A relativistic retrocausal model violating Bell’s inequality

Dustin Lazarovici

Ludwig-Maximilians University, Department of Mathematics, Theresienstrasse 39, Munich 80333, Germany

We present a simple model demonstrating that time-symmetric (advanced + retarded) relativistic interactions can account for statistical correlations violating the Bell inequalities while avoiding conspiracies as well as the commitment to instantaneous (direct space-like) influences. We provide an explicit statistical analysis of the model while highlighting some of the difficulties arising from retrocausal effects. Finally, we discuss how this account fits into the framework of Bell’s theorem.

1. Introduction

In his article ‘Speakable and unspeakable in quantum mechanics’ John S. Bell, discussing the implications of his seminal non-locality theorem [1–3] concludes

For me then this is the real problem with quantum theory: the apparently essential conflict between any sharp formulation [of quantum theory] and fundamental relativity. That is to say, we have an apparent incompatibility, at the deepest level, between the two fundamental pillars of contemporary theory…. It may be that a real synthesis of quantum and relativity theories requires not just technical developments but radical conceptual renewal. [4, cited after 5, p. 172]

From today’s perspective, there is only little to add to this assessment. For a start, one could emphasize that it is now well understood (and well established by various no-signalling theorems) that quantum non-locality does not imply the possibility of faster-than-light communication or any other way to violate the principles of relativity operationally. One could also note that in the three decades that have passed since Bell’s statement, modest yet significant progress has been made towards generalizing ‘sharp formulations’ of non-relativistic quantum mechanics to the relativistic regime.
In fact, Bell himself in a later publication [6] suggested that the GRW collapse theory may lend itself to a precise relativistic formulation, a feat that was indeed accomplished by Tumulka [7], albeit only in a non-interacting setting (but see [8] and references therein for further developments; see [9] for a critical discussion of Tumulka’s model). Also, Lorentz invariant generalizations of Bohmian mechanics can be formulated, although for the price of introducing a preferred foliation of space–time (which, however, can be generated by a Lorentz invariant law and shown to be empirically undetectable [10]). All in all, the understanding that has grown over the last few years is that there is no incompatibility but a distinct tension between relativity and quantum non-locality and that this tension is not primarily characterized by the thread of superluminal signalling but, more subtly, by the fact that relativistic space–time—having no structure of absolute simultaneity and no objective temporal order between space-like separated events—is not particularly hospitable to the kind of instantaneous influences that, according to Bell’s theorem, the explanation of certain non-local correlations seems to require. (For a comprehensive discussion of this issue, see [11,12]; see [13] for a simple argument substantiating said tension.)

In this paper, we want to discuss a possible route to alleviating the tension between relativity and non-locality by explaining non-local correlations as a result of interactions that are both retarded and advanced. Despite the technical and conceptual difficulties that come with advanced interactions, the hope associated with such models is that they could provide a complete physical account of quantum correlations while drawing exclusively on the resources of relativistic space–time with particles interacting along their past and future light cones. In fact, it can be argued on the basis of time-symmetry that advanced + retarded interactions (possibly mediated by a field or bosonic particles) are really the generic case of relativistic interactions and that it is the empirical violation of this symmetry—e.g. the absence of advanced electromagnetic radiation—that requires explanation. And it has been speculated that such an explanation could parallel, or even reduce to, the statistical explanation of the thermodynamic asymmetry, accounting for the absence of advanced effects on macroscopic scales, while the time-symmetry of the fundamental laws may still reveal itself on the microscopic level in the quantum phenomena [14]. (Indeed, such an analysis, trying to establish the empirical adequacy of a theory with advanced interactions, is not without precedence. In their ‘absorber theory’, Wheeler & Feynman [15,16] outline a statistical account of the radiative asymmetry based on a time-symmetric formulation of classical electrodynamics; see [17] for a recent clarifying discussion.)

The idea of accounting for quantum non-locality by admitting some form of retrocausation is not new but has been advanced by various authors for many decades (e.g. [18–26]). Nevertheless, it is rarely acknowledged as a possible implication of Bell’s theorem and even rarer to pass the threshold from a logical possibility to a serious option. On the one hand, this is quite understandable, considering how drastically the proposal contradicts our ordinary sense of time and causation. On the other hand, it would not be the first time that new physics require a radical revision of our respective (pre)conceptions. Hence, the best way to challenge the status quo is to develop and discuss further physical models that can serve as an intuition pump and demonstrate that retrocausation, in the appropriate context, must be neither as incomprehensible nor as unbecoming as generally assumed. This is precisely what this paper aims to accomplish.

By means of a simple toy model, tailored, in particular, to the EPRB experiment, we want to demonstrate in a quantitative manner that advanced relativistic interactions can in principle account for violations of the Bell inequalities without presupposing conspiratorial correlations between the experimental parameters and the variables related to the preparation of the system. While retrocausal models are in general very difficult to analyse, we will present an explicit statistical treatment, showing how and in what sense relevant predictions can be extracted.

