Bell Inequalities with Postselection*

Asher Peres

Department of Physics, Technion – Israel Institute of Technology, 32 000 Haifa, Israel

Abstract

Experimental tests of Bell inequalities ought to take into account all detection events. If the latter are postselected, and only some of these events are included in the statistical analysis, a Bell inequality may be violated, even by purely classical correlations. The paradoxical properties of Werner states, recently pointed out by Popescu, can be explained as the result of a postselection of the detection events, or, equivalently, as due to the preparation of a new state by means of a nonlocal process.

* Dedicated to Professor Abner Shimony, on the occasion of his 70th birthday.
Quantum mechanics is a statistical theory. It does not describe physical processes that are happening in nature, but merely predicts probabilities of events. Namely, if a physical system is prepared in a definite way (mathematically represented by a Hermitian matrix $\rho$), and that system is then subjected to a definite test (represented by a projection operator $P$), the probability of passing that test is equal to $\text{Tr} (\rho P)$. A natural question is whether there is more to say than that. Can there be a more elaborate theory, requiring a more detailed description of preparations and tests, such that the outcomes of tests would be definite, rather than probabilistic?

In 1964, John Bell proved that quantum mechanics is incompatible with the existence of such a theory, if the latter has to obey the principle of local causes. This principle (also called Einstein locality, but conjectured well before Einstein) asserts that events occurring in a given spacetime region are independent of external parameters that may be controlled, at the same moment, by agents located in distant spacetime regions. Bell’s theorem [1] states that, as a consequence of the principle of local causes, the outcomes of tests performed on spatially distant systems cannot have arbitrarily large correlations: the latter must satisfy a certain inequality. Bell also showed that this inequality does not hold for spin correlations of a pair of spin-$\frac{1}{2}$ particles in a singlet state. That is, quantum mechanics is incompatible with the principle of local causes.

At first, Bell’s momentous discovery attracted only scant attention [2], perhaps because in its original form, Bell’s inequality had a restricted domain of validity and could not be directly tested. However, in 1969, Clauser, Horne, Shimony, and Holt [3] derived a more useful inequality, valid under more general assumptions, and amenable to experimental tests (it is called the CHSH inequality). Actual tests soon followed. The most remarkable were those by Aspect and his collaborators [4, 5], involving pairs of correlated photons originating from $SPS$ atomic cascades.

Ironically, although the experiments fully agreed with the predictions of quantum mechanics, doubts have been expressed whether a violation of the CHSH inequality had actually been observed [6]. While no one denies that the quantum mechanical formalism permits the existence of states that violate the inequality, the interpretation of experimental results is problematic: not all the particle detections are taken into account in the statistical analysis. Some undesirable data are rejected. For example, if only one of the two distant detectors is excited, the unpaired detection event is ignored (this must often happen indeed, since it is only occasionally that the two photons have nearly opposite directions).

This selection of “good” data and rejection of “bad” ones is suspicious. The use of biased statistical protocols is notorious for producing outright fraud. In another quantum context, “postselected” measurement data (namely, data sifted according to a procedure
carried out after completion of the measuring act) can sometimes yield average values which are larger than the largest eigenvalue of an observable [7]. To further illustrate this point, I shall now show how a simple classical model can lead to a gross violation of the CHSH inequality, if not all data are included in the statistical analysis.

Consider a massive classical object, initially at rest, which splits into two parts carrying angular momenta \( J \) and \(-J\). Let \( \mathbf{n} = J/J \) be the unit vector in the direction of \( J \). The direction of \( \mathbf{n} \) is random (it is isotropically distributed on the unit sphere). Two distant observers, conventionally called Alice and Bob, independently choose unit vectors \( \mathbf{a} \) and \( \mathbf{b} \), respectively. Alice measures \( \mathbf{n} \cdot \mathbf{a} \) and records a result, \( \alpha \), as follows:

\[
\begin{align*}
\alpha &= 1 & \text{if} & & \mathbf{n} \cdot \mathbf{a} > 1/\sqrt{2}, \\
\alpha &= -1 & \text{if} & & \mathbf{n} \cdot \mathbf{a} < -1/\sqrt{2},
\end{align*}
\]

and \( \alpha = 0 \) in any other case. Likewise Bob measures \(-\mathbf{n} \cdot \mathbf{b}\) and records \( \beta = \pm 1 \) or 0, according to the same rule. (You can easily visualize these rules by thinking of \( \mathbf{n} \) as the time axis in a Minkowski spacetime. Then Alice records 1 or \(-1\) when \( \mathbf{a} \) lies in the future or past light cone, respectively, and Bob follows the same rule for \(-\mathbf{b}\).)

