philosophies

MDPI

Article

The EPR-Bell Experiments: The Role of Counterfactuality and Probability in the Context of Actually Conducted Experiments

Anthony J. Leggett

Illinois Quantum Information Science and Technology Center, Department of Physics, Anthony J. Leggett Institute of Condensed Matter Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA; aleggett@illinois.edu

Abstract: Some aspects of the concepts of counterfactuality and probability are explored as they apply to the specific example of the famous "EPR-Bell" experiments realized by physicists over the last half-century. In particular the question is raised: what hypotheses about **actually conducted** experiments do the results exclude? It is argued that the answer depends on both whether these hypotheses are deterministic or stochastic, and on the "cardinality" of the experiment relative to the theory.

Keywords: counterfactuality; probability; EPR-Bell; nonlocality

check for updates

Citation: Leggett, A.J. The EPR-Bell Experiments: The Role of Counterfactuality and Probability in the Context of Actually Conducted Experiments. *Philosophies* **2024**, *9*, 133. https://doi.org/10.3390/philosophies9050133

Academic Editor: Rafael A. Alemañ

Received: 9 June 2024
Revised: 12 August 2024
Accepted: 12 August 2024
Published: 27 August 2024

Copyright: © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

This year will be the 60th anniversary of John Bell's famous paper [1], which showed that a class of eminently "commonsensical-seeming" theories about the physical world makes experimental predictions at variance with those of standard quantum mechanics (hereafter QM), it is also over a half-century since the first experiments [2] which provided clear evidence against this class of theories, and a decade since the first set of such experiments claimed to be "loophole-free" [3–5][1]. And yet the relevant community—crudely speaking, the wide body of physicists, philosophers, mathematicians and popular science writers interested in the foundations of QM—has apparently still not reached a universally agreed or even large-majority verdict on the conclusions which should be drawn from this state of affairs about the physical world and/or our knowledge of it. Crudely speaking, as a first pass one may divide this community into those who believe that Bell's theorem and the subsequent related work [6] of Clauser et al. (hereafter CHSH) is conceptually or technically flawed, and those who accept the conclusions of these and related papers. Since the former category seems to be, numerically, a fairly small (though vocal) minority and I do not count myself among them, I shall not address their arguments further in this essay. Among those who accept the Bell-CHSH conclusions, the belief of a large fraction seems to be that we are confronted with the choice of rejecting "locality" or rejecting "realism"; my subjective impression is that while popular writers on the subject tend to opt for the former choice, professional physicists by and large prefer the latter. However, there is also a point of view, espoused particularly trenchantly in a 2012 paper [7] by N. Gisin, that we do not in fact have this choice, since the celebrated "CHSH inequality" (see below) can be proved on the basis of locality alone, so that it is the latter which must be rejected.

In this essay I shall attempt to use this controversy as a template on which to explore some of the questions which arise in a more general context concerning two related concepts of quintessentially philosophical interest, namely counterfactuality and probability. While it is needless to say that this is not at all a new idea, I believe that there is a particular aspect which may not have received all the attention which it deserves, namely the way that the issue plays out when one compares a particular type of (non-QM) theory with a particular set of experiments, and it is this aspect which I shall attempt to emphasize.

To flesh this out a little:let us do a thought-experiment in which we have never heard of quantum mechanics and its predictions, but are familiar with the experimental situation regarding single photons (i.e., Malus's law). In anticipation of experiments on correlated photon pairs, we have devised a _specific type_ of "local hidden-variable (HV)" theory to explain the data; as we shall see, the two most important qualitative characteristics of this theory which we need to specify are whether the relevant "hidden variable" λ is continuous or discrete, and whether the predictions it makes for given λ are are deterministic or probabilistic. We then conduct the first set of experiments of the EPR-Bell type; needless to say, the number of runs in this experiment is finite, and we specify that number. Assuming that the experimental data found in our experiments are typical of those which we (now) know to be correct, what can we infer about the validity or not of our HV theory? Note that this formulation is a little different from that of many of the discussions in the literature, which tend to assume (a) that the "right" theory is quantum-mechanical, and/or (b) that the number of experimental runs can be imagined to tend to infinity. It is hoped that formulating the problem in the way done here may allow us to sharpen up some of the relevant considerations.

More specifically, the question which I shall attempt to address in this essay is whether any further light can be shed on these issues by replacing the notion of (microscopic) "realism" by an idea which some writers have preferred, namely that of "macroscopic counterfactual definiteness" (MCFD). In particular I will try to explore whether the proof of the CHSH theorem indeed requires only the assumption of locality, or whether some notion of MCFD or some related notion is also implicitly required. In addition I will briefly address a somewhat tangential issue, namely whether MCFD can be shown to be violated independently of the EPR-Bell experiments.

The plan of the paper is as follows: in Section 2 I very briefly review, first, some basic experimental facts about single photons (the system which has been most widely used in the EPR-Bell experiments), then describe the physical setup which generically defines this series of experiments, and finally summarize the results obtained to date. In Section 3 I define, following ref. [8], the notion of an "objective local theory (OLT)" (a formalized version of the class of "commonsensical" theories mentioned above) and the associated notion of MCFD, and give a concise (but not at all original) proof of the CHSH theorem using the latter. Section 4 makes the case that it is possible to explicate the concept of MCFD in terms of **actually conducted** experiments. By contrast, Section 5 considers the major claim made in ref. [7], namely the sufficiency of the locality condition, and argues that this cannot be explicated in terms of actually conducted experiments but needs an implicit assumption somewhat related to (but weaker than) MCFD. Section 6 is somewhat tangential to the main thrust of the paper: it discusses how far the postulate of MCFD may be refuted by experiments in an area quite different from EPR-Bell. Section 7 is a conclusion which tries to draw together the results of the paper. An appendix supplies the proofs of various statements made in Sections 4 and 5.

Some notes:

(1) Although I believe that the simple equation of "realism" (or MCFD) in general with determinism asserted in ref. [7] may need further discussion, in the context of the present work I am content to concede it; indeed at various points I will use the three words interchangeably.

(2) This essay is written on the assumption that a fair proportion of its readers may not be professional physicists or even familiar with the EPR-Bell experiments, so those who are may wish to skip Sections 2 and 3, except for the definition of MCFD if they are not already familiar with it.

(3) To conclude this brief introduction, let me emphasize that despite the title of the collection of which it forms a part, and the historical motivation of the EPR-Bell experiments, this paper is really not about quantum mechanics at all. In fact, it can be read and understood under the fiction of the alternative history sketched above, in which for some reason the first experiments were done with no knowledge of QM or

its predictions for them; we can still ask what conclusions we would be entitled to draw from the results about the structure of the physical world.

2. Photons: The Relevant Experimental Facts

In this section I will briefly review the salient facts about the properties of photons and their interactions with a standard measuring apparatus, and the basic physical setup of an experiment of the "EPR-Bell" class. As mentioned above, this is standard material and can be skipped by readers already familiar with it.

According to our current conceptions of physics, a "photon" is an irreducible minimal quantum[2] of electromagnetic radiation (light), which is typically emitted when an atom makes a transition between two of its allowed states. (The emission may or may not be "heralded", that is, we may or may not have independent evidence that it has occurred; this is unimportant for most of the ensuing discussion).

The photon carries various properties: energy, momentum and, what is crucial for the present discussion, polarization; just as in the case of a classical (linearly polarized[3]) light wave the electric field associated with the photon can lie in any direction in the plane perpendicular to the direction of propagation (hereafter denoted z). Most (though not all) of the EPR-Bell experiments have to do with the measurement of polarization, so for our present purpose we can forget about most of the other properties; however, one fact which will be implicitly used in Section 3 is that photon wave packets can be (and typically are) spatially localized over a distance small compared to the geometrical dimensions of a typical EPR-Bell experiment.