Our model, we emphasize, is neither able nor intended to reproduce all quantum spin statistics (e.g. it will fail to do so for repeated spin-measurements in varying directions) and cannot make very ambitious claims regarding empirical accuracy, in general. Instead, it is used as a toy model for exploring the range of logical possibilities that is restricted by Bell’s theorem and for emphasizing the chances and challenges associated with retrocausal accounts of non-locality.
Finally, it should be needless to say, though we want to clarify nonetheless, that we do not engage in any futile attempt to evade the consequences of Bell’s theorem. The question today is not if the principle of local causality is violated in nature (unless, maybe, one is willing to endorse a many-worlds picture), but how it is violated—and what this means for us as we move forward.

2. The model

Our model is defined by the following assumptions:

— The particles have an internal degree of freedom (a ‘hidden variable’) represented by a vector \( S \) on the 2-sphere \( S^2 \subset \mathbb{R}^3 \). We refer to this degree of freedom as the particle’s spin.

— A spin measurement in the direction \( \mathbf{a} \in S^2 \) does not determine the exact value of \( S \), but only its orientation relative to \( \mathbf{a} \). The result of a spin measurement is thus given by

\[
\text{sgn}(\mathbf{a}, S) \in \{\pm 1\},
\]

where \( \text{sgn}(x) \) denotes the sign of \( x \).

We say that the particle has *a-spin up* if \( \langle \mathbf{a}, S \rangle > 0 \), i.e. \( \text{sgn}(\mathbf{a}, S) = +1 \) and *a-spin down* if \( \langle \mathbf{a}, S \rangle < 0 \), i.e. \( \text{sgn}(\mathbf{a}, S) = -1 \).

We can disregard the special case \( \langle \mathbf{a}, S \rangle = 0 \), as it will have probability 0.

— A spin measurement is not a purely passive process but influences the particle’s hidden spin by projecting it onto the direction in which it is being measured. That is, if the state of the particle undergoing a spin measurement in the \( \mathbf{a} \)-direction is \( S \), its state immediately after passing the measurement apparatus (Stern–Gerlach magnet) will be

\[
\text{sgn}(\mathbf{a}, S) \mathbf{a} = \frac{\langle \mathbf{a}, S \rangle}{|\langle \mathbf{a}, S \rangle|} \mathbf{a}.
\]

A spin measurement usually ends with the particles being absorbed in a screen or a detector shortly after passing the magnet.

— We consider an ensemble of pairs of particles whose initial spin variables are prepared with opposite orientation in a random direction, which is equi-distributed on the unit sphere \( S^2 \). For any pair, we denote

\[
S^A(t = 0) = -S^B(t = 0) = S_0.
\]

— The spin state \( S \) is subject to a *pair interaction* whose effect is such that a particle continuously rotates the spin of its partner towards the orientation antipodal to its own. This effect is manifested by an advanced and retarded action of one particle on the other. This could mean, for example, that the interaction is transmitted by a medium—like a field or a massless particle—propagating with the speed of light towards past and future, or that the particles interact directly along their past and future light cones.

To acknowledge the fact that EPR correlations persist over very long distances [27]—indeed, if quantum mechanics is correct, over arbitrary long distances—we will have to assume that the interactions in a particle pair are unattenuated, i.e. unaffected by distance, and discriminating, i.e. unfazed and unscreened by other matter, thus realizing two essential properties of what Maudlin describes as the ‘quantum connection’ [11, p. 22] or what, in other words, could be understood as a form of entanglement.

To realize this model, we could for instance consider a pair interaction of the following type: For two unit vectors \( \mathbf{X}, \mathbf{Y} \in S^2 \subset \mathbb{R}^3 \), their distance vector is given by

\[
D(\mathbf{X}, \mathbf{Y}) = \frac{\arccos(\mathbf{Y}, \mathbf{X})}{\sqrt{1 - \langle \mathbf{Y}, \mathbf{X} \rangle^2}}(\mathbf{Y} - \langle \mathbf{Y}, \mathbf{X} \rangle \mathbf{X}) \in T_X S^2,
\]

on September 16, 2017
where $T_X S^2$ denotes the tangent bundle of the sphere at point $X$. Then, assuming the particles have world-lines $z_1^\mu(t)$, $z_2^\mu(t)$, we can set
\[
\frac{d}{dt} S_i(t) \propto D(S_i(t), -S_j(\tau_{\text{ret}})) + D(S_i(t), -S_j(\tau_{\text{adv}})),
\] (2.5)
where $\tau_{\text{ret}}, \tau_{\text{adv}}$ are the advanced and retarded time, i.e. the solutions of
\[
(z_i^\mu(t) - z_j^\mu(\tau))(z_{i,\mu}(t) - z_{j,\mu}(\tau)) = 0.
\] (2.6)

Hence, the spin of particle $i$ at time $t$ is ‘repelled’ by the spin of particle $j$ at the advanced and retarded times. This is an example of a direct particle interaction along past and future light cones (figure 1). (cf. the above-mentioned Wheeler–Feynman theory of classical electrodynamics [15,16].)

