Entanglement re-examined: If Bell got it wrong, then maybe Einstein had it right

Antony R. Crofts

Department of Biochemistry and Center for Biophysics and Quantitative Biology
University of Illinois at Urbana-Champaign, Urbana IL 61801

Correspondence:
A.R. Crofts
Department of Biochemistry
University of Illinois at Urbana-Champaign
417 Roger Adams Lab
600 S. Mathews Ave
Urbana, IL 61801
Phone: (217) 333-2043
Email: crofts@illinois.edu
Bell’s treatment of the entanglement question and the outcome of subsequent experiments to test his hypothesis are generally taken as demonstrating that at the quantum level nature is non-local (NL). From standard quantum mechanical (QM) predicates, when coincidences are counted as a function of polarizer settings measuring vectors at separate stations, tests using photon pairs are expected to show the same amplitude (full “visibility”) sinusoidal curves independent of orientation of the photon frame. This NL behavior, claimed to demonstrate non-locality, cannot be expected from local realistic (LR) treatments. Experimental confirmation of this difference is presented as one of the major achievements of 20th century physics because it constrains all subsequent theoretical discussion of process at this level to a non-local perspective.

In this paper, I argue that this edifice is built on sand. Einstein, Podolsky and Rosen (EPR) started the entanglement debate by framing it through a model, necessarily local, in which quantum entities have intrinsic properties (“elements of reality”). Since with stochastic populations outcomes would not depend on vectors, Bell suggested an LR model in which vectorial properties had been replaced by the scalar sign of the spin. This model generated a zigzag outcome rather than the sinusoid expected from NL models. In Part A of the paper, I explore through discrete math simulation, a vectorial model (vLR) following Einstein’s expectations. The simulation sets up a population of photon pairs with defined polarization and shows in typical experimental tests that when the photon and polarizer frames are aligned, the outcome is the same as in the NL model. Bell’s claim that his inequalities excluded such an outcome cannot apply to this case. Analysis of the discrepancy shows that no model that excludes vectors can be realistic. Experimental tests based on the limits of ≤2 from this model failed to find the expected LR outcome not because nature is necessarily non-local, but because the model was wrong.

Nevertheless, the full visibility on rotation of frames expected from the predicates of the NL treatment, cannot be explained by the vLR model. In Part B of the paper, I show that the NL predicates are poorly justified and not well supported by experiment. Three predicates lead to the NL expectation: (i) that the uncertainty principle requires evolution to a measurement context in a superposition of states; (ii) that the pairs are in a dichotomy of \(H\) and \(V\) spin states that determines their behavior at discriminators; (iii) that the matrix operations describes a real process through which pairs are actualized with vectors aligned in frame with the reference polarizer. In the first predicate, the superposition bears no relation to the state studied in recent experimental protocols to test Bell’s theorem using well-determined PDC sources. It is a poor starting point
because the uncertainty argument, originally applied to conjugate properties of a single electron, has no logical justification when applied to the two photons of a pair, interrogated through refraction and detected separately. Correlations are represented in the superposition by spin states in an ontic dichotomy, but electrons are fundamental particles and are all the same, so such a dichotomy cannot be intrinsic. Nevertheless, the dichotomic spin states are essential to the treatment. In the wavefunction, they have no vectorial properties, but in the matrix operations of the third predicate, the spin states are invested with vectorial agency and become aligned with the polarizer frame, the only state vectorially defined. This alignment generates an ordered from a disordered state, in contravention with the second law. Because stations are space-like separated, the simultaneous actualization implemented in the operation must also contravene relativistic constraints. It is claimed that the NL outcome is clearly demonstrated experimentally, but descriptions of state preparation are ambiguous, and the ambiguities allow alternative interpretations that, with trivial adjustment to protocols, support LR explanations that are simpler and more plausible than the conventional NL ones. All outcomes report can be simulated and explained in terms of the LR state modeled.

The difficulties for the NL treatment might be ameliorated by later theoretical developments in the wavy realm. Dirac developed quantum electrodynamics (QED) by co-opting Maxwell’s electromagnetic wave treatment to the interaction of photons with electrons through their fields, with the potential for action at a distance. Some theoretical sophistication is required in applying the wavy treatment to process at the quantum level because photons are neutral. Feynman showed that Maxwell’s approach could be used to calculate phase delays arising from differences in path length using his path integral approach. Photons travelled different paths considered as separately quantized. In effect, interference effects then deselect paths more convoluted. However, quantized interactions necessarily involve processes at that level, and neutral entities cannot act through fields. A more natural approach is to treat photonic interactions in terms of momentum transfer. Work can be applied either in generation of a photon by displacement of an electron, or in displacement of an electron by a photon, both constrained by the two properties that can be measured, frequency and orientation. In effect, this momentum-based scenario is that implemented in my simulation. The simpler treatment brings new insights into processes on scales from the quantum to the cosmic, which might provide an alternative path forward.
INTRODUCTION

Einstein, Podolsky and Rosen\textsuperscript{1} started the entanglement debate by introducing the so-called “EPR-paradox” (Einstein’s most cited paper\textsuperscript{2}). This crystallized earlier arguments\textsuperscript{3-7} with Bohr over the foundations of quantum mechanics (QM), and focused the debate through consideration of how correlations between two quantum objects should be represented when they are generated through a common transition. The EPR model was of a pair of entities, correlated through a common cause, which then separated to space-like distance before measurement. EPR pointed out that if the correlated properties were intrinsic to each entity (“elements of reality”), local measurement of one would allow the observer to infer a property of the other, allowing direct determination of pairwise correlations in a population. The predicates of the Copenhagen interpretation precluded such a picture\textsuperscript{8-12}. Measurements at the quantum scale introduce uncertainties arising from quantized energy exchanges, requiring superposition of entangled pairs in indeterminate orientation. The math has then to deal with that non-local (NL) state, and with the resolution of discrete entities with real properties from a population conjoined over space-like distance on measurement. These features require a sophisticated math, unavailable from the local realistic (LR) nature of Einstein’s model\textsuperscript{11}. In the common wavefunction describing correlations in superposition, partners are represented by spin states (the phase difference) in ontic dichotomy. Vectorial properties are initially indeterminate, but entities with real vectors are actualized on measurement, modelled in operations of the Pauli matrices. These align the photon (or electron) frame with the discriminator frame. Measurement of an object at one location then appears to cause simultaneous actualization of a partner with complementary properties at a distant other. The math allows this, but in reality it would require transfer of energy and/or information faster than light, - that “spooky action at a distance” Einstein so disliked\textsuperscript{13,14}. On the other hand, the LR nature of Einstein’s model limited the treatment of the measurement context to standard probability theory\textsuperscript{11} which cannot account for the NL expectations, and this has been widely seen as excluding such LR models from further consideration.