How to measure the polarization of a single photon? The standard procedure, which is idealized in Figure 1a, is to direct it on to a "measuring station" of the type shown, which consists of a polarizer followed by a detector. The (idealized) polarizer is a crystal with strongly anisotropic electromagnetic properties in the xy-plane, which we arrange to be oriented perpendicular to the photon's direction of propagation: it has the fundamental property that if a particular photon incident on it has electric field parallel to (say) the x-axis of the crystal, it will be transmitted and enter the detector Y (for "yes"), while if the electric field is oriented parallel to the y-axis, the photon will be reflected and enter the detector N ("no"). In either case, the incidence of the photon on the relevant detector triggers a shower of electrons, which can then be amplified and used to trigger further events of a macroscopic nature, such as the flashing of a light, ringing of a bell or printout of the result on a computer tape; at least for now, I will assume (despite the reservations of some of my theoretical colleagues (cf. Section 6), but apparently in agreement with just about all the experimental ones), that this collocation of events does in fact occur, and moreover, within a timescale which can be made very short (typically $\lesssim$ **1** $\boldsymbol{\mu sec}$); as a shorthand, I will characterize it by saying that the relevant detector "clicks". It is an experimental fact that whenever a photon (say a heralded one) is incident on the measuring station, a click occurs **either** in detector Y **or** in detector N (never in both, never in neither), and moreover that when we have an independent means of inferring the photon polarization, then a "vertical" polarization indeed triggers detector Y and a "horizontal" polarization detector N.

But what if the photon is linearly polarized in some arbitrary direction in the xy-plane, let us say at an angle θ with respect to the x-axis? Then the experimentally observed result is that if a large number of photons with this property is directed at the measuring station, a fraction of them roughly equal to $\mathbf{cos^2}\theta$ trigger detector Y and the remainder (a fraction $\mathbf{sin^2}\theta$) detector N. (This is the single-photon version of Malus's law for classical light, of which a fraction $\mathbf{cos^2}\theta$ is transmitted). At least within our present understanding, it is impossible to predict which photons will trigger Y and which N, but the statistics of the result is an experimental fact.

One feature of the above arrangement (or indeed any measurement of polarization) which is at first sight rather trivial, but is essential to the arguments examined in this essay, is that for any given photon the orientation of the polarizer in the measuring station, while arbitrary, must be specified in advance; we cannot simultaneously measure: "yes"

versus "no" for more than one orientation. However, a convenient possibility used in many of the actual experiments is to set up two measuring stations M_a and $M_{a'}$ which have transmission axes oriented in different directions, say $\boldsymbol{a}$ and $\boldsymbol{a}'$ (see Figure 1b), and to arrange a switch, activated either by the experimenter or, more usually, by some random process, which will direct each particular photon into either M_a or $M_{a'}$; thus for any given photon we measure either $\boldsymbol{a}$ or $\boldsymbol{a}'$ but not both. We will call this measurement setup, shown in Figure 1b, a "measurement superstation".

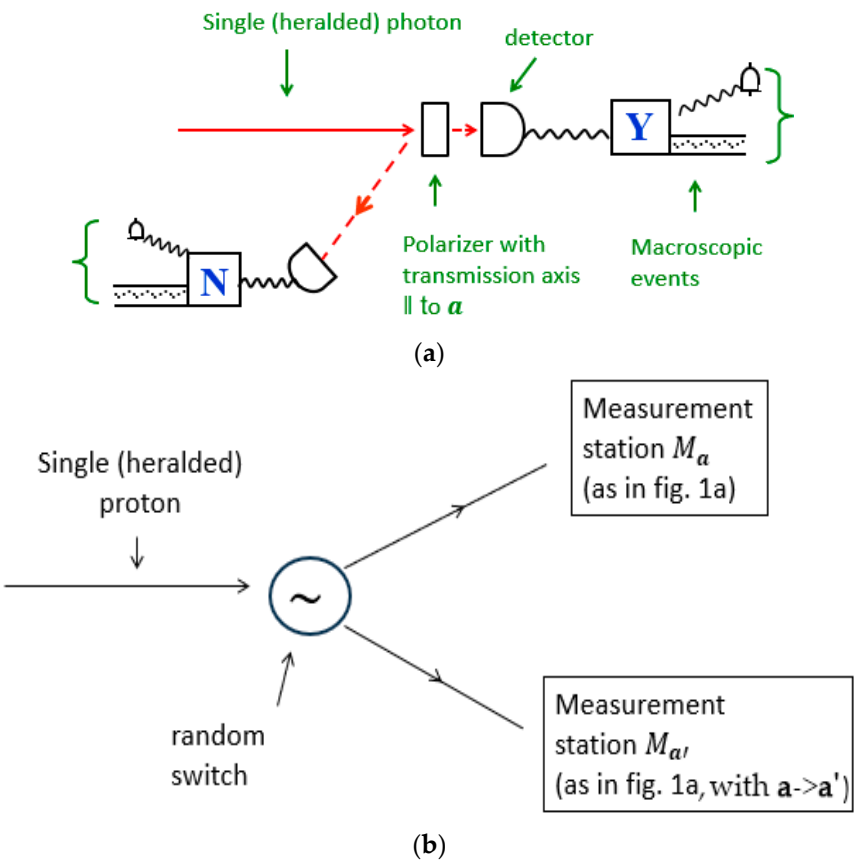

Figure 1. (**a**): a measurement station, (**b**): a measurement superstation.

The above is really all we need to know about single photons: in sum, for any given photon we can ask one (but not more than one) of two (or more) "yes/no" questions, and the photon is guaranteed to give to the question asked either the answer "yes" or the answer "no". These are experimental facts independent of any theory.

Now I turn to the description of a typical "EPR-Bell" experiment: again, this description will be somewhat idealized.

The basic desideratum is a source of back-to-back photon **pairs** (e.g., a set of atoms which are promoted into an excited state and can decay back to the groundstate by a two-step process) and two "distant" measurement superstations L ("left") and R ("right") (originally at opposite ends of the laboratory, nowadays sometimes in different satellites etc.) which will each conduct measurements of the polarization of one photon from each pair: see Figure 2. The essential feature is that the distances involved, and the time interval for the two-photon emission process, should be such that there is no time for the choice of the transmission axis of superstation L ($\boldsymbol{a}$ versus $\boldsymbol{a}'$) to be transmitted non-superluminally to measurement superstation R in advance of the measurement event (switching and "click") there. To make the experimental result "interesting", it is also required that it should be possible to identify which of the photons detected at superstation L was emitted in conjunction with a particular photon detected at superstation R (in practice, this means that the frequency of "joint" detections in L and R, meaning detection of both events within

some time period of the order of the full decay time of the atomic species involved in the emission, should be a reasonable fraction of all detections).

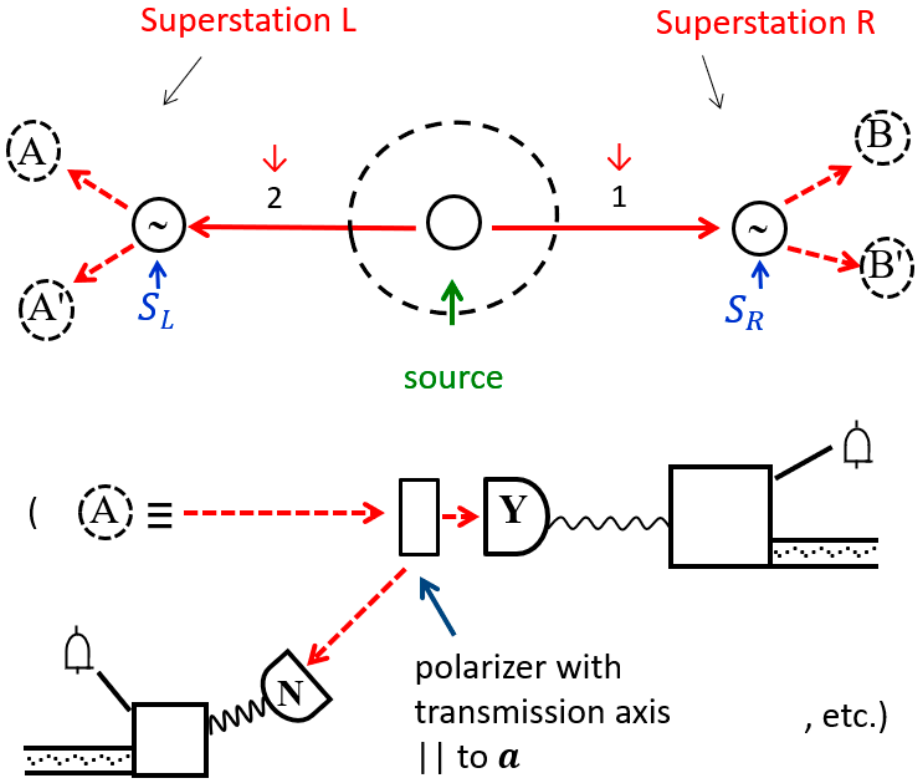

Figure 2. THE EPR-BELL EXPERIMENTS (idealized).