(a) Remarks

We have to note one issue with respect to this particular (most simple) example of advanced + retarded ‘spin’-interactions. Being defined by a first-order differential equation, the law is actually not time-symmetric in the sense of time-reversal invariant, thus undermining our best a priori argument for admitting advanced interactions in the first place. In principle, this flaw can be resolved, for example, by considering the analogous second-order equation, while adding appropriate constraints on the initial conditions, i.e. that $\dot{S}_1(0) = \dot{S}_2(0) = 0$ (at least on average). The analysis of this interaction would be more tedious, though, and our arguments will be significantly simplified if we have in mind the first-order law of motion as defined by equation (2.4).

Nevertheless, if the reader is pursued by the following discussion, she should appreciate the limited degree to which the main message depends on the details of the model. In fact, there is a variety of conceivable interactions that could realize our assumptions in the relevant (qualitative) respects, the most interesting of which might involve something akin to a quantum wavefunction. Indeed (lacking a different revolutionary idea) such an object would seem necessary to make manifest the structure of entanglement that is assumed in our discussion and would arguably play a central role in explaining how this ‘quantum connection’ (and thus the possibility of further retrocausal influences) brakes down in the course of an irreversible, macroscopic ‘measurement’. However, at this point, these are all speculations (well-founded ones,
but speculations nonetheless) and expanding on them would go beyond the scope of the present discussion and, in fact, beyond the scope of our simple toy model per se.

3. A heuristic analysis of the time-evolution

Figure 2 shows a sketch of the space–time diagram of a typical EPRB experiment. A pair of particles prepared with opposite (hidden) spin move in different directions towards an apparatus, where they undergo a spin measurement in directions \(a\) and \(b\), respectively. These parameters can be freely chosen by the experimentalists right before measurement.

In our model, the measurement outcomes are determined by the particles’ spin variables as they pass the device. We denote the corresponding values by \(S_A\) and \(S_B\) and the resulting measurement outcomes by \(A := \text{sgn}(a, S_A)\) and \(B := \text{sgn}(b, S_B)\), respectively. In the end, we are interested in the probabilities of the coincidences \(A = B\), respectively, the anti-coincidences \(A \neq B\). We recall that the Bell inequality (in its simplest version, assuming ‘perfect’ anti-correlations for spin measurements in the same direction as predicted by QM as well as by our model) reads

\[
\mathbb{P}(A \neq B | a, b) + \mathbb{P}(A \neq B | b, c) + \mathbb{P}(A \neq B | a, c) \geq 1,
\]

for arbitrary parameter-settings \(a, b, c\) [28]. This inequality is violated for the correlations observed in the EPRB experiment (for certain choices of \(a, b, c\)) thus implying, according to Bell’s theorem, that they cannot be reproduced by any local theory.

Let’s consider, for starters, the predictions of our hidden variable model without advanced interactions. This will serve as our point of reference as a completely local model. Without advanced interactions, the initial configuration is a stationary point of the (retarded) dynamics and we have \(S_A = S_0\) and \(S_B = -S_0\). Hence, the outcomes of the spin measurement depend only on the orientation of the initial hidden spin \(S_0\) relative to those of the measurement devices. Concretely, we find

\[
A = B \iff \text{sgn}(a, S_0) = \text{sgn}(b, -S_0)
\]

and

\[
A \neq B \iff \text{sgn}(a, S_0) = \text{sgn}(b, +S_0).
\]

Of course, we now know from Bell that there is no probability distribution for \(S_0\) for which such a model will reproduce statistical correlations violating the above inequality. And indeed, a short calculation shows that for the assumed equi-distribution of \(S_0\) on \(S^2\) and arbitrary (coplanar) angles with \(a + b + c = 360^\circ\):

\[
\mathbb{P}(A \neq B | a, b) + \mathbb{P}(A \neq B | b, c) + \mathbb{P}(A \neq B | a, c) = 1,
\]

and in general

\[
\mathbb{P}(A \neq B | a, b) + \mathbb{P}(A \neq B | b, c) + \mathbb{P}(A \neq B | a, c) \geq 1,
\]

so that Bell’s inequality is always satisfied. (More precisely, one can see by a few geometric considerations that \(P(A \neq B | a, b) = 1 - 2(\angle a, b/360^\circ)\), where \(\angle a, b \in [0, +180^\circ]\) is the acute angle between the vectors \(a\) and \(b\).)

(a) ‘Zig-Zag causality’ and advanced effects

Let’s now turn to the more interesting case and consider our model with retarded and advanced interactions. The key difference, of course, is that now the particles post measurement, whose states were affected by the external intervention of the measurement process, can have a retrocausal effect on the particles before measurement. To extract statistical predictions for this case, we will have to understand how the final spin states \(S_A\) and \(S_B\), determining the experimental outcomes, result from the advanced-r-retarded interactions between the particles. This will be essential a combination of three effects, due to what is sometimes described as zig-zag causality:
Figure 2. Space–time diagram of the EPRB experiment. For the analysis of the interactions, we distinguish four increments of the particles world-lines in different slices of space–time.