Some thirty years later, in “probably the most celebrated result in the whole of twentieth-century physics”\textsuperscript{15,16}, Bell\textsuperscript{17,18} explored a model suggested by Bohm and Aharonov\textsuperscript{9}. They treated a population of electron pairs generated with opposite spin in which vectors were probed by interaction with photons, and showed that in this vectorial model, coincidence counts from an LR treatment would be half those predicted by QM. Bell suggested an alternative model, based on an analysis that showed that with a stochastic population of pairs, terms for vectorial properties cancelled from equations leading to outcome expectations. In his LR model, vectorial properties were therefore replaced by the sign of the spin. Expectations from this treatment were compared with those from an NL
model of the same population in superposition and shown to be different. Specifically, the LR treatment generated a linear zigzag outcome curve as a function of angle difference, while the NL model generated a sinusoidal curve. The difference opened the possibility of experimental tests. Later treatments extended this approach to populations of photons pairs correlated in polarization but with similar expectations\textsuperscript{19-21}, and subsequent experiments, mostly with photons, all showing the NL expectations and demonstrating ‘violations of Bell-type inequalities’, have provided the main underpinning for the non-local picture\textsuperscript{4,20-23}. The resulting world-picture, with properties of entangled quantum states shared and dispersed though space-time, has excited physicists, puzzled philosophers, inspired sci-fi lovers, and emboldened mystics over the last half-century\textsuperscript{24}. Much of the excitement comes from the prospect of new physics, because, as is widely recognized\textsuperscript{5,9,11,19,25-33}, current explanations imply an observer-dependent non-local reality governed by laws that allow information-sharing over space-like distance forbidden under classical constraints\textsuperscript{11,34,35}.

In the first part of this commentary, I demonstrate by simulation of an alternative model using populations of correlated vectorial photon pairs (vLR, a vectorial application of Einstein’s model), that contrary to expectations of Bell’s theorem\textsuperscript{7-9,11,13,20}, an LR model can give outcome curves that match those of the NL treatment applied to the same initial state. I show that this behavior is natural and fully compliant with locality constraints, and then examine why the LR model and treatment that Bell chose\textsuperscript{17} have been taken to excluded such outcomes.

Bell’s treatment of electrons extended earlier work\textsuperscript{9,11} and included new insights. Importantly, the differences in expectations were easier to test than with previous efforts, but the results supported earlier models in excluding local realism. The model was later adapted by Clauser and colleagues (CHSH) to cover experiments with photon pairs\textsuperscript{21}. They also provided an important more general treatment that extended the LR limits from Bell’s treatment to all LR models\textsuperscript{20,21}. With the strong case from the earlier papers, these extensions cemented the non-local view, now widely accepted. Against this, I show that both Bell’s analysis and the generalization from the CHSH extension involved choice of an unrealistic model and an invalid treatment of the limit for all LR models. The conclusions have prompted a re-examination of the claim that violations of Bell-type inequalities exclude local realism.

In the second part, I address the questions now opened as to whether the experimental results claimed to support the non-local interpretation really do, and whether the orthodox NL approach adequately explains the results claimed.

I frame my argument in terms of the recognition by Furry\textsuperscript{11} that LR models are constrained by standard probability theory\textsuperscript{36}, which cannot deal with models based on states in superposition.
This same conclusion emerges from my model and simulation. Although this is commonly taken to exclude local realism, I suggest a different perspective, - that we agree that the two models are irreconcilable and ask how to test them on their own merits. Protocols based on parametric down-conversion (PDC) sources are ideal for testing Einstein’s view because the determinate nature of the photon source conforms to the model tested, which can also be explored in my simulation. Unfortunately, this determinate state is then in direct contradiction with the state in superposition. Furthermore, the NL outcome claimed requires generation of an ordered state from a disordered state, not an insignificant problem.

Since Heisenberg uncertainties demand such a state, and the superposition demands the sophisticate math of the standard Copenhagen interpretation, I question the justification for the conventional treatment, and demonstrate that many experimental outcomes claimed to support non-locality come up against the problem above, and/or with relativistic constraints. However, I show that the natural properties of the vLR model can account for experimental results without any contradiction. In all protocols examined, trivial adjustments to specific configurations allowed, within ambiguities in description of protocols, a simulation that demonstrated the experimental result claimed.

I discuss extensions of the NL treatment by co-option of Maxwell’s electromagnetic theory to quantum field approaches, which might ameliorate the problems of the standard approach. I suggest that these cannot be practically applied at the elemental level because photons are neutral and cannot interact through fields. An alternative treatment that represents photons as carriers of momentum has been commonly invoked when appropriate and can better represent the quantized actions of photons in processes at that level, but then requires a particulate representation that is both local and realistic.

**PART A. WHERE BELL WENT WRONG**

The locality constraints introduced by EPR came from their assignment of properties to discrete entities (cf.9,22); if properties are intrinsic, information pertinent to measurement can only be accessed locally, measurement on one of a pair can have no effect on the distant other. This constraints the mathematical treatment11 as noted above, which then cannot account for the QM outcome expected, taken as excluding all LR models. In contrast, in NL treatments, Heisenberg uncertainties require evolution of entities in a superposition of all possible states, vectorial properties are neither intrinsic nor discrete, and can become known only when real entities are actualized on measurement. Then, for entangled partners, measurement of one partner of a pair simultaneously actualizes that and the distant other in appropriate correlation. Bell was well aware of the antecedent work 9,11,12,37. His own treatment extended that by consideration of a novel LR model which generated differences in
the properties of outcome curves more easily tested experimentally, expressed through the ‘inequalities’ \(^{17,19,38,39}\), which also supported the earlier consensus excluding LR models.

Bell’s original treatment\(^{17}\) considered electrons with opposite spin\(^9\), determined on discrimination using Stern-Gerlach magnets. The treatment was refined and extended to photons and polarizers by Clauser, Horne, Shimony and Holt\(^{20,21}\) (CHSH), and later adopted by Bell, leading to the Bell/CHSH (BCHSH\(^{19,22}\)) consensus that inspired the early experiments\(^{23,40-43}\). An atomic beam (Ca or Hg) was excited by arc discharge\(^{23,43}\), or, in later work, by lasers\(^{40-42,44}\), and pairs of photons with correlated orientation were generated in a two-step decay process (cascade). The pairs evolved to separate measurement stations for analysis using polarizers set to differentiate between expectations. In more recent tests, lasers have been used to excite PDC in non-linear crystals\(^{45-49}\), using setups like that in Fig. 1. These latter sources have also been exploited in more complex arrangements following similar protocols, to explore applications in quantum encryption, teleportation and computation, and to test the finer points (loopholes) arising from the entanglement debate\(^{22,48,50-54}\).

