# Bell vs Bell: a ding-dong battle over quantum incompleteness

# Michael J. W. Hall

Theoretical Physics, Research School of Physics, Australian National University

#### Abstract

Does determinism (or even the incompleteness of quantum mechanics) follow from locality and perfect correlations? In a 1964 paper John Bell gave the first demonstration that quantum mechanics is incompatible with local hidden variables. Since then a vigorous debate has rung out over whether he relied on an assumption of determinism or instead, as he later claimed in a 1981 paper, derived determinism from assumptions of locality and perfect correlation. This paper aims to bring clarity to the debate via simple examples and rigorous results. It is first recalled, via quantum and classical counterexamples, that the weakest statistical form of locality consistent with Bell's 1964 paper (parameter independence) is insufficient for the derivation of determinism. Attention is then turned to critically assess Bell's appealing to the Einstein-Rosen-Podolsky incompleteness argument to support his claim. It is shown this argument is itself incomplete, via counterexamples that expose two logical gaps. However, closing these gaps via a strong "counterfactual" reality criterion enables a rigorous derivation of each of quantum incompleteness, determinism and parameter independence, and in this sense justifies Bell's claim. Consequences for quantum interpretations are briefly discussed.

"I have made an effort to present the deduction ... shorn of all superfluous mathematical technicalities and woolly interpretative commentary. (A reader as yet unfamiliar with the literature will be astounded to see the incredible metaphysical extravaganzas to which this subject has led.)" — van Fraassen [1]

# 1 Introduction

Bell's theorem, that there are no local hidden variable models that can reproduce all quantum correlations [2, 3], is one of the most surprising results of twentieth century physics. It has important ramifications for both physics and metaphysics, not only underlying the security of a number of device-independent communication protocols such as quantum key distribution and random number generation, but also ruling out naive classical interpretations of quantum probability [3].

The mathematics required to prove Bell's theorem is remarkably straightforward. However, there is one aspect of the assumptions used in Bell's first exposition that has generated differing opinions and vigorous debate [1, 4]–[27]: do the assumptions of locality and perfect correlation made in his 1964 paper [2] necessarily imply that certain measurement outcomes are predetermined, or must such determinism be postulated?

It may be noted that the answer to this question is unimportant in one sense: later generalisations of Bell's theorem do not rely on assuming perfect correlations nor determinism [3]. Nevertheless, the question remains of strong interest for several reasons. First, there is no general consensus on the answer: while early commenters such as Wigner wrote that Bell postulated both deterministic hidden variables and locality [4] (see also [5, 6]), Bell himself later claimed that determinism was in fact inferred rather than assumed in the 1964 paper [7], and the debate has only intensified since then [1, 8]–[27]. Second, the validity of Bell's claim would suggest that there is no choice between giving up determinism or giving up locality in interpreting quantum phenomena: in the presence of perfect correlations it would be compulsory to give up locality [9, 25, 28]—a conclusion that adds further fuel to the fire and is strongly contested in its own right [29]–[36]. Finally, noting that Bell appeals to the famous Einstein-Podolsky-Rosen (EPR) argument for the incompleteness of quantum mechanics [37] to support his claim, the debate puts the latter argument itself into question.

With the aim of making a clear and concise contribution to these issues, this paper is guided by the quote given at the beginning (taken from an excellent early discussion on the subject by van Fraassen [1]), which continues to resonate decades later. In particular, attention will be focused on what can be proved in a simple yet rigorous and general manner, and what can be disproved via simple counterexamples, whilst avoiding woolly assertions and metaphysical extravaganzas!

In Section 2 it is recalled that the weakest statistical sense of locality consistent with Bell's 1964 paper, parameter independence, is too weak for deriving determinism for perfectly correlated measurement outcomes (Proposition 1). This still leaves open the possibility of an alternative route for inferring determinism, based on Bell's appeal to the EPR incompleteness argument to support his claim. However, two gaps in the EPR logic are identified via simple counterexamples in Section 3, i.e., the argument itself is incomplete, so that determinism and quantum incompleteness do not strictly follow (Propositions 2 and 3). Fortunately, as shown in Section 4, a suitable strengthening of the EPR reality criterion to a "counterfactual reality" criterion, combined with a (usually left implicit) "free choice" condition, closes these gaps and leads to rigorous derivations not only of determinism and incompleteness but also, in a closing of the circle, of parameter independence (Propositions 4 and 5). Implications for quantum interpretations are briefly discussed in Section 5, and conclusions given in Section 6.

# 2 Statistical locality, perfect correlation and determinism

## 2.1 Statistical formulation of locality: parameter independence

Bell spells out his intended sense of locality in several parts of his 1964 paper [2]:

"It is the requirement of locality, or more precisely that the result of a measurement on one system be unaffected by operations on a distant system, with which it has interacted in the past ..."

"Now we make the hypothesis, and it seems one at least worth considering, that if the two measurements are made at places remote from on another, the orientation of one magnet does not influence the result obtained with the other."

"In a theory in which parameters are added to quantum mechanics to determine the results of individual measurements, without changing the statistical predictions, there must be a mechanism whereby the setting of one device can influence the reading of another instrument, however remote."

These quotes strongly suggest a deterministic element is involved (since results must preexist before they can be "affected" or "influenced"). However, Bell later claimed that this element is a direct consequence of locality and perfect correlations [7] (giving rise to the debate mentioned in Section 1), and this claim will be examined in detail in this paper. For now we need only note a purely logical point: if determinism is able to be derived rather than assumed, then it must be possible to mathematically formulate Bell's sense of locality without reference to determinism. Thus a statistical formulation of locality is desirable.