1) *Feed-forward.* Particle A, after undergoing measurement in space–time region $\mathcal{A}$, exhibits an advanced action on its partner, reaching particle B in space–time region 3. The change in the spin variable that particle B experiences as a result is thus directed towards $-a$ if the measurement on A yields a-spin up and $+a$ if the measurement on A yields a-spin down. The same holds vice versa, with particle B in $\mathcal{B}$ exhibiting an advanced action on an earlier version of particle A, rotating its spin towards the direction opposite to B’s state after measurement. We call this effect a feed-forward since, intuitively spoken, every particle receives a ‘feedback’ of its partner’s future measurement result.

2) *Preinforcement.* After being affected by the feed-forward, the advanced back-reaction of particle B in space–time region 4 on particle A in region 2 will tilt the spin of particle A slightly towards the direction in which it’s going to be measured in the future. The advanced action of particle A in space–time region 4 has an analogous effect on particle B in space–time region 2. For further reference, we call this effect preinforcement—so to speak, a pre-emptive reinforcement of the future measurement result.

Of course, in a similar way, particle A in space–time region 2 has an advanced interaction with an earlier version of particle B, and so on... But further down the worldlines, the advanced effects originating from the particles after measurement are more and more diluted and will affect the state of the early particles only marginally.

3) *Inertia.* As the early spins are largely unaffected by the retrocausal influences, the effect of the retarded action exhibited by the particles in space–time regions 1 and 2 is essentially to rotate the spins back towards their initial values $S_0$ and $-S_0$, respectively.

An analytic treatment of such a time-evolution, even in our rather primitive example, is notoriously difficult, requiring the solution of two coupled delay differential equations. Nevertheless, we can make the following important observation: with the boundary conditions set by the initial configuration and the final measurement results, the spin of particle A (let’s say) will experience an attraction towards the vector $-Bb$ as a result of feed-forward, towards $Aa$ due to the effect of preinforcement and back towards its initial state $S_0$ due to what we called inertia.
More precisely, in the course of the advanced-retarded spin interactions described by (2.4), the vector $S_1(t)$ will vary within the convex set spanned by the points $\{S_0, -Bb, Aa\}$, while the spin $S_2(t)$ of particle B will vary within the antipodal triangle, spanned by $\{-S_0, Bb, -Aa\} \subset S^2$. That is because for the first-order equation, these configurations form an invariant set, as the ‘velocities’ $\dot{S}_i$ always point towards the interior of the respective triangles.

On this basis, we can make the following ansatz for the final spin state—which will be exact for the first-order law, while for more general interactions we can take it as an approximation valid for short interaction times. For the spin configuration of particle A by the time it enters the measurement device, we write

$$S_A = \frac{\alpha S_0 - \beta Bb + \gamma Aa}{\|\alpha S_0 - \beta Bb + \gamma Aa\|}, \quad (3.2)$$

with positive parameters $\alpha, \beta, \gamma$ with $\alpha + \beta + \gamma = 1$. Correspondingly, for the second particle, the final spin state $S_B$ will be of the form

$$S_B = \frac{-\alpha' S_0 - \beta' Aa + \gamma' Bb}{\|\alpha' S_0 - \beta' Aa + \gamma' Bb\|}, \quad (3.3)$$

where for simplicity we shall set $\alpha' = \alpha$, $\beta' = \beta$ and $\gamma' = \gamma$. The precise values of the parameters $\alpha, \beta, \gamma$ in every run of the experiment will of course depend on the details of the interaction and the experimental setting. The parameters $\beta$ and $\gamma$ can thereby be thought of as reflecting the effect of feed-forward and preinforcement, respectively. As feed-forward will, in general, act for a short period of time before the particles are absorbed in a detector and as preinforcement, in turn, is only the ‘echo’ of feed-forwards, we assume $\alpha > \beta > \gamma$.

Finally, for the results of the spin measurements, we have to consider the projections of $S_A$ and $S_B$ onto $a$ and $b$, respectively. This yields

$$A = \text{sgn}(a, S_A) = \text{sgn}(\alpha (a, S_0) - B\beta (a, b) + A\gamma)$$

and

$$B = \text{sgn}(b, S_B) = \text{sgn}(-\alpha (b, S_0) - A\beta (a, b) + B\gamma), \quad (3.4)$$

where we can neglect the normalization constants.

4. Statistical analysis

(a) The underdetermination problem

At this point, we encounter a difficulty that is characteristic for theories admitting retrocausal influences: the underdetermination of the time-evolution by Cauchy data. This is to say that the specification of initial conditions at a single moment in time (respectively, on a space-like hypersurface) is not sufficient to distinguish a unique solution of the equations of motion and thus to determine the system’s history unambiguously. In our model, this problem is manifested in the fact that for certain values of $S_0$ (when $(a, S_0)$ or $(b, S_0)$ is small compared with the advanced effects), two or more different prescriptions for $A$ and $B$ correspond to consistent solutions of equation (3.4) and thus to possible solutions of the time-evolution. Moreover, in these cases, there is a sense in which one could say that the measurement outcomes are retro-causally responsible for their own occurrence. For further reference, we call this phenomenon a self-fulfilling prophecy (SFP).