1. The BCHSH consensus.

Before discussing where Bell went wrong, it is necessary to understand both the historical context sketched out above, and his logic. A summary of the treatment for photons is provided in the SI (Section 1). I also discuss, mainly in the second part, how Bell’s thinking was determined by his earlier analysis of hidden variables. Even those familiar with this field might benefit from a re-cap, so here, a brief synopsis will serve to frame the argument.

As already noted, Bell was following a number of formative earlier treatments of the EPR paradox. Most importantly, Furry\(^{11}\) had pointed out soon after publication of the EPR paper that, in representing the measurement process, their choice of discrete entities with intrinsic properties enforced an LR treatment that constrained probabilities for coincidence between stations to the product of probabilities at separate stations. In contrast, the Copenhagen interpretation required a superposition of states evolving to a measurement context at separate stations where actualization of real entities from the superposition had to be handled jointly. Mathematical treatment requires representation of vectors in all dimension in the superposition, dealt with in Hilbert space. Since superposition excluded assignment of vectorial properties, correlations had to be represented by spin states; however, photons with the real vectorial properties tested by discriminators then had to be actualized, a process modeled in operation of the Pauli matrices. Since no LR state could generate the outcome expected from the Copenhagen formalism, Furry concluded that participation of such states had to be excluded. Building on these insights, Bohm and Aharonov\(^9\) considered how a stochastic LR population
would be expected to behave under the different regimes. Their treatment used a model in which electron pairs correlated by generation from a common source were interrogated by interaction with photons to test vectorial properties and showed that the outcome of the LR treatment would be of a reduced amplitude compared to the QM treatment. In his seminal paper, Bell\textsuperscript{17} built on this perspective by adopting the Bohm-Aharonov model in considering electrons pairs, but suggested that spin states could be tested by Stern-Gerlach magnets to provide the framework for the NL model he presented. However, he explored an alternative LR treatment. He noted that with a stochastic source, vectorial properties in the photon population did not appear in the equations determining the outcome and suggested a treatment in which the photon vectors were replaced by the sign of the spin. Later Clauser and colleagues (CHSH)\textsuperscript{20,21} adapted this treatment to photon pairs. These provided a more tractable medium for experimental work as shown in earlier pioneering experiments\textsuperscript{43}, but could be discussed through essentially the same framework to give a consensus approach, - the BCHSH treatment. CHSH also extended the treatment to provide a more general expression of LR constraints, and Bell used both features in his later treatments of photon pairs\textsuperscript{19,39,55,56}. Bell’s model provided a sharper difference in outcome than the previous treatments, the shape of the outcome curve; a zigzag for the LR model compared the sinusoidal outcome expected from QM. This difference was expressed in terms of inequalities which provided experimental targets.

To proceed further, we need to examine the BCHSH treatments and early experiments\textsuperscript{19-22}. These considered photons with spin states designated by \(H\) (horizontal) or \(V\) (vertical) in pairs correlated in Bell-states \(HH/\overline{VV}\) or \(HV/\overline{VH}\), and measured coincidence in detection at two stations, tested at four settings of the polarizers. A setup suitable for discussion of later developments is shown in Fig. 1, which is similar to that in a seminal paper introducing use of PDC sources\textsuperscript{45}. With \(\alpha\) or \(\alpha'\) and \(\beta\) or \(\beta'\) as polarizer settings respectively at stations 1 and 2, the approach considers possible coincidences at settings \(\alpha, \beta; \alpha', \beta; \alpha, \beta'\); and \(\alpha', \beta'\), chosen to discriminate between outcomes expected, - the LR zigzag and the NL sinusoid. The expectations from coincidences in pairwise measurements at separate stations, \((E_{a,\beta}, \text{etc.})\)\textsuperscript{22}, then give outcomes summed through:

\[
S_{\text{BCHSH}} = E_{a,\beta} + E_{a',\beta'} + E_{a,\beta'} - E_{a',\beta} \quad (\text{eq. 1a}),
\]

where \(E\) values reflect expectations from normalized coincidence yields at each of the four settings. At the elemental level, a value of +1 was assigned to a photon detected in the ordinary ray, and of -1 if in the extraordinary ray, and each of the four \(E_{a,\beta}, \text{etc.}\) terms can have a value ±1. The sum \(S_{\text{BCHSH}}\) is then constrained to the range ±2 so that \(S_{\text{BCHSH}} \leq 2\) (eq. 1b). This LR outcome was compared with the expectations from the NL treatment.
Bell’s NL approach with electron pairs was conventional\textsuperscript{9,11}, summarized in his equation,
\[ (\mathbf{s}_1 \cdot \mathbf{\alpha} \mathbf{s}_2 \cdot \mathbf{\beta}) = -\mathbf{\alpha} \cdot \mathbf{\beta} = -\cos \sigma \text{ (eq. 2)}, \]
(describing the projections from the matrix operations. In the BCHSH consensus, essentially the same QM approach was applied to photon pairs\textsuperscript{9,21}. Using photon pairs and standard QM principles, the projections of the matrix operation generate expectation of a sinusoidal curve of coincidence as a function of angle difference, \( \sigma \), between polarizers at different stations, with an amplitude of \( \pm 2 \), invariant on rotation of the photon and polarizer frames. The full visibility, invariant on rotation of the photon frame, depends on the dichotomic correlation in spin states in the superposition, and actualization as photons with real vectors in alignment with the polarizer frame on operation of the matrices as above. Then, the projections lead to yields at the polarizers given by Malus’ law, the asymmetry of the matrix operations leads to cancellation of contributions from the photon vectors, and coincidences are dependent only on angle difference, \( \sigma \). Fig. 2 (see legend for explanation) shows a unit circle representation of such projections, leading to the observable Malus’ law values at particular settings of the variable polarizer.