The raw data collected in the experiment is simply the correlations of "joint" detections in the L and R superstations. To make this explicit, let us first identify, out of all the possible pairs of detections, those which constitute "joint" detections in the sense of the last paragraph. (If we do not take this precaution, the subsequent steps will still be meaningful but the results will not be particularly striking). Then we look up our record of the positions of the switches S_L, S_R at the time of the joint measurement in question, so that we know whether superstation L measured polarization with respect to $\boldsymbol{a}$ or $\boldsymbol{a}'$ (similarly whether superstation R measured with respect $\boldsymbol{b}$ or $\boldsymbol{b}'$). To simplify the subsequent discussion, we introduce the notation that if on a given joint event i superstation L measured with respect to $\boldsymbol{a}$ and gave the result "yes" (i.e., detector Y of station M_a clicked) then we say that for this joint event the variable $\boldsymbol{A_i}$ takes the value $+1$, whereas if the result was "no" (click in detector N of $\boldsymbol{M_a}$) then $\boldsymbol{A_i} = -1$. Similarly, if for the event j in question superstation L measured with respect to $\boldsymbol{a}'$, if detector Y of $\boldsymbol{M_{a'}}$ clicked we say that the variable. A'_j takes the value +1 for that event, whereas if detector N of $M_{a'}$ clicked we assign $\boldsymbol{A_j}' = -1$. Similar definitions are made for superstation R. Note that if it is the polarization with respect to $\boldsymbol{a}$ which is measured on a particular event, then neither the Y nor the N detector of measurement station $\boldsymbol{M_{a'}}$ clicks, so at least for the moment $\boldsymbol{A'}$ **is simply not defined**; similarly if polarization with respect to a' is measured for a particular event, the quantity A is not defined; and similarly for the outputs at superstation R. Thus, for any given joint event i one can measure exactly **one** of the four quantities $\boldsymbol{AB}, \boldsymbol{A'B}, \boldsymbol{AB'}$ and $\boldsymbol{A'B'}$, the other three remaining at least for the moment undefined. It is this table of numbers which constitutes the raw data acquired in the experiment, and we can now see that the general scheme extends to cases where it is some (binary) variable other than polarization which is measured, or even where the microscopic objects in question are not photons at all; a minority of the existing EPR-Bell experiments in fact has this property.

Having collected the raw data, we can now organize it in various ways. It turns out that a particularly useful procedure runs as follows: first, we define the quantity $\boldsymbol{N_{ab}}$ to be the total number of photon pairs for which photon L was switched into measurement station M_a and photon R into measurement station M_b, irrespective of the results of the measurement. Then we define the experimentally measured average of the quantity AB, which we will denote $\langle AB \rangle_{exp}$, by the formula

$$\langle AB \rangle_{exp} \equiv \sum_i (A_i B_i) / N_{ab}$$

Note carefully that this is the average over the subensemble of photon pairs for which L was switched into **a** and R into **b**. We similarly define experimentally measured averages of the quantities $A'B$, AB' and $A'B'$ over the corresponding (different) subensembles. Finally we add and subtract the experimental averages to obtain various combinations of them; for simplicity, in the rest of this paper we will specialize to the particular combination

$$K_{exp}(\boldsymbol{a}, \boldsymbol{a'}, \boldsymbol{b}, \boldsymbol{b\prime}) \equiv \langle AB \rangle_{exp} + \langle A'B \rangle_{exp} + \langle AB' \rangle_{exp} - \langle A'B' \rangle_{exp} \tag{2.1}$$

As we will see in the next section, the class of "objective local theories" (OLTs) makes a specific prediction for the quantity K_{exp}, and the punch-line is that over the last half-century a sequence of experiments, of progressively increasing sophistication and refinement, have been carried out and show that this prediction is violated by many standard deviations. (The actual values of K_{exp} obtained approach the QM prediction, which for suitable choices of the settings, is $2\sqrt{2}$).

To conclude, let me re-emphasize that the description of the experimental protocol given above is idealized and has not been followed in all of the EPR-Bell experiments conducted to date. In particular, one ("Eberhard" [9]) variant, which places substantially weaker requirements on the experimental setup and has been used in some of the most "spectacular" experiments, requires the detection, in each measurement station, only of the Y outcome. The discussion of these variants proceeds in parallel to that which will be given in the rest of this paper, and I do not believe introduces any new conceptual points, so for simplicity I shall specialize from now on to an experiment of the idealized type described above, with an output given by the quantity K_{exp} defined by Equation (2.1).

3. Definition of an Objective Local Theory and of Macroscopic Counterfactual Definiteness: The CHSH Inequality

In this section I shall define the class of OLT theories, introducing in the process the concept of MCFD, and show that any theory of this class predicts an upper limit (the "CHSH inequality"[4]) on the quantity K_{exp}.

To the best of my knowledge, the idea of an "objective local theory" was first formulated explicitly in the EPR-Bell context by Clauser and Horne [8], and has been developed by many subsequent writers. I will paraphrase the rather qualitative definition given in ref. [8] by the conjunction of three postulates, of which the first two are not usually regarded as controversial or in need of further discussion:

(1) Einstein locality: No causal effect can propagate faster than the speed of light in vacuum.
(2) No retrocausality: future events cannot causally affect present or past ones.

A couple of notes:

(i) Of course within the framework of special relativity as standardly formulated, postulate (1) entails postulate (2). However, since we wish to consider very general classes of theories about the world, it is convenient to regard these two postulates as independent.
(ii) The significance of postulate (2) in the context of the EPR-Bell experiments is that it guarantees (given, of course, an appropriate experimental setup) that the statistical properties of the ensemble of photon pairs on which (say) A and B were measured are identical to those of the total ensemble of emitted pairs.

We now focus on the third postulate which is usually held to define, in conjunction with postulates (1) and (2), the class of objective local theories. In the original Clauser-Horne paper, which considers the photonic "wave packets" in transit from the source to the measurement superstations, the postulate is essentially that e.g., photon L carries sufficient information to determine the probability of the Y and N responses of either of the measurement stations M_a and $M_a{}'$ into which that photon may be switched, whichever that is. A similar assumption is made for photon R. In other words, even before the measurement, each L photon "possesses" a definite value of both A and A'; similarly, each photon R "possesses" a definite value of both B and B'. It seems natural to rephrase this by saying that for each L photon the quantities A and A' are both "real". Thus we have for one possible form of the third postulate

(3a) Realism: Any property of a microscopic object which may be measured on it possesses a definite value even in the absence of measurement.

A couple of things about this definition should be noticed. First, in terms of the three "species" of realism which are sometimes distinguished in the philosophical literature (see e.g., ref. [10]), this would seem to fall most naturally under "entity realism", since by referring to the properties of a "microscopic object" we have presumably implied the existence of such an object. (We could perhaps modify the definition by ascribing the "property" to the experimental setup as a whole, but this takes us close to the notion of MCFD, see below). Secondly, the definition makes no explicit reference to determinism; cf. however Section 4.

However, it is sometimes argued that the whole idea of a "photon" is essentially quantum-mechanical, so that when trying to formulate general hypotheses about the nature of the physical world one should not rely on such notions. That has led to the idea that one should formulate one's definitions and hypotheses entirely in terms of macroscopic events ("clicks in counters", etc.). In the context of the EPR-Bell experiments, the most obvious and natural way of doing this is to formulate the notion of macroscopic counterfactual definiteness (MCFD). Probably the most succinct discussion of this idea (though not the name) was by the late A. Peres [11], and it has been extensively explored by Stapp and others (see e.g., [12]). To explain it, let us consider a single heralded photon incident on the apparatus shown in Figure 1b. Suppose, first, that it is switched into the measuring station $\boldsymbol{M_a}$. Then it is an experimental fact that either counter Y clicks (and counter N does not) or counter N clicks (and Y does not). Now consider the case that this photon was switched into measuring station $M_{a'}$ and thus does not enter M_a; thus the sequence of events which had it entering M_a is counterfactual. Consider now two possible statements which we might consider making about this counterfactual situation:

(i) Had it been the measuring station M_a which the (macroscopic) switch S_L activated, then it is a fact that **either** counter Y would have clicked (and not N) **or** counter N would have clicked (and not Y). This seems a rather plausible, indeed obvious, statement (cf. Section 4), since as we have seen, on every event where the switch actually worked in this way one (and only one) of these outcomes was realized.