We emphasize that this underdetermination of the time-evolution by initial data is not a result of our analysis being too coarse, but an intrinsic feature of theories admitting retrocausal influences. (For a mathematical discussion of this issue in the context of Wheeler–Feynman electromagnetism, e.g. [15,29].) We also note that while the possibility of SFPs may be mind boggling, it need not imply the possibility of logical paradoxes—there is nothing inconsistent about the time-evolutions we consider here—though the threat of potential inconsistencies is something to be addressed in the context of a more mature theory. (In our discussion, the assumption that the particles are destroyed after the first measurement is still a crucial one, for
within our limited model that does not make precise the nature and origin of ‘entanglement’ and how it might break down, we cannot dispel worries that subsequent measurements could lead to wrong—or even paradoxical—results, as we admit retrocausal influences.

Anyway, in addition to the philosophical headaches that might be caused by SFP, a very concrete difficulty we have to face here is that theories in which solutions are not parametrized by Cauchy data are in general not statistically transparent, in the sense that there is no obvious notion of a state-space on which one could implement a statistical hypothesis or define a measure of typicality. In our case, this is to say that \((A, B)\) cannot be realized as a random variable on \(S^2\) in order to describe the outcome statistics.

For this reason, we will have to resort to a more unconventional form of statistical analysis. While we cannot assign to each \((A, B)\in\{±1\}^2\) a set of initial conditions \(S_0\) sufficient to produce that outcome—thus implementing \((A,B)\) as a random variable—we will instead consider the necessary conditions in terms of \(S_0\), that is, the range of \(S_0\) corresponding to a consistent evaluation of equation (3.4), thus determining upper and lower bounds on its probability.

**b** A case-by-case analysis

To this end, we will focus on the case \((a, b) < 0\), as quantum mechanics predicts that Bell’s inequality (3.1) is always violated if the scalar product between any two of the (coplanar) vectors \(a, b, c\) is negative, with the most pronounced violation for \(<a, b = <b, c = <a, c = 120^\circ\).

Furthermore, observing that the problem is symmetric under \(S_0 \leftrightarrow -S_0\) together with an exchange of the particle labels (and assuming that SFP respects this symmetry), we can w.l.o.g. assume \((a, S_0) > 0\). Now we can brake down all remaining possibilities leading to consistent solutions of equation (3.4). In the table below, we have listed for all \((A, B)\in\{±1\}^2\) the range of initial conditions \(S_0\) that can produce the corresponding outcome.

Admittedly, this may still seem quite confusing, but the dust will settle in a minute. Note that unless \((a, b)\) is very close to 0 or the ratio \(\gamma/\beta\) unreasonably large, we will always find that \(\beta|\langle a, b\rangle|\) is greater or equal \(\gamma\), meaning that we can disregard all the cases requiring \(\beta|\langle a, b\rangle| < \gamma\). Moreover, we recall that what we are ultimately interested in, are the probabilities of the coincidences \(A = B\), respectively, the anti-coincidences \(A \neq B\). To recover those, only the following cases remain to be distinguished:

1. If \(\text{sgn}(a, S_0) \neq \text{sgn}(b, S_0)\), then \(A = B\) occurs (almost surely).
2. If \(\text{sgn}(a, S_0) = \text{sgn}(b, S_0)\), then:

\[
A = B \text{ possible if } (\alpha|\langle a, S_0\rangle| < \beta|\langle a, b\rangle| + \gamma \quad \lor \quad \alpha|\langle b, S_0\rangle| < \beta|\langle a, b\rangle| + \gamma) \]

and

\[
A \neq B \text{ possible if } (\alpha|\langle a, S_0\rangle| + \gamma > \beta|\langle a, b\rangle| \quad \land \quad \alpha|\langle b, S_0\rangle| + \gamma > \beta|\langle a, b\rangle|). \tag{4.1}
\]

Note, in particular, that our model predicts perfect (anti-)correlations if the spins are measured in the exact opposite (respectively, the same) direction.

Comparing this to the local model without advanced interactions, where \(A \neq B\) occurred if and only if \(\text{sgn}(a, S_0) = \text{sgn}(b, S_0)\), we see that now, in certain cases that would have been ‘on the edge’ (i.e. where \((a, S_0)\) and \((b, S_0)\) have the same sign but either one of the terms is small in absolute value), the spins of the particles are rotated just enough by the retrocausal feed-forward to produce the coinciding event \(A = B\) instead.

