How Much Measurement Independence Is Needed to Demonstrate Nonlocality?

Barrett, Jonathan; Gisin, Nicolas

How to cite

BARRETT, Jonathan, GISIN, Nicolas. How Much Measurement Independence Is Needed to Demonstrate Nonlocality? In: Physical review letters, 2011, vol. 106, n° 10, p. 4. doi: 10.1103/PhysRevLett.106.100406

This publication URL: https://archive-ouverte.unige.ch/unige:15877
Publication DOI: 10.1103/PhysRevLett.106.100406

How much measurement independence is needed in order to demonstrate nonlocality?

Jonathan Barrett
Department of Mathematics, Royal Holloway, University of London, Egham, Surrey TW20 0EX, United Kingdom

Nicolas Gisin
Group of Applied Physics, University of Geneva, 1211 Geneva 4, Switzerland

(Dated: January 28, 2011)

If nonlocality is to be inferred from a violation of Bell's inequality, an important assumption is that the measurement settings are freely chosen by the observers, or alternatively, that they are random and uncorrelated with the hypothetical local variables. We study the case where this assumption is weakened, so that measurement settings and local variables are at least partially correlated. As we show, there is a connection between this type of model and models which reproduce nonlocal correlations by allowing classical communication between the distant parties, and a connection with models that exploit the detection loophole. We show that even if Bob's choices are completely independent, all correlations obtained from projective measurements on a singlet can be reproduced, with the correlation (measured by mutual information) between Alice's choice and local variables less than or equal to a single bit.

PACS numbers:

Quantum nonlocality, whereby particles appear to influence one another instantaneously even though they are widely separated in space, is one of the most remarkable phenomena known to modern science [1–4]. Historically, this peculiar prediction of quantum theory triggered many debates and even doubts about its validity. Today, it is a well established experimental fact [5].

The profound implications of quantum nonlocality for our world view remain controversial. But it is no longer considered as suspicious, or of marginal interest. It is central to our understanding of quantum physics, and in particular it is essential for the powerful applications of quantum information technologies. In 1991, A. Ekert showed how shared entanglement could enable distant partners to establish a shared cryptographic key [8]. Ekert's intuition is quite simple – if there are no local variables, then no adversary could possibly hold a copy of these variables – yet it came decades after the birth of quantum mechanics. In 2005, Barrett and co-workers used this intuition to show how, with no further assumptions about quantum theory, nonlocal correlations alone could ensure security of a secret key [9]. Acín and coworkers extended this result, showing how to generate secret keys using simple nonlocal correlations that violate the well-known CHSH inequality [10]. Simultaneously, it was realized that Bell inequalities are the only entanglement witnesses that can be trusted in cases where the dimensions of the relevant Hilbert spaces are unknown [10].

In an experimental demonstration of quantum nonlocality, measurements are performed on separated quantum systems in an entangled state, and it is shown that the measurement outcomes are correlated in a manner that cannot be accounted for by local variables. In order to conclude that nonlocality is exhibited, it is crucial for the analysis that the choices of which measurement to perform are freely made by the experimenters. Alternatively, they must be random and uncorrelated with the hypothetical local variables. It is well known that if the measurement settings are not random, but are in fact determined by the local variables, then arbitrary correlations can be reproduced [11].

Here we reverse the argument. Taking for granted that quantum nonlocal correlations are produced, how independent must the measurement settings be assumed to be in order to rule out an explanation in terms of local variables [6]? We show that all correlations obtained from projective measurements on a singlet state can be reproduced with the mutual information between local variables and the measurement setting on one side less than or equal to one bit. This is not only of fundamental interest. If a cryptographic protocol is relying on the nonlocality of correlations for security, it is vital that there is no underlying local mechanism that an eavesdropper may be exploiting. The result shows that random number generators used to determine settings must be assumed to have a very high degree of independence from the particle source.

Bell Experiments. In an experimental test of a Bell inequality, two experimenters – traditionally named Alice and Bob – are spatially separated. They each control a quantum system, and the joint state of these two systems is an entangled state. Alice and Bob each perform a measurement on their quantum system, in such a way that the measurements are spacelike separated. Alice's measurement is chosen from a finite set of possibilities, as is Bob's. Call these measurement settings the *inputs* and label them x for Alice's choice and y for Bob's choice. The outcome of Alice's measurement (her *output*) is labeled a and Bob's b. By repeating this procedure many times and collating the data, Alice and Bob can estimate conditional probabilities of the form $P(a, b|x, y)$.

