P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\pm a_i\lambda)P(\beta_\pm|b_{i_\perp}\lambda) \tag{17}$$

$$P(\alpha_\pm\beta_\pm|a_{i_\perp}b_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\pm a_{i_\perp}\lambda)P(\beta_\pm|b_i\lambda) \tag{18}$$

Since probabilities are non-negative (and since without loss of generality we can 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 i and λ we must have:

$$\begin{aligned}
& [\ P(\alpha_+|\beta_-a_i\lambda) = 0 \quad \vee \quad P(\beta_-|b_i\lambda) = 0\ ] && (19)\\
\wedge\ & [\ P(\alpha_-|\beta_+a_i\lambda) = 0 \quad \vee \quad P(\beta_+|b_i\lambda) = 0\ ] && (20)\\
\wedge\ & [\ P(\alpha_+|\beta_-a_{i_\perp}\lambda) = 0 \quad \vee \quad P(\beta_-|b_{i_\perp}\lambda) = 0\ ] && (21)\\
\wedge\ & [\ P(\alpha_-|\beta_+a_{i_\perp}\lambda) = 0 \quad \vee \quad P(\beta_+|b_{i_\perp}\lambda) = 0\ ] && (22)\\
\wedge\ & [\ P(\alpha_+|\beta_+a_i\lambda) = 0 \quad \vee \quad P(\beta_+|b_{i_\perp}\lambda) = 0\ ] && (23)\\
\wedge\ & [\ P(\alpha_-|\beta_-a_i\lambda) = 0 \quad \vee \quad P(\beta_-|b_{i_\perp}\lambda) = 0\ ] && (24)\\
\wedge\ & [\ P(\alpha_+|\beta_+a_{i_\perp}\lambda) = 0 \quad \vee \quad P(\beta_+|b_i\lambda) = 0\ ] && (25)\\
\wedge\ & [\ P(\alpha_-|\beta_-a_{i_\perp}\lambda) = 0 \quad \vee \quad P(\beta_-|b_i\lambda) = 0\ ] && (26)
\end{aligned}$$

From these conditions one can infer that all involved probabilities must be 0 or 1 (determinism). More precisely, for every i and λ one of the following two cases holds:

Case 1: $P(\alpha_+|\beta_- a_i \lambda) = 0$

$$\overset{(CE)}{\Rightarrow} P(\alpha_-|\beta_- a_i \lambda) = 1 \quad \overset{(24)}{\Rightarrow} P(\beta_-|b_{i_\perp}\lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\beta_+|b_{i_\perp}\lambda) = 1 \quad \underset{(23)}{\overset{(22)}{\Rightarrow}} P(\alpha_-|\beta_+ a_{i_\perp}\lambda) = 0 \wedge P(\alpha_+|\beta_+ a_i \lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\alpha_+|\beta_+ a_{i_\perp}\lambda) = 1 \wedge P(\alpha_-|\beta_+ a_i \lambda) = 1 \quad \underset{(20)}{\overset{(25)}{\Rightarrow}} P(\beta_+|b_i \lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\beta_-|b_i \lambda) = 1 \quad \overset{(26)}{\Rightarrow} P(\alpha_-|\beta_- a_{i_\perp}\lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\alpha_+|\beta_- a_{i_\perp}\lambda) = 1$$

NB: (CE) stands for the following theorem of probability theory: $P(A|B)+P(\bar{A}|B) = 1$.

Case 2: $P(\alpha_+|\beta_- a_i \lambda) > 0$

$$\overset{(19)}{\Rightarrow} P(\beta_-|b_i \lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\beta_+|b_i \lambda) = 1 \quad \underset{(25)}{\overset{(20)}{\Rightarrow}} P(\alpha_-|\beta_+ a_i \lambda) = 0 \wedge P(\alpha_+|\beta_+ a_{i_\perp}\lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\alpha_+|\beta_+ a_i \lambda) = 1 \wedge P(\alpha_-|\beta_+ a_{i_\perp}\lambda) = 1 \quad \underset{(22)}{\overset{(23)}{\Rightarrow}} P(\beta_+|b_{i_\perp}\lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\beta_-|b_{i_\perp}\lambda) = 1 \quad \underset{(24)}{\overset{(21)}{\Rightarrow}} P(\alpha_+|\beta_- a_{i_\perp}\lambda) = 0 \wedge P(\alpha_-|\beta_- a_i \lambda) = 0$$

$$\overset{(CE)}{\Rightarrow} P(\alpha_-|\beta_- a_{i_\perp}\lambda) = 1 \wedge P(\alpha_+|\beta_- a_i \lambda) = 1$$

Since in each case we have

$$P(\alpha_+|\beta_+a_i\lambda) = P(\alpha_+|\beta_-a_i\lambda) \tag{27}$$
$$P(\alpha_-|\beta_+a_i\lambda) = P(\alpha_-|\beta_-a_i\lambda) \tag{28}$$
$$P(\alpha_+|\beta_+a_{i_\perp}\lambda) = P(\alpha_+|\beta_-a_{i_\perp}\lambda) \tag{29}$$
$$P(\alpha_-|\beta_+a_{i_\perp}\lambda) = P(\alpha_-|\beta_-a_{i_\perp}\lambda) \tag{30}$$

it is true that

$$\forall\alpha,\beta,a,\lambda: \quad P(\alpha|\beta a\lambda) = P(\alpha|a\lambda). \tag{31}$$

By this statistical independence the product form (H_{16}^{α})

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda) \tag{32}$$

loses its dependence on the outcome β in the first factor, i.e. it reads

$$P(\alpha\beta|ab\lambda) = P(\alpha|a\lambda)P(\beta|b\lambda). \tag{33}$$

This, however, is the well known local product form (local factorization), contradicting the assumption that we have the non-local product form (H_{16}^{α}).

Note that this proof makes essential use of the perfect correlations and perfect antiocorrelations (15–18), i.e. the probabilities $P(\alpha_\pm\beta_\mp|a_ib_i)$, $P(\alpha_\pm\beta_\mp|a_{i_\perp}b_{i_\perp})$, $P(\alpha_\pm\beta_\pm|a_ib_{i_\perp})$ and $P(\alpha_\pm\beta_\pm|a_{i_\perp}b_i)$ have to be strictly 0. If these conditions are only slightly relaxed, i.e. if any of these probabilities takes on a positive value, even if very small, the conclusion does not follow. q.e.d.

Proof of lemma 2

By autonomy and (H_{16}^{α}) we rewrite (some of) the conditions for nearly perfect correlations and for nearly perfect anti-correlations:

$$P(\alpha_\pm\beta_\mp|a_ib_i) = \delta_{ii} \quad = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\mp a_i\lambda)P(\beta_\mp|b_i\lambda) \tag{34}$$
$$P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = \delta_{ii_\perp} = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\pm a_i\lambda)P(\beta_\pm|b_{i_\perp}\lambda), \tag{35}$$

where δ_{ii} and $\delta_{ii_\perp}$ are positive and small. Since probabilities are non-negative, all summands are non-negative; so each summand must be less or equal than the total value of the sum:

$$P(\lambda)P(\alpha_\pm|\beta_\mp a_i\lambda)P(\beta_\mp|b_i\lambda) \leq \delta_{ii} \tag{36}$$
$$P(\lambda)P(\alpha_\pm|\beta_\pm a_i\lambda)P(\beta_\pm|b_{i_\perp}\lambda) \leq \delta_{ii_\perp} \tag{37}$$

In order to facilitate the following considerations, let us define

$$\epsilon := \max_{i=1,2,3}(\sqrt[3]{\delta_{ii}}, \sqrt[3]{\delta_{ii_\perp}}) \tag{38}$$

where $i = 1, 2, 3$ represent three distinct measurement directions. We can then write:

$$P(\lambda)P(\alpha_\pm|\beta_\mp a_i\lambda)P(\beta_\mp|b_i\lambda) \leq \epsilon^3 \tag{39}$$

$$P(\lambda)P(\alpha_\pm|\beta_\pm a_i\lambda)P(\beta_\pm|b_{i_\perp}\lambda) \leq \epsilon^3. \tag{40}$$

Since a product of three non-negative factors is never smaller than the cube root of its smallest factor, each product must contain (at least) one factor that is less or equal than ϵ, i.e. for all values i and λ we must have:

$$\big[P(\lambda) \leq \epsilon \quad \vee \quad P(\alpha_+|\beta_- a_i\lambda) \leq \epsilon \quad \vee \quad P(\beta_-|b_i\lambda) \leq \epsilon\big] \tag{41}$$

$$\wedge \quad \big[P(\lambda) \leq \epsilon \quad \vee \quad P(\alpha_-|\beta_+ a_i\lambda) \leq \epsilon \quad \vee \quad P(\beta_+|b_i\lambda) \leq \epsilon\big] \tag{42}$$

$$\wedge \quad \big[P(\lambda) \leq \epsilon \quad \vee \quad P(\alpha_+|\beta_+ a_i\lambda) \leq \epsilon \quad \vee \quad P(\beta_+|b_{i_\perp}\lambda) \leq \epsilon\big] \tag{43}$$

$$\wedge \quad \big[P(\lambda) \leq \epsilon \quad \vee \quad P(\alpha_-|\beta_- a_i\lambda) \leq \epsilon \quad \vee \quad P(\beta_-|b_{i_\perp}\lambda) \leq \epsilon\big] \tag{44}$$

There are three cases that solve these conditions:

Case 1: $P(\lambda) > \epsilon \;\wedge\; P(\alpha_+|\beta_- a_i\lambda) \leq \epsilon$

$$\overset{\text{(CE)}}{\Rightarrow} P(\alpha_-|\beta_- a_i\lambda) > 1-\epsilon \qquad \overset{(44)}{\Rightarrow} P(\beta_-|b_{i_\perp}\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\beta_+|b_{i_\perp}\lambda) > 1-\epsilon \qquad \overset{(43)}{\Rightarrow} P(\alpha_+|\beta_+ a_i\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\alpha_-|\beta_+ a_i\lambda) > 1-\epsilon \qquad \overset{(42)}{\Rightarrow} P(\beta_+|b_i\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\beta_-|b_i\lambda) > 1-\epsilon$$

Case 2: $P(\lambda) > \epsilon \;\wedge\; P(\alpha_+|\beta_- a_i\lambda) > \epsilon$

$$\overset{(41)}{\Rightarrow} P(\beta_-|b_i\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\beta_+|b_i\lambda) > 1-\epsilon \qquad \overset{(42)}{\Rightarrow} P(\alpha_-|\beta_+ a_i\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\alpha_+|\beta_+ a_i\lambda) > 1-\epsilon \qquad \overset{(43)}{\Rightarrow} P(\beta_+|b_{i_\perp}\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\beta_-|b_{i_\perp}\lambda) > 1-\epsilon \qquad \overset{(44)}{\Rightarrow} P(\alpha_-|\beta_- a_i\lambda) \leq \epsilon$$

$$\overset{\text{(CE)}}{\Rightarrow} P(\alpha_+|\beta_- a_i\lambda) > 1-\epsilon$$

Case 3: $P(\lambda) \leq \epsilon$

(no particular restrictions for other probabilities)

The three cases are disjunct and define a partition of the values of $\boldsymbol{\lambda}$:

$$\Lambda_1(i) := \{\lambda | P(\lambda) > \epsilon \ \wedge \ P(\alpha_+|\beta_- a_i \lambda) \leq \epsilon)\}$$

$$\Lambda_2(i) := \{\lambda | P(\lambda) > \epsilon \ \wedge \ P(\alpha_+|\beta_- a_i \lambda) \geq 1-\epsilon)\}$$

$$\Lambda_3(i) := \{\lambda | P(\lambda) \leq \epsilon\} = \Lambda_3$$

Note that each value i defines a different partition, but that $\Lambda_3(i) = \Lambda_3$ is independent of i.