To obtain such a formulation, consistent with the above quotes, let x and y label possible measurements which may be made in two separate regions of spacetime, with respective outcomes labeled by a and b and joint probability density p(a,b|x,y) for these outcomes. If  $\lambda$  denotes any additional parameters of interest—arising, e.g., from a physical or mathematical model—then it follows from the basic rules of probability that

$$p(a,b|x,y) = \sum_{\lambda} p(a,b,\lambda|x,y) = \sum_{\lambda} p(a,b|x,y,\lambda) p(\lambda|x,y).$$
 (1)

Here summation is replaced by integration over any continuous range of  $\lambda$ . A natural definition of statistical locality for any such model can now be rigorously formulated.

**Definition 1** (Parameter independence). The probability distribution of the result of a measurement in one region is unaffected by measurement operations in a distant region, i.e.,

$$p(a|x, y, \lambda) = p(a|x, \lambda), \qquad p(b|x, y, \lambda) = p(b|y, \lambda),$$
 (2)

for all  $x, y, \lambda$ .

In particular, the first equality in Eq. (2) implies that the setting y of one measurement device cannot affect or influence the underlying statistics of the reading a of a remote device, and vice versa for the second inequality. This definition appears to be the weakest possible formulation of statistical locality that is consistent with the above quotes from Bell's 1964 paper (see also Proposition 5 below). It is, moreover, consistent with a supporting quote from Einstein given by Bell in a footnote to his 1964 paper [2]:

"But on one supposition we should, in my opinion, absolutely hold fast: the real factual situation of system S2 is independent of what is done with the system S1, which is spatially separated from the former."

The formulation in Definition 2 has also been called 'hidden locality' [1] and 'locality' [38], with the now common term 'parameter independence' being introduced by Shimony [39].

# 2.2 Determinism does not follow from parameter independence and perfect correlations

In 1981 Bell made the following claim about his 1964 paper [7]:

"My own first paper on this subject ...starts with a summary of the EPR argument from locality to deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables."

"It was only in the context of perfect correlation (or anticorrelation) that determinism could be inferred [...] (for any indeterminism would have spoiled the correlation)."

Thus, it is claimed that determinism is inferred as a logical consequence of locality and perfect correlation. The ensuing debate in the literature arises in large part from the simple observation that the form of locality in Definition 1 is inconsistent with such an inference, as may be demonstrated by simple counterexamples.

**Proposition 1.** If locality is formulated as parameter independence, then determinism does not follow from locality and perfect correlations, i.e.,

parameter independence + perfect correlations 
$$\Longrightarrow$$
 determinism. (3)

*Proof.* A first counterexample of interest is a version of the position and momentum correlations considered in the EPR incompleteness argument [37]. In particular, suppose that  $x_0$  and  $x_1$  correspond to measurements of conjugate position and momentum components of a first particle, and  $y_0$  and  $y_1$  to measurement of the corresponding components of a second particle (where the particles maybe be classical or quantum), with

$$p(a,b|x_0,y_0|\lambda) = q(a|\lambda)\delta(a+b), \qquad p(a,b|x_1,y_1) = r(a|\lambda)\delta(a-b), \tag{4}$$

$$p(a, b|x_0, y_1|\lambda) = q(a|\lambda)r(b|\lambda), \qquad p(a, b|x_1, y_0) = r(a|\lambda)q(-b|\lambda), \tag{5}$$

for position and momentum probability densities  $q(a|\lambda)$  and  $r(b|\lambda)$  conditioned on some underlying parameter  $\lambda$ . Thus the positions and momenta are perfectly correlated (the sum of the position components is zero, as is the difference of the momentum components); parameter independence conditions (2) are easily checked to hold; and yet the outcomes are not deterministic whenever q and r are not delta-functions. A second noteworthy counterexample is based on the perfect correlations of "PR-boxes" [3, 40]. In particular, suppose that there are two possible measurements  $x_0, x_1$  and  $y_0, y_1$  in each region, each having outcomes  $a, b = \pm 1$ , with  $p(a, b|x_j, y_k, \lambda) = \frac{1}{2}\delta_{ab,(-1)^{jk}}$  for all values of some underlying parameter  $\lambda$ . Thus the outcomes are perfectly correlated for all choices of settings, with  $b = (-1)^{jk}a$ , and it is easily checked that parameter independence holds, with  $p(a|x_j, y_k, \lambda) = \frac{1}{2} = p(b|x_j, y_k, \lambda)$  in Eq (2), despite the outcomes not being deterministic.

It is worth noting that both counterexamples in the proof of Proposition 1 (as well as the counterexample of standard quantum mechanics given by Wiseman [15]) have underlying models (with additional underlying parameters) that satisfy both parameter independence and determinism. This is easily seen for the first counterexample, which is compatible with deterministic outcomes corresponding to members of a joint phase space ensemble. More generally, local deterministic models exist for any set of correlations  $\{p(a,b|x,y)\}$ , including for the second counterexample in the proof, provided that the dependence of  $p(\lambda|x,y)$  in Eq. (1) on x and y is sufficiently nontrivial [41].

However, Bell's claim is that determinism, not just the mere possibility of determinism, can be inferred from perfect correlations and locality. Proposition 1 immediately implies that this claim fails if parameter independence is used as a proxy for locality. Hence, while it appears to be the weakest form of statistical locality consistent with his 1964 paper, it is too weak to play the role of locality in his claim.

Fortunately, Bell's claim as quoted above suggests that Bell's intended form of locality is indeed stronger, and is directly related to the EPR incompleteness argument. Hence, if the claim can be shown to rely on a suitable formulation of locality based on this argument, then it is logically valid. And this can indeed be shown — but only if the EPR argument is suitably strengthened to close two gaps. This is the subject of the next two sections.

# 3 Can the EPR argument be considered complete?

As noted above, Bell's 1981 claim, to have inferred determinism from locality and perfect correlations in his 1964 paper, is based on a reference to the EPR incompleteness argument (see the quote at the beginning of Section 2.2). And indeed, in the 1964 paper he gives a very brief summary of this argument, in the context of two perfectly correlated spins  $\sigma_1$  and  $\sigma_2$ , to conclude that [2]

Since we can predict in advance the result of measuring any chosen component of  $\sigma_2$ , by previously measuring the same [perfectly correlated] component of  $\sigma_1$ , it follows that the result of any such measurement must be predetermined."