The question is then, when can the correlations produced in an experiment like this be explained by underlying local variables? Locality here implies that the outcome of Alice's measurement cannot be directly influenced by Bob's choice of which measurement to perform, and vice versa. More precisely, let the hypothetical un-

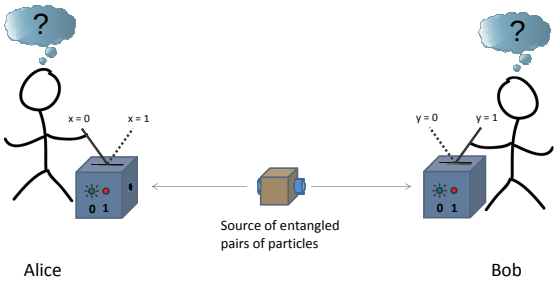

FIG. 1: *Schematic illustration of an experiment for testing quantum nonlocality.*

derlying variables be denoted λ, and assume a distribution $P(\lambda)$. Then *Bell locality* is the condition that

$$P(a,b|x,y,\lambda) = P(a|x,\lambda)\cdot P(b|y,\lambda). \qquad (1)$$

If the correlations could in principle be explained as arising from underlying local variables, then they can be written in the form

$$P(a,b|x,y) = \sum_\lambda P(\lambda)\cdot P(a|x,\lambda)\cdot P(b|y,\lambda). \qquad (2)$$

(Here and throughout, sums should be replaced with integrals when the variable is not discrete.) The correlations cannot be written in this form if and only if they violate a Bell inequality.

Suppose that correlations are obtained which do violate a Bell inequality. What assumptions are necessary to conclude that this is indicative of nonlocality? There is a large literature on this topic, and a reader could do a lot worse than refer to Bell's original papers [1]. But it is uncontroversial that the standard argument needs to assume that the inputs x, y are freely, or at least independently and randomly chosen by Alice and Bob.

Non-independent measurement choices. We wish to analyze the case in which measurement settings can be correlated with local variables. Hence in addition to a set of correlations $P(a,b|x,y)$, assume a distribution over inputs $P(x,y)$. A *correlated-settings* model is defined by a variable λ, and an overall joint probability distribution $P_{CS}(a,b,x,y,\lambda)$. The correlated-settings model *reproduces* the given correlations and input distribution if $P_{CS}(a,b|x,y) = P(a,b|x,y)$ and $P_{CS}(x,y) = P(x,y)$. The model is Bell-local if the distribution P_{CS} also satisfies an equation of the form of Eq (1). In this case

$$P_{CS}(a,b|x,y) = \sum_\lambda P_{CS}(\lambda|x,y)\cdot P_{CS}(a,b|x,y,\lambda)$$
$$= \sum_\lambda P_{CS}(\lambda|x,y)\cdot P_{CS}(a|x,\lambda)\cdot P_{CS}(b|y,\lambda). \qquad (3)$$

Note the similarity with Eq. (2). The usual kind of model, in which Alice's and Bob's inputs are assumed independent of the local variables, is a special case with $P_{CS}(x,y|\lambda) = P_{CS}(x,y)$. By Bayes' rule, this also gives $P_{CS}(\lambda|x,y) = P_{CS}(\lambda)$, and Eq. (3) reduces to an equation of the form of Eq. (2).

Another special case is the extreme case where the inputs are completely determined by λ, so that $x = f(\lambda)$ and $y = g(\lambda)$ for functions f and g. Here, for all x, y, λ, $P_{CS}(x,y|\lambda) = 0$ or 1. Any correlations can be produced by a Bell-local model of this form [11].

In other models, x and y will be neither independent nor determined by λ. In general there are many numerical measures of the degree to which they are correlated. One particularly natural and simple measure is the *mutual information*. Thus a correlated-settings model can be characterized by the mutual information between the measurement choices x, y and λ:

$$I(x,y:\lambda) = H(x,y) + H(\lambda) - H(x,y,\lambda), \qquad (4)$$

where H is the Shannon entropy. When x and y are independent of λ, $I(x,y:\lambda) = 0$. If x and y are functions of λ on the other hand, then $I(x,y:\lambda) = H(x,y)$. Since in general $I(x,y:\lambda) \leq H(x,y)$, this means that the mutual information is *maximal* (with respect to a fixed distribution over inputs $P(x,y)$).