**c** Minimal and maximal probabilities

Now a quantitative statistical analysis will require some information about the distribution of the variables \(\beta\) and \(\gamma\) parametrizing the final spin states \(S_A\) and \(S_B\). Recall that, realistically, these parameters would vary in every run of the experiment, depending, in particular, on \(S_0\) and the control parameters \(a\) and \(b\). Going forward, we will however make the simplest possible ansatz which is that \(\beta\) and \(\gamma\) are not only the same for both particles in each pair, but constant throughout
and where the factor of 2 accounts for the case $+ +$ that this assumption is not only simple but somewhat simplistic, it will be fairly obvious that our main results do not depend on it too strongly. We now define as follows:

The **maximal probability** $\mathbb{P}_{\text{max}}(A \neq B|a, b)$ is the probability of $A \neq B$ assuming that $A \neq B$ will occur whenever it is possible for an initial configuration $S_0$ (or, in other words, that SFP always favours $A \neq B$).

The **minimal probability** $\mathbb{P}_{\text{min}}(A \neq B|a, b)$ is the probability of $A \neq B$ assuming that it will occur only if $A = B$ is impossible for an initial configuration $S_0$ (or, in other words, that SFP always favours $A = B$).

From equation (4.1), we can conclude

$$\mathbb{P}_{\text{max}}(A \neq B|a, b) = 2\mathbb{P}\left(\frac{\langle a, S_0 \rangle > 0}{\langle a, b, S_0 \rangle > 0}, \frac{\langle b, S_0 \rangle > 0}{\langle b, S_0 \rangle < 0}\right)$$

and

$$\mathbb{P}_{\text{min}}(A \neq B|a, b) = 2\mathbb{P}\left(\frac{\langle a, S_0 \rangle + \gamma > \langle a, b \rangle}{\langle a, b, S_0 \rangle + \gamma > \langle a, b \rangle}, \frac{\langle b, S_0 \rangle + \gamma > \langle b, a \rangle}{\langle b, S_0 \rangle + \gamma > \langle b, a \rangle}\right),$$

where the factor of 2 accounts for the case $(a, S_0), (b, S_0) < 0$. This can be evaluated by means of the following identity that we derive in the appendix:

$$|S^2|^{-1}|\int S \in S^2 : \langle a, S \rangle > C \wedge \langle b, S \rangle > C| = \frac{1}{\pi} \int_C \left[ \frac{\sqrt{1/(1+C)} + \sqrt{1/(1+b)}}{2} \right] \frac{\sqrt{2 - C^2}}{\sqrt{2 - z^2}} dz. \quad (4.2)$$

For better illustration of the results, we will simplify things a bit further, still, by estimating the ratio of $\beta$ to $\gamma$, thus obtaining a parametrization of the final spin states $S_A$ and $S_B$ in terms of a single affine parameter $\nu \in [0, 1]$. This is to say that we differentiate the analysis only by the strength of the advanced effects—represented by the parameter $\nu$—rather than their dependence on distance or the duration of action which could be dissected in a more fine-grained analysis by fitting the parameters $\beta$ and $\gamma$ independently (cf. figure 4 below). Setting $\beta = \nu$, a reasonable estimate is $\gamma = \nu^2$, as preinforcement is a result of two advanced interactions, and hence $\nu = 1 - (\beta + \gamma) = 1 - \nu - \nu^2$.

### 5. Results

In figure 3, we have plotted the minimum and maximum probabilities for varying values of $\nu$ and parameter settings $\langle a, b \rangle = 120^\circ$, i.e. $\langle a, b \rangle = \frac{1}{2}$, which is the case for which quantum mechanics predicts the greatest violation of the Bell inequality (3.1). The upper curve represents the highest
possible probability, the lower curve the lowest possible probability of the anti-coincidence $A \neq B$ depending on the value of $\nu$ and we see that both are decreasing with $\nu$, i.e. as retrocausal effects get more pronounced. The shaded area in between thus corresponds to the range of possible probabilities that could result from the present model, depending on how SFP’s are fixed.

We see that for $\nu \equiv 0$, i.e. in the absence of advanced interactions, the account reduces to the local model discussed above, that is, $P_{\min}$ and $P_{\max}$ coincide (since $(A, B)$ now is a random variable on $S^2$) yielding, in particular, $P(A \neq B | a = 120^\circ) = \frac{1}{3}$ and thus for $a = 0^\circ, b = 120^\circ, c = 240^\circ$:

$$P(A \neq B | a, b) + P(A \neq B | b, c) + P(A \neq B | a, c) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1.$$ 

Hence, the Bell inequality is (of course) satisfied, though it is important to note that this case is already critical, i.e. that equality holds in equation (3.1). For now, as we consider values of $\nu$ greater 0, we observe that the statistical effect of the advanced interactions is to lower the probability of the anti-coincidence $A \neq B$, thus leading to a violation of the Bell inequality. Notabene, the fact that the maximal probability $P_{\max}$ already is smaller than $\frac{1}{3}$ shows that the violation of the Bell inequality is not because of self-fulfilling prophecies, i.e. is not achieved by exploiting the underdetermination of the outcomes by the initial state, but would have to occur in any case, regardless of how this underdetermination is resolved and what the resulting statistical description would look like in detail.