Hence his claim is perfectly justified, providing that the EPR argument can in fact be used to derive determinism in this way. But is this the case?

In this regard that Bell appears to have (e.g., in contrast to Niels Bohr [42]) fully accepted the EPR argument, and so does not spell out any details of how determinism rigorously follows from it. He is content for example, when revisiting and motivating the example of perfect spin correlations in Section 3 of his 1981 paper [7], to simply state that "if we do not accept the intervention on one side as a causal influence on the other, we seem obliged to admit that the results on both sides are determined in advance" (my italics). Likewise, in earlier papers he similarly argues that "This strongly suggests that the outcomes of such measurements, along arbitrary directions, are actually determined in advance" [43]; or asks "Is it not more reasonable to assume that the result was somehow predetermined all along?" [44] (my italics).

Hence, also noting the discussion in Section 2 and given the long-running debate mentioned in the Introduction, there is a need for a critical assessment of the EPR argument,

to see whether it can indeed be used to give a rigorous derivation of determinism. Such an assessment is the subject of this Section. It is found that there are in fact two logical gaps in the argument. These gaps will be closed in Section 4, in part via a nontrivial strengthening of the EPR reality criterion, allowing both determinism and parameter dependence to be derived. This supports the essence of Bell's claim, as well as showing the consistency of parameter independence with the concept of locality used in Bell's 1964 paper.

#### 3.1 Logic of the EPR argument

The derivation of determinism via perfect correlation in the EPR paper begins with a simple sufficient condition for an element of reality, quoted here [37]:

**Definition 2** (EPR reality criterion). If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity.

This condition is then applied to derive determinism from perfect correlations (and in particular from perfect position and momentum correlations), using the following logic:

- 1. Assume measurement  $x_0$  is made in a first region, with result a (assumption).
- 2. The outcome of a measurement  $y_0$  in a distant second region can then be predicted as b = f(a) with certainty, for some 1:1 function f (perfect correlation).
- 3. This prediction can be obtained without disturbing the distant region in any way (assumption).
- 4. Hence, the value of b is an element of physical reality, prior to any actual measurement of  $y_0$  in the distant region (EPR reality criterion).

The above steps are all that EPR explicitly use to derive a deterministic value, i.e., a predetermined real value of b, for a given perfect correlation. However, when one tries to make the above steps mathematically rigorous, an (easily fixed) "asymmetry" gap shows up in the above logic, as shown via a counterexample in Section 3.2.1 below. Hence the EPR reality criterion is insufficient for concluding the reality of b via Steps 1–4.

Moreover, EPR then go on to consider what can be said if there are two or more perfectly correlated pairs of measurements, and add the following additional step:

5. If  $x_1$  and  $y_1$  are a second pair of perfectly correlated measurements, for the first and second regions respectively, then applying the same steps as above implies that the outcomes of both  $y_0$  and  $y_1$  are real and predetermined.

EPR rely on this last step to further conclude that quantum mechanics is incomplete [37] (since quantum mechanics allows both the positions and momenta of two particles to be perfectly correlated without assigning any deterministic values), and Bell similarly relies on it for the derivation of his 1964 inequality [2] (since this inequality requires determinism for multiple pairs of perfectly correlated measurements). However, as is shown via several counterexamples in Section 3.2.2 below, Step 5 does not strictly follow as a logical consequence of Steps 1–4: there is a significant "joint measurement" gap.

Two propositions follow immediately from these gaps and counterexamples:

**Proposition 2.** The logic of the EPR argument for the incompleteness of quantum mechanics is itself incomplete, i.e.,

EPR reality criterion 
$$\implies$$
 incompleteness of QM. (6)

**Proposition 3.** The EPR reality criterion is insufficient to derive determinism for perfect correlations, i.e.,

EPR reality criterion + perfect correlations 
$$\implies$$
 determinism. (7)

Not all is lost, however. In particular, as previewed above, it is not difficult to formulate a stronger form of the reality criterion that (together with a "free choice" condition) allows the EPR argument to be completed and determinism to be derived for perfect correlations, as will be done in Section 4.

## 3.2 Two gaps in the logic

Propositions 2 and 3 above follow from counterexamples that highlight two gaps in the EPR logic, preventing Steps 4 and 5 from going through without additional assumptions. These gaps are explicitly identified and discussed below.

#### 3.2.1 The asymmetry gap

The first gap in the EPR logic is relatively minor, and arises from the inherent asymmetry of the argument: it requires only that a measurement made in the first region does not disturb the system in the second region in any way, as per Step 3 (e.g., "at the time of measurement ... no real change can take place in the second system in consequence of anything that may be done to the first system" [37]). This does not rule out, however, the possibility that making a measurement in the second region disturbs the system in the first region. Note this is consistent with no direct interaction between the two systems if it is the measurement device that is responsible for the disturbance.

A class of counterexamples that exploits this gap is to allow (possibly superluminal) signalling from the second region to the first region, but not vice versa.

Counterexample. Suppose that each region contains a single spin- $\frac{1}{2}$  particle prepared in a totally mixed state, and that if a device measures spin in the  $\mathbf{y}$  direction of the second particle, with result  $b=\pm 1$ , it sends a signal (e.g., superluminally or along its backward lightcone) that puts the first particle into the -b eigenstate of spin in the  $\mathbf{y}$  direction. Hence, if a measurement of spin in the  $\mathbf{x}$  direction is made in the first region, the statistics of the singlet state are reproduced, i.e.,  $p(a,b|\mathbf{x},\mathbf{y})=\frac{1}{4}(1-ab\,\mathbf{x}\cdot\mathbf{y})$ .