Obviously, Alice and Bob’s results are correlated: they are controlled by the common “hidden” variable \( \mathbf{n} \). This is what Mermin [8] calls a CLASS situation: Correlation Locally Attributable to the Situation at the Source. (This term is meant to signify a degree of virtue.)

It is easily shown that the correlation \( \langle \alpha \beta \rangle \) is a continuous function of \( \mathbf{a} \cdot \mathbf{b} \), and takes values from \( \langle \alpha \beta \rangle = 1 - 2^{-1/2} \approx 0.3 \), for \( \mathbf{a} \cdot \mathbf{b} = 1 \), to \( \langle \alpha \beta \rangle \approx -0.3 \), for \( \mathbf{a} \cdot \mathbf{b} = -1 \). If Alice and Bob consider other possible testing directions, say \( \mathbf{a}' \) and \( \mathbf{b}' \), and likewise define results \( \alpha' \) and \( \beta' \), the CHSH inequality [3]

\[
|\langle \alpha \beta \rangle + \langle \alpha \beta' \rangle + \langle \alpha' \beta \rangle - \langle \alpha' \beta' \rangle| \leq 2,
\]

is satisfied, as it should be for any CLASS model.

Suppose, however, that Alice and Bob consider a null result as a failure, and retain in their statistics only those events where both results differ from zero. It is easily seen that, in the events postselected in that way, \( \alpha \beta = 1 \) if the angle between \( \mathbf{a} \) and \( \mathbf{b} \) is less than \( 90^\circ \) (this is a necessary condition, not a sufficient one), and \( \alpha \beta = -1 \) if that angle is more than \( 90^\circ \). Consider now four directions, making angles of \( 45^\circ \), as shown in the figure.

The four directions used in Eq. (2) make angles of \( 45^\circ \).
We then have
\[
\langle \alpha \beta \rangle = \langle \alpha' \beta' \rangle = \langle \alpha' \beta \rangle = -\langle \alpha' \beta' \rangle = 1,
\] (3)
and the left hand side of Eq. (2) is equal to 4, so that the CHSH inequality is grossly violated in this CLASS model. Even the Cirel’son inequality [9, 10], which is respected by quantum mechanics, is violated by the postsel ected results.

Why is this ridiculous example instructive? Some time ago, Werner [11] constructed a density matrix \( \rho_W \) for a pair of spin-\( j \) particles, with paradoxical properties. Werner’s state \( \rho_W \) cannot be written as a sum of factorable density matrices, \( \sum_j c_j \rho_{A_j} \otimes \rho_{B_j} \), where \( \rho_{A_j} \) and \( \rho_{B_j} \) belong to the two particles. Therefore, genuinely quantal correlations are involved in \( \rho_W \).

For example, in the simple case of a pair of spin-\( \frac{1}{2} \) particles, Werner’s state is
\[
\rho_W = \frac{1}{8} \mathbb{1} + \frac{1}{2} \rho_{\text{singlet}},
\] (4)
namely, an equal weight mixture of a totally uncorrelated random state, and of a singlet state (the latter maximally violates the CHSH inequality). A definitely nonclassical property of this \( \rho_W \) was discovered by Popescu [12], who showed that such a particle pair could be used for teleportation of a quantum state [13], albeit with a fidelity lesser than if a perfect singlet were employed for that purpose.