The BCHSH LR treatments giving equations 1a) and b) have provided the framework for later debate, and on this basis, values for \( S_{\text{BCHSH}} \) then depend on estimation of values for the \( E_{\alpha,\beta} \), etc. terms through the conventional application of the probabilistic principles pointed out by Furry\textsuperscript{11}. The treatment as adapted for photons\textsuperscript{19,20,39,55} first asks what would be expected from local measurement on a stochastic population. This starts with an elemental probability, \( p_1(\lambda, \alpha) \), derived from the vectors of the photon polarization, \( \lambda \), and of the polarizer, \( \alpha \). The outcome with a population is then given by integration,
\[ p_1(\lambda, \alpha) = \int \rho(\lambda)p_1(\lambda, \alpha) d\lambda \text{ (eq. 3a)}, \]
where the function \( \rho(\lambda) \) is a normalized probability distribution for \( \lambda \) over the circle intercepted. With the stochastic population considered, terms in \( \lambda \) would disappear, because on integration over allowed orientations vectorial contributions would cancel through \[ \int \rho(\lambda) d\lambda = 1 \text{ (eq. 3b)} \] (the isotropic condition). As a consequence, no information about \( \lambda \) would be available from measurement on the population at either station (any kid experimenting with polarizers could confirm this).

LR expectation of coincidences could be obtained by similar integration using the product of elemental probabilities at the two stations, - the LR constraint from Furry\textsuperscript{11}. This is represented in the first equation on the right. With polarizers set at \( \alpha \) (station 1) or \( \beta \) (station 2):
\[ E_{\alpha,\beta(O)} = \int \rho(\lambda)p_{12}(\lambda, \alpha)(\lambda, \beta) d\lambda = \int \rho(\lambda)p_{12}(\lambda, \alpha, \beta) d\lambda = p_{1,2}(\alpha, \beta) \text{ (eq. 4)}, \]
with similar equations for other polarizer settings.

In equating the first and the second equations on the right (RHE1 and RHE2), Bell recognized that, as implicit in RHE2 and the isotropic condition of a stochastic source, eq. 3b would come
into play, so that terms in $\lambda$ would also cancel here. Then, $p_{1,2}(\alpha, \beta)$, and $E_{a,\beta}(LR)$, etc., and hence $S_{\text{BCHSH}}$, would depend only on polarizer orientations to give the rotationally invariant outcome of reduced amplitude. This was anticipated by Bohm and Aharonov and is demonstrated in my simulation (Figs. 2 and 3B). Bell noted that the outcome appeared to be independent of the photon vectors, and as an alternative to the vectorial model, he developed an LR model in which the photon correlations were represented by the scalar sign of the spin instead of by the polarization vectors. His model has been likened to an “exploding penny”, in which the explosion separates the head side from the tail side, with the two parts hurtling apart in opposite directions. The equation he suggested,

$$E_{a,\beta(LR)} = p_{1,2}(\alpha, \beta) = -1 + 2\sigma/\pi \quad (eq. 5),$$

yielded a linear dependence of expectations on $\sigma$ in the range $\pm 1$, giving a zigzag curve as a function of angle difference, invariant on rotation of frames. From the four settings expected from eq. 1a, this gave an amplitude in the range $\pm 2$ for $S_{LR}$. Since curves from the two models were different, they predicted different outcomes from coincidence measurements at canonical angle differences along the curves, giving “inequalities” expressed as limits of either $S_{NL} \leq 2.83$ from the NL sinusoid, or $S_{LR} \leq 2$ from the LR zigzag.

As first suggested by CHSH, the same limiting value can alternatively be derived generically for any LR model from substitution of elemental counts of $\pm 1$ in eq. 1a; for the four terms, this gives $S_{\text{CHSH}}$ as $\pm 2 \quad (eq. 1b)$. Since the mean from integration over a population cannot exceed the maximal elemental value, the same limit of $S_{LR} \leq 2$ for all LR models follows from this approach. These are the first and second Bell-type inequalities. Despite spectacular technical and theoretical advances, this treatment continues to provide the consensus framework for discussion of the inequalities in the entanglement debate. In this framework, a “violation of Bell’s inequalities” means that the result of an entanglement test significantly exceeds the $\leq 2$ limit expected under the LR model either from the zigzag or the CHSH elemental counts.

Bell’s third inequality came from an LR model based on a stochastic population retaining vectorial properties, for which the expectation value was also $\leq 2$. In effect, this was similar enough to the Bohm-Aharonov model, that Bell did not expand on the earlier treatment or comment further. All three of these models conformed to LR limitations and Bell’s LR limits $\leq 2$. However, since my focus in this part of the paper is on the limits commonly used to discriminate, - those derived from Bell’s zigzag and the CHSH treatment, - I will defer further discussed of the third case till later (Part B 1).
2. The critical experimental results.

Louisa Gilder\textsuperscript{4} has a nice account of developments in the years following the first results\textsuperscript{23,43}, and reports an early consensus based on outcomes testing readily distinguished features; the difference between Bell’s zigzag and the NL sinusoidal curves, and the invariance of the sinusoidal outcome on rotation of the polarizer frame\textsuperscript{20,23,40,43}. With cascade sources, technical difficulties and inefficiency of detectors (at <20% efficiency, <4% of coincidences could be detected), required normalization of coincidence counts, and correction for “accidental” counts\textsuperscript{23}. The ‘inequalities’ were represented in the $\delta$-function, which scored the difference between coincidences measured at canonical angle differences, and that expected from the LR limit of $\leq 2^{23,40,42}$.

In terms of the above criteria, the NL case wins hands-down, because the LR model was shown to fail; with one exception (cf.\textsuperscript{59}, but later disavowed), early tests\textsuperscript{23,40,41,43,44} showed the rotational invariance expected from either model, but the sinusoidal curves expected from NL rather than the LR zigzag. After normalization and correction, amplitudes were interpreted as showing the full visibility expected from NL models\textsuperscript{4,19,20,23,29,40,57}, and these results steered the debate towards a consensus supporting Bell’s theorem\textsuperscript{58}. Subsequent experiments of increasing sophistication\textsuperscript{45,48-50,60-62} using PDC sources, have refined results so as to approach closer and closer to the NL expectation of 2.83, with the ‘violation of Bell-type inequalities’ scored through standard deviations from the LR expectation of $\leq 2$ (cf.\textsuperscript{40,42,45,54,63}). In recent work with high-efficiency detectors (~91%), outcomes measured have exceed the LR limit of $\leq 2$ (or equivalent) without any special corrections\textsuperscript{64-66}. Based on such results, the quantum theorists have concluded that no explanation in LR terms could be tenable. For Bell\textsuperscript{17}, “…the quantum mechanical expectation value cannot be represented, either accurately or arbitrarily closely…” by any equation in the form of \textit{eq}. 4. Since in all reports the results matched the NL sinusoid, they could only be accounted for in LR terms by invoking “hidden variables”, requiring “conspiracies” between source and stations to explain the vectorial outcome, a view summarized in “…if [a hidden variable theory] is local it will not agree with quantum mechanics, and if it agrees with quantum mechanics it will not be local…” (from Bell\textsuperscript{57}, as qualified by Shalm et al.\textsuperscript{65}). Various scenarios (“loopholes”\textsuperscript{22}), most notably detection, freedom of choice, and communication loopholes, are discussed through which such conspiracies could be enabled. Experiments, some of spectacular sophistication (cf.\textsuperscript{64-66,68-70}), have successfully eliminated such loopholes, apparently
solidifying the NL case. Shimony’s conclusion\textsuperscript{22} that “the \textit{prima facie} nonlocality of Quantum Mechanics will remain a permanent part of our physical world view, in spite of its apparent tension with Relativistic locality” is widely accepted (cf. \textsuperscript{39}).