We can use the fact that the $\boldsymbol{\lambda}$-partitions depend on just *one* setting i to estimate values for 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 settings (see table 4). Note that the table only comprises five of the nine combinatorially possible cases; the ignored cases are empty sets ($\Lambda_1(i) \wedge \Lambda_3 = \emptyset$ because $\Lambda_1(i)$ requires $P(\lambda) > \epsilon$, whereas Λ_3 implies $P(\lambda) \leq \epsilon$; and analogously $\Lambda_2(i) \wedge \Lambda_3 = \emptyset$, $\Lambda_3 \wedge \Lambda_1(j) = \emptyset$, $\Lambda_3 \wedge \Lambda_2(j) = \emptyset$). The last column is defined as $\Lambda_3(i) \cap \Lambda_3(j) = \Lambda_3$, and the label 'n.r.' means 'no restriction', i.e. the value of the hidden joint probability is not confined to any specific interval; rather, in this set it is the case that $P(\lambda) \leq \epsilon$.

Table 4: Values of the hidden joint probability

	$\lambda \in$				
	$\Lambda_1(i) \cap \Lambda_1(j)$	$\Lambda_1(i) \cap \Lambda_2(j)$	$\Lambda_2(i) \cap \Lambda_1(j)$	$\Lambda_2(i) \cap \Lambda_2(j)$	Λ_3
$P(\alpha_+\beta_+|a_ib_j\lambda)$	$\leq \epsilon^2$	$\leq \epsilon$	$\leq \epsilon$	$> (1-\epsilon)^2$	n.r.
$P(\alpha_+\beta_-|a_ib_j\lambda)$	$\leq \epsilon$	$\leq \epsilon^2$	$> (1-\epsilon)^2$	$\leq \epsilon$	n.r.
$P(\alpha_-\beta_+|a_ib_j\lambda)$	$\leq \epsilon$	$> (1-\epsilon)^2$	$\leq \epsilon^2$	$\leq \epsilon$	n.r.
$P(\alpha_-\beta_-|a_ib_j\lambda)$	$> (1-\epsilon)^2$	$\leq \epsilon$	$\leq \epsilon$	$\leq \epsilon^2$	n.r.

Given the estimations for the hidden joint probability in table 4 one can derive a generalised Wigner-Bell inequality. Consider the inequality

$$P(X \cap \overline{Z}) \leq P(X \cap \overline{Y}) + P(Y \cap \overline{Z}), \tag{45}$$

which in general holds for any events X, Y, Z of a measurable space, as can easily be seen by rewriting the involved probabilities:

$$P(X \cap \overline{Z}) = P(X \cap Y \cap \overline{Z}) + P(X \cap \overline{Y} \cap \overline{Z}) \tag{46}$$
$$P(X \cap \overline{Y}) = P(X \cap \overline{Y} \cap Z) + P(X \cap \overline{Y} \cap \overline{Z}) \tag{47}$$
$$P(Y \cap \overline{Z}) = P(X \cap Y \cap \overline{Z}) + P(\overline{X} \cap Y \cap \overline{Z}) \tag{48}$$

Assuming $X = \Lambda_1(1) \cup \Lambda_3$, $Y = \Lambda_1(2) \cup \Lambda_3$ and $Z = \Lambda_1(3)$ gives the inequality

$$P([\Lambda_1(1) \cup \Lambda_3] \cap \overline{\Lambda_1(3)}) \leq P([\Lambda_1(1) \cup \Lambda_3] \cap \overline{[\Lambda_1(2) \cup \Lambda_3]}) + P([\Lambda_1(2) \cup \Lambda_3] \cap \overline{\Lambda_1(3)}). \tag{49}$$

We calculate the sets involved in the inequality:

$$\begin{aligned}
[\Lambda_1(i) \cup \Lambda_3] \cap \overline{\Lambda_1(j)} &= [\Lambda_1(i) \cup \Lambda_3] \cap [\Lambda_2(j) \cup \Lambda_3] = \\
&= [\Lambda_1(i) \cap \Lambda_2(j)] \cup [\underbrace{\Lambda_1(i) \cap \Lambda_3}_{\emptyset}] \cup [\underbrace{\Lambda_3 \cap \Lambda_2(j)}_{\emptyset}] \cup [\underbrace{\Lambda_3 \cap \Lambda_3}_{\Lambda_3}] \\
&= [\Lambda_1(i) \cap \Lambda_2(j)] \cup \Lambda_3 \qquad (50)
\end{aligned}$$

$$\begin{aligned}
[\Lambda_1(i) \cup \Lambda_3] \cap [\overline{\Lambda_1(j) \cup \Lambda_3}] &= [\Lambda_1(i) \cup \Lambda_3] \cap \Lambda_2(j) \\
&= [\Lambda_1(i) \cap \Lambda_2(j)] \cup [\underbrace{\Lambda_3 \cap \Lambda_2(j)}_{\emptyset}] \\
&= \Lambda_1(i) \cap \Lambda_2(j) \qquad (51)
\end{aligned}$$

If we further define the shorthand

$$\Lambda_{kl}(i,j) := \Lambda_k(i) \cap \Lambda_l(j), \qquad (52)$$

we can rewrite inequality (49) as

$$P(\Lambda_{12}(1,3) \cup \Lambda_3) \leq P(\Lambda_{12}(1,2)) + P(\Lambda_{12}(2,3) \cup \Lambda_3). \qquad (53)$$

This inequality can be transformed to yield a generalized Wigner-Bell inequality. We have to rewrite the inequality such that it only involves empirically accessible probabilities, i.e. probabilities that do not involve the hidden state $\boldsymbol{\lambda}$, and this can be done by using the estimates for the hidden joint probability from table 4. Especially, we have to find a lower estimate for the left hand side of the inequality and an upper estimate for its right hand side. We start by deriving the former:

$$\begin{aligned}
P(\Lambda_{12}(1,3) \cup \Lambda_3) &\overset{(\sigma\text{-additivity})}{=} \sum_{\lambda \in \Lambda_{12}(1,3) \cup \Lambda_3} P(\lambda) \\
&\geq \sum_{\lambda \in \Lambda_{12}(1,3) \cup \Lambda_3} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) \\
&= \sum_{\lambda \in \Lambda} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) - \sum_{\lambda \in \Lambda \setminus [\Lambda_{12}(1,3) \cup \Lambda_3]} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) \\
&= P(\alpha_- \beta_+ | a_1 b_3) - \sum_{\lambda \in \Lambda_{11}(1,3)} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) \\
&\quad - \sum_{\lambda \in \Lambda_{21}(1,3)} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) - \sum_{\lambda \in \Lambda_{22}(1,3)} P(\lambda) P(\alpha_- \beta_+ | a_1 b_3 \lambda) \\
&\overset{(\text{table 4})}{\geq} P(\alpha_- \beta_+ | a_1 b_3) - \epsilon \sum_{\lambda \in \Lambda_{11}(1,3)} P(\lambda) - \epsilon^2 \sum_{\lambda \in \Lambda_{21}(1,3)} P(\lambda) - \epsilon \sum_{\lambda \in \Lambda_{22}(1,3)} P(\lambda) \\
&\geq P(\alpha_- \beta_+ | a_1 b_3) - 2\epsilon - \epsilon^2 \qquad (54)
\end{aligned}$$

An upper estimate for the right hand side of (53) can be calculated as follows:

$$P(\Lambda_{12}(1,2)) + P(\Lambda_{12}(2,3) \cup \Lambda_3) = \tag{55}$$

$$\overset{(\sigma\text{-additivity})}{=} \sum_{\lambda \in \Lambda_{12}(1,2)} P(\lambda) + \sum_{\lambda \in \Lambda_{12}(2,3) \cup \Lambda_3} P(\lambda)$$

$$\leq \sum_{\lambda \in \Lambda_{12}(1,2)} P(\lambda) \frac{P(\alpha_-\beta_+|a_1b_2\lambda)}{(1-\epsilon)^2} + \sum_{\lambda \in \Lambda_{12}(2,3) \cup \Lambda_3} P(\lambda) \frac{P(\alpha_-\beta_+|a_2b_3\lambda)}{(1-\epsilon)^2}$$

$$\leq \sum_{\lambda \in \Lambda} P(\lambda) \frac{P(\alpha_-\beta_+|a_1b_2\lambda)}{(1-\epsilon)^2} + \sum_{\lambda \in \Lambda} P(\lambda) \frac{P(\alpha_-\beta_+|a_2b_3\lambda)}{(1-\epsilon)^2}$$

$$= \frac{P(\alpha_-\beta_+|a_1b_2\lambda) + P(\alpha_-\beta_+|a_2b_3\lambda)}{(1-\epsilon)^2} \tag{56}$$

The resulting inequality

$$P(\alpha_-\beta_+|a_1b_3) - 2\epsilon - \epsilon^2 \leq \frac{P(\alpha_-\beta_+|a_1b_2\lambda) + P(\alpha_-\beta_+|a_2b_3\lambda)}{(1-\epsilon)^2} \tag{57}$$

is the Wigner-Bell inequality we have been looking for. It generalizes usual Wigner-Bell inequalities such as

$$P(\alpha_-\beta_+|a_1b_3) \leq P(\alpha_-\beta_+|a_1b_2\lambda) + P(\alpha_-\beta_+|a_2b_3\lambda) \tag{58}$$

in that it introduces correction terms with the parameter ϵ. It is an inequality of fourth order in ϵ and one can check numerically that it is violated by the empirical measurement results

$$P(\alpha_-\beta_+|a_1b_3) = 0.375, \qquad P(\alpha_-\beta_+|a_1b_2) = 0.125, \qquad P(\alpha_-\beta_+|a_2b_3) = 0.125, \tag{59}$$

(which are a maximal violation of the usual Wigner-Bell inequality and occur e.g. for the measurement settings being chosen as $1 = 0°$, $2 = 30°$, $3 = 60°$ given the quantum state ψ_0), if

$$0 < \epsilon < 0.048328 \tag{60}$$

Hence, the maximal deviation of perfect correlations for the generalized Wigner-Bell inequality still to be violated is

$$\delta = (\epsilon_{\max})^3 = 0.048328^3 = 1.1280 \cdot 10^{-4}, \tag{61}$$

i.e. at least 99.989% of the photons must be perfectly correlated and anti-correlated.

q.e.d.