On a more quantitative note, we have to keep in mind that we have made a series of simplifications and assumptions along the way, so it may or may not be significant that we find the range of possible probabilities to be fairly close to the quantum mechanical prediction of 0.25 for reasonable values of $\nu$.

To settle on more precise predictions, we could consider the median of $P_{\min}(A \neq B)$ and $P_{\max}(A \neq B)$, corresponding to the probability of the anti-coincidence assuming that SFP is not biased between $A = B$ and $A \neq B$. These values are plotted in figure 4—now, for greater generality, with the parameters $\beta$ and $\gamma$ varying independently—to be compared with the predictions of the local model ($P = \frac{1}{2}$) and of quantum mechanics ($P = \frac{1}{4}$). One of the most interesting things to note about this plot is that it does not look particularly interesting—the values are rather steady except for the upward slope as $\gamma$ and $\beta$ go to 0 and the downward slope as they become very large—indicating a remarkable robustness of the predictions of the retrocausal model against the details of the interactions (figure 5).

Finally, we note that there is no fixed choice of $\beta$ and $\gamma$ for which our statistical evaluation reproduces the quantum statistics (of the spin-singlet state) for all values of $a$ and $b$, which would be

$$P(A \neq B | a, b) = \frac{1}{2} + \frac{1}{2} \langle a, b \rangle.$$ 

Figure 4. Possible values for $P(A \neq B | a = 0^\circ, b = 120^\circ)$. (Online version in colour.)
Figure 5. Mean probability for different values of $\beta$ and $\gamma$. (Online version in colour.)

Of course, it would have been quite miraculous if it did, given the overall crudeness of our statistical hypothesis. An interesting question might thus be whether by taking into consideration the possible dependence of $\beta$ and $\gamma$ on the relevant physical variables, it is possible to find a probability distribution $p(\beta, \gamma|S_0, a, b)$ for which this model reproduces equation (5.1) exactly.

In conclusion, what our analysis shows is that the quantum non-locality we observe in nature in form of statistical correlations violating Bell’s inequality could really be understood as the signature of retrocausal effects due to time-symmetric relativistic interactions, rather than instantaneous (or superluminal) influences between space-like separated events.

6. Retro-causality and Bell’s theorem

Having seen that time-symmetric (advanced + retarded) relativistic interactions can in principle account for violations of the Bell inequality, it is instructive to reflect on how exactly such a model fits into the framework of Bell’s theorem. We recall that the most general derivation of a Bell inequality, more specifically the CHSH inequality \[2,30\], is based on two assumptions:

(i) The locality assumption. The statistical correlations are _locally explainable_. This means that conditioning on all the physical data in the (causal) past of A and B that, according to any candidate theory, could be relevant to the prediction of A and B will _screen off_ the correlations, so that the specification of A and a becomes redundant for the prediction of the probability of B, and vice versa. Formally, comprising all possible ‘common causes’ of A and B in a set of variables $\lambda$, this locality condition reads

$$
\mathbb{P}(A|B, a, b, \lambda) = \mathbb{P}(A|a, \lambda)
$$

and

$$
\mathbb{P}(B|A, a, b, \lambda) = \mathbb{P}(B|b, \lambda).
$$

(ii) The no-conspiracy assumption. The explanation of the correlations must not be _conspiratorial_, meaning that the experimental parameters $a$ and $b$ can be chosen _freely_ or _randomly_, independent of each other and of any other physical process that might be relevant to the system before measurement, hence independent of $\lambda$. Formally,

$$
\mathbb{P}(\lambda|a, b) = \mathbb{P}(\lambda).
$$

As our model violates the Bell (as well as the CHSH) inequality, it must violate at least one of these assumptions. In many discussions, a ‘retrocausal explanation’ of EPR correlations is understood basically as synonymous with a ‘conspiracy’. However, the description provided by our model seems reasonable enough to show that the issue deserves a second look, and indeed, on that second look, things turn out to be a bit more subtle:
If we condition the measurement outcomes on the relevant physical configurations in the past of \( A \) and \( B \)—that is, on the particles’ initial spin variables \( \pm S_0 \)—we find that the locality condition (6.1) (in form of parameter independence) is violated, i.e.

\[
\mathbb{P}(A|a, b, S_0) \neq \mathbb{P}(A|a, S_0) \quad \text{and} \quad \mathbb{P}(B|a, b, S_0) \neq \mathbb{P}(B|b, S_0).
\]

This initial data, in other words, are not sufficient to screen off the correlations between the measurement outcomes on one side of the experiment and the parameter settings on the other. Hence, a physicist seeking to explain the correlations between the distant events, as we usually do, by looking for a common cause in their past, is bound to fail and may reasonably conclude that there must be some sort of direct instantaneous influence between the two sides of the experiment. However, if we lived in a world guided, on the microscopic level, by time-symmetric relativistic laws, this physicist, in doing so, would literally miss half the story since the physical laws would be such that the outcomes of the spin measurements were actually determined by the physical configurations in both past and future of \( A \) and \( B \).