3. \textit{Simulation of a vectorial LR model (vLR) shows that the BCHSH LR model is inadequate}

The BCHSH algebra seems impeccable, but \textit{something} is not right. Prediction of the sinusoidal outcome of the NL treatment depends on the behavior of pairs actualized with real vectors on discrimination by polarizers (the Malus’ law behavior). In the derivation of Bell’s LR zigzag above, vectorial information about the photon source had been excluded, so Malus’ law was never invoked. His model could never generate sinusoidal curves because the vectors needed at the polarizers had been replaced by scalar spin states.

In order to better understand what factors might be important, typical tests\textsuperscript{45,48,62} using photon pairs (Fig. 1) have been modeled through discrete math simulation. The initial intent was to address the naïve question implicit in the above paragraph: \textit{What would be the outcome if a vectorial LR population of correlated photon pairs was analyzed using Malus’ law compliant polarizers?} The short answer is – \textit{the sinusoidal outcome expected from NL predicates} (Figs. 2, 3A).

\textit{(i) Simulation program.}

The core of the program is simple. For each ‘experiment’ a population of ‘photon’ pairs is generated with each partner defined by an explicit polarization vector appropriate to program settings for source, Bell-state, correlation, etc. (the Make Light subroutine). Uncertainties are implemented where appropriate (value of photon angle for each pair in a stochastic population, variable polarizer settings, allocation of photons to beams, etc.) by use of a random number generator. Orientations of the polarizers are set after the photon population has been established (“with photons in flight”), but before discrimination, measurement and plotting of coincidences and singles counts (the Measurement and Plot Point subroutines). The Measurement subroutine includes a For…Next loop that, in the default (Malus LR) mode, implements at the elemental level the Malus’ law distribution to stations, and, through invocation of the Plot Points subroutine, the integration of \textit{eq. 4}. This is achieved by discrete summation, using the same code for all photon populations. Since \textit{eq. 4} includes the fundamental statistical constraint on an LR treatments recognized earlier\textsuperscript{11}, my implementation should
generate all features implicit in that treatment. The most important difference from Bell’s scalar dichotomic model is that the photons of the vLR population measured have the intrinsic vectors Einstein would have expected.

As in real experiments, yields at the ‘polarization analyzers’ are derived from discrete measurement of elemental outcomes; a photon appears in either the ordinary or extraordinary ray of a polarization analyzer. Discrimination at the polarizers is assumed to follow Malus’ law \( I/I_0 = \cos^2 \phi \), with \( I/I_0 \), the normalized transmission, \( \theta \) the orientation of the polarizer, \( \lambda \) the polarization vector of the photon, and \( \phi = \theta - \lambda \). At the elemental level, Malus’ law applies statistically; the yield calculated is a probability for transmission corresponding to the BCHSH term, \( p_1(\lambda, \alpha) \) above. The sampling of the statistical spread is implemented in default mode by incrementing counts at detectors (\( Q \) or \( R \) at station 1, \( S \) or \( T \) at station 2, Fig. 1), based on comparison of the normalized Malus’ law yield to a random number between 0 and 1; if the yield is greater, the photon goes to the ordinary, else to the extraordinary ray. The distribution gives the single counts, which within the statistical limits, are the same at all four detectors, as expected. There are two exceptions to this distribution rule; in simulation of Bell’s LR model, and in mimicking the QM expectations. Simulation of Bell’s LR model is by comparing the calculated yield to 0.5 to give a binary outcome (see below); simulation of the QM expectations is by implementing the distributions implicit in the outcome of the matrix operations as summarized by Shimony \(^{22} \) (see Part B, Section 7). In the default mode, for each pair, coincidences between stations are scored, with the mean plotted to give points on the coincidence curve. Scoring of coincidences can be selected from a choice among four different algorithms. With the two exceptions noted, all these components are fully compliant with strict locality constraints.

At any particular setting of controls, a point on the outcome curve is determined from a population of pairs by a simple count of pairwise coincidences. The menu bar (Fig. 2, top) includes different Run options that implement a complete set of experiments generating an outcome curve. Run 1 generates a different photon population for each experiment (for each point plotted) and is the default. Run 2 uses the same population for each set of experiments (for all points on the curve plotted). Run 3 uses successive Run 1 experiments and takes a mean so as to improve signal to noise ratio. These Run options iterate through three subroutines (Make Light, Measurement, Plot Points in the Menu bar) that implement the functions suggested, to generate a point on the curve.

The user can choose from several different algorithms for the count. These differ in the assignment of elemental values (count increment, true/false, ±1) appropriate to the different algorithms (anti-correlation count, Boolean coincidence count, CHSH count, respectively). An emulation of the
Freedman and Clause (FC) δ-count, can also be chosen (see Program Notes for implementation). Whichever is used, the outcome is, as appropriate to the settings, essentially independent of the algorithm.

Useful features of the program are diagnostic aids. Elemental counts for the last population tested, and vectors relevant to measurement of each pair, can be displayed pair-by-pair in the Gadgets (top right, Fig. 3A). As can be seen there, counts at the elemental level follow the complementary symmetry expected, seen experimentally, and discussed at length by Mermin. The pairwise counts generated at a particular setting of the polarizers show elemental values, either +2 or -2 for the CHSH count, 0 or 4 (for the anti-correlation) or 4 or 0 (for Boolean counts), which accumulate when integrated in a population to give points on a curve plotted as a function of polarizer angle difference, with the Boolean count 90° out of phase from the anti-correlation count.

The ranges above are shown in the Gadgets (+2 or 0-4), but might lead to some confusion, because, on a single run, the fixed polarizer is set at a particular value, and the range of the count could reflect only two (for example α, β and α, β') of the four $E$ terms contributing to $S_{BCHSH}$ (a range ±1 or 0–2). The range shown in the Gadgets comes from scoring both detection and non-detection, which is practicable with perfect detectors, and facilitates taking cross-products, but counts each coincidence twice (half of the coincidences are redundant because a non-detection always mirrors a detection). However, after normalization to the two photons of a pair, the range at a particular setting (±1 for the CHSH count, etc.) is the same as the conventional scoring, and this is the default for outcome curves shown in the display panels.