Proof of theorem 1.1

We split the theorem up into three partial claims:

Claim 1: Autonomy, perfect correlations, perfect anti-correlations and a class of probability distributions (H_i^α) form an inconsistent set if (i) the product form of (H_i^α) involves at most one of the settings.

Claim 2: Autonomy, perfect correlations, perfect anti-correlations and a class of probability distributions (H_i^α) form an inconsistent set if (ii) the product form of (H_i^α) involves both settings but its first factor involves the distant outcome and at most one setting.

Claim 3: A class (H_i^α) is consistent with autonomy, perfect correlations and perfect anti-correlations if (¬i) the product form of (H_i^α) involves both settings and (¬ii) in case the distant outcome appears in the first factor of (H_i^α)'s product form, also both settings appear in that factor.

Proof of claim 1

Condition (i), that the product form involves at most one of the settings, is fulfilled by the classes $\{(\mathrm{H}_{17}^\alpha), \ldots, (\mathrm{H}_{32}^\alpha)\} \setminus \{(\mathrm{H}_{22}^\alpha), (\mathrm{H}_{29}^\alpha)\}$. Here we have to show the inconsistency of these classes with the set of assumptions autonomy, perfect correlations and perfect anti-correlations.

Consider, for instance,

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|a\lambda) = P(\alpha\beta|a\lambda), \tag{H_{17}^α}$$

which fails to involve the setting $\boldsymbol{b}$. It is easy to show that this product form can neither account for the perfect correlations nor for the perfect anti-correlations. The perfect correlations read:

$$P(\alpha_\pm\beta_\pm|a_ib_i) = \frac{1}{2} \qquad\qquad P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = 0. \tag{62}$$

Now, the value of these empirical probabilities depends crucially on the value of the setting $\boldsymbol{b}$. However, one can demonstrate without much effort that (H_{17}^α)'s failure to involve the setting $\boldsymbol{b}$ on a hidden level, extends to the empirical level, if one assumes autonomy:

$$\begin{aligned} P(\alpha\beta|ab) = \sum_\lambda P(\lambda|ab)P(\alpha\beta|ab\lambda) &\overset{(\mathrm{A})}{=} \sum_\lambda P(\lambda|ab')P(\alpha\beta|ab\lambda) = \\ &\overset{(\mathrm{H}_{17}^\alpha)}{=} \sum_\lambda P(\lambda|ab')P(\alpha\beta|ab'\lambda) = P(\alpha\beta|ab') \end{aligned} \tag{63}$$

This implies that according to (H_{17}^α) all empirical probabilities $P(\alpha\beta|ab)$ that only differ by their value for the setting $\boldsymbol{b}$ must equal another—which obviously contradicts (62). For the same reason, (H_{17}^α) contradicts the perfect anti-correlations

$$P(\alpha_\pm\beta_\mp|a_ib_{i_\perp}) = \frac{1}{2} \qquad\qquad P(\alpha_\pm\beta_\mp|a_ib_i) = 0. \tag{64}$$

In the same way, all other product forms that do not involve the setting $\boldsymbol{b}$ are in conflict with the perfect correlations (62) and perfect anti-correlations (64), and, similarly, all product forms that fail to involve the setting $\boldsymbol{a}$ are in conflict with the perfect correlations

$$P(\alpha_\pm\beta_\pm|a_i b_i) = \frac{1}{2} \qquad\qquad P(\alpha_\pm\beta_\pm|a_{i_\perp} b_i) = 0. \tag{65}$$

or the perfect anti-correlations

$$P(\alpha_\pm\beta_\mp|a_{i_\perp} b_i) = \frac{1}{2} \qquad\qquad P(\alpha_\pm\beta_\mp|a_i b_i) = 0. \tag{66}$$

<u>Proof of claim 2</u>

Condition (ii), that the product form involves both settings but its first factor involves the distant outcome and at most one setting, is fulfilled by the product forms (H_4^α), (H_5^α), (H_{10}^α), (H_{15}^α) and (H_{16}^α). Here we have to show the inconsistency of these classes with the set of assumptions autonomy, perfect correlations and perfect anti-correlations.

By lemma 1 we have already proven that (H_{16}^α)

$$P(\alpha\beta|ab\lambda) = P(\alpha|\beta a\lambda)P(\beta|b\lambda) \tag{67}$$

forms an inconsistent set with autonomy, perfect correlations and perfect anti-correlations. Mutatis mutandis, also the classes (H_{10}^α) and (H_{15}^α) lead to a similar inconsistency. In each case the product form looses its dependence on the distant outcome in the first factor, i.e. (H_{10}^α) reduces to (H_{14}^α), whereas (H_{15}^α) reduces to (H_{22}^α).

The proofs against the classes (H_4^α) and (H_5^α) work in a similar way, but require a little more care due to an additional case differentiation. Let me shortly demonstrate this for class (H_5^α). As for (H_{16}^α) one starts with expressing the perfect (anti-)correlations in terms of the product form,

$$P(\alpha_\pm\beta_\mp|a_i b_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\mp a_i\lambda)P(\beta_\mp|a_i b_i\lambda) \tag{68}$$

$$P(\alpha_\pm\beta_\mp|a_{i_\perp} b_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda)P(\beta_\mp|a_{i_\perp} b_{i_\perp}\lambda) \tag{69}$$

$$P(\alpha_\pm\beta_\pm|a_i b_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\pm a_i\lambda)P(\beta_\pm|a_i b_{i_\perp}\lambda) \tag{70}$$

$$P(\alpha_\pm\beta_\pm|a_{i_\perp} b_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|\beta_\pm a_{i_\perp}\lambda)P(\beta_\pm|a_{i_\perp} b_i\lambda). \tag{71}$$

In the case of (H_{16}^α) there were two cases, defined by $P(\alpha_\pm|\beta_\mp a_i\lambda) = 0$ or $P(\alpha_\pm|\beta_\mp a_i\lambda) > 0$, respectively, and all other probabilities followed from each of these defining probabilities. In the present case, however, when, accordingly, we assume $P(\alpha_\pm|\beta_\mp a_i\lambda) = 0$ or $P(\alpha_\pm|\beta_\mp a_i\lambda) > 0$, respectively, only the factors of the hidden joint probabilities on the right hand side of equations (68) and (70) are implied, i.e. the probabilities in (69) and (71) remain undetermined by these assumptions (due to the fact that there are *two* settings in the second factor of the product form). The latter probabilities have to be determined by further assumptions, e.g. by setting $P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) = 0$ or $P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) > 0$, respectively. These assumptions introduce two new cases, that are logically independent of the former two. In total, this makes four cases (instead of two):

$$P(\alpha_\pm|\beta_\mp a_i\lambda) = 0 \qquad \wedge \qquad P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) = 0 \tag{72}$$
$$P(\alpha_\pm|\beta_\mp a_i\lambda) = 0 \qquad \wedge \qquad P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) > 0 \tag{73}$$
$$P(\alpha_\pm|\beta_\mp a_i\lambda) > 0 \qquad \wedge \qquad P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) = 0 \tag{74}$$
$$P(\alpha_\pm|\beta_\mp a_i\lambda) > 0 \qquad \wedge \qquad P(\alpha_\pm|\beta_\mp a_{i_\perp}\lambda) > 0 \tag{75}$$

While this renders the proof slightly more complex, the crucial fact to mention here is that in all four cases we have

$$\forall \alpha, \beta, a, \lambda: \quad P(\alpha|\beta a\lambda) = P(\alpha|a\lambda), \tag{76}$$

i.e. (H^α_5) reduces to (H^α_{12}). Similarly, one can show that (H^α_4) reduces to (H^α_{11}).

Proof of claim 3

Condition (¬i) *and* (¬ii), that the product form involves both settings and in case the distant outcome appears in the first factor, also both settings appear in that factor, is fulfilled by the product forms $\{(\mathrm{H}^\alpha_1), \ldots, (\mathrm{H}^\alpha_{14})\}\backslash\{(\mathrm{H}^\alpha_4), (\mathrm{H}^\alpha_5), (\mathrm{H}^\alpha_{10})\}$, ($\mathrm{H}^\alpha_{22}$) and ($\mathrm{H}^\alpha_{29}$). Here we have to show the consistency of these classes with the set of assumptions autonomy, perfect correlations and perfect anti-correlations.

Since a class being inconsistent with certain assumptions means that every distribution of a class contradicts the assumptions, a class being consistent means that there is at least one probability distribution in that class which is compatible with the assumptions. Hence, what we need for each of these classes in order to show their consistency with the background assumptions, is one example of a probability distribution belonging to that class that respects the background assumptions. In fact, such examples are easy to construct. Let me demonstrate the procedure with one of the weakest classes in that group, (H^α_{29}), whose product form is local factorization.