(ii) Had it been the measuring station M_a which was activated by S_L, then **either** it is a fact that counter Y would have clicked **or** it is a fact that counter N would have clicked. In other words, the outcome of an experiment which we have not conducted is definite, a "fact about the world".

Despite their grammatical similarity, the two statements (i) and (ii) are not logically equivalent, and it is the second which is the assertion of macroscopic counterfactual definiteness. Thus, as an alternative to postulate (3a), we can formulate the third defining postulate of an objective local theory as

(3b) MCFD: Counterfactual statements about macroscopic events possess truth-values. Note that here again no explicit reference is made to determinism (but cf. Section 4 below)

Once given a definition of the class of objective local theories via postulates (1), (2), and either (3a) or (3b), it is straightforward to construct a proof of the CHSH theorem, which with our simplifications reads simply

For all possible choices of the xy-plane vectors $\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'$, the quantity defined by Equation (2.1) satisfies the inequality

$$K_{exp}\left(\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'\right) \leqq 2 \tag{3.2}$$

The proof I now give uses the definition with (3b) and is essentially just a more formal version of that given by Peres [11]: that using (3a) will be discussed in Section 4.

Step 1: By postulate (3b), for any given pair event i all four quantities A_i, A'_i, B_i and B'_i exist, whether or not they are measured on that event, and each take a value $\pm\mathbf{1}$.

Step 2: By postulate (1), all four products $(AB)_i, (A'B)_i, (AB')_i, (A'B')_i$ exist, with (e.g.,) the value of A_i in the product $(AB)_i$ **the same** as that in $(AB')_i$, and take values $\pm\mathbf{1}$.

Step 3: By simple exhaustion of the 2^4 = 16 possibilities, the quantity $K_i = (AB)_i + (A'B)_i + (AB')_i - (A\prime B\prime)_i$ exists for each pair event i and satisfies the inequality $K_i \leq 2$.

Step 4: Thus, provided the expectation values of the four quantities $AB, A'B, AB'$, and $A'B'$ are evaluated **on the same ensemble** of pair events (e.g., the total ensemble) the corresponding expectation value K_{ens} of K satisfies

$$K_{ens} \leqq 2,$$

Step 5: By postulate (2) (see the explication (ii)) following it), the statistical properties of the total ensemble of pair events are identical to those of each of the subensembles ab, ab', etc., and hence $\langle AB \rangle_{ens} = \langle AB \rangle_{ab}$.

Step 6: Thus the quantity $\langle AB \rangle_{ab} + \langle A'B \rangle_{a'b} + \langle AB' \rangle_{ab'} - \langle A'B' \rangle_{a'b'} \equiv K_{exp}\left(\boldsymbol{a}, \boldsymbol{a}', \boldsymbol{b}, \boldsymbol{b}'\right) \leq 2$, QED.

4. Can We Give MCFD a Meaning?

Like many of the concepts studied by philosophers, MCFD is an idea which we routinely apply in everyday life without a second thought. If I say, for example, "Had I woken up ten minutes earlier this morning, I would have caught the 8.0 bus", it does not normally occur to you to doubt that this is a meaningful statement. Similarly, if the prosecuting counsel says to the jury "Had the accused not pushed his victim down the stairs, she would be alive today", the jury may have to decide whether this is a true or a false statement (any juror who refused to answer on the grounds that s/he did not believe that counterfactuals have truth-values would, I imagine, be promptly disqualified!). Also, again like many concepts dear to philosophers, on closer examination it rather seems to disappear into thin air. Of course, there are some circumstances under which counterfactual statements about macroscopic events trivially possess truth values: for example, if for some reason the 8.0 bus did not run on the particular morning on which the statement is made, then the latter is trivially false. But what if the 8.0 bus did run as usual?

A famous "first-pass" attempt at defining the meaning of counterfactual statements is given by David K. Lewis as the first paragraph of his book "Counterfactuals" [13]:

"'If kangaroos had no tails, they would topple over' seems to me to mean something like this: 'in any possible state of affairs in which kangaroos have no tails, and which resemble our actual state of affairs as much as kangaroos having no tails permits it to, the kangaroos topple over'". This concise statement is followed, not unnaturally, by a whole chapter's worth of qualification. However, it is convenient to take it in its "bare" form as a template for the following discussion. Moreover, while as formulated it applies to generic statements, we may transpose it so as to apply to particular events (such as the (non-)catching of the bus) on a particular morning. With these qualifications the direct transposition to the EPR-Bell context, for an event in which (say) $M_{\boldsymbol{a}'}$ is in fact triggered,

the meaning of the statement (say) "Had it in fact been instead measuring station M_a which is triggered, the Y detector would have been triggered" is roughly

"In all possible worlds for which the switching is into M_a and all other relevant circumstances are the same, detector Y is triggered". (4.1)

and thus, abbreviating "world" by W and the class of possible worlds by PW, the statement of MCFD becomes "either for every W in PW, $A = +1$ or for every W in PW, $A = -1$" (by contrast with statement (i) of Section 3, which comes out as "for every W in PW, either $A = +1$ or $A = -1$" which looks trivially true. (We must of course use a "common-sense" definition of "relevant circumstances", e.g., exclude the possibility that a meteor has destroyed measuring station M_a while photon L is on its way there. Cf. also Section 6)).

Statement (4.1) needs a certain amount of unpacking. First, since we wish to maintain postulate (1) in the definition of an OLT, "relevant" has to include as a necessary component "local"; in particular, neither the choice at the distant measurement superstation R between M_b and $M_{b'}$ nor the output of the measurement at that station are to be counted among the "relevant" variables. Secondly, in the absence of theoretical preconception or further experiment we do not know whether or not (4.1) is true, since we do not know what are the "relevant" local considerations (remember, these can include information associated with the process of generation of the pair in question at the source, which is certainly not light-like separated from the detection events). Such information may include "hidden variables", or be more general in nature, e.g., the quantum description of the two-photon wave function. Below I shall follow the convention in the literature by denoting this general state description by λ, and for convenience only will assume it can be modelled by a one-dimensional real variable whose possible values fall between 0 and 1 and whose distribution is given by a density function ρ(λ). I do not commit myself at this point as to whether the possible values of λ are continuous or discrete; in the latter case I will denote the number of different possible values of λ by M. It will turn out (see below and the Appendix) that an important quantity, for any given model of the state description and any given set of experiments, is the ratio of N, the order of magnitude of the quantities N_{ab}, and M, the (order of magnitude of) the number of possible values of λ; below I will denote this ratio by C ("cardinality"), and note that in the case of continuous λ C is by definition zero. Needless to say, we do not know the value of C; however, this lack of understanding in no way affects the claim that for any given model of the photon state and any given set of experiments the concept is **meaningful**.

But the trickiest question relates to the meaning, in this proposed statement of MCFD, of the phrase "all possible worlds" (or "any possible state of affairs"). We could, no doubt, at this point follow the argument of Lewis's ch. 4 and explore the related concept of "similarity". However, the discrete character of the measured variables in the EPR-Bell experiments and possibly of the state description λ suggests an alternative tactic: We first note that, whatever else we may believe about its meaning, the class PW of all possible worlds with a specified property undoubtedly includes the class AW of all actually existing worlds with that property. In the present context, if we consider a particular pair of photons of which A' is measured on L (L switched into $M_{a'}$) the relevant class of actual worlds AW is the class of all (actually examined) pair events in which photon L is switched into M_a **and all other relevant variables are the same**. We note that since the "relevant variables" include the state variable λ describing the relevant photon pair, with this definition for continuous λ the class AW is empty. So let's specialize for the moment to the case of finite and reasonably small M (large C), and ask what would be the effect of replacing, in the definition of MCFD, the word "possible" by "actual". It is then clear that the statement in question is **meaningful**; but the interesting question now is

With "realism" in the definition of an OLT interpreted as MCFD and the latter redefined as indicated, can the CHSH theorem be proved? Explicitly, is the following statement,

when combined with OLT postulates (1) and (2), adequate to allow a proof of the CHSH inequality? (3.2)

$$\begin{gathered}\text{Either for every actually conducted experiment in which } A \text{ is} \\ \text{measured andin which the value of } \lambda \text{ is } \lambda_i, \text{ the measured value of } A \text{ is } +1, \\ \text{or for everyactually conducted experiment with those properties the measured} \\ \text{value of } A \text{ is } -1 \text{ (and similar statements with } A \text{ replaced by } A', B, B'\text{).}\end{gathered} \tag{4.2}$$

Rather surprisingly, at the time of writing I have not found this question explicitly posed or answered in the existing literature. The answer is in fact a qualified yes (the qualification is a standard "fair-sampling" one and N→infinity); the proof is given (as a special case of the more general one discussed in the following) in the Appendix.