In each of the counting algorithms, the summation over a population represents an integration in the form of eq. 4, in which RHE1 is considered as the expectation. The outcome is determined solely from the elemental counts. The program functions at the level of observables, defined by the use of Malus’ law probabilities. No algebraic sophistication is involved in the simulation, - refractive components are treated naïvely in terms of their empirical behavior, and only linear polarization is considered. Elementary trigonometry is used, but in only two contexts; (i) to determine from Malus’ law what yields would be expected at the settings used, so as to implement the statistics above; and (ii) to calculate Malus’ law compliant theoretical curves. Links to the executable program (Bell_Ineq_VB8.exe), source codes (in Microsoft Visual Basic 12), and Program Notes (also accessible as Help in the program), are available at http://www.life.illinois.edu/crofts/Bell_Ineq/
(ii) Program outcome. The correlations observed depend on the model for the ‘photon’ population, on how the ‘polarizers’ are chosen to respond (Fig. 3 A-E), and on choice of settings. When the polarizers are set to implement Malus’ law, photon pairs are oriented in a common frame, and share an angle with the fixed polarizer, coincidence counts from a \( v \)LR population follow the full-amplitude sinusoidal curve expected from the NL treatment (Fig. 3A, LR0 curve), not the linear zigzag expected from Bell’s LR model. Compared to the coincidence count, the anti-correlation count is 90\(^\circ\) out of phase, and the CHSH count gives essentially the same curve (Fig. 3C) but offset because of the choice of ±1 for elemental values. The curves show the full amplitude (‘visibility’) expected from the NL treatment. This LR0 outcome applies at all angles for the common frame (Fig. 3 A, C, E), - a classical rotational invariance.

When the photon pair orientation is random (isotropic in the plane of measurement), and the polarizer function is natural, the outcome is a sinusoidal curve, which shows the invariant behavior expected by NL on rotation between the photon and polarizer frames, but with amplitude half that expected from NL treatments (Fig. 3B). Such behavior (LR2 curve) is fully compliant with locality constraints and earlier analysis\(^9,11,72,73\); the rotational invariance reflects the isotropic source, and the reduced amplitude represents the entropic penalty on measurement of a stochastic source. As discussed at greater length in Part B 1, this outcome also corresponds to expectations from the Bohm-Aharonov-Bell vectorial treatment\(^9\) (the third inequality).

With the photon source randomized and the polarizers set to Bell binary mode (equivalent in effect to Bell’s scalar choice), the correlations follow the linear zigzag expected from his LR model\(^19\), with a partitioning proportional to \( \sigma \) (Fig. 3D, LR1 curve). This outcome is also invariant to rotation of frames. However, in contrast to the LR2 outcome, it shows full amplitude. Note that the photon population of the simulation differs from Bell’s original model in retaining its vectorial character, so the outcome here depends on forcing the polarizers to behave (unnaturally) as binary rather than Malus’ law discriminators (see above for implementation).

Outcomes from other combinations are unremarkable, though not without interest. The user can set parameters for modifications to the photon state on insertion of additional refractory elements into one or the other path, or both. Different coincidence counts give the result appropriate to the phase relationships tested, with different Bell states showing the expected behavior. However, with oriented photon pairs, when the polarizer frame is rotated away from alignment with the photon frame (or vice versa), the amplitude of the sinusoidal curve, as expected, decreases, and is zero when the rotation is by 45\(^{\circ}\) (Figs. 2, 3E). This is in contrast with NL expectations, where full-amplitude is
expected at any orientation of the photon frame, including stochastic. The NL outcome can be generated by implementing a “QM simulation” option to match the effective outcome summarized by Shimony\textsuperscript{22}. As discussed later in Part B, this, in effect, implements a Maxwellian demon that generates all the expectations of the NL treatment.

For settings at which Bell-correlated pairs were analyzed, the local mean yields (the ‘singles-counts’) at the four outcome rays have the same property, - a mean probability of 0.5 (within the statistical limitations of the sample), independent of photon source or polarizer settings (right-panels in Figs 3A, B, D). This is what Bell expected and is the natural behavior. An option (PDC_EO photon mode) is provided allowing one or other of the separate beams output from a PDC source to be directed to one or other of the two stations. Since HWP0 inserted as shown in Fig. 1 would convert the signal and idler beams from $HV$ to $VH$, this allows simulation of the two separate PDC processes of the Sagnac interferometric source introduced recently (see SI section 3 2 b) and Figs. SI_1A-D).

Since in this set-up, polarized beams are sent to the two stations, the singles-counts at the 4 detectors reflect the polarization. However, the normalized sum of counts from the four stations still gives the value of 0.5.

An Analog option allows display in the right panel of the fractional yield differences given by Malus’ law (the sinusoidal curves in Fig. 3C, right panel, explained in the legend). The theoretical curves in the left panel can be derived from the analog distributions in several ways, but the points plotted always reflect the statistical outcome from counts of pairwise coincidences at the two stations. The full visibility theoretical curves expected from the NL treatment (in effect the LR0 curves) are displayed in the left panel when frames are aligned to match that expected under NL predicates.

In summary, the outcome shown in Figs. 3A and 3C are the full-amplitude sinusoidal curves expected under NL predicates (cf.\textsuperscript{29,49,50,57}), generally presented as impossible from any LR treatment, but which here demonstrate a violation of Bell’s inequalities from a vectorially correlated population of photon pairs, fully consistent with LR constraints.

4. What can we learn from the simulation?

Local realistic models are often presented as classically constrained and are contrasted with NL models following QM predicates. However, in the argument between Einstein and Bohr both models were QM inspired, though interpreted from different perspectives. A critical difference lay in the statistical constraints pointed out by Furry\textsuperscript{11} and discussed above. The behaviors extracted in my vLR simulation follow those constraints, which are essentially classical but are quantized as expected
by Einstein; no model following his expectations would be antithetical to the fundamental QM principle. However, the Copenhagen interpretation required that QM treatments start with states in superposition, leaving the measurement math to sort out the conjoined fates of the entangled pairs. Furry\textsuperscript{11} was correct in pointing out that the mathematic options opened in the QM treatment were inaccessi-
ble to LR states, and it is this ‘failure’ that underlies the general conclusion that Einstein’s model must be ruled out.