Requiring just *any* example we can presuppose a minimal setup, i.e. the hidden variable as well as each setting can be assumed to have only two possible values: $\boldsymbol{\lambda} = \lambda_1, \lambda_2$, $\boldsymbol{a} = a_i, a_{i_\perp}$ and $\boldsymbol{b} = b_i, b_{i_\perp}$ with $a_i = b_i$ and $a_{i_\perp} = b_{i_\perp}$. We start by writing down the perfect correlations and perfect anti-correlations, and express the probabilities on the empirical level by the probabilities on the hidden level using the product form and autonomy:

$$P(\alpha_\pm\beta_\mp|a_ib_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|a_i\lambda)P(\beta_\mp|b_i\lambda) \tag{77}$$

$$P(\alpha_\pm\beta_\mp|a_{i_\perp}b_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|a_{i_\perp}\lambda)P(\beta_\mp|b_{i_\perp}\lambda) \tag{78}$$

$$P(\alpha_\pm\beta_\pm|a_{i_\perp}b_i) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|a_{i_\perp}\lambda)P(\beta_\pm|b_i\lambda) \tag{79}$$

$$P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = 0 = \sum_\lambda P(\lambda)P(\alpha_\pm|a_i\lambda)P(\beta_\pm|b_{i_\perp}\lambda). \tag{80}$$

$$P(\alpha_\pm\beta_\pm|a_ib_i) = \frac{1}{2} = \sum_\lambda P(\lambda)P(\alpha_\pm|a_i\lambda)P(\beta_\pm|b_i\lambda) \tag{81}$$

$$P(\alpha_\pm\beta_\pm|a_{i_\perp}b_{i_\perp}) = \frac{1}{2} = \sum_\lambda P(\lambda)P(\alpha_\pm|a_{i_\perp}\lambda)P(\beta_\pm|b_{i_\perp}\lambda) \tag{82}$$

$$P(\alpha_\pm\beta_\mp|a_{i_\perp}b_i) = \frac{1}{2} = \sum_\lambda P(\lambda)P(\alpha_\pm|a_{i_\perp}\lambda)P(\beta_\mp|b_i\lambda) \tag{83}$$

$$P(\alpha_\pm\beta_\mp|a_ib_{i_\perp}) = \frac{1}{2} = \sum_\lambda P(\lambda)P(\alpha_\pm|a_i\lambda)P(\beta_\mp|b_{i_\perp}\lambda). \tag{84}$$

Then choose a value for any of the probabilities on the right hand side that does not lead into inconsistencies, e.g.

$$P(\alpha_+|a_i\lambda_1) = 0. \tag{85}$$

By (77)–(80) this entails the following probabilities:

$$\overset{\text{(CE)}}{\Rightarrow} \quad P(\alpha_-|a_i\lambda_1) = 1 \quad \underset{(80)}{\overset{(77)}{\Rightarrow}} \quad P(\beta_+|b_i\lambda_1) = 0 \tag{86}$$

$$\wedge\ P(\beta_-|b_{i_\perp}\lambda_1) = 0 \tag{87}$$

$$\overset{\text{(CE)}}{\Rightarrow} \quad P(\beta_-|b_i\lambda_1) = 1 \quad \underset{(78)}{\overset{(79)}{\Rightarrow}} \quad P(\alpha_-|a_{i_\perp}\lambda_1) = 0 \tag{88}$$

$$\wedge\ P(\beta_+|b_{i_\perp}\lambda_1) = 1 \tag{89}$$

$$\overset{\text{(CE)}}{\Rightarrow} \quad P(\alpha_+|a_{i_\perp}\lambda_1) = 1 \tag{90}$$

Similarly, choose a value for the corresponding probability conditional on λ_2, e.g.

$$P(\alpha_+|a_i\lambda_2) = 1 \tag{91}$$

and draw the appropriate consequences:

$$\overset{(77)}{\underset{(80)}{\Rightarrow}} \quad P(\beta_-|b_i\lambda_2) = 0 \tag{92}$$

$$\wedge \; P(\beta_+|b_{i_\perp}\lambda_2) = 0 \tag{93}$$

$$\overset{(CE)}{\Rightarrow} \quad P(\beta_+|b_i\lambda_2) = 1 \quad \overset{(77),(80)}{\underset{(78),(79)}{\Rightarrow}} \quad P(\alpha_-|a_i\lambda_2) = 0 \tag{94}$$

$$\wedge \; P(\beta_-|b_{i_\perp}\lambda_2) = 1 \qquad P(\alpha_+|a_{i_\perp}\lambda_2) = 0 \tag{95}$$

$$\overset{(CE)}{\Rightarrow} \quad P(\alpha_-|a_{i_\perp}\lambda_2) = 1 \tag{96}$$

These probabilities determine the values of the hidden joint probabilities consistent with equations (77)–(80). Note that we have

$$\forall\alpha,\lambda: \quad P(\alpha|a_i\lambda) \neq P(\alpha|a_{i_\perp}\lambda) \qquad \forall\alpha,a: \quad P(\alpha|a\lambda_1) \neq P(\alpha|a\lambda_2) \tag{97}$$

$$\forall\beta,\lambda: \quad P(\beta|b_i\lambda) \neq P(\beta|b_{i_\perp}\lambda) \qquad \forall\beta,b: \quad P(\beta|b\lambda_1) \neq P(\beta|b\lambda_2), \tag{98}$$

which means that the product form does not reduce to any other product form (i.e. the product form is consistent with the assumptions so far).

Inserting the determined values of the hidden joint probability into equations (81)–(84) yields:

$$P(\lambda_1) = \frac{1}{2} \qquad P(\lambda_2) = \frac{1}{2} \tag{99}$$

Finally we can freely choose, say,

$$P(a_i) = \frac{1}{2} = P(a_{i_\perp}) \qquad P(b_i) = \frac{1}{2} = P(b_{i_\perp}) \tag{100}$$

such that by the formula

$$P(\alpha\beta ab\lambda) = P(\alpha|a\lambda)P(\beta|b\lambda)P(\lambda)P(a)P(b) \tag{101}$$

we arrive at the following probability distribution:

$$\begin{array}{llll}
P(\alpha_+\beta_+a_ib_i\lambda_1) = 0 & P(\alpha_+\beta_-a_ib_i\lambda_1) = 0 & P(\alpha_-\beta_+a_ib_i\lambda_1) = 0 & P(\alpha_-\beta_-a_ib_i\lambda_1) = \frac{1}{8} \\
P(\alpha_+\beta_+a_ib_{i_\perp}\lambda_1) = 0 & P(\alpha_+\beta_-a_ib_{i_\perp}\lambda_1) = 0 & P(\alpha_-\beta_+a_ib_{i_\perp}\lambda_1) = \frac{1}{8} & P(\alpha_-\beta_-a_ib_{i_\perp}\lambda_1) = 0 \\
P(\alpha_+\beta_+a_{i_\perp}b_i\lambda_1) = 0 & P(\alpha_+\beta_-a_{i_\perp}b_i\lambda_1) = \frac{1}{8} & P(\alpha_-\beta_+a_{i_\perp}b_i\lambda_1) = 0 & P(\alpha_-\beta_-a_{i_\perp}b_i\lambda_1) = 0 \\
P(\alpha_+\beta_+a_{i_\perp}b_{i_\perp}\lambda_1) = \frac{1}{8} & P(\alpha_+\beta_-a_{i_\perp}b_{i_\perp}\lambda_1) = 0 & P(\alpha_-\beta_+a_{i_\perp}b_{i_\perp}\lambda_1) = 0 & P(\alpha_-\beta_-a_{i_\perp}b_{i_\perp}\lambda_1) = 0 \\
P(\alpha_+\beta_+a_ib_i\lambda_2) = \frac{1}{8} & P(\alpha_+\beta_-a_ib_i\lambda_2) = 0 & P(\alpha_-\beta_+a_ib_i\lambda_2) = 0 & P(\alpha_-\beta_-a_ib_i\lambda_2) = 0 \\
P(\alpha_+\beta_+a_ib_{i_\perp}\lambda_2) = 0 & P(\alpha_+\beta_-a_ib_{i_\perp}\lambda_2) = \frac{1}{8} & P(\alpha_-\beta_+a_ib_{i_\perp}\lambda_2) = 0 & P(\alpha_-\beta_-a_ib_{i_\perp}\lambda_2) = 0 \\
P(\alpha_+\beta_+a_{i_\perp}b_i\lambda_2) = 0 & P(\alpha_+\beta_-a_{i_\perp}b_i\lambda_2) = 0 & P(\alpha_-\beta_+a_{i_\perp}b_i\lambda_2) = \frac{1}{8} & P(\alpha_-\beta_-a_{i_\perp}b_i\lambda_2) = 0 \\
P(\alpha_+\beta_+a_{i_\perp}b_{i_\perp}\lambda_2) = 0 & P(\alpha_+\beta_-a_{i_\perp}b_{i_\perp}\lambda_2) = 0 & P(\alpha_-\beta_+a_{i_\perp}b_{i_\perp}\lambda_2) = 0 & P(\alpha_-\beta_-a_{i_\perp}b_{i_\perp}\lambda_2) = \frac{1}{8}
\end{array}$$

This distribution is in accordance with the axioms of probability theory; by construction its hidden joint probability has the product form that is characteristic for class (H_{29}^{α}), and it reproduces the perfect correlations and anti-correlations. This explicit example shows that class (H_{29}^{α}) is consistent with the assumptions autonomy, perfect correlations and perfect (anti-)correlations.

In a similar way one can construct examples of probability distributions for the other classes fulfilling (¬i) and (¬ii). Since (H_{22}^{α}) is symmetric to (H_{29}^{α}) under interchanging the settings, it is clear that the constructed distribution for the latter class can easily be turned into an example for the former if in each total probability one swaps the values of the settings. Furthermore, it is straightforward to modify the construction such that it yields distributions for the classes $\{(H_1^{\alpha}), \ldots, (H_{14}^{\alpha})\} \backslash \{(H_4^{\alpha}), (H_5^{\alpha}), (H_{10}^{\alpha})\}$. Note that in these classes there are more degrees of freedom than in the presented example, so one might freely choose more values of probabilities. This completes our proof of theorem 1.1.

q.e.d.

Proof of theorem 1.2

We split the theorem up into two partial claims:

Claim 1: Given autonomy, perfect correlations and perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i) and (¬ii)) implies Bell inequalities if (iii) each factor of its product form involves at most one setting.