This nonclassical property is surprising, because, for any pair of ideal local measurements performed on the two particles, the correlations derived from \( \rho_W \) satisfy the CHSH inequality. Moreover, as Werner showed, it is possible to introduce a “hidden-variable” model, which correctly produces all the observable correlations for such ideal measurements. In this model, the hidden variable is a unit vector \( \mathbf{r} \) in Hilbert space, and the quantum probability rules are correctly reproduced if \( \mathbf{r} \) is isotropically distributed. Werner’s prescription for the results of measurements of projection operators is the following: if Alice considers a complete set of orthonormal vectors \( \mathbf{v}_\mu \), and measures the corresponding projection operators \( P_\mu \), the result is \( P_\mu = 1 \) for the \( \mathbf{v}_\mu \) having the smallest value of \( |\mathbf{r} \cdot \mathbf{v}_\mu| \) (that is, the one most orthogonal to \( \mathbf{r} \)) and \( P_\mu = 0 \) for all the other \( \mathbf{v}_\mu \). For Bob, the rule is different and the results are only probabilistic: the expectation value of \( P_\mu \), for given \( \mathbf{r} \), is \( \langle \mathbf{r}, P_\mu \mathbf{r} \rangle \).

Werner’s algorithm for Alice’s result becomes ambiguous for spin \( > \frac{1}{2} \), and it must be supplemented by further rules, when we consider projection operators of rank 2 or higher. For any projection operator on a multi-dimensional subspace, Alice has to introduce, in an arbitrary way, orthogonal frames which span that subspace and its orthogonal complement. This defines a privileged complete orthogonal basis, for which all the \( P_\mu \) are defined as above. Then, the value of a projection operator on any subspace is taken as
equal to the sum of the values, 0 or 1, of the projection operators on all the privileged orthogonal vectors spanning that subspace. This rule is unambiguous (once we have decided how to choose the privileged vectors), but it has curious consequences.

Consider for example a 3-dimensional Hilbert space, with an orthogonal basis \{x, y, z\}. Let \{u, v, z\} be another orthogonal basis, so that \{x, y\} and \{u, v\} span the same subspace, orthogonal to z. Let \{x, y\} be our choice of privileged orthogonal basis for defining the value of the projection operator, \(P_{xy} = P_{uv}\), on that subspace. It is then always possible to find “hidden” vectors \(r\) such that

\[
|r \cdot u| < |r \cdot z| < |r \cdot v|,
\]
and

\[
|r \cdot z| < |r \cdot x| < |r \cdot y|.
\]

In that case, Werner’s rules imply that \(P_u = 1\) if Alice simultaneously measures \(P_v\) and \(P_z\), but, on the other hand, the value of \(P_{xy} = P_{uv}\) is zero! This looks paradoxical, and yet, after averaging over all \(r\), we still have

\[
\langle P_u \rangle + \langle P_v \rangle = \langle P_{uv} \rangle,
\]

in agreement with quantum mechanics.

We thus see that the phrase “a measurement of \(P_u\)” is ambiguous. We may have, for some values of the hidden variable \(r\), different outcomes depending on whether we measure \(P_u\) directly, or we first perform a coarser measurement for \(P_{uv}\), which is then refined for \(P_u\). This ambiguity was exploited by Popescu [14], as follows. Instead of measuring complete sets of projection operators of rank 1, Alice and Bob measure suitably chosen (and mutually agreed) projection operators of rank 2, say \(P_A\) and \(P_B\). If one of them gets a null result, the experiment is considered to have failed, and they test another Werner pair. Only if both Alice and Bob find the result 1, they proceed by independently choosing projection operators of rank 1, on vectors that lie in the subspaces spanned by \(P_A\) and \(P_B\), respectively. Popescu then shows that if the initial Hilbert space (for each particle) has dimension 5 or higher, the correlation of the final results violates the CHSH inequality. In other words, Werner’s hidden variable model, which worked for single ideal measurements, is incapable of reproducing the results of several consecutive measurements (and of course no other hidden variable model would do it).

How can we understand this paradoxical result? We had what appeared to be a CLASS model, similar to the classical model described at the beginning of this essay. In the former case, the CHSH inequality was violated as a result of faulty (postselected) statistics—all
the failures were discarded. The present case is subtler: Alice and Bob can, if they wish, discard their failures before proceeding to the final measurements. In other words, they can select a subensemble out of the original ensemble, and it is this subensemble that violates the CHSH inequality. The paradox is that the selection of this subensemble apparently involves only local operations. How can it destroy the CLASS property?