In particular, whenever both spins are measured in the same direction, i.e., x = y, then (i) the outcome in the second region can be predicted with probability unity from the result of a measurement in the first region, and (ii) there is no disturbance of the second region by any measurement carried out in the first region, thus fulfilling the conditions for the EPR reality criterion. Nevertheless, contrary to Step 4 of the EPR logic, there is no pre-existing element of reality for the outcome of the measurement in the second region:

a totally random and unpredictable result  $b = \pm 1$  is obtained for any spin direction y. Thus, use of the EPR reality criterion to derive a predetermined value for b fails, providing a counterexample to the EPR logic, and Proposition 2 immediately follows.

The obvious way to close the asymmetry gap is simply to symmetrise the EPR logic, by further assuming that measurements in the second region do not disturb the first region in any way (which has the additional advantage that one can also obtain the reality of the outcome of  $x_0$ , via Steps 1–4 with the roles of  $x_0$  and  $y_0$  reversed). Thus both a and b are predetermined if measurements of both  $x_0$  and  $y_0$  are made. It turns out, however, that closing the asymmetry gap in this way is not sufficient for completing the EPR argument: there is a second gap!

#### 3.2.2 The joint measurement gap

To expose the second (and most important) gap in the EPR logic (see also the analysis in Section 2 of [45]), observe that it starts with the assumption in Step 1 that  $x_0$  is measured. Without this assumption the value of b cannot be predicted with certainty, as required for Steps 2 and 4. It follows that, even if the asymmetry gap is closed via symmetrisation as above, one cannot proceed directly from Step 4 to Step 5. In particular, if there is also a second pair of perfectly correlated measurements  $x_1$  and  $y_1$ , then it follows from Steps 1–4 that the outcomes of both  $y_0$  and  $y_1$  are real and predetermined only if both  $x_0$  and  $x_1$  are measured without disturbing the predicted perfect correlations. Conversely, if they cannot be jointly measured in this way then Step 5 does not logically follow—there is a "joint measurement" gap.

There are various classes of counterexamples that exploit the joint measurement gap. In his response to the EPR paper, Bohr gives a quantum counterexample [42] (note that the 4th and 5th pages of his paper describing this counterexample are printed in reverse order). In particular, one cannot jointly measure the position and momentum of a first particle as they require physically incompatible experimental arrangements, and hence one cannot jointly infer the position and momentum of a distant second particle via the EPR logic. The spin measurements considered by Bell [2] supply a similar counterexample (since a Stern-Gerlach magnet cannot simultaneously have two measurement orientations). One can also construct classical counterexamples, in which the required joint measurement of position and momentum is forbidden by an epistemic restriction [46]. Proposition 3 immediately follows from these counterexamples.

EPR allude to the joint measurement gap in the penultimate paragraph of their paper, where they reject a possible replacement of their reality criterion by a more restrictive criterion (in reference to the reality of the momentum P and position Q of a particle in the second region) [37]:

"Indeed, one would not arrive at our conclusion if one insisted that two or more physical quantities can be regarded as simultaneous elements of reality only when they can be simultaneously measured or predicted. ... This makes the reality of P and Q depend upon the process of measurement carried out on the first system, which does not disturb the second system in any way. No reasonable definition of reality could be expected to permit this."

This quote suggests that a "reasonable" way to close the joint measurement gap is by requiring that the reality of a physical quantity in a given region is prior to any process of making a measurement that does not disturb that region (Clauser and Shimony call such independence of reality from measurement "physical realism" [45]). This is a counterfactual requirement (an element of reality is assumed to exist whether or not some measurement is made from which it can be inferred), that will be seen to allow Step 5 of the EPR logic in Section 3.1 to go through if it is possible to measure either one of  $x_0$  and  $x_1$  in the first region. The latter condition is essentially a "free choice" assumption (usually left implicit in discussions of the EPR argument), that rules out "superdeterministic" counterexamples in which there is only a single predetermined choice of measurement, fixed by initial conditions in the far past.

# 4 Completing the EPR incompleteness argument

## 4.1 Closing the gaps with a stronger reality criterion and free choice

It is clear from the counterexamples discussed in Section 3.2 that the EPR logic is not sufficient to rigorously support Bell's claim or to derive the incompleteness of quantum mechanics, as per Propositions 2 and 3: stronger assumptions are needed. It is further clear from the discussion of the counterexamples that natural choices for these further assumptions are a stronger (symmetrised) non-disturbance assumption to close the first gap, and a stronger (counterfactual) reality assumption and a free choice assumption to close the second gap.

The first two assumptions are conveniently implemented by strengthening the EPR reality criterion in Definition 2 as follows.

**Definition 3** (Counterfactual reality criterion). If, without in any way disturbing or being disturbed by a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists, prior to us making the prediction, an element of physical reality corresponding to this physical quantity.

Here the bolded phrases indicate additions to the EPR reality criterion in Definition 2, directed to closing the asymmetry and joint measurement gaps respectively. This leaves the third assumption to be implemented as an explicit condition:

**Assumption** (Free choice condition). *Measurements may be freely chosen.* 

This condition ensures that two elements of physical reality, corresponding to making two different measurements, can be both be inferred (see Section 3.2.2). It is not particularly controversial (and indeed Bohr emphasised in his reply to EPR that "we are ...left with a *free choice* whether we wish to know the momentum of the particle or its initial position" [42]), although it has become of more interest in recent times as a possible loophole for Bell inequalities [41, 47].

The counterfactual reality criterion and the free choice condition above allow modified versions of Steps 4 and 5 of the EPR logic in Section 3.1 to go through, yielding positive counterparts to the no–go results in Propositions 1, 2 and 3, as shown below.