Claim 2: Given autonomy, perfect correlations and perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i) and (¬ii)) does not imply Bell inequalities if (¬iii) at least one factor of its product form involves both settings.

Proof of claim 1

The set of classes fulfilling (¬i), (¬ii) and (iii) consists of (H_{22}) and (H_{29}). Here we have to show that, given autonomy, perfect correlations and perfect (anti-)correlations, each of these classes implies Bell inequalities.

By usual derivations of Wigner-Bell inequalities, it is well-known that local factorisation (H_{29}) implies Bell inequalities (given autonomy and perfect correlations; cf. premise (P4) of the Bell argument above). Now, it is easy to see that in a very similar way one can use (H_{22}^{α}) to derive Bell inequalities. For, as we have said, (H_{22}) differs from local factorisation only in that the settings in the product form are swapped: instead of a dependence of each outcome on the local settings each factor involves a dependence on the *distant* setting. Accordingly, the derivation from (H_{22}) results from the usual one by interchanging the settings in each expression.

Proof of claim 2

The classes fulfilling conditions (¬i) and (¬ii) while violating (iii) are $(\mathrm{H}_1^\alpha)..(\mathrm{H}_{14}^\alpha)\setminus\{(\mathrm{H}_4^\alpha),(\mathrm{H}_5^\alpha),(\mathrm{H}_{10}^\alpha)\}$. Here we have to show that given the background assumptions autonomy, perfect correlations and perfect anti-correlations, these classes do *not* imply the Bell inequalities. Since a class implying the inequalities means that all distributions of a class imply the inequalities, demonstrating that a class does not imply the inequalities amounts to showing that there is at least one distribution in that class that violates the inequalities. In other words, we have to show that there is at least one distribution for each class that fulfills autonomy, perfect correlations, perfect anti-correlations and violates the Bell inequalities.

One way to find such examples is to look at existing hidden-variable theories that successfully explain the statistics of EPR/B experiments. In our overview of the classes we have seen that the de-Broglie-Bohm theory falls under different classes depending on which temporal order the experiment has, (H_6^α), (H_9^α) or (H_{12}^α). For each of these classes, the probability distribution of the theory provides an example with the desired features. Moreover, the example for (H_9^α) can be turned into one for (H_8^α) by reversing the dependence on the settings. And similarly, the example for (H_{12}^α) can be turned into one for (H_{11}^α). Since (H_1^α), (H_2^α), (H_3^α) and (H_7^α) are stronger product forms (involve more dependences) than one or several of the classes (H_6^α), (H_8^α), (H_9^α), (H_{11}^α) or (H_{12}^α), by small modifications of the available examples one can construct examples for these classes as well.

It remains to find examples for classes (H_{13}^α) and (H_{14}^α). Since there are no theories available for these classes, here the construction has to be from scratch. Let me demonstrate how the construction works for class (H_{14}^α). We first of all take into account the perfect correlations and perfect anti-correlations. This goes, mutatis mutandis, very similar to finding a probability distribution from class (H_{29}^α) that is compatible with perfect (anti-)correlations (see proof of claim 3 in the proof of theorem 1.1). By similar equations to (77)–(80) (exchange the product form of (H_{29}^α) on the right hand side with the product form of (H_{14}^α)), for any i and λ there are two possible cases:

Case I:

$$P(\alpha_+|\lambda)=0 \qquad P(\alpha_-|\lambda)=1 \tag{102}$$

$$P(\beta_+|a_ib_i\lambda)=0 \qquad P(\beta_-|a_ib_i\lambda)=1 \qquad P(\beta_+|a_{i_\perp}b_{i_\perp}\lambda)=0 \qquad P(\beta_-|a_{i_\perp}b_{i_\perp}\lambda)=1 \tag{103}$$

$$P(\beta_+|a_ib_{i_\perp}\lambda)=1 \qquad P(\beta_-|a_ib_{i_\perp}\lambda)=0 \qquad P(\beta_+|a_{i_\perp}b_i\lambda)=1 \qquad P(\beta_-|a_{i_\perp}b_i\lambda)=0 \tag{104}$$

Case II: (replace all 0's in case I by 1 and vice versa)

Requiring just *any* example we can assume a toy model with only two possible hidden states ($\boldsymbol{\lambda}=1,2$). Then we might, for instance, choose case I for λ_1 and case II for λ_2 for all i's. Then, by equations similar to (81)–(84) it follows

$$P(\lambda_1)=\frac{1}{2} \qquad P(\lambda_2)=\frac{1}{2} \tag{105}$$

In this way we have accounted for the perfect correlations as well as for the perfect anti-correlations.

Now it remains to reproduce the EPR/B correlations for non-parallel and non-perpendicular settings. A minimal set of such probabilities, which can violate the Bell inequalities (both the usual ones as well as the Wigner-Bell inequalities), can be found if each of the settings $\boldsymbol{a}$ and $\boldsymbol{b}$ has two possible values, e.g. $a_1 = 0°$, $a_2 = 30°$, $b_1 = 30°$ and $b_2 = 60°$. Measuring the quantum state $\psi_0 = (|++\rangle + |--\rangle)/\sqrt{2}$ at these settings yields the following observable probabilities:

$$P(\alpha_\pm\beta_\pm|a_1b_1) = \tfrac{3}{8} \quad P(\alpha_\pm\beta_\mp|a_1b_1) = \tfrac{1}{8} \quad P(\alpha_\pm\beta_\pm|a_1b_2) = \tfrac{1}{8} \quad P(\alpha_\pm\beta_\mp|a_1b_2) = \tfrac{3}{8} \tag{106}$$

$$P(\alpha_\pm\beta_\pm|a_2b_1) = \tfrac{1}{2} \quad P(\alpha_\pm\beta_\mp|a_2b_1) = 0 \quad P(\alpha_\pm\beta_\pm|a_2b_2) = \tfrac{3}{8} \quad P(\alpha_\pm\beta_\mp|a_2b_2) = \tfrac{1}{8} \tag{107}$$

These are sixteen equations, and any of the probabilities on their left hand sides can be expressed by the product form of the hidden joint probability:

$$P(\alpha\beta|ab) = \sum_\lambda P(\lambda)P(\alpha|\lambda)P(\beta|ab\lambda) \tag{108}$$

$P(\lambda)$ and $P(\alpha|\lambda)$ are already completely determined by the perfect (anti-)correlations, $P(\beta|ab\lambda)$ partly so (namely only for the parallel settings $a_2 = b_1$):

$$P(\lambda_1) = \tfrac{1}{2} \quad P(\lambda_2) = \tfrac{1}{2} \tag{109}$$

$$P(\alpha_+|\lambda_1) = 0 \quad P(\alpha_-|\lambda_1) = 1 \quad P(\alpha_+|\lambda_2) = 1 \quad P(\alpha_-|\lambda_2) = 0 \tag{110}$$

$$P(\beta_+|a_2b_1\lambda_1) = 0 \quad P(\beta_-|a_2b_1\lambda_1) = 1 \quad P(\beta_+|a_2b_1\lambda_2) = 1 \quad P(\beta_-|a_2b_1\lambda_2) = 0 \tag{111}$$

Inserting these values in the respective equations yields the following consistent values for the missing probabilities $P(\beta|ab\lambda)$:

$$P(\beta_+|a_1b_1\lambda_1) = \tfrac{1}{4} \quad P(\beta_-|a_1b_1\lambda_1) = \tfrac{3}{4} \quad P(\beta_+|a_1b_1\lambda_2) = \tfrac{3}{4} \quad P(\beta_-|a_1b_1\lambda_2) = \tfrac{1}{4} \tag{112}$$

$$P(\beta_+|a_1b_2\lambda_1) = \tfrac{3}{4} \quad P(\beta_-|a_1b_2\lambda_1) = \tfrac{1}{4} \quad P(\beta_+|a_1b_2\lambda_2) = \tfrac{1}{4} \quad P(\beta_-|a_1b_1\lambda_2) = \tfrac{3}{4} \tag{113}$$

$$P(\beta_+|a_2b_2\lambda_1) = \tfrac{1}{4} \quad P(\beta_-|a_2b_2\lambda_1) = \tfrac{3}{4} \quad P(\beta_+|a_2b_2\lambda_2) = \tfrac{3}{4} \quad P(\beta_-|a_2b_2\lambda_2) = \tfrac{1}{4} \tag{114}$$

Finally, choosing, say,

$$P(a_i) = \frac{1}{2} \quad P(a_{i_\perp}) = \frac{1}{2} \quad P(b_i) = \frac{1}{2} \quad P(b_{i_\perp}) = \frac{1}{2} \tag{115}$$

the formula

$$P(\alpha\beta ab\lambda) = P(\alpha|\lambda)P(\beta|ab\lambda)P(\lambda)P(a)P(b) \tag{116}$$

entails the following total probabilities, which constitute the searched for probability distribution:

$$\begin{array}{llll}
P(\alpha_+\beta_+a_1b_1\lambda_1)=0 & P(\alpha_+\beta_-a_1b_1\lambda_1)=0 & P(\alpha_-\beta_+a_1b_1\lambda_1)=\frac{1}{32} & P(\alpha_-\beta_-a_1b_1\lambda_1)=\frac{3}{32}\\
P(\alpha_+\beta_+a_1b_2\lambda_1)=0 & P(\alpha_+\beta_-a_1b_2\lambda_1)=0 & P(\alpha_-\beta_+a_1b_2\lambda_1)=\frac{3}{32} & P(\alpha_-\beta_-a_1b_2\lambda_1)=\frac{1}{32}\\
P(\alpha_+\beta_+a_2b_1\lambda_1)=0 & P(\alpha_+\beta_-a_2b_1\lambda_1)=0 & P(\alpha_-\beta_+a_2b_1\lambda_1)=0 & P(\alpha_-\beta_-a_2b_1\lambda_1)=\frac{1}{8}\\
P(\alpha_+\beta_+a_2b_2\lambda_1)=0 & P(\alpha_+\beta_-a_2b_2\lambda_1)=0 & P(\alpha_-\beta_+a_2b_2\lambda_1)=\frac{1}{32} & P(\alpha_-\beta_-a_2b_2\lambda_1)=\frac{3}{32}\\
P(\alpha_+\beta_+a_1b_1\lambda_2)=\frac{3}{32} & P(\alpha_+\beta_-a_1b_1\lambda_2)=\frac{1}{32} & P(\alpha_-\beta_+a_1b_1\lambda_2)=0 & P(\alpha_-\beta_-a_1b_1\lambda_2)=0\\
P(\alpha_+\beta_+a_1b_2\lambda_2)=\frac{1}{32} & P(\alpha_+\beta_-a_1b_2\lambda_2)=\frac{3}{32} & P(\alpha_-\beta_+a_1b_2\lambda_2)=0 & P(\alpha_-\beta_-a_1b_2\lambda_2)=0\\
P(\alpha_+\beta_+a_2b_1\lambda_2)=\frac{1}{8} & P(\alpha_+\beta_-a_2b_1\lambda_2)=0 & P(\alpha_-\beta_+a_2b_1\lambda_2)=0 & P(\alpha_-\beta_-a_2b_1\lambda_2)=0\\
P(\alpha_+\beta_+a_2b_2\lambda_2)=\frac{3}{32} & P(\alpha_+\beta_-a_2b_2\lambda_2)=\frac{1}{32} & P(\alpha_-\beta_+a_2b_2\lambda_2)=0 & P(\alpha_-\beta_-a_2b_2\lambda_2)=0
\end{array}$$

Note that here we have not explicitly noted the probabilities for parallel or perpendicular settings, but by constructing the distribution in the indicated way we have implicitly taken account of the perfect (anti-)correlations at these settings and it is straight forward to extent the distribution to include these settings as well (the distribution just becomes much longer, when for each measurement setting at one side one includes a parallel and a perpendicular setting at the other side).