However, this result is restricted to the case of large *C*. What of the case of *C* small or zero (continuous λ)? At this stage it is interesting to cast ourselves off from the question of MCFD which originally motivated this section, partition the whole set λ into 16 mutually exclusive and exhaustive subsets Λ_k, and replace in postulate (4.2) the words "is λ_i" by "lies within a given subset Λ_k". Can we still prove the CHSH inequality? Again, as shown in the Appenidx A, the answer is a qualified yes. Thus for any **deterministic** OLT the CHSH inequality (3.2) can be proved on the basis of hypotheses which are exclusively about **actually conducted experiments**. It should be noted that, while with our present state of knowledge (or ignorance) we obviously cannot determine the truth or falsity of either the restricted or the more general postulate, certain types of OLT leave open the possibility of future experiments which may be able to determine the value of λ for each actually conducted event, and thus resolve the question.

5. Is Locality Enough?

In ref. [7], Gisin makes the following two claims, which I reproduce verbatim:

(1) "Realism is nothing but a fancy word for determinism".
(2) "Bell's inequality can be stated and proved without any assumption about determinism".

In (2) presumably no harm is done by replacing "Bell's inequality" by the more general term "CHSH inequality"; and since in view of (1) we can replace in (2) the word "determinism" by "realism", and my objective in this essay is to examine the consequences of interpreting the latter term as equivalent to MCFD, in this section I will ask the question

"Can the CHSH inequality (3.2) be stated and proved from the assumption of locality alone, without any implicit assumption about either MCFD or some closely related concept?"

To examine this question, let us follow Gisin (who was himself following the original work of Bell's second paper [14] and of Clauser and Horne [8]) in supposing that the "state" of any given photon pair *i* is characterized by some parameter ('state of affairs") λ_i (which need not be a classical "hidden variable" in the usual sense—it could for example be a quantum state), and define "conditional probabilities" which in our notation are, given that switch L (R) directs photon 1 (2) into measuring apparatus $M_a(M_b)$, denoted by $p(A, B : \lambda)$; it is noteworthy that Gisin himself takes these "conditional probabilities" as not in need of any further definition or discussion. He then assumes (this is implied by the words of his p. 81, para. 3 rather than explicitly stated as an equation) that (in our notation) the joint probability of A and B over the ensemble of pairs is given by the formula

$$p(A, B) = \int d\lambda \rho(\lambda) p(A, B : \lambda) \tag{5.1}$$

So far everything is general. Gisin then states that

"the general assumption underlying all Bell's inequalities reads [his Equation (1)]", which in our notation reads (again given the appropriate switchings)

$$p(A, B, \lambda) = p(A, \lambda) p(B, \lambda) \tag{5.2}$$

which is (one) standard statement of the locality assumption. The proof that the conjunction of Equations (3) and (4) does in fact lead to the CHSH inequality is a matter of straightforward algebra and is given e.g., in ref. [8], section III and Appendix B; there seems no point in reproducing it here.

So, locality alone is necessary to establish the CHSH inequalities, and so the experiments force us to reject it, right?

Well, maybe not quite. Let's look a little more closely at (say) the quantity $p(A, \lambda)$. While, as we have noted, ref. [7] does not define this explicitly, ref. [8] does, in the last sentence of p. 527, col. 2, para. 1: "Let the probabilities of a count being triggered at 1 and 2 be $p_1(\lambda, a)$ and $p_2(\lambda, b)$ respectively..." [their $a \Rightarrow$ our A, etc.]. In the present context CH's footnote 11 is also crucial:

"even though we have introduced λ as the state of a specific single system, the assumed objectivity of the system described by this state allows us to consider an ensemble of these, physically identical to the extent that they are all characterized by the same λ. **The probabilities are to be associated with this ensemble**. Clearly, this procedure is conceptually sound, even in cases where we cannot in practice prepare the pure-**λ** ensemble." [emphasis supplied]

So what is the nature of this ensemble? In particular, is it real or counterfactual? Of course, as is widely recognized, this is a problem more generally with the concept of probability: if, for example, the weather service announces that there is a 60% chance of rain tomorrow, it is presumably legitimate for-a philosophically-minded listener to interpret this statement as equivalent to saying that if we consider the total of recorded days in which the temperature, humidity, cloud cover, etc, were approximately what they are today, then in 60% of these cases it rained on the next day. On the other hand, in everyday life we regularly sling around (perhaps illegitimately) the concept of probability in contexts where the relevant reference ensemble does not actually exist; e.g., we talk about the probability that the next election in [the imaginary country of] Euphoria will be won by the National Party, even though the circumstances of each Euphorian election are presumably sufficiently unique that no reference ensemble actually exists. (I am told that in some countries one can place bets on this kind of unique event.) So what is the situation in the EPR-Bell context?

This is where the previously perhaps under-appreciated quantity $C = N/M$ defined in Section 3 plays a role. As shown in the Appendix, for $C \gg 1$ the relevant reference ensemble needed for the proof of the CHSH inequality exists (which is of course not to say that we can prepare it!), while for $C \ll 1$, and in particular the case of continuous **λ**, it is intrinsically counterfactual. The implications of this state of affairs are briefly addressed in the Conclusion.

6. Independent Evidence against MCFD

This section is somewhat off the main theme of the paper. In Section 4 we saw that we can, if we wish, maintain a belief in macroscopic counterfactual definiteness (MCFD) even in the face of the most recent EPR-Bell experiments, provided that we are prepared to abandon the idea of locality. However, this raises the question: do we have reasons to believe or not in MCFD independently of the EPR-Bell experiments?

To introduce this question, let us consider one objection which may be raised against the discussion I have given in the previous sections of this paper, namely that I have assumed that a definite outcome of a particular trial in the EPR-Bell setup is realized as soon as there has been time for bells to ring, lights to flash etc.; in other words, for the state of some reasonably macroscopic system to register that of the photons (or other microscopic entity). It has been pointed out by various people[5], including the present author, that this assumption (sometimes called that of "collapse locality") is not self-evident and, were it to be violated in a given experiment, might allow the results to remain consistent with local realism. While in principle it should be possible to evade this "collapse locality loophole" by using ultra-long baselines, direct human observation etc., another fruitful direction may be to look at the postulate of collapse locality as just a special case of the more general

postulate of **macrorealism** [16], which we now define following the somewhat revised statement in ref. [17]. (For an extended discussion of the considerations of the next few paragraphs, see this reference, especially Section 6):

(1) Macrorealism per se: A macroscopic object which has available to it two or more macroscopically distinct states is at any given time in a definite one of those states. (6.1)

(2) Non-invasive measurability. It is possible in principle to determine which of these states the system is in without any effect on the state itself or on the subsequent system dynamics. (6.2)

(3) Induction. The properties of ensembles are determined exclusively by initial conditions (and in particular not by final conditions). (6.3)

It is clear that postulate 1 is (at least prima facie: but see below) just a rewording and generalization of the postulate (3b) of Section 3 above which in the EPR-Bell context we have called macroscopic counterfactual definiteness and abbreviated MCFD. Similarly, postulate 3 is essentially postulate 2 of Section 3. However, in the experimental situations which we shall be considering there is no question of spatial separations sufficient to bring in considerations of locality, so to obtain a conjunction of postulates which may be capable of experimental test we need to replace postulate 2 of Section 3 by a different kind of assumption, and this is what is done in (6.2) above. It is clear that some further qualifications are needed: for example, in postulate 1 we need to qualify the words "all times" by something like "except for brief transit times whose measure can be made to approach zero", and the term "macroscopic" needs some discussion; for these and other complications, see ref. [17].

It should be emphasized that a denial of the postulates of macrorealism as applied to the counters, etc., in an EPR-Bell experiment as postulated e.g., in ref. [15] (or for that matter to single-photon polarization experiments) is considerably stronger than a denial of the postulate of MCFD in the same context:since it is conceived to apply to actual (conducted) experiments, it should apply a fortiori to counterfactual ones, so that it is equivalent to a denial not only of statement (ii) of the comments on OLTs, in Section 3, but even of statement (i) of that section.