In the context of correlated-settings models, the question is no longer whether quantum correlations violate a Bell inequality. The question is, how large must $I(x,y:\lambda)$ be if the correlations are to be reproduced within a Bell-local model? Alternatively, for a fixed degree of correlation $I(x,y:\lambda)$, do quantum correlations violate a Bell inequality by a sufficiently large margin that the model cannot possibly be Bell-local? This depends on the precise experiment. For example, if the input alphabet consists of only two choices, as with the CHSH inequality, then any correlations can be reproduced by a correlated-settings model with $I(x,y:\lambda) \leq 1$. But the question remains open for larger input alphabets. Intuitively one may think that for large alphabets, if Alice's choice is only mildly correlated with the variable λ, then a sufficient violation of a suitable Bell inequality should still rule out locality. But we shall see that this intuition is wrong, at least for maximally entangled qubits.

A connection to communication cost. One way to simulate quantum correlations, including nonlocal correlations, is for Alice and Bob to communicate with one another after inputs are chosen, but before outputs are produced. Any correlations can be produced this way if Alice and Bob share random data, and Alice simply informs Bob of her input choice. Hence, if Alice is choosing from n_A possible inputs, then an obvious upper bound on how many classical bits need to be communicated is given by $\log n_A$. The minimum number of bits that must be communicated, on the other hand, provides a natural way of quantifying the amount of nonlocality inherent in a given set of correlations. This interesting line of research was started by T. Maudlin [2] and independently by Tapp et al. and by Steiner [12, 13]. It culminated with a model of Toner and Bacon, who proved that one single

bit of communication from Alice to Bob is sufficient to simulate all correlations obtained from arbitrary projective (i.e. Von Neumann) measurements on two qubits in a maximally entangled state [14]. It was extended to all two-party correlations, though ignoring the marginals, in [15].

A slightly more precise description of a communication model for simulating quantum correlations is as follows. The random data shared between Alice and Bob is a variable μ, with distribution $P_C(\mu)$. Assuming that Alice communicates first, Alice sends a bit string c_1 to Bob. The string c_1 can depend on μ and on Alice's input x. It may also depend on further random data held by Alice, but this data can without loss of generality be absorbed into μ. Hence assume that $c_1 = c_1(\mu, x)$. Depending on the protocol Bob may now send a bit string c_2 to Alice, where again without loss of generality, we assume that c_2 is a function of μ, c_1 and y. This continues for a total of k messages. The entire conversation between Alice and Bob on a particular round is the sequence $m = c_1, c_2, \ldots, c_k$, and is a function $m = f(x, y, \mu)$ of x, y, and μ. Finally, Alice outputs a, where again absorbing all randomness into μ, a can be assumed to be a function $a = g(x, \mu, m)$ of her input, the shared random data and the conversation. Similarly, Bob outputs b, which is a function $b = h(y, \mu, m)$. The protocol then terminates.

For each pair of inputs, a communication model defines a distribution $P_C(a, b, m, \mu|x, y)$, and reproduces correlations $P(a, b|x, y)$ if $P_C(a, b|x, y) = P(a, b|x, y)$. If a distribution over inputs $P(x, y)$ is also given, the conversation m has a well-defined distribution $P(m) = \sum_{x,y} P_C(m|x, y) \cdot P(x, y)$. Let the entropy of this distribution be $H(m)$.

The connection with correlated-settings models is expressed in the following

Theorem 1 *Consider a fixed input distribution $P(x, y)$, and a communication model with total conversation m. Then there exists a Bell-local correlated-settings model, which reproduces the same correlations and $P(x, y)$, with $I(x, y : \lambda) \leq H(m)$.*

Proof. Given $P(x, y)$ and a communication model as above, construct a correlated-settings model as follows. Define λ as the pair $\lambda = (\mu, m)$. In the communication model, m is the conversation between Alice and Bob, but in the correlated-settings model there is no communication, and λ represents the underlying variable. Define the correlated-settings model by $P_{CS}(a, b, x, y, \lambda) = P(x, y)P_C(a, b, \mu, m|x, y)$. It is easy to see that the resulting model reproduces $P(x, y)$ and the same correlations as the communication model. Bell locality is satisfied since

$$\begin{aligned} P_{CS}(a, b|x, y, \lambda) &= P_C(a, b|x, y, \mu, m) \\ &= P_C(a|x, \mu, m) \cdot P_C(b|y, \mu, m) \\ &= P_{CS}(a|x, \lambda) \cdot P_{CS}(b|y, \lambda). \end{aligned}$$

The mutual information between λ and (x, y) is

$$\begin{aligned} I(\lambda : x, y) &= I(\mu : x, y) + I(m : x, y|\mu) && (5) \\ &= 0 + H(m|\mu) - H(m|\mu, x, y) && (6) \\ &= H(m|\mu) && (7) \\ &\leq H(m). && (8) \end{aligned}$$

The first line uses the chain rule [16], the second line follows from the fact that μ is assumed to be independent of the inputs x, y, and the third line holds because m is a function of μ and x, y. This concludes the proof.