## 4.2 Deriving quantum incompleteness

The stronger criterion of counterfactual reality in Definition 3, combined with the free choice condition, restores the desired logical rigour of the EPR argument for the incompleteness of quantum mechanics. In particular, in contrast to the negative result in Proposition 2:

**Proposition 4.** The incompleteness of quantum mechanics follows from the counterfactual reality criterion and free choice of measurement, i.e.,

counterfactual reality criterion + free choice 
$$\implies$$
 incompleteness of QM. (8)

*Proof.* The proposition is shown by modifying the EPR logic in Section 3.1, for two pairs of perfectly correlated measurements  $x_0, y_0$  and  $x_1, y_1$ , as follows (with primes and italics indicating modified steps):

- 1. Assume measurement  $x_0$  is made in a first region, with result a (assumption).
- 2. The outcome of a measurement  $y_0$  in a distant second region can then be predicted as b = f(a) with certainty, for some 1:1 function f (perfect correlation).
- 3'. This prediction can be obtained without disturbing or being disturbed by the distant region in any way (assumption).
- 4'. Hence, the value of b is an element of physical reality, prior to any actual measurement of  $y_0$  in the distant region, and prior to making the prediction via an actual measurement of  $x_0$  (counterfactual reality criterion).
- 5'. If  $x_1$  and  $y_1$  are a second pair of perfectly correlated measurements, for the first and second regions respectively, and one can freely choose to make either of  $x_0$  or  $x_1$  in the first region, then applying the same steps as above implies that the outcomes of both  $y_0$  and  $y_1$  are real and predetermined prior to any actual measurement of  $x_0$  or  $x_1$  (free choice).

Noting that quantum mechanics allows, e.g., the positions and momenta of two separated particles and the spin components of two separated spins to be perfectly correlated without assigning any real and predetermined values to them, the Proposition immediately follows.

It is seen that Step 3' of the proof of Proposition 4 closes the asymmetry gap in Section 3.2.1, and Steps 4' and 5' close the joint measurement gap in Section 3.2.2.

#### 4.3 Deriving determinism and parameter independence

Note that Step 4' of the modified EPR logic in the proof of Proposition 4 is sufficient to demonstrate that the outcome of  $y_0$  is predetermined for two perfectly correlated measurements  $x_0$  and  $y_0$ . However, Step 5' is also needed below to further demonstrate predetermined values for the outcomes of multiple pairs of perfectly correlated measurements, as required for the derivation of Bell's 1964 inequality [2]. With some additional work one

can also obtain a rigorous derivation of parameter independence, i.e., of the weak form of statistical locality in Definition 2, as per Proposition 5 below. First, however, a simple lemma is required.

**Lemma.** If the conditional distribution of a random variable u is deterministic for given prior information v (i.e., p(u|v) = 0 for all but one value of u), then it remains deterministic when conditioned on any further information w compatible with v (i.e., p(u|v,w) = 0 for all but one value of u).

Proof. We have p(u|v) = 0 for all  $u \neq u_0$ , for some  $u_0$  (determinism), and p(w|v) > 0 (compatibility). Hence  $0 \leq p(u, w|v) \leq p(u|v) = 0$  for  $u \neq u_0$ , and so, using the standard rules of probability, p(u|v, w) = p(u, w|v)/p(w|v) = 0 for all  $u \neq u_0$ .

This lemma is used several times below to prove that, in contrast to the negative results in Propositions 1 and 3, we have:

**Proposition 5.** Both determinism and parameter independence, for the outcomes of any sets of possible measurements  $\{x_s\}$  and  $\{y_s\}$  in respective separated regions, follow from the counterfactual reality criterion and the free choice condition if each pair of measurements  $(x_s, y_s)$  is perfectly correlated, i.e.,

counterfactual reality criterion + free choice + perfect correlations 
$$\implies$$
 determinism + parameter independence.

Proof. Note first that, if no assumptions are made, the probability of outcome b of some measurement  $y_0$  could conceivably depend not only on  $y_0$ , but also on a distant measurement  $x_0$  and its outcome a, as well as on some set of underlying variables  $\lambda'_0$ . Hence, in general, this probability has the form  $p(b|a, x_0, y_0, \lambda'_0)$ . For the case of two perfectly correlated measurements  $x_0$  and  $y_0$ , it follows from the counterfactual reality criterion in Definition 3, as per Step 4' in the proof of Proposition 4, that the outcome of  $y_0$  is predetermined, and hence that  $p(b|a, x_0, y_0, \lambda'_0) = 0$  for all but the predetermined value of b, prior to making a measurement of  $x_0$ . Further, from Step 3' in the proof of Proposition 4, actually making a measurement of  $x_0$  to obtain an outcome a in the first region does not disturb (nor is disturbed by) the predetermined value of b in the second region, i.e.,  $x_0$  and a are redundant for determining the value of b, and hence  $p(b|a, x_0, y_0, \lambda'_0) = p(b|y_0, \lambda'_0) = 0$  for all but one value of b.

Second, suppose that some other measurement,  $x_1$  say, rather than  $x_0$ , is actually then made in the first region. Then, since we have  $p(b|y_0, \lambda'_0) = 0$  for all but one value of b, and noting  $x_1$  must be compatible with  $y_0$  and  $\lambda'_0$  if it is possible to make it (as guaranteed by the free choice condition), it follows from the Lemma that

$$p(b|x_1, y_0, \lambda_0') = p(b|y_0, \lambda_0') = 0$$
(10)

(9)

for all but one value of b. Interchanging the roles of the first and second regions in the argument thus far, and letting  $\lambda_0''$  denote any additional set of underlying variables needed to determine the outcome of  $x_0$ , it similarly follows that

$$p(a|x_0, y_1, \lambda_0'') = p(a|x_0, \lambda_0'') = 0$$
(11)

for all but one value of a, for any measurement  $y_1$  that can be made in the second region.