With a stochastic source my simulation generates the outcome expected from the con-
straints\textsuperscript{11}. This was anticipated by Bohm and Aharonov\textsuperscript{9} in their half-amplitude outcome. Neverthe-
less, the simulation also finds an outcome that Bell thought to be impossible; it generates from an LR model the full-visib-
ility sinusoid curve expected by QM, albeit under restricted conditions. Previous
analyses have likely failed to find this result because the photon state is ordered and vectorial, and no such sources were available in the formative days. However, all recent experiments have used
sources generated by PDC, where the pairs are clearly determinate and ordered, as appropriate to
tests of the vLR model. That the model does generate the same result as QM means that the Furry’s distinction from the measurement math\textsuperscript{11} is not the only one that needs to be considered. Since both
in my simulation and in the NL treatment, on measurement the photon source is aligned with the po-
larizer frame, another significant difference is in how that alignment is achieved.

The ‘problem’ for the vLR treatment is that it is constrained by classical conservation laws.
The NL treatment apparently is not, because it generates an ordered from a disordered (stochastic, or
misaligned) state without an input of work. It is this difference, rather than constraints from the math,
which determines that my model cannot match the expectations of QM. In the simulation the math
conforms to the LR constraints, and the physics is represented realistically; all interrogative interac-
tions (refractive discrimination at polarizers, half-wave plates (HWPs) etc.) involve discrete photons
and full commitment of their action. In effect, for entities with intrinsic properties, all exchanges are
quantized and local, but all behaviors from the simulation (except the QM simulation and Bell binary
options) follow empirically justified laws (see Part B for discussion).

The factorizability of cross-probabilities\textsuperscript{11,22,74} (eq. 4, RHE1) is generally considered as char-
acteristic of valid implementations of the locality constraints\textsuperscript{23}. The LR0 outcome shows, contrary to
conventional expectations\textsuperscript{11}, the same full amplitude sinusoidal curve as the orthodox NL treatment.
Such an outcome should not come as a surprise. For any particular setting of the fixed polarizer,
there must always be an ordered vectorial LR population (readily available using PDC) that generates
the same curve as NL, - that in which the photon pairs of the population are correlated by orthogonal
orientation, and the reference frame is aligned with the polarizer frame, - the configuration expected
under NL predicates on actualization at the time of measurement. The outcomes are then the same because the vector projections are the same (see Fig. 2). As long as the photon vectors are represented, the only condition needed for prediction of such an outcome is the alignment of frames. As also noted above, an important conclusion from this is that the difference between treatments must lie in how alignment is achieved; it is implicit in the NL matrix operations but has to be explicit under vLR constraints. With this proviso, since the curves are the same, the differences in yields at canonical values (~2.83) ‘violate Bell-type inequalities’ just as does the NL value. In this limited case, the conclusion that no LR outcome can match the NL expectation value must then be invalid; in particular, the LR limit of ≤2 must be artificial. Since the summation of outcomes giving the LR0 curves is equivalent to integration in the form of eq. 4 (the RHE1), Bell’s conclusion that “…the quantum mechanical expectations cannot be represented…” in that form is also clearly wrong for this case.

I make no claim for originality in introducing the vLR model; it is simply Einstein’s perspective applied in a vectorial context. Several previously published efforts (cf.72,75-82) have arrived at similar conclusions, most explicitly in Thompson’s work72,77,83. The earliest of these by Angelidis75, a protégé of Popper84, was dismissed by Garg and Leggett58, essentially in terms of the limit of ≤2 from the second (BCHSH) inequality (eq. 1 above). Their brief paper was selected by the editors as representative of a much wider community that responded similarly. This consensus reflected a confidence in the conclusion from earlier influences and from Bell’s theorem, reinforced by its apparent validation in contemporary experiments23,40-42,44. The rejection was perhaps understandable in a historical context, reflecting wide acceptance of the Copenhagen interpretation and the mathematic constraints9-12. Similar dismissals of all later claims have been justified by the same rationale. However, from the above, perhaps confidence in this dismissal was misplaced.

5. Three mistakes by Bell and two additional ones by CHSH

Note that in Bell’s analysis summarized above, the value of $S_{BCHSH}$ depends only on the angle difference between polarizers, $\sigma$. The vectorial properties of the photons in the source population were not included because, with the stochastic source he was treating, they appeared to have no role in determining the outcome. With the vectors omitted, their engagement would have to be treated as involving “hidden variables”38. With or without vectors for the photons, discussion was constrained to the Copenhagen framework inspired by Bohr, Heisenberg, Born, Dirac, and von Neumann, and by the probabilistic constraints11 accepted both by Bohm and by Bell, - a mental box that seems to have effectively precluded consideration of Einstein’s model. I discuss Bell’s contribution as involving some mistakes below, but in the context of the prevailing opinion he was blameless; perhaps he’d
just been dealt a losing hand.

a) **Local measurements.** For stochastic sources, as Bell pointed out, the mean local yields would necessarily be independent of polarizer orientation because of the isotropic nature of the source. All early treatments involved such sources; the outcome was measured in the $R_1/R_0$, etc., terms of Freedman and Clauser, and their equation was also used in later reports using Ca-cascade sources, and by Fry and Thompson with a $^{200}$Hg-cascade excited by laser. It is simulated in the singles-counts of 0.5 (see right panels of Figs. 2A, B, D), and is diagnostic of an isotropic condition in the plane of measurement. Bell’s **first mistake** was to draw the wrong conclusion from this behavior. I cannot know his thoughts, but since the vectors cancel, he seems to have inferred that such properties would not be involved in determining the behavior seen and could therefore be excluded from analysis.

b) **The singles-count at each station are accounted for by natural behavior at the quantum level.** I have called Bell’s inference that vectorial properties would not contribute to processes determining the above behavior mistake because it flies in the face of the Malus’ law behavior demonstrated in 200 years of experimental work exploring polarization. Although values for individual photon vectors are lost in the mean, the behavior observed locally must access them at the elemental level. It depends on what happens at the polarizers, where the behavior requires vectors for both photon and polarizer. The experimental outcome can then be explained naturally in terms of local vectorial properties. For elemental measurements, $I/I_0$ (the Malus’ law expectation, see Section 3, (i)) gives the BCHSH probability, $p_1(\lambda, \alpha)$, and on integration over a population at $\lambda$, the Malus’ law yield. With a stochastic population, and a polarizer at any setting $\theta$, sampling $\lambda$ by integration over the hemisphere (as in eqs. 3a, b) would show a distribution of values for $\cos^2\varphi$ varying with $\lambda$, centered at the polarizer vector, $\theta$, and with the mean yield of 0.5. Since with an isotropic source the same curve is found at all values of $\theta$, this accounts in terms of local properties for the behavior Bell took as demanding exclusion of such properties. There are no “hidden variables” in this treatment, so their invocation in further discussion would not be useful.