To test the conjunction of postulates (6.1-3) by which we have defined macrorealism (hereafter abbreviated MR), we need a different kind of ensemble from that employed in the EPR-Bell tests, namely a "time ensemble" of repeated measurements conducted on a single object which by our criteria is not only macroscopic but possesses two (or more) macroscopically distinct states; we will characterize these states by the values $Q = \pm 1$ of some appropriate macroscopic variable Q (a specific example is given below). Moreover, the theory whose predictions we wish to oppose to those of MR (typically QM) should predict a reasonable rate of transition between these two states as a function of time. The time variable then formally plays a role similar to that of the polarizer settings in the EPR-Bell case, and it is possible [16] to derive CHSH-like inequalities on the basis of (6.1-3). However, it turns out that there is actually a simpler protocol, based on the intermediate Equation(1) of ref. [16], which has no analog for the EPR-Bell case and is illustrated in Figure 3: namely, we repeatedly start the system at an initial time t_{in} from (say) state $Q = +1$ and let it evolve until a specified time t_{int}, at which we partition the total set of runs into two sets distinguished by the experimental protocol: in protocol 1 we do nothing at time t_{int}, while in protocol 2 we conduct (in principle!) an ideally noninvasive measurement of Q. We then wait until a final time t_f and measure the value of Q. The difference $\langle Q \rangle_1 - \langle Q \rangle_2$ between the final expectation value of Q on the two sets of runs is denoted $\delta\langle Q \rangle$. Any theory defined by the conjunction of postulates (6.1-3), i.e., any theory of the MR class, will predict that the statistics of the two runs are the same within standard statistical error and thus $\delta\langle Q \rangle = 0$ (whereas theories not embodying MR, such as QM, will in general predict a nontrivial difference between them, so that $\delta\langle Q \rangle \neq 0$). The proof is a straightforward consequence of the definitions.

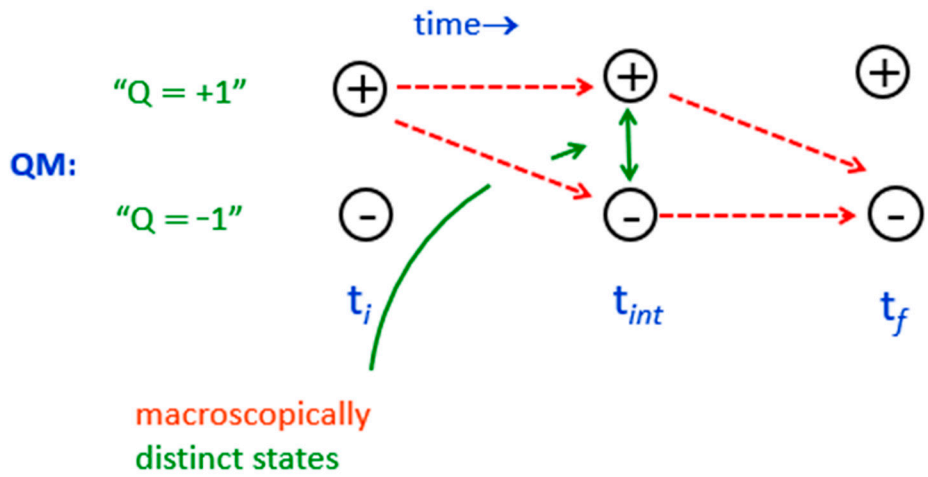

Figure 3. MACROSCOPIC QUANTUM COHERENCE (MQC).

There is one major loophole [17,18] in this protocol which has no analog in the EPR-Bell case, namely that we cannot be sure *a priori* that on a given run of set 2 our measurement of Q at the intermediate time t_{int} is genuinely noninvasive (this is sometimes called [18] the "clumsiness loophole"). Should this be so, then a nonzero value of $\delta\langle Q\rangle$ need not indicate a breakdown of MR. A way to partially eliminate this loophole [17,18] is to conduct an auxiliary set of experiments in which t_{int} is chosen so that Q_{int} is known (from prior measurements) to be certainly (say) +l, and then to compare two sets of runs 1 and 2 as above; if the result for $\delta\langle Q\rangle_f$, which is a measure of the "clumsiness" of the measurement technique employed, is denoted $\delta\langle Q\rangle_{f,cl}$, then this needs to be subtracted from the $\delta\langle Q\rangle_f$ measured in the main experiment, i.e., the significant figure is

$$\delta\langle Q\rangle_{f,true} = \delta\langle Q\rangle_f - \delta\langle Q\rangle_{f,cl} \tag{6.1}$$

If a nonzero value of this quantity which exceeds the standard statistical error is obtained in the experiment, we may regard this as evidence for a breakdown of MR. It is worth noting that this protocol cannot exclude a scenario in which MR is satisfied but the measurement at t_{int} is noninvasive only when the state at that time is definitely $Q = +1$ (or when it is definitely -1); however, this hypothesis seems so contorted as to water down the notion of MR beyond recognition.

The only existing test along these lines known to me which uses states which can reasonably be called "macroscopically distinct"[6] is that of Knee et al. [20]. The system used in this experiment is a so-called "flux qubit", namely a superconducting ring interrupted by a Josephson junction: see Figure 4. For this system the macroscopic variable Q corresponds to the current circulating in the ring (or the total flux trapped through it), and differs for the two states by $\sim$ 170 μA; for a discussion of whether or not this separation can legitimately count as a "macroscopic" difference, see Supplementary Note 3 of ref. [20] and the literature cited therein. The result of the experiment was that $\delta\langle Q\rangle_{f,true}$ not only agrees with the predictions of QM but differs from zero by more than 80 standard deviations, thus contradicting (within the assumptions of the experiment) the hypothesis of MR and thus of MCFD at the level of flux qubits.

One amusing point about the flux-qubit system used in this experiment is that devices of this type were originally developed in the 60's of the last century as sensitive magnetometers, and at the time the current (or flux) degree of freedom was thought of as a purely classical variable. Nowadays the sensitivity is such that complexes of as few as 300 electron spins scan be detected, and it seems probable that in the not so distant future single·spins will be within reach. Additionally, to convert the qubit from a "quantum object" (as it shows up in protocol 1) to effectively a "classical detector" (which requires measurement of the flux/current as in protocol 2), all that is necessary is a tiny (apparently ~ 0.1 (mW)$^{1/2}$

for 12 ns) pulse applied to a so-called Josephson bifurcation amplifier coupled to the system. Such is the infamous "quantum-classical boundary"!

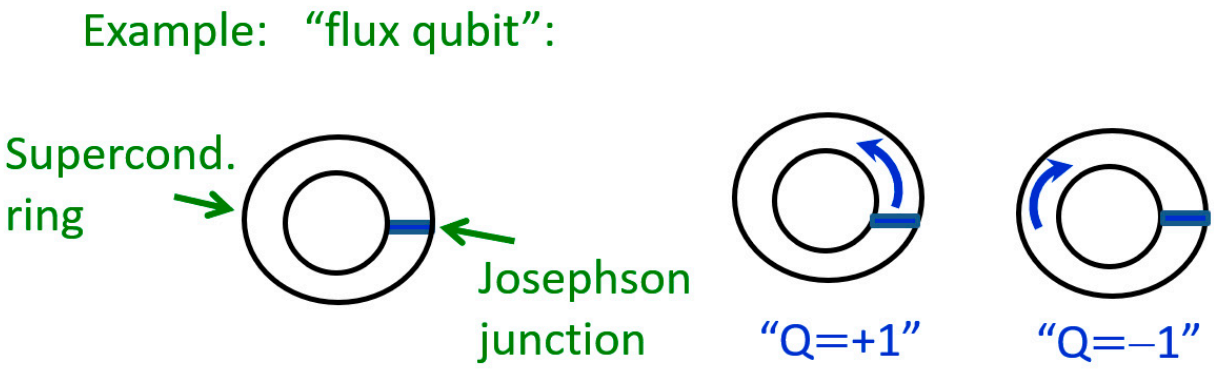

Figure 4. A "flux qubit", the system used in ref. [20] for a test of MR (idealized).