Third, suppose there is some set of perfectly correlated pairs of measurements,  $\{(x_s, y_s) : s \in S\}$ , for some index set S. The above two equations then generalise to, replacing subscripts 0 and 1 by r and s,

$$p(b|x_r, y_s, \lambda'_s) = p(b|y_s, \lambda'_s) = 0,$$
  $p(a|x_s, y_r, \lambda''_s) = p(a|x_r, \lambda''_s) = 0,$ 

for all but one value of a and b and all  $r, s \in S$ . Finally, letting  $\lambda$  denote the set of pairs  $\{(\lambda'_s, \lambda''_s) : s \in S\}$ , and swapping r and s in the second equation, application of the Lemma to each of the distributions gives

$$p(a|x_r, y_s, \lambda) = p(a|x_r, \lambda) = 0, \qquad p(b|x_r, y_s, \lambda) = p(b|y_s, \lambda) = 0, \tag{12}$$

for all but one value of a and b and for all  $r, s \in S$ . Thus, the outcomes of all measurement pairs  $(x_r, y_s)$  are deterministic, and satisfy parameter independence as per Eq. (2), as required.

Proposition 5 provides a rigorous justification for Bell's claim that his 1964 paper, in relying on the EPR argument, infers determinism from locality and perfect correlations — provided that the EPR argument itself is suitably strengthened via the counterfactual reality criterion and free choice condition in Sec. 4.1. The Proposition also demonstrates that parameter independence follows from the strengthened EPR argument, and hence is consistent with the notion of locality used in Bell's 1964 paper. Note that the free choice condition used in Proposition 5 is also consistent with the 1964 paper, and in particular with the measurement independence condition  $p(\lambda|x,y) = p(\lambda)$  used therein [2, 3] (note that, in the context of Eq. (1), the free choice condition corresponds to  $p(x,y|\lambda) > 0$  for all x and y, and hence is weaker than measurement independence).

# 5 Consequences for quantum interpretations

It is of interest to briefly discuss the significance of Propositions 4 and 5 for several interpretations of quantum phenomena (despite the risk of providing "woolly interpretive commentary" as per the van Fraassen quote at the beginning of this paper!).

First, any interpretation of quantum mechanics that relies exclusively on the standard Hilbert space description (such as the Copenhagen [49, 50], many worlds [51, 52], consistent histories [53], spontaneous collapse [54, 55], relational quantum mechanics [56] and QBist [57] interpretations), must necessarily, in the light of Proposition 4, reject the counterfactual reality criterion in Definition 3 if it is to maintain free choice and completeness. One way to do this is to reject the existence of an element of reality *prior* to being able to make a prediction with probability unity. For example, in the many worlds interpretation such a prediction cannot be made until a measurement establishes the relevant branch (or memory sequence) of the observer making the prediction [52], while for the Copenhagen interpretation, in the context of the EPR argument, "no well defined use of the concept of 'state' can be made as referring to the object separate from the body with which it has been in contact, until the external conditions involved in the definition of this concept are

unambiguously fixed" [50] (see also Section 3.2.2). Another approach is to assert the reality of wave function collapse, thus implying a disturbance of one region by measurement in another, as in spontaneous collapse interpretations [54, 55] (and in some variants of the Copenhagen interpretation, albeit explicitly rejected by Bohr [42]: "Of course there is in a case like that just considered no question of a mechanical disturbance of the system under investigation"—see also [58]).

In contrast, interpretations which either supplement or replace standard quantum theory with a deterministic description of outcomes, such as the deBroglie-Bohm [59, 60], many-interacting-worlds [61, 62] and superdeterministic [63, 64] interpretations, treat standard quantum mechanics as incomplete and deterministic from the outset, and hence are untroubled by Propositions 4 and 5. Such interpretations may or may not satisfy the counterfactual reality criterion and the free choice condition in Section 4.1. For example, for the case of perfect position and momentum correlations considered by EPR [37], both the deBroglie-Bohm and many-interacting-worlds interpretations satisfy the counterfactual reality criterion for position measurements (the positions of all particles are 'real'), but not for momentum measurements (such measurements on one particle necessarily disturb the other particle, directly via a nonlocal influence in the deBroglie-Bohm interpretation [60], and indirectly via interaction with other worlds in the many-interacting-worlds approach [62]). In contrast, the counterfactual reality criterion is always satisfied in superdeterministic interpretations and the free choice condition always fails, as in these interpretations not only are all measurement outcomes predetermined, but also which measurements will actually be made [63, 64].

#### 6 Conclusions

Bell's 1981 claim that determinism is inferred from locality and perfect correlations in his 1964 paper is seen to be justified, to a significant extent, by his appeal to the EPR argument for the incompleteness of quantum mechanics. In particular, while the EPR argument has some logical gaps in its original form, a strengthened argument based on additional assumptions (Proposition 4) leads to a rigorous derivation of determinism as per Proposition 5. This provides a solid foundation for a consensus position with respect to the Bell vs Bell debate referred to in the Introduction, that may help reconcile the two main camps to some degree (see [15] for a campanology analysis of the debate).

In particular, while it supports the essence of Bell's claim quoted at the beginning of Section 2.2, this foundation nevertheless relies on the use of stronger assumptions than in the original EPR argument — the counterfactual reality criterion and free choice condition in Section 4.1 — that go beyond the general concept of locality. It follows that the predicted violation of the 1964 Bell inequality by perfect quantum correlations does not imply that that the only option explain these correlations is to give up locality, as sometimes asserted [9, 25, 28] (see also Section 5).

In addition to determinism, Proposition 5 also provides a rigorous derivation of parameter independence for perfect correlations, demonstrating that the strengthened EPR argument is consistent with weakest formulation of statistical locality consonant with Bell's 1964 paper. Noting further that the combination of determinism and parameter

dependence implies that the property of local causality,  $p(a, b|x, y, \lambda) = p(a|x, \lambda)p(b|y, \lambda)$ , is satisfied, it follows that this property is also consistent with the strengthened EPR argument. This is of interest because, while local causality is equivalent to this combination for the case of perfect correlations (e.g., [1, 9, 11, 15, 25]), it may be used more generally (together with the measurement independence condition  $p(\lambda|x,y) = p(\lambda)$ ) to derive more general Bell inequalities that apply even in the absence of perfect correlations [3, 7, 48], thus sidestepping the EPR argument altogether.