This completes our construction of a distribution from class (H^{α}_{14}) which respects, autonomy, perfect correlations, perfect anti-correlations and violates the Bell inequalities. In a similar way, one can construct an example for class (H^{α}_{13}), which differs from (H^{α}_{14}) just in that the dependence on both settings is not in the second but in the first factor of its product form. q.e.d.

Proof of theorem 2.1

We split the theorem up into two partial claims:

Claim 1: Autonomy, nearly perfect correlations, nearly perfect anti-correlations, and a class of probability distributions (H^{α}_{i}) form an inconsistent set if (i) the product form of (H^{α}_{i}) involves at most one of the settings.

Claim 2: A class (H^{α}_{i}) is consistent with autonomy, nearly perfect correlations and nearly perfect anti-correlations if (¬i) the product form of (H^{α}_{i}) involves both settings.

<u>Proof of claim 1</u>

Condition (i), that the product form involves at most one of the settings, is fulfilled by the classes $\{(\mathrm{H}^{\alpha}_{17}),\ldots,(\mathrm{H}^{\alpha}_{32})\}\backslash\{(\mathrm{H}^{\alpha}_{22}),(\mathrm{H}^{\alpha}_{29})\}$. Here we have to show the inconsistency of these classes with the set of assumptions autonomy, nearly perfect correlations and nearly perfect anti-correlations.

The proof runs very similar to our demonstration of claim 1 in the proof of theorem 1.1. On the one hand, the nearly perfect correlations and anti-correlations involve dependences on each of the settings, e.g. the nearly perfect correlations

$$P(\alpha_\pm\beta_\pm|a_ib_i) = \frac{1}{2} - \delta_{ii} \qquad\qquad P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = \delta_{ii_\perp} \tag{117}$$

reveal a dependence on the setting $\boldsymbol{b}$, while e.g. the nearly perfect correlations

$$P(\alpha_\pm\beta_\pm|a_ib_i) = \frac{1}{2} - \delta_{ii} \qquad\qquad P(\alpha_\pm\beta_\pm|a_{i_\perp}b_i) = \delta_{i_\perp i} \tag{118}$$

show a dependence on the setting $\boldsymbol{a}$. On the other hand, any hidden joint probability that does not involve the setting $\boldsymbol{b}$ (i.e. is independent of $\boldsymbol{b}$), cannot account for changing values in the empirical joint probability with changing values of $\boldsymbol{b}$ (cf. (63)); so it necessarily contradicts the set of equations (117). And similarly, hidden joint probabilities that are independent of $\boldsymbol{a}$ contradict the set of equations (118).

Note that condition (ii) from theorem 1.1 is not a criterion for inconsistency according to theorem 2.1, because the inconsistency in question essentially relies on *strictly* perfect (anti-)correlations, which are not assumed in theorem 2.1.

Proof of claim 2

The classes fulfilling criterion (¬i) to involve both settings in their product forms are (H_1^α)...(H_{16}^α), (H_{22}^α) and (H_{29}^α). Here we have to show the consistency of these classes with the set of assumptions autonomy, nearly perfect correlations and nearly perfect anti-correlations.

As in the proof of claim 3 in the proof of theorem 1.1 one can demonstrate the present claim by providing an example of a probability distribution for each class that is consistent with these assumptions. Since nearly perfect correlations are a weaker requirement than strictly perfect ones, it is clear that for all classes which we have shown to be consistent with the latter—viz. $(\mathrm{H}_1^\alpha)..(\mathrm{H}_{14}^\alpha)\backslash\{(\mathrm{H}_4^\alpha), (\mathrm{H}_5^\alpha), (\mathrm{H}_{10}^\alpha)\}$—are also consistent with the former. Therefore, what still needs to be proven here is that autonomy and nearly perfect (anti-)correlations are consistent with those classes fulfilling criterion (¬i) that are inconsistent with the strictly perfect ones (because they fulfill (ii)). The classes in question are (H_4^α), (H_5^α), (H_{10}^α), (H_{15}^α) and (H_{16}^α).

Again, the best way to find such examples is by constructing them such that the conditions are fulfilled. Here we show how to construct a distribution for class (H_{10}^α). The starting point are the equations for nearly perfect (anti-)correlations:

$$P(\alpha_\pm\beta_\mp|a_ib_i) = \delta_{ii} \quad P(\alpha_\pm\beta_\mp|a_{i_\perp}b_{i_\perp}) = \delta_{i_\perp i_\perp} \quad P(\alpha_\pm\beta_\pm|a_{i_\perp}b_i) = \delta_{i_\perp i} \quad P(\alpha_\pm\beta_\pm|a_ib_{i_\perp}) = \delta_{ii_\perp}$$
$$P(\alpha_\pm\beta_\pm|a_ib_i) = \tfrac{1}{2} - \delta_{ii} \quad P(\alpha_\pm\beta_\pm|a_{i_\perp}b_{i_\perp}) = \tfrac{1}{2} - \delta_{i_\perp i_\perp} \quad P(\alpha_\pm\beta_\mp|a_{i_\perp}b_i) = \tfrac{1}{2} - \delta_{i_\perp i} \quad P(\alpha_\pm\beta_\mp|a_ib_{i_\perp}) = \tfrac{1}{2} - \delta_{ii_\perp}$$

Replacing the empirical probability on the left hand side of each equation by an equivalent expression involving hidden probabilities of the product form,

$$P(\alpha\beta|ab) = \sum_{\lambda} P(\lambda)P(\alpha|\beta\lambda)P(\beta|ab\lambda), \tag{119}$$

yields a set of equations, whose solutions determine probability distributions with the required features.

The δ's in these equations indicate the deviation from strictly perfect correlations. One might use realistic empirical values for them but since the task here is a merely conceptual one, one might as well just stipulate any small, positive values. Due to the lacking perfectness, the resulting set of equations is more complicated than that in theorem 1.2, and solutions are best determined by appropriate computer algorithms. Here, we shall present a solution for the special case

$$\delta_{ii} = \delta_{i_\perp i_\perp} = \delta_{i_\perp i} = \delta_{ii_\perp} =: \delta, \tag{120}$$

which reads:

$$P(\lambda_1) = \tfrac{1}{2} \qquad P(\lambda_2) = \tfrac{1}{2} \tag{121}$$

$$P(\alpha_+|\beta_+\lambda_1) = 0 \qquad P(\alpha_-|\beta_+\lambda_1) = 1 \qquad P(\alpha_+|\beta_+\lambda_2) = 1-2\delta \qquad P(\alpha_-|\beta_+\lambda_2) = 2\delta \tag{122}$$

$$P(\alpha_+|\beta_-\lambda_1) = 2\delta \qquad P(\alpha_-|\beta_-\lambda_1) = 1-2\delta \qquad P(\alpha_+|\beta_-\lambda_2) = 1 \qquad P(\alpha_-|\beta_-\lambda_2) = 0 \tag{123}$$

$$P(\beta_+|a_ib_i\lambda_1) = 0 \qquad P(\beta_-|a_ib_i\lambda_1) = 1 \qquad P(\beta_+|a_ib_i\lambda_2) = 1 \qquad P(\beta_-|a_ib_i\lambda_2) = 0 \tag{124}$$

$$P(\beta_+|a_ib_{i_\perp}\lambda_1) = \tfrac{4\delta-1}{2\delta-1} \qquad P(\beta_-|a_ib_{i_\perp}\lambda_1) = \tfrac{2\delta}{1-2\delta} \qquad P(\beta_+|a_ib_{i_\perp}\lambda_2) = \tfrac{2\delta}{1-2\delta} \qquad P(\beta_-|a_ib_{i_\perp}\lambda_2) = \tfrac{4\delta-1}{2\delta-1} \tag{125}$$

$$P(\beta_+|a_{i_\perp}b_i\lambda_1) = \tfrac{4\delta-1}{2\delta-1} \qquad P(\beta_-|a_{i_\perp}b_i\lambda_1) = \tfrac{2\delta}{1-2\delta} \qquad P(\beta_+|a_{i_\perp}b_i\lambda_2) = \tfrac{2\delta}{1-2\delta} \qquad P(\beta_-|a_{i_\perp}b_{i_\perp}\lambda_2) = \tfrac{4\delta-1}{2\delta-1} \tag{126}$$

$$P(\beta_+|a_{i_\perp}b_{i_\perp}\lambda_1) = 0 \qquad P(\beta_-|a_{i_\perp}b_{i_\perp}\lambda_1) = 1 \qquad P(\beta_+|a_{i_\perp}b_{i_\perp}\lambda_2) = 1 \qquad P(\beta_-|a_{i_\perp}b_{i_\perp}\lambda_2) = 0 \tag{127}$$