However, an important objection to the consequences suggested for MCFD remains to be discussed. Assuming that the experimental data of ref. [20] and the theoretical interpretation given in this section are accepted, we can conclude that MR is falsified in the context of (appropriately prepared) flux qubits. Does it follow that it is false also for the counters, etc., used in photon detection, including the EPR-Bell experiments? There are important differences between the two cases. In the first place, at first sight the variable in the former case is a single one-dimensional variable (the flux), whereas the counter has a macroscopic number of different degrees of freedom. I would argue that this is not in itself an important distinction, since in the former case not only does the difference between the two "macroscopically distinct" states in question involve an arguably macroscopic number of Cooper pairs (see Supplementary Note 3 to ref. [20]), but the flux is strongly coupled to other degrees of freedom such as phonons, nuclear spins etc. (The reason this result does not destroy the QM predictions is that this coupling is overwhelmingly adiabatic in nature).

A more worrying distinction is that the behavior of the counters, unlike that of the flux qubit, involves a macroscopic degree of irreversibility due to coupling of the "principal" degree of freedom to a (macroscopic) environment, or (non-adiabatically) to a myriad of internal degrees of freedom. (as does the amplification of the tiny "measuring" pulse in the latter case). It is a well-known theme of the quantum measurement literature that the resulting decoherence effects the transition from a quantum to a classical description. Elsewhere (e.g., in ref. [17]) I have argued that within the hypothesis that QM is the whole truth about the physical world this "solution" to the measurement (or "realization") problem is unviable. However, since we have determined in the context of this essay to pretend we have no knowledge of QM, and cannot necessarily assume that an extension of our chosen OLT to the macroscopic world contains any analog of the notion of decoherence, in the context of this section the point would appear to be moot.

7. Conclusions

What, if anything, is new in this paper? First, let me emphasize that I have given no reason at all to question the "standard" wisdom that all OLTs, as defined for example in ref. [8], make experimental predictions which satisfy the CHSH inequality (3.2). However, I have shown (Section 4) that in the definition of a **deterministic** OLT (i.e., one satisfying MCFD) the assumption of locality can be formulated entirely in terms of **actually conducted** experiments. Secondly, in the case of a **stochastic** OLT, I have shown (Section 5) that this is also true in the case of a subclass of OLT's and associated experiments, namely under the condition that the "cardinality" index C defined in Section 3 is $\gg 1$. However, I have been unable to extend the proof to the experimentally interesting case that $C \ll 1$, and while it cannot be excluded that a more sophisticated argument may be able to do so, my instinct is to doubt it. Thus, it looks as if we have to conclude that for a general stochastic

OLT the concept of probability needs to be taken either as a primitive (as in ref. [7]), or defined, as is done in footnote 11 of ref. [8], in terms of, an ensemble which not only cannot be "prepared", but does not actually even exist.

What are we doing when we implicitly appeal to such an ensemble? In effect, if we attempt to extend Lewis's analysis to this case, we are postulating the existence of a whole set of "possible worlds" which are similar to the world of our real experiment and in which the occurrence of (e.g.,) the value +1 of $A(\lambda)$ behaves according to the probability $p(A, \lambda)$. In other words, we may not be postulating "macroscopic counterfactual definiteness", but we are certainly postulating "macroscopic counterfactual stochasticity"! Whether or not this state of affairs is held to be of any significance will no doubt depend on the reader's pre-existing conception of the notion of probablility.

Finally, I have noted that while in the standard view of the EPR-Bell experiments one can, by accepting nonlocality, avoid having to exclude MCFD as a property of the physical world, the experiments described in Section 6 may constitute evidence for the falsity of the latter postulate. Thus, at the end of the day the simplest attitude which is consistent with all the experimental evidence to date may be that **both locality and realism** are false!

I would like to acknowledge the benefit I have derived from recent correspondence with Karl Hess and Juergen Jakumeit, who have independently emphasized the importance of discussing actually conducted experiments and in particular the role of "cardinality" (cf. ref. [21]), even though at least at the time of writing we seem to disagree about the conclusions to be drawn from this.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing is not applicable.

Conflicts of Interest: The author declares no conflict of interest.

Appendix

In this appendix I discuss what extra assumptions (besides locality and non-retrocausality) we need to make **about actually conducted experiments** in order to prove the CHSH inequality or some variant of it. The notation is as in the main text. I consider separately the deterministic and stochastic cases:

A. Deterministic case:

For orientation we first discuss the case of a small number M of discrete values of the state description variable λ_i, with moreover the (unrealistic) assumption that the number $N_{kl}(i)$ of events in which $\lambda = \lambda_i$ and k and l are measured is independent of $k = a, a'$ and $l = b, b'$. Then the situation is clear: for each possible value of $\boldsymbol{\lambda_i}$ the experimentally measured value of each of the quantities A, A', B, B' (call it A_i, etc.) is the same for all pairs with this λ_i and equal to either +1 or -1. It is then immediately obvious that with the above assumption the contribution K_i of K from this value of i satisfies the inequality

$$K_i \equiv A_i B_i + A'_i B_i + A_i B_i - A'_i B'_i \leqq 2 \tag{A1}$$

so summing over i, and dividing by $N_{kl} = N$, we obtain the CHSH inequality (Equation (3.2) of the main text). Note that this conclusion is independent of both N and M (and of the distribution $\rho(\lambda_i)$ of λ_i).

Next, we relax the assumption that for given λ_i the quantity $N_{kl}(i)$ the number of events for which a given λ is realized and the measurement is of k and l, is independent of k and l. The ratio $C \equiv N/M$ is now crucial, since this is the order of magnitude of $N_{kl}(i)$ and the fluctuations of this quantity are thus of order $C^{1/2}$. Now, while the inequality (A1) is still valid, the subsequent steps in the argument of the last paragraph are not. However, we can proceed as follows: Write $\delta \boldsymbol{N}_{kl}(i) = N_{kl}(i) - N(i)$ where N_i is the total number of

pairs with $\lambda = \lambda_i$ and as we have seen is of order N/M. Then it is clear that when we sum over the pair events in class i, the extra contribution δK_i to the LHS from the terms $\delta \boldsymbol{N}_{kl}(i)$ is of order $C^{1/2}$. Then, recalling that the **total** number of pairs switched into the (k,l) channel is independent of[7] k and l and given simply by $N/4$, the CHSH inequality is modified to

$$K \leqq 2 + N^{-1} \sum_i \delta K_i \tag{A2}$$

Let us try to estimate the order of magnitude of the second expression on the RHS of Formula (A2). Since each of the $N_{kl}(i)$ fluctuates by $O\left(\sqrt{N/M}\right)$, the rms value of δK_i is similarly of this order. Hence the second term on the RHS of (A2) is of order $M^{1/2}N^{-1}$ times this, i.e., of order $N^{-1/2}$ independently of M.

However, the above argument is implicitly based on the assumption that $N_i \sim N/M$ is large. What if it is not (and in particular if λ is continuous so that N_i is not defined)? We can generalize the above argument to this (still deterministic) case by partitioning the set λ into 16 mutually exclusive and exhaustive subsets each corresponding to a different choice of at least one of the variables A, A', B, B'. The above argument then goes through, with i now interpreted as labelling not individual values of λ but a particular one of the 16 subsets, and the output is the same. Thus we reach the conclusion that for this (deterministic) case the CHSH theorem is always valid in the "thermodynamic limit" $N \to \infty$, and this can be shown using only assumptions about **actually measured** quantities. While this conclusion may seem unsurprising, we wish to contrast it with what happens in the "stochastic" case discussed next.