Finally, in addition to the Bell vs Bell debate referred to in the Introduction, there is a related debate over whether local causality is purely a locality concept or, in contrast, incorporates elements not directly related to locality (e.g., [28, 32, 33, 34, 38, 65, 66, 67]). Since local causality is equivalent to the combination of parameter independence and the outcome independence condition  $p(a, b|x, y, \lambda) = p(a|x, y, \lambda)p(b|x, y, \lambda)$  [38, 39], where the former is generally acknowledged to be an unambiguous locality concept (see Section 2.1), this debate essentially boils down to whether the same can be said of outcome independence. While the debate is outside the scope of this paper, it is of interest to note here that outcome independence is far stronger than parameter independence with respect to the main focus of this paper, i.e., to the derivation of determinism via perfect correlations. In particular, in contrast to the no–go results in Propositions 1– 3, it may be shown that outcome independence and perfect correlations, by themselves, are sufficient to derive both determinism and the incompleteness of quantum mechanics [1, 68].

# References

- [1] B. C. van Fraassen, The Charybdis of realism: epistemological implications of Bell's inequality, Synthese **52**, 25–38 (1982)
- [2] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Physics 1, 195–200 (1964)
- [3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. 86, 419–478 (2014)
- [4] E.P. Wigner, On hidden variables and quantum mechanical probabilities, Am. J. Phys. **38**, 1005–1009 (1970)
- [5] P. Suppes and M. Zanotti, On the determinism of hidden variable theories with strict correlation and conditional statistical independence of observables, in: P. Suppes (ed.), Logic and Probability in Quantum Mechanics, pp. 445–455, D. Reidel Publishing Company, Dordrecht Holland (1976)
- [6] W. Demopoulos, Boolean representations of physical magnitudes and locality, Synthese **42**, 101–119 (1979)
- [7] J.S. Bell, Bertlmann's socks and the nature of reality, Journal de Physique Colloques 42, 41–62 (1981)
- [8] G. Grasshof, S. Portmann and A. Wüthrich, Minimal assumption derivation of a Bell-type inequality, Brit. J. Phil. Sci. **56**, 663–680 (2005)

- [9] T. Norsen, Bell locality and the nonlocal character of nature, Found. Phys. Lett. 19, 633–655 (2006)
- [10] G. Blaylock, The EPR paradox, Bell's inequality, and the question of locality Am. J. Phys. 78, 111–120 (2010)
- [11] T. Maudlin, What Bell proved: a reply to Blaylock, Am. J. Phys. 78, 121–125 (2010)
- [12] G.C. Ghirardi, On a recent proof of nonlocality without inequalities, Found. Phys. 41, 1309–1317 (2011)
- [13] R.B. Griffiths, EPR, Bell, and quantum locality, Am. J. Phys. **79**, 954–965 (2011)
- [14] A. Wüthrich, Local acausality, Found. Phys. 44, 594–609 (2014)
- [15] H.M. Wiseman, The two Bell's theorems of John Bell, J. Phys. A 47, 424001 (2014)
- [16] T. Maudlin, What Bell did, J. Phys. A 47, 424010 (2014)
- [17] R.F. Werner, Comment on 'What Bell did', J. Phys. A 47, 424011 (2014)
- [18] T. Maudlin, Reply to Comment on 'What Bell did', J. Phys. A 47, 424012 (2014)
- [19] R.F. Werner, What Maudlin replied to, Eprint arXiv:1411.2120 (2014)
- [20] T. Norsen, Are there really two different Bell's theorems?, Int. J. Quantum Found. 1, 65–84 (2015)
- [21] H.M. Wiseman and E.G. Rieffel, Reply to Norsen's paper "Are there really two different Bell's theorems?", Int. J. Quantum Found. 1, 85–99 (2015)
- [22] E.J. Gillis, On the Analysis of Bell's 1964 Paper by Wiseman, Cavalcanti, and Rieffel, Int. J. Quantum Found. 1, 199–214 (2015)
- [23] H.M. Wiseman, E.G. Rieffel, and E.C.G. Cavalcanti, Reply to Gillis's "On the Analysis of Bell's 1964 Paper by Wiseman, Cavalcanti, and Rieffel", Int. J. Quantum Found. 2, 143–154 (2016)
- [24] H.R. Brown and C.G. Timpson, Bell on Bell's theorem: the changing face of nonlocality, in: M. Bell and S. Gao (eds), Quantum Nonlocality and Reality—50 Years of Bell's Theorem, pp. 91–123, Cambridge University Press (2016)
- [25] R. Tumulka, The assumptions of Bell's proof, in: M. Bell and S. Gao (eds), Quantum Nonlocality and Reality—50 Years of Bell's Theorem, pp. 79–90, Cambridge University Press (2016)
- [26] T. Norsen, Quantum solipsism and non-locality, in: M. Bell and S. Gao (eds), Quantum Nonlocality and Reality—50 Years of Bell's Theorem, pp. 204–237, Cambridge University Press (2016)