On the other hand, if we condition \( (A, B) \) on the initial states \( \pm S_0 \) and the states after measurement—which, in a retrocausal model like ours, can actually be regarded as ‘causing’ the outcomes—the probability, trivially, factorizes and the locality condition is, formally, satisfied.

Note that no assumption about the localization of \( \lambda \) actually enters the derivation of the CHSH inequality, though it is only in the case that \( \lambda \) refers to configurations or events in the causal past of \( A \) and \( B \) that we would speak of a local explanation in Bell’s sense. This is to say, in particular, that the concept of locality or local causality presupposes a distinguished direction of time. From a time-symmetric perspective, a more natural desideratum (at least for theories that are not intrinsically stochastic) would simply be the absence of direct influences between space-like separated events. And while our model is explicitly non-local in the conventional sense, it does satisfy the latter requirement which makes it arguably as relativistic as an explicitly non-local theory can get.

In any case, from this point of view—admitting relevant \( \lambda \)’s in the past and in the future—our model violates the Bell inequalities, as it cannot be otherwise, by violating the no-conspiracy condition. Obviously, the physical variables screening off the correlations are not independent of the parameter settings, as they include the particle states post measurement which are collinear with the chosen orientations \( a \) and \( b \). However, we see no reason to deem such an account ‘conspiratorial’—at least not in the devastating sense argued by Bell [2,3] and Shimony et al. [31] to essentially render futile the scientific enterprise. The fact that the state of the system after measurement reflects our experimental choices is hardly mysterious and in no way different from what we have anyway come to expect. More importantly, the advanced effects of the microscopic interactions do in no way infringe upon our freedom to choose the parameters of the experiment as we please, or make the choice completely random, nor on the possibility to prepare a system (its initial state, that is) according to our liking and practical abilities. In other words, while formally violating the no-conspiracy condition, the account does not presuppose any dependence between the parameters associated with the preparation of the system and the parameters associated with the set up of the measuring apparatus. Hence, it does not involve a conspiracy.

Acknowledgements. I am grateful to Christian Beck as well as anonymous referees for insightful comments on earlier versions of this manuscript.

Appendix A. Derivation of equation (4.2)

We want to compute the measure of the set \( \{S \in S^2 \mid \langle a, S \rangle > C, \langle b, S \rangle > C\} \), for \( C > 0 \) and \( \langle a, b \rangle < 0 \). In spherical coordinates, the variable \( S \in S^2 \subset \mathbb{R}^3 \) is parametrized as

\[
S = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad \theta \in [0, \pi), \phi \in [0, 2\pi).
\]
W.l.o.g. we can locate \(a\) and \(b\) in the \(x-y\)-plane and set
\[
a = (0, 1, 0) \quad \text{and} \quad b = (\sin \chi, \cos \chi, 0),
\]
where \(\chi\) is the angle between \(a\) and \(b\) and hence \((a, b) = \cos \chi\). We thus have
\[
(a, S) = \sin \theta \sin \phi > C
\]
and
\[
(b, S) = \sin \theta (\sin \chi \cos \phi + \cos \chi \sin \phi) > C.
\]
As the set is symmetric under interchange of \(a\) and \(b\), it suffices to consider the case \((b, S) > (a, S)\) (and then double the measure).

As the set is symmetric under reflection on the \(x-y\)-plane, it suffices to consider the case \(\theta < \pi/2\), i.e. \(S_z > 0\) (and then double the measure).

Then we have, for once,
\[
0 < (a, S) < (b, S) \iff 0 < \sin \phi < \sin \chi \cos \phi + \cos \chi \sin \phi
\]
which yields \(\cos \phi > 0\) and, after a little bit of algebra,
\[
0 < \sin \phi < \sqrt{\frac{1}{2}(1 + \cos \chi)} = \sqrt{\frac{1}{2}(1 + (a, b))} =: D.
\]
And we compute for \(M := \{S \in S^2 | (b, S) > (a, S) > C, S_z > 0\} \subset S^2\):
\[
|M| = \int_{\pi/2}^{2\pi} \int_0^{2\pi} 1_{M(\theta, \phi)} \sin \theta \sin \phi \, d\theta \, d\phi = \int_0^{2\pi} \int_0^1 1 \left\{ \sin \theta > \frac{C}{\sin \phi}, \sin \phi < D, \phi \in \left(0, \frac{\pi}{2}\right) \right\} \, d\cos \phi \, d\phi
\]
\[
= \int_0^{\pi/2} 1 \{C < \sin \phi < D\} \sqrt{1 - \frac{C^2}{\sin^2 \phi}} \, d\phi = \int_C^D \sqrt{\frac{z^2 - C^2}{z^2 - z^4}} \, dz,
\]
where in the last step, we substituted \(z := \sin \phi\). Together with \(|S^2| = 4\pi\), equation (4.2) follows.

References