B. Stochastic case:

We suppose now that the quantities A, A' etc. are not uniquely determined functions of the state description λ, but all we have to go on are the probabilities $\boldsymbol{p}(\boldsymbol{A}|\boldsymbol{\lambda})$ (etc.). Of course, if we treat these probabilities as primitive, as Gisin does, and invoke the locality assumption Equation (5.2), then we can prove the CHSH inequality (3.2) along the lines of ref. [8]. However, it is convenient for our purpose to recast the proof along the lines of the discussion A above. We first consider the discrete case, and as there, we partition the whole set λ into 16 subsets, corresponding now to the 16 possibilities for the 4 probabilities p_j for $k = A, A', B, B'$ to be > or <1/2. While it seems intuitively plausible that the ensuing argument can be carried out for a general distribution of p_j within each subset, for the sake of a concise argument I will at this point make the simplifying assumption that within each subset $p_j(\lambda)$ has a unique value, which is hereafter referred to simply as p. Then we can consider a particular one i of the subsets and a quadruple of measurement on this subset, one each of the four possibilities k, l, and ask for the statistics of the contribution K_q made to K from this quadruple. If we consider for example the subensemble with all p's > 1/2, it is a matter of simple algebra to show that the probability of finding $K_q = 4$ (the only allowed value greater than 2) is given by

$$p\left(K_q = 4\right) = 2p(1-p)\left(p^2 + (1-p)^2\right) \tag{A3}$$

and a more tedious but straightforward calculation yields for the average value

$$\langle K \rangle = 2 - 8p \cdot (1-p) \tag{A4}$$

Since the maximum value of the expression (A3) is 27/256, and is reached at $p = (1/2) \cdot \left(1 + \left((1/2)^{1/2}\right)\right. \sim= 0.85$, we may use this value of p to estimate a maximum contribution to δK from this one of the 16 subsets: since from Equation (A4) $\langle K \rangle$ is 1, it follows that to obtain a value of δK greater than some positive number x one needs the fraction of $f(x)$ of quadruples within this subemesemble to satisfy the inequality

$$f(x) \geqq (27 + 229x)/795 \tag{A5}$$

(note that even in the limit $x \to 0$ the RHS is still of order (1/4), a result which is perhaps intuitively plausible).

Evidently the probability of such a result in an overall ensemble of size N is proportional to

$$C_{Nn}(27/256)^{nf(x)}(229/256)^{N-nf(x)} \tag{A6}$$

where C_{Nn} is the binomial coefficient and is exponentially small in the limit $N \to \infty$. We should be able to give a similar though more complicated argument for the contribution to δK from the other 15 subensembles (in this case p becomes in general a 4-dimensional variable), and thus, assuming this is right, we unsurprisingly recover the result of ref. [8], that even in the stochastic case the CHSH inequality is satisfied in the presence of locality and the thermodynamic limit.

However, while this result is reassuring, it does not absolve us from answering the question: what is the ensemble we are implicitly referencing when we talk about the conditional probability $p(A/\lambda)$? In the limit of $C \to \infty$ this ensemble can be treated as all the actual events with this λ (and the given k, l), and since this number is equal to C, this ensemble is adequate for our purposes. What of the case $C \to 0$ (including continuous λ)? Here, if we use the argument of the last paragraph, it is tempting to argue that since by definition the different λ falling within the same subset give the same (probabilistic) results for any measurable quantity, it should be legitimate to treat them as effectively having the same λ, so that we recover the discrete case with $M = 16$ and thus (for $N \gg 16$) $C \gg 1$. However, it is clear that this argument fails the moment that we relax the simplifying assumption that p is constant within any given one of the 16. Thus, at least within the context of the argument of this appendix, we are forced to conclude that we cannot escape the use of "counterfactual" ensembles—ensembles which include not just the experiments that we have actually done, but those which we might have done. The implications of this state of affairs are briefly discussed in Sections 5 and 7.

Notes

1 In the following, I will use the term "EPR-Bell experiments" as a catch-all phrase for all the experimental work over the last 60 years which has been motivated by Bell's and related work.

2 At least for present purposes, we do not need to consider the apparent "splitting" which photons undergo e.g., in a Mach-Zehnder interferometer.

3 Circular and even elliptical polarization is also possible, but does not need to be discussed here.

4 Or "inequalities"; since the original CHSH paper [6] proves several inequalities, the plural is often used.

5 An extended early discussion is ref. [15].

6 To avoid possible confusion, it should be noted that in the beautiful EPR-Bell test of Storz et al. [19] using superconducting qubits, the two qubit states involved are the lowest two states of a transmon and thus are not macroscopically distinct.

7 Any errors introduced by violation of this assumption are manifestly of order $N^{-1/2}$.

References

1. Bell, J.S. On the Einstein Podolsky Rosen Paradox. *Physics* **1964**, *1*, 195–200. Available online: https://journals.aps.org/ppf/pdf/10.1103/PhysicsPhysiqueFizika.1.195 (accessed on 12 August 2024). [CrossRef]
2. Freedman, S.J.; Clauser, J.F. Experimental Test of Local Hidden-Variable Theories. *Phys. Rev. Lett.* **1972**, *28*, 938. [CrossRef]
3. Hensen, B.; Bernien, H.; Dréau, A.E.; Reiserer, A.; Kalb, N.; Blok, M.S.; Ruitenberg, J.; Vermeulen, R.F.L.; Schouten, R.N.; Abellán, C.; et al. Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres. *Nature* **2015**, *526*, 682. [CrossRef] [PubMed]
4. Shalm, L.K.; Meyer-Scott, E.; Christensen, B.G.; Bierhorst, P.L.; Wayne, M.A.; Hamel, D.; Stevens, M.J.; Gerrits, T.; Glancy, S.C.; Allman, M.S.; et al. Strong Loophole-Free Test of Local Realism. *Phys. Rev. Lett.* **2015**, *115*, 250402. [CrossRef] [PubMed]
5. Giustina, M.; Versteegh, M.A.M.; Wengerowsky, S.; Handsteiner, J.; Hochrainer, A.; Phelan, K.; Steinlechner, F.; Kofler, J.; Larsson, J.A.; Abellán, C.; et al. Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons. *Phys. Rev. Lett.* **2015**, *115*, 250401. [CrossRef]
6. Clauser, J.F.; Horne, M.A.; Shimony, A.; Holt, R.A. Proposed Experiment to Test Local Hidden-Variable Theories. *Phys. Rev. Lett.* **1969**, *23*, 880. [CrossRef]
7. Gisin, N. Non-realism: Deep Thought or a Soft Option? *Found. Phys.* **2012**, *42*, 80–85. [CrossRef]
8. Clauser, J.F.; Horne, M.A. Experimental consequences of objective local theories. *Phys. Rev. D* **1974**, *10*, 526. [CrossRef]

9. Eberhard, P.H. Background level and counter efficiencies required for a loophole-free Einstein-Podolsky-Rosen experiment. *Phys. Rev. A* **1993**, *47*, R747. [CrossRef] [PubMed]
10. "Scientific Realism" in Stanford Encyclopedia of Philosophy. Available online: https://plato.stanford.edu/entries/scientific-realism/ (accessed on 12 August 2024).
11. Peres, A. Unperformed experiments have no results. *Am. J. Phys.* **1978**, *46*, 745–747. [CrossRef]
12. Stapp, H.P. Meaning of counterfactual statements in quantum physics. *Am. J. Phys.* **1998**, *66*, 924–926. [CrossRef]
13. Lewis, D.K. *Counterfactuals*; Harvard University Press: Cambridge, UK, 1973. [CrossRef]
14. Bell, J.S. The theory of local beables. *Epistemol. Lett.* 1976. Available online: https://cds.cern.ch/record/980036/files/197508125.pdf (accessed on 12 August 2024).
15. Franson, J.D. Bell's theorem and delayed determinism. *Phys. Rev. D* **1985**, *31*, 2529. [CrossRef] [PubMed]
16. Leggett, A.J.; Garg, A. Quantum mechanics versus macroscopic realism: Is the flux there when nobody looks? *Phys. Rev. Lett.* **1985**, *54*, 857. [CrossRef]
17. Leggett, A.J. Testing the limits of quantum mechanics: Motivation, state of play, prospects. *J. Phys. Condens. Matter* **2002**, *14*, R415. [CrossRef]
18. Wilde, M.M.; Mizel, A. Addressing the Clumsiness Loophole in a Leggett-Garg Test of Macrorealism. *Found. Phys.* **2012**, *42*, 256. [CrossRef]
19. Storz, S.; Schär, J.; Kulikov, A.; Magnard, P.; Kurpiers, P.; Lütolf, J.; Walter, T.; Copetudo, A.; Reuer, K.; Akin, A.; et al. Loophole-free Bell inequality violation with superconducting circuits. *Nature* **2023**, *617*, 265. [CrossRef] [PubMed]
20. Knee, G.C.; Kakuyanagi, K.; Yeh, M.C.; Matsuzaki, Y.; Toida, H.; Yamaguchi, H.; Saito, S.; Leggett, A.J.; Munro, W.J. A strict experimental test of macroscopic realism in a superconducting flux qubit. *Nat. Commun.* **2016**, *7*, 12353. [CrossRef] [PubMed]
21. Jakumeit, J.; Hess, K. Breaking a Combinatorial Symmetry Resolves the Paradox of Einstein-Podolsky-Rosen and Bell. *Symmetry* **2024**, *16*, 255. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.