- [27] H. M. Wiseman and E. C. Cavalcanti, Causarum investigatio and the two Bell's theorems of John Bell, in: R. Bertlmann and A. Zeilinger (eds), *Quantum [Un]Speakables II*, pp. 119–142, Springer (2017)
- [28] A. Shimony, Controllable and uncontrollable non-locality, in: S. Kamefuchi et al. (eds.), Foundations of Quantum Mechanics in the Light of New Technology, pp. 225–230, Physical Society of Japan, Tokyo Japan (1984)
- [29] D. Albert and B. Loewer, Interpreting the many-worlds interpretation, Synthese 77, 195–213 (1988)
- [30] N.D. Mermin, What do these correlations know about reality? Nonlocality and the absurd, Found. Phys. **29**, 571–587 (1999)
- [31] A. Kent, Locality and reality revisited, in: T. Placek and J. Butterfield (eds.) Quantum Locality and Modality, pp. 163-171. Kluwer, Dordrecht (2002)
- [32] N. Gisin, Non-realism: deep thought or a soft option?, Found. Phys. 42, 80–85 (2012)
- [33] M.J.W. Hall, Comment on 'Non-realism: deep thought or a soft option?' by N. Gisin, Eprint arXiv:0909.0015
- [34] M. Żukowski and C. Brukner, Quantum nonlocality—it ain't necessarily so..., J. Phys. A 47, 424009 (2014)
- [35] C.A. Fuchs, N.D. Mermin and R. Schack, An introduction to QBism with an application to the locality of quantum mechanics, Am. J. Phys. 82, 749–754 (2014)
- [36] R.B. Griffiths, Nonlocality claims are inconsistent with Hilbert space quantum mechanics, Phys. Rev. A **101**, 022117 (2020)
- [37] A. Einstein, B. Podolsky and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. 47, 777–780 (1935)
- [38] J. Jarrett, On the physical significance of the locality conditions in the Bell arguments, Noûs 18, 569–589 (1984)
- [39] A. Shimony, Events and processes in the quantum world, in: R. Penrose and C.J. Isham (ed.), Quantum Concepts in Space and Time, pp. 182–203, Clarendon Press, Oxford UK (1986)
- [40] P. Rastall, Locality, Bell's theorem and quantum mechanics, Found. Phys. **15**, 963–972 (1985)
- [41] M. J. W. Hall, The significance of measurement independence for Bell inequalities and locality, in *At the Frontier of Spacetime*, edited by T. Asselmeyer-Maluga (Springer, Cham, 2016), Chap. 11
- [42] N. Bohr, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. 48, 696–702 (1935)

- [43] J.S. Bell, Introduction to the hidden-variable question, in: J.S. Bell, Speakable and Unspeakable in Quantum Mechanics, pp. 29–39, Cambridge University Press, Cambridge UK (1987)
- [44] J.S. Bell, Einstein-Podolsky-Rosen experiments, in: J.S. Bell, *Speakable and Unspeakable in Quantum Mechanics*, pp. 81–92, Cambridge University Press, Cambridge UK (1987)
- [45] J.F. Clauser and A. Shimony, Bell's theorem: Experimental tests and implications, Rep. Prog. Phys. 41 1881–1927 (1978)
- [46] S.D. Bartlett, T. Rudolph, R.W. Spekkens, Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction, Phys. Rev. A 86, 012103 (2012)
- [47] S. Hossenfelder and T. Palmer, "Rethinking superdeterminism", Front. Phys. 8, 139 (2020)
- [48] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden–variable theories, Phys. Rev. Lett. 23, 880–884 (1969)
- [49] N. Bohr, Atomic Physics and Human Knowledge, pp. 32–101, New York, Wiley (1958)
- [50] N. Bohr, The causality problem in atomic physics, in: J. Kalckar (ed.), Foundations of Quantum Physics II (1933–1958), pp. 299–322, Elsevier (1996)
- [51] H. Everett III, Rev. Mod. Phys. 29, 454 (1957)
- [52] H. Everett III, The theory of the universal wavefunction, in: B.S. DeWitt and N. Graham (eds.), The Many-Worlds Interpretation of Quantum Mechanics, pp. 3–140, Princeton University Press, New Jersey USA (1973)
- [53] R.B. Griffiths, Consistent histories and the interpretation of quantum mechanics, J. Stat. Phys. **36**, 219—272 (1984)
- [54] G.C. Ghirardi, A. Rimini and T. Weber, Unified dynamics for microscopic and macroscopic systems, Phys. Rev. D **34**, 470–491 (1986)
- [55] P. Pearle, Combining stochastic dynamical state-vector reduction with spontaneous localization, Phys. Rev. A **39**, 2277–2289 (1989)
- [56] C. Rovelli, Relational quantum mechanics, Int. J. Theoret. Phys. 35, 1637—1678 (1996)
- [57] C.A. Fuchs, C.M. Caves and R. Shack, Phys. Rev. A 65, 022305 (2002)
- [58] D. Howard, Who invented the "Copenhagen Interpretation"? A study in mythology, Phil. Sci. 71, 669–682 (2004)

- [59] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables I, Phys. Rev. 85, 166–179 (1952)
- [60] D. Bohm, A suggested interpretation of the quantum theory in terms of "hidden" variables II, Phys. Rev. **85**, 180–193 (1952)
- [61] M.J.W. Hall, D.-A. Deckert and H.M. Wiseman, Quantum phenomena modeled by interactions between many classical worlds, Phys. Rev. X 4, 041013 (2014)
- [62] M. Ghadimi, M.J.W. Hall and H.M. Wiseman, Nonlocality in Bell's theorem, in Bohm's theory, and in many interacting worlds theorising, Entropy **20**, 567 (2018)
- [63] C.H. Brans, Bell's theorem does not eliminate fully causal hidden variables, Int. J. Theoret. Phys. 27, 219–226 (1988)
- [64] G. S. Ciepielewski, E. Okon and D. Sudarsky, On superdeterministic rejections of settings independence, Brit. J. Phil. Sci. **74**, 435–467 (2023)
- [65] F. M. Kronz, Hidden locality, conspiracy and superluminal signals, Phil. Sci. 57, 420–444 (1990)
- [66] J. Butterfield, "Bell's Theorem: What It Takes", Brit. J. Phil. Sci. 43, 41–83 (1992)
- [67] T. Norsen, "Local Causality and Completeness: Bell vs Jarrett", Found. Phys. 39, 273–294 (2009)
- [68] M. J. W. Hall, Eprint arXiv:2009.14223v1