Note that according to this solution all dependences of the product form (H_{10}^{α}) are preserved, because, for instance, we have

$$P(\alpha_+|\beta_+\lambda_1) \neq P(\alpha_+|\beta_-\lambda_1) \qquad P(\alpha_+|\beta_+\lambda_1) \neq P(\alpha_+|\beta_+\lambda_2) \tag{128}$$

$$P(\beta_+|a_ib_i\lambda_1) \neq P(\beta_+|a_{i_\perp}b_i\lambda_1) \qquad P(\beta_+|a_ib_i\lambda_1) \neq P(\beta_+|a_ib_{i_\perp}\lambda_1) \tag{129}$$

$$P(\beta_+|a_ib_i\lambda_1) \neq P(\beta_+|a_ib_i\lambda_2) \tag{130}$$

Finally, when we further assume, say,

$$P(a_i) = \frac{1}{2} \qquad P(a_{i_\perp}) = \frac{1}{2} \qquad P(b_i) = \frac{1}{2} \qquad P(b_{i_\perp}) = \frac{1}{2} \tag{131}$$

by the equation

$$P(\alpha\beta ab\lambda) = P(\alpha|\beta\lambda)P(\beta|ab\lambda)P(\lambda)P(a)P(b) \tag{132}$$

the results so far determine the values of the total probability distribution:

$$P(\alpha_+\beta_+a_ib_i\lambda_1) = 0 \qquad P(\alpha_-\beta_+a_ib_i\lambda_1) = 0 \tag{133}$$
$$P(\alpha_+\beta_+a_ib_{i_\perp}\lambda_1) = 0 \qquad P(\alpha_-\beta_+a_ib_{i_\perp}\lambda_1) = \tfrac{1-4\delta}{8(1-2\delta)} \tag{134}$$
$$P(\alpha_+\beta_+a_{i_\perp}b_i\lambda_1) = 0 \qquad P(\alpha_-\beta_+a_{i_\perp}b_i\lambda_1) = \tfrac{1-4\delta}{8(1-2\delta)} \tag{135}$$
$$P(\alpha_+\beta_+a_{i_\perp}b_{i_\perp}\lambda_1) = 0 \qquad P(\alpha_-\beta_+a_{i_\perp}b_{i_\perp}\lambda_1) = 0 \tag{136}$$

$$P(\alpha_+\beta_-a_ib_i\lambda_1) = \tfrac{\delta}{4} \qquad P(\alpha_-\beta_-a_ib_i\lambda_1) = \tfrac{1}{8}(1-2\delta) \tag{137}$$
$$P(\alpha_+\beta_-a_ib_{i_\perp}\lambda_1) = \tfrac{\delta^2}{2(1-2\delta)} \qquad P(\alpha_-\beta_-a_ib_{i_\perp}\lambda_1) = \tfrac{\delta}{4} \tag{138}$$
$$P(\alpha_+\beta_-a_{i_\perp}b_i\lambda_1) = \tfrac{\delta^2}{2(1-2\delta)} \qquad P(\alpha_-\beta_-a_{i_\perp}b_i\lambda_1) = \tfrac{\delta}{4} \tag{139}$$
$$P(\alpha_+\beta_-a_{i_\perp}b_{i_\perp}\lambda_1) = \tfrac{\delta}{4} \qquad P(\alpha_-\beta_-a_{i_\perp}b_{i_\perp}\lambda_1) = \tfrac{1}{8}(1-2\delta) \tag{140}$$

$$P(\alpha_+\beta_+a_ib_i\lambda_2) = \tfrac{1}{8}(1-2\delta) \qquad P(\alpha_-\beta_+a_ib_i\lambda_2) = \tfrac{\delta}{4} \tag{141}$$
$$P(\alpha_+\beta_+a_ib_{i_\perp}\lambda_2) = \tfrac{\delta}{4} \qquad P(\alpha_-\beta_+a_ib_{i_\perp}\lambda_2) = \tfrac{\delta^2}{2(1-2\delta)} \tag{142}$$
$$P(\alpha_+\beta_+a_{i_\perp}b_i\lambda_2) = \tfrac{\delta}{4} \qquad P(\alpha_-\beta_+a_{i_\perp}b_i\lambda_2) = \tfrac{\delta^2}{2(1-2\delta)} \tag{143}$$
$$P(\alpha_+\beta_+a_{i_\perp}b_{i_\perp}\lambda_2) = \tfrac{1}{8}(1-2\delta) \qquad P(\alpha_-\beta_+a_{i_\perp}b_{i_\perp}\lambda_2) = \tfrac{\delta}{4} \tag{144}$$

$$P(\alpha_+\beta_-a_ib_i\lambda_2) = 0 \qquad P(\alpha_-\beta_-a_ib_i\lambda_2) = 0 \tag{145}$$
$$P(\alpha_+\beta_-a_ib_{i_\perp}\lambda_2) = \tfrac{1-4\delta}{8(1-2\delta)} \qquad P(\alpha_-\beta_-a_ib_{i_\perp}\lambda_2) = 0 \tag{146}$$
$$P(\alpha_+\beta_-a_{i_\perp}b_i\lambda_2) = \tfrac{1-4\delta}{8(1-2\delta)} \qquad P(\alpha_-\beta_-a_{i_\perp}b_i\lambda_2) = 0 \tag{147}$$
$$P(\alpha_+\beta_-a_{i_\perp}b_{i_\perp}\lambda_2) = 0 \qquad P(\alpha_-\beta_-a_{i_\perp}b_{i_\perp}\lambda_2) = 0 \tag{148}$$

By construction this distribution has the product form that is characteristic for class (H^α_{10}), and it involves autonomy, nearly perfect correlations for parallel settings and nearly perfect anti-correlations for perpendicular settings. This explicitly shows class (H^α_{14}) to be consistent with these assumptions. In a similar way, one can find examples for classes (H^α_4), (H^α_5), (H^α_{15}) and (H^α_{16}) consistent with the mentioned assumptions.

q.e.d.

Proof of theorem 2.2

We split the theorem up into two partial claims:

Claim 1: Given autonomy, nearly perfect correlations and nearly perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i)) implies Bell inequalities if (iii) each factor of its product form involves at most one setting.

Claim 2: Given autonomy, nearly perfect correlations and nearly perfect anti-correlations a consistent class (i.e. a class that fulfills (¬i)) does not imply Bell inequalities if (¬iii) at least one factor of its product form involves both settings.

Proof of claim 1

The set of classes fulfilling (¬i) and (iii) consists of (H_{15}), (H_{16}), (H_{22}) and (H_{29}). It has to be shown that given autonomy, nearly perfect correlations and nearly perfect anti-correlations, each of these class implies Bell inequalities.

In lemma 2 we have already demonstrated that under these conditions (H^{α}_{16}) implies Bell inequalities. Since (H^{α}_{15}) only differs from (H^{α}_{16}) in that the settings are swapped in the product form, mutatis mutandis also (H^{α}_{15}) implies the inequalities. Finally, since local factorization (H^{α}_{29}) is a weaker product form than (H^{α}_{16}), and since (H^{α}_{22}) is a weaker form than (H^{α}_{15}), it is clear that also these other two product forms imply the Bell inequalities in the given circumstances.

Note that though it might seem obvious that local factorization implies the inequalities, it is a non-trivial claim that it does imply the Wigner-Bell inequalities with only *nearly* perfect (anti-)correlations, because usual derivations so far did have to assume strictly perfect correlations; however, our derivation with (H^{α}_{16}) can easily be adopted to derive the inequalities from local factorization and just the nearly perfect (anti-)correlations. q.e.d.

Proof of claim 2

The classes fulfilling condition (¬i) while violating (iii) are (H^{α}_{1})...(H^{α}_{14}). Here we have to show that given the background assumptions autonomy, nearly perfect correlations and nearly perfect anti-correlations, these classes do *not* imply the Bell inequalities. This amounts to showing that there is at least one distribution for each class that fulfills the background assumptions and violates the Bell inequalities.

We know already from theorem 1.2 that the classes $(H^{\alpha}_{1})..(H^{\alpha}_{14})\backslash\{(H^{\alpha}_{4}),(H^{\alpha}_{5}),(H^{\alpha}_{10})\}$ can violate the inequalities given the assumptions of autonomy and *strictly* perfect correlations. Since the latter are a stronger condition than *nearly* perfect correlations, it is clear that these classes can violate the Bell inequalities also in the present case. It remains to show that the classes $(H^{\alpha}_{4}),(H^{\alpha}_{5}),(H^{\alpha}_{10})$ can violate the inequalities under the given assumptions. Here we explicitly construct an example for class (H^{α}_{10}).

In the proof of claim 2 of theorem 2.1 we have constructed a toy example of a probability distribution for this class that is compatible with autonomy and nearly perfect (anti-)correlations. When, for any setting i, we use the resulting probabilities (121)–(127) we can be sure that the distribution we are about to construct is consistent with the nearly perfect (anti-)correlations. What remains to be done is to reproduce the EPR/B correlations for non-parallel and non-perpendicular settings. We again choose the settings $a_1 = 0°$, $a_2 = 30°$, $b_1 = 30°$ and $b_2 = 60°$ as well as the quantum state $\psi = (|++\rangle + |--\rangle)/\sqrt{2}$. Then the observable probabilities read:

$$P(\alpha_\pm\beta_\pm|a_1b_1) = \tfrac{3}{8} \quad P(\alpha_\pm\beta_\mp|a_1b_1) = \tfrac{1}{8} \quad P(\alpha_\pm\beta_\pm|a_1b_2) = \tfrac{1}{8} \quad P(\alpha_\pm\beta_\mp|a_1b_2) = \tfrac{3}{8} \tag{149}$$
$$P(\alpha_\pm\beta_\pm|a_2b_1) = \tfrac{1}{2} - \delta \quad P(\alpha_\pm\beta_\mp|a_2b_1) = \delta \quad P(\alpha_\pm\beta_\pm|a_2b_2) = \tfrac{3}{8} \quad P(\alpha_\pm\beta_\mp|a_2b_2) = \tfrac{1}{8} \tag{150}$$

(Note the difference to the probabilities with the same settings and quantum state in (106)–(107), which involve strictly perfect anti-correlations for parallel settings $a_2 = b_1$

($P(\alpha_\pm\beta_\pm|a_2b_1) = \frac{1}{2}$ and $P(\alpha_\pm\beta_\mp|a_2b_1) = 0$) instead of nearly perfect ones ($P(\alpha_\pm\beta_\pm|a_2b_1) = \frac{1}{2} - \delta$ and $P(\alpha_\pm\beta_\mp|a_2b_1) = \delta$).
These are sixteen equations, and any of the probabilities on their left hand sides can be expressed by the product form of the hidden joint probability:

$$P(\alpha\beta|ab) = \sum_\lambda P(\lambda)P(\alpha|\beta\lambda)P(\beta|ab\lambda) \tag{151}$$

$P(\lambda)$ and $P(\alpha|\beta\lambda)$ are completely determined by the requirements of the perfect (anti-)correlations (121)–(123), $P(\beta|ab\lambda)$ partly so (namely only for the parallel settings, (124)).