The point is that, in addition to the local measurements of $P_A$ and $P_B$, an exchange of classical information is needed for the selection of the CHSH-violating subensemble. That classical information is not just an abstract notion: it is conveyed by physical agents, such as electromagnetic pulses. It is customary to consider information carriers as exophysical systems [15], but this can only be an approximation, which now raises suspicion. To further sharpen the issue, let us promote the information carriers to endophysical status, by attributing to them dynamical properties. This leads to a new difficulty: there is no consistent hybrid dynamical formalism for interacting classical and quantum systems. We must therefore treat the information carriers as quantum systems, whose interaction with the Werner particles is generated by a Hamiltonian, as usual. These additional quantum systems are manifestly nonlocal, since their role is to propagate between Alice and Bob. It now becomes obvious that the selection of the CHSH-violating subensemble involves a nonlocal operation, and it is the latter that violates the CLASS property of the original ensemble.

Popescu’s construction [14] did not work for spaces with fewer than $5^2$ dimensions, but similar protocols have been found [16–19] for Werner pairs of spin-$\frac{1}{2}$ particles. If each one of these pairs is tested separately, the CHSH inequality is satisfied, as we know. We may, however, test several pairs together. For example, two pairs are described by a $4^2$-dimensional space, in which there are nontrivial rank-2 projection operators for each observer. Then, suitable subensembles can be selected, that violate the CHSH inequality. It is even possible to distill, from a large set of Werner pairs, a subset of almost pure singlets [16–19]. Here again, no hidden variable model can reproduce the results of collective measurements performed on several Werner pairs.

In conclusion, we see that the notion of quantum nonlocality is subtler than we may have thought. The conversion of a CLASS model into one that violates the CHSH inequality can be explained in two equivalent ways: by the use of biased statistics (postselected data), or by the introduction of a nonlocal agent carrying information between the observers, before completion of their measurements. Further investigations are needed, for which the advice of Abner Shimony will be most precious.

I am grateful to the authors of Refs. [16–19] for advance copies of their articles, and to N. D. Mermin and S. Popescu for an illuminating exchange of correspondence. This work was supported by the Gerard Swope Fund, and the Fund for Encouragement of Research.
1. J. S. Bell, Physics 1 (1964) 195.

2. L. E. Ballentine, Am. J. Phys. 55 (1987) 785.

3. J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. 23 (1969) 880.

4. A. Aspect, P. Grangier, and G. Roger, Phys. Rev. Lett. 47 (1981) 460; 49 (1982) 91.

5. A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. 49 (1982) 1804.

6. E. Santos, Phys. Rev. Lett. 66 (1991) 1388.

7. Y. Aharonov, D. Z. Albert, and L. Vaidman, Phys. Rev. Lett. 60 (1988) 1351.

8. N. D. Mermin, unpublished Bielefeld lecture notes (1995).

9. B. S. Cirel’son, Lett. Math. Phys. 4 (1980) 93.

10. L. J. Landau, Phys. Lett. A120 (1987) 54.

11. R. F. Werner, Phys. Rev. A40 (1989) 4277.

12. S. Popescu, Phys. Rev. Lett. 72 (1994) 797.

13. C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Phys. Rev. Lett. 70 (1993) 1895.

14. S. Popescu, Phys. Rev. Lett. 74 (1995) 2619.

15. A. Peres, Quantum Theory: Concepts and Methods (Kluwer, Dordrecht, 1993) p. 173.

16. C. H. Bennett, H. Bernstein, S. Popescu, and B. Schumacher, Concentrating Partial Entanglement by Local Operations (submitted to Phys. Rev. A, 1995).

17. C. H. Bennett, Quantum and Clasical Information Transmission Acts and Reducibilities (to appear in proceedings of EPR60 conference, Haifa, 1995).

18. C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. Smolin, and W. K. Wootters, Purification of Noisy Entanglement, and Faithful Teleportation via Noisy Channels (submitted to Phys. Rev. Lett., 1995).

19. D. Deutsch, A. K. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, Security of Quantum Cryptography over Noisy Channels (preprint, 1995).