Intuitively, $\lambda = (\mu, m)$ restricts Alice's and Bob's choices to inputs (x, y) such that $m = f(x, y, \mu)$. We stress the generality of this result: the communication model may have involved two way communication, and may have been such that the communication on a given run is unbounded. As long as $H(m)$ is finite, the bound is useful.

We now apply the theorem with reference to the Toner-Bacon communication model. Here, x and y are arbitrary projective measurements. The shared random data μ takes the form of two vectors on the Bloch sphere and the distribution $P(\mu)$ is such that the two vectors are independent, with each uniformly distributed. For our purposes most of the details of the model do not matter, and we refer the reader to [14] for a description and proof that it reproduces quantum correlations for projective measurements on a singlet. It suffices to know that each round, the conversation m is a single bit communicated from Alice to Bob. The bit $m = f(x, \mu)$ is a function of Alice's input x and μ.

The resulting correlated-settings model also reproduces all correlations from projective measurements on a singlet. In general, the actual value of $I(x, y : \lambda)$ depends on the distribution over inputs $P(x, y)$. The theorem tells us that for any such distribution, $I(x, y : \lambda) \leq H(m) \leq 1$, since m is a single bit. In particular, if Alice's and Bob's inputs are independent, with each chosen from a uniform distribution over all possible directions, it is easy to verify that $I(x, y : \lambda) \approx 0.85$.

In a typical Bell-type experiment with a singlet, Alice and Bob will be choosing from finite sets of measurements. The Toner-Bacon model can still be applied, with the distribution $P(x, y)$ having support only on these finite sets. In this case too, $I(x, y : \lambda) \leq 1$. The Bell-local correlated-settings model reproduces the quantum correlations no matter how large the input alphabets, and no matter what Bell inequality is being tested.

Finally, in the Toner-Bacon model, Bob's setting is of course independent from the communication he receives from Alice. It follows that in the derived correlated-settings model, Bob's setting is independent from λ, i.e., $I(y : \lambda) = 0$.

The detection loophole. In a real Bell experiment, detection is inefficient. A typical analysis of the experiment estimates the correlations $P(a, b|x, y)$ from those runs on which both Alice's and Bob's detectors clicked, and simply ignores all the other runs. Such an analysis is valid, as long as it is assumed that the probability of a detector clicking is independent of the hypothetical local variables. This is sometimes called the *fair sampling assumption*.

It is possible to reproduce nonlocal correlations with

a model that is Bell-local, but which violates the fair sampling assumption [17, 18]. Denote the event that Alice's (Bob's) detector clicks by D_A (D_B). A *detection-efficiency* model is defined by a variable λ, with distribution $P_{DE}(\lambda)$, independent of x, y, and for each x, y a distribution $P_{DE}(a, b, D_A, D_B|x, y, \lambda)$. Bell-locality is the condition that $P_{DE}(a, b, D_A, D_B|x, y, \lambda) = P_{DE}(a, D_A|x, \lambda) \cdot P_{DE}(b, D_B|y, \lambda)$. The model reproduces correlations $P(a, b|x, y)$ if

$$P(a, b|x, y) = P_{DE}(a, b|x, y, D_A, D_B). \tag{9}$$

The efficiencies of Alice's detectors in this model are given by $P_{DE}(D_A|x) = \sum_\lambda P_{DE}(\lambda) P_{DE}(D_A|x, \lambda)$, and similarly for Bob's. If the detection efficiencies in a real experiment are high enough, and if a Bell inequality is violated by a large enough margin, then a Bell-local detection-efficiency model reproducing the correlations can be ruled out.

Given a distribution $P(x, y)$, and a detection-efficiency model reproducing correlations $P(a, b|x, y)$, it is easy to construct a correlated-settings model which also reproduces $P(x, y)$ and $P(a, b|x, y)$. Simply define

$$P_{CS}(a, b, x, y, \lambda) = P(x, y) \cdot P_{DE}(a, b, \lambda|x, y, D_A, D_B). \tag{10}$$

It is easy to show, using Eq. (9), that the correlated-settings model reproduces $P(x, y)$ and $P(a, b|x, y)$.