Inserting these predetermined probabilities into equations (149)–(150) yields the following consistent values for the missing probabilities $P(\beta|ab\lambda)$:

$$P(\beta_+|a_1b_1\lambda_1) = \tfrac{1-8\delta}{4(1-2\delta)} \quad P(\beta_-|a_1b_1\lambda_1) = \tfrac{3}{4(1-2\delta)} \quad P(\beta_+|a_1b_1\lambda_2) = \tfrac{3}{4(1-2\delta)} \quad P(\beta_-|a_1b_1\lambda_2) = \tfrac{1-8\delta}{4(1-2\delta)} \tag{152}$$

$$P(\beta_+|a_1b_2\lambda_1) = \tfrac{3-8\delta}{4(1-2\delta)} \quad P(\beta_-|a_1b_2\lambda_1) = \tfrac{1}{4(1-2\delta)} \quad P(\beta_+|a_1b_2\lambda_2) = \tfrac{1}{4(1-2\delta)} \quad P(\beta_-|a_1b_1\lambda_2) = \tfrac{3-8\delta}{4(1-2\delta)} \tag{153}$$

$$P(\beta_+|a_2b_2\lambda_1) = \tfrac{1-8\delta}{4(1-2\delta)} \quad P(\beta_-|a_2b_2\lambda_1) = \tfrac{3}{4(1-2\delta)} \quad P(\beta_+|a_2b_2\lambda_2) = \tfrac{3}{4(1-2\delta)} \quad P(\beta_-|a_2b_2\lambda_2) = \tfrac{1-8\delta}{4(1-2\delta)} \tag{154}$$

Finally, choosing, say,

$$P(a_i) = \frac{1}{2} \qquad P(a_{i_\perp}) = \frac{1}{2} \qquad P(b_i) = \frac{1}{2} \qquad P(b_{i_\perp}) = \frac{1}{2} \tag{155}$$

the formula

$$P(\alpha\beta ab\lambda) = P(\alpha|\lambda)P(\beta|ab\lambda)P(\lambda)P(a)P(b) \tag{156}$$

entails the following total probabilities:

$$\begin{array}{llll}
P(\alpha_+\beta_+a_1b_1\lambda_1) = 0 & P(\alpha_+\beta_-a_1b_1\lambda_1) = \frac{3\delta}{16(1-2\delta)} & P(\alpha_-\beta_+a_1b_1\lambda_1) = \frac{1-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_-a_1b_1\lambda_1) = \frac{3}{32} \\
P(\alpha_+\beta_+a_1b_2\lambda_1) = 0 & P(\alpha_+\beta_-a_1b_2\lambda_1) = \frac{\delta}{16(1-2\delta)} & P(\alpha_-\beta_+a_1b_2\lambda_1) = \frac{3-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_-a_1b_2\lambda_1) = \frac{1}{32} \\
P(\alpha_+\beta_+a_2b_1\lambda_1) = 0 & P(\alpha_+\beta_-a_2b_1\lambda_1) = \frac{\delta}{4} & P(\alpha_-\beta_+a_2b_1\lambda_1) = 0 & P(\alpha_-\beta_-a_2b_1\lambda_1) = \frac{1-2\delta}{8} \\
P(\alpha_+\beta_+a_2b_2\lambda_1) = 0 & P(\alpha_+\beta_-a_2b_2\lambda_1) = \frac{3\delta}{16(1-2\delta)} & P(\alpha_-\beta_+a_2b_2\lambda_1) = \frac{1-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_-a_2b_2\lambda_1) = \frac{3}{32} \\
P(\alpha_+\beta_+a_1b_1\lambda_2) = \frac{3}{32} & P(\alpha_+\beta_-a_1b_1\lambda_2) = \frac{1-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_+a_1b_1\lambda_2) = \frac{3\delta}{16(1-2\delta)} & P(\alpha_-\beta_-a_1b_1\lambda_2) = 0 \\
P(\alpha_+\beta_+a_1b_2\lambda_2) = \frac{1}{32} & P(\alpha_+\beta_-a_1b_2\lambda_2) = \frac{3-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_+a_1b_2\lambda_2) = \frac{\delta}{16(1-2\delta)} & P(\alpha_-\beta_-a_1b_2\lambda_2) = 0 \\
P(\alpha_+\beta_+a_2b_1\lambda_2) = \frac{1-2\delta}{8} & P(\alpha_+\beta_-a_2b_1\lambda_2) = 0 & P(\alpha_-\beta_+a_2b_1\lambda_2) = \frac{\delta}{4} & P(\alpha_-\beta_-a_2b_1\lambda_2) = 0 \\
P(\alpha_+\beta_+a_2b_2\lambda_2) = \frac{3}{32} & P(\alpha_+\beta_-a_2b_2\lambda_2) = \frac{1-8\delta}{32(1-2\delta)} & P(\alpha_-\beta_+a_2b_2\lambda_2) = \frac{3\delta}{16(1-2\delta)} & P(\alpha_-\beta_-a_2b_2\lambda_2) = 0
\end{array}$$

This completes our construction of a distribution from class (H_{10}^α) which respects, autonomy, nearly perfect correlations, nearly perfect anti-correlations and violates the Bell inequalities. Similarly, one can construct examples of distributions for class (H_4^α) and (H_5^α).
q.e.d.

Proof of theorem 3

One can demonstrate the equivalence between product forms and conjunctions of independences for each hidden joint probability separately (analogous to how Jarrett derived (P7)), but the following constructive method is more elegant: in the case that the hidden joint probability factorises according to the product rule, (H_1^α), none of the relevant independences holds (and vice versa). Then we consider the five cases in which exactly *one* independence holds (H_2^α)–(H_6^α). Here is the proof of (H_2^α) $\leftrightarrow$ (ℓPI_2^β):

$$\boxed{\leftarrow} \quad P(\alpha\beta|ab\lambda) \quad = \quad P(\alpha|\beta ba\lambda)P(\beta|ab\lambda) \quad \overset{(\ell PI_2^\beta)}{=} \quad P(\alpha|\beta ba\lambda)P(\beta|a\lambda) \tag{157}$$

$$\boxed{\rightarrow} \quad P(\beta|ab\lambda) = \sum_\alpha P(\alpha\beta|ab\lambda) \overset{(H_2^\alpha)}{=} P(\beta|a\lambda)\sum_\alpha P(\alpha|\beta ba\lambda) = P(\beta|a\lambda) \tag{158}$$

The equivalence (H_3^α) $\leftrightarrow$ (PI_2^β) can be shown mutatis mutandis (just swap the local with the distant parameter). (H_4^α) $\leftrightarrow$ (ℓPI_1^α) can be derived as follows:

$$\boxed{\leftarrow} \quad P(\alpha\beta|ab\lambda) \quad = \quad P(\alpha|\beta ba\lambda)P(\beta|ab\lambda) \quad \overset{(\ell PI_1^\alpha)}{=} \quad P(\alpha|b\beta\lambda)P(\beta|ab\lambda) \tag{159}$$

$$\boxed{\rightarrow} \quad P(\alpha|\beta ba\lambda) = \frac{P(\alpha\beta|ab\lambda)}{P(\beta|ab\lambda)} \overset{(H_4^\alpha)}{=} \frac{P(\alpha|\beta b\lambda)\cancel{P(\beta|ab\lambda)}}{\cancel{P(\beta|ab\lambda)}} = P(\alpha|\beta b\lambda) \tag{160}$$

The equivalences (H_5^α) $\leftrightarrow$ (PI_1^α) and (H_6) $\leftrightarrow$ (OI_1) are proved similarly. Then, by pairs of these five equivalences involving one independence, we prove equivalences with two independences, and subsequently, equivalences with three independences, and so on. Here is an example how to derive an equivalence with two independences, (H_7^α) $\leftrightarrow$ (PI_2^β)$\wedge$(ℓPI_2^β), on the basis of the corresponding equivalences with one independence respectively:

$$\boxed{\leftarrow} \quad (PI_2^\beta) \wedge (\ell PI_2^\beta) \overset{(157),\,(158)}{\longleftrightarrow} (PI_2^\beta) \wedge (H_2^\alpha) \overset{(161)}{\rightarrow} (H_7^\alpha)$$

$$P(\alpha\beta|ab\lambda) \quad \overset{(H_2^\alpha)}{=} \quad P(\alpha|\beta ba\lambda)P(\beta|ab'\lambda) \quad \overset{(PI_2^\beta)}{=} \quad P(\alpha|\beta ba\lambda)P(\beta|a'b'\lambda) \tag{161}$$

$$\boxed{\rightarrow} \quad (H_7^\alpha) \overset{(*)}{\rightarrow} (H_3^\alpha) \leftrightarrow (PI_2^\beta); \quad (H_7^\alpha) \overset{(*)}{\rightarrow} (H_2^\alpha) \leftrightarrow (\ell PI_2^\beta)$$

($*$): (H_7^α) is a common special case of (H_2^α) and (H_3^α); if (H_7^α) holds, then a forteriori (H_2^α) and (H_3^α):

$$\forall a, a', b, b' : \quad P(\alpha\beta|ab\lambda) = P(\alpha|\beta ba\lambda)P(\beta|a'b'\lambda) \tag{H_7^α}$$

$$\forall a = a', b, b' : \quad P(\alpha\beta|ab\lambda) = P(\alpha|\beta ba\lambda)P(\beta|a'b'\lambda) \tag{H_2^α}$$

$$\forall a, a', b = b' : \quad P(\alpha\beta|ab\lambda) = P(\alpha|\beta ba\lambda)P(\beta|a'b'\lambda) \tag{H_3^α}$$

Similarly, one can derive step by step the other equivalences between product forms and independences in table 1.