The construction can be applied, for example, to the Gisin-&-Gisin detection-efficiency model [19], which reproduces correlations from arbitrary projective measurements on a singlet. The efficiencies are independent of x and y and are given by $P(D_A) = 1/2$, $P(D_B) = 1$. In this model λ is a vector on the Bloch sphere, and $P_{DE}(\lambda|x, y, D_A, D_B) = |\lambda.\vec{x}|/2\pi$ where $\vec{x}$ denotes the Bloch vector representing the measurement x. In the resulting correlated-settings model, $I(x, y : \lambda)$ depends on $P(x, y)$. If x and y are independently and uniformly distributed, it is easy to show that $I(x, y : \lambda) = I(x : \lambda) \approx 0.28$. This is an even lower value than was obtained above using the Toner-Bacon model.

Finally, note that in the case of communication models, there was a connection between $I(x, y : \lambda)$ and the amount of communication (as measured by the entropy of the conversation's distribution). With detection efficiency models, we expect there to be a connection between $I(x, y : \lambda)$ and the detection efficiencies. Exploring this connection is an interesting direction for future work.

Conclusion. It is well known that in order to derive nonlocality from violation of a Bell inequality, one has to assume that there is no correlation between the hypothetical local variables λ and the experimenters' measurement choices x and y. We have investigated how much correlation could be allowed if quantum predictions are to remain incompatible with local variables. Surprisingly, no matter how large the alphabet size of the inputs, correlations from projective measurements on a singlet can be reproduced, with the mutual information between Alice's input and local variables not more than one bit, and Bob's input completely independent. This deserves further investigation. Future work should analyze more general scenarios, involving generalized measurements, partial entanglement [20], higher dimensional Hilbert spaces and more than two parties. It would also be instructive to consider measures of correlation other than the mutual information.

Finally, let us emphasize the change in paradigm since the old EPR paper [4]. If, contrary to EPR, one accepts nonlocality as a fact, then not only can one develop powerful applications in quantum information science, like device-independent quantum key distribution [9, 10], but moreover one can upper bound the lack of free choice of the players !

Acknowledgment

JB is supported by an EPSRC Career Acceleration Fellowship. NG acknowledges support from the ERC advanced grant QORE and the Swiss NCCR-QP. After completion of this work M. J. W. Hall posted a related interesting paper: arXiv:1007.5518 (PRL 105, 250404 (2010)). We thank him and C. Branciard for constructive discussions.

[1] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected papers on quantum philosophy*, Cambridge University Press, Cambridge, 1987, revised edition 2004.

[2] T. Maudlin, *Quantum Non-Locality and Relativity*, Blackwell Publishers, 2nd edition, 2002.

[3] L. Gilder, *The Age of Entanglement*, Alfred A. Knopf, 2008.

[4] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. **47**, 777 (1935).

[5] A. Aspect, Nature **398**, 189 (1999).

[6] see also: J. Kofler et al., Phys. Rev. A **73**, 022104 (2006); M. Pawlowski et al., arXiv:0903.5042.

[7] A. Aspect, Nature **466**, 866 (2007).

[8] A. K. Ekert, Phys. Rev. Lett. **67** 661 (1991).

[9] J. Barrett, L. Hardy and A. Kent, *Phys. Rev. Lett.* **95**, 010503 (2005).

[10] A. Acin, N. Gisin and Ll. Masanes, *Phys. Rev. Lett.* **97**, 120405 (2006).

[11] C. Brans, Int. J. Theoret. Phys. **27**, 219 (1988)

[12] G. Brassard, R. Cleve and A. Tapp, Phys. Rev. Lett. **83**, 1874 (1999).

[13] M. Steiner, Phys. Lett. A **270**, 239 (2000).

[14] B. F. Toner and D. Bacon, Phys. Rev. Lett. **91**, 187904 (2003).

[15] O. Regev and B. F. Toner, Proceedings of 48th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2007).

[16] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley, New York, 1991.

[17] P. Pearle, Phys. Rev. D **2**, 1418 (1970).
[18] E. Santos, Phys. Rev. A **46**, 3646 (1992).
[19] B. Gisin, N. Gisin, Phys. Lett. A **260**, 323 (1999).
[20] For two partially entangled qubits, there is a 2 bits communication model [14]; surprisingly, however, it is still unknown whether this is optimal or whether a single bit would suffice.