When using an oriented source of photon pairs with dichotomic distribution of spin states represented by $H$ and $V$ in each population (as in Bell states $HV/VH$ or $HH/VV$ from PDC), the mean yield in any ray is given by $0.5(\cos^2(\theta - \lambda) + \cos^2(\theta - (\lambda + 90^\circ))) = 0.5(\cos^2\theta + \sin^2\theta) = 0.5$ (eq.
where $\theta$ is the orientation of the polarizer and $\lambda$ that of the photon reference frame (the $H$ photon)\textsuperscript{57,69}. The partition of $H$ and $V$ photons is isotropic because symmetrical about the reference axis, and gives in the mean, the same yield as Bell’s integration for a stochastic source.

Within the above constraints the outcome is therefore independent of source model, or specific values for $\theta$ and $\lambda$. A similar mix could be expected from either a $\nu$LR or a NL state, so this outcome is of little interest in distinguishing between models. Nevertheless, the vectors determine the outcome.

c) **Expectations from comparison between stations.** As noted above, with an isotropic source, no vectorial correlations between stations could be predicted from the mean yields from local measurements because all information that would allow comparison of each photon to its partner is lost in the mean. Experimentally, correlations are determined from pairwise comparison of elemental measurement outcomes at separate stations. This selection is important because with pairs in stochastic orientation, correlations are conserved only on a pairwise basis. With cascade sources, the protocol was designed to select pairs by temporal coincidence, by use of color filters to select pairs based on the different energies expected from the cascade, and on opposite directions of flight. Coincidences were maximal when polarizers were aligned, demonstrating that both photons of a pair had close to the same orientation; since these last two properties are expected from conservation of angular momentum in the source process, the protocol was predicated on determinate properties. Experimentally, correlations were either detected on-the-fly by coincidence counters or determined from data recorded and time-tagged on-the-fly and analyzed later. Further analysis requires the pairwise data, their time of measurement, knowledge of polarizer settings, etc., but all this information is exchanged subluminally\textsuperscript{68,69}.

Different approaches providing justification for the LR limit of $\leq 2$ were outlined above:

(i) **Bell’s derivation from the zigzag.** In deriving the zigzag\textsuperscript{17}, Bell noted that, in light of RHE2, on applying eq. 4 to a stochastic source, and on the integration through eq. 3b, the contributions from vectorial properties cancelled. When these were replaced by the sign of the spin, the partition to different hemispheres could only be scalar (eq. 5), equivalent in effect to the binary discrimination giving my LR1 curve. Peres\textsuperscript{57} suggests that “Bell’s theorem is not a property of quantum theory. It applies to any physical system with dichotomic variables, whose values are arbitrarily called 1 and -1”. While this is correct, EPR certainly considered their model to be consistent with the fundamental QM principles established by Einstein over the previous three decades. If Bell’s LR treatment was “…not a property of quantum theory…”, it was because Furry\textsuperscript{11}
had eliminated consideration of Einstein’s model, and because, contrary to EPR, Bell had stripped the entities of their intrinsic vectorial properties. Bell’s second mistake was that he didn’t acknowledge that he had thereby excluded the model he was meant to be testing. The zigzag (Fig. 3D) would be expected for radiating scalar partners from dichotomic pairs (the “exploding penny” model), but scalar correlations could never lead to a sinusoidal curve.

(ii) Bell’s LR expectations depend on the degree of order in the source. With a stochastic source, the outcome expected from eq. 4 is dependent only on the polarizer difference; it is independent of the setting of the reference polarizer. This behavior is a consequence of the isotropic condition, which justifies invocation of eq. 3b, and leads to an outcome, $p_{1,2}(\alpha, \beta)$, which would be rotationally invariant. Bell’s math was impeccable here. However, the behavior is not a consequence of loss of vectorial properties. This can be seen in the fact that all experimental reports have found sinusoidal outcome curves. What these show is simply Malus’ law behavior at the polarizers, which requires that photons have vectors. This natural behavior is also reflected in the model of Bohm and Aharonov, which was vectorial and Malus’ law compliant. The same point is demonstrated in the half-amplitude LR2 curves of the simulation, and in their rotational invariance. The apparent “loss” of vectorial consequence is a trivial epistemological issue arising from the stochastic nature of the source. In this light, Bell’s third mistake lay in extending the wrong conclusion from his first mistake (above) to an interpretation of what count to expect from coincidences. Both for the singles counts and for the coincidence counts from an isotropic population (eq, 4), he seems to have interpreted the rotationally invariant behavior as showing that the vectorial properties of the photons need not be considered as determining that behavior. In effect the replacement of the vectorial property by the sign of the spin was an ontic surgery, removing the vectorial property from consideration.

The counts from integration at a single station and the integration of the pairwise differences between stations involve different operations. The singles-counts come from integration at one station of elemental responses from an isotropic mix of $V$ and $H$ photons (eq. 1-3). Then each population measured at separate detectors gives the same singles count, no matter how the polarizer is set (Section 5 b)). In contrast, pairwise measurement involves detectors at two separate stations, and analysis in which each photon is compared (via Malus law) to its space-like separated partner. While in the singles counts, summing the elemental probabilities at a local station (eqs. 3a and 3b, or 3c) gives in the mean the same 0.5 local yields at any polarizer setting, when applied in pairwise counts, the same elemental probabilities lead to Malus’ law differences...
which accumulate to yield, in the mean, points on a sinusoidal curve. The Malus’ law outcome is given by $E_{\alpha,\beta} = \cos^2\sigma - \sin^2\sigma = \cos2\sigma$, etc., but at the elemental level, the yields are probabilities expressed in the distribution of $\pm1$ values on detection at each station. In the mean from a population in a stochastic distribution of vectors, this generates a curve of half-amplitude (LR2), as analyzed in detail in Part B.

With an ordered population, eq. 3b does not come into play. The RHE1 still holds, but the RHE2 term is no longer relevant, expectation of rotational invariance of $p_{1,2}(\alpha,\beta)$ no longer holds, and the vectors have to be included in analysis. With an ordered and aligned photon source, the mean from integration of pairwise coincidence (in effect using RHE1) gives the LR0 curve of the simulation. There are no tricks in the simulation; the same outcome can also be derived analytically from that equation, using the four cross-products between yields calculated at the two stations ($QS, RS, RT,$ and $QT$) taken for each configuration of the pair (for example $VH$ or $HV$). The outcome (the green curves) depends on the degree of order in the population. For ordered populations, with the photon and polarizer frames aligned, projections from the eight comparisons above lead to the same Malus’ law outcomes as in the matrix operations of the NL approach, and differences give points following the full-visibility curve (LR0) expected from this (see Fig. 2 and legend). For a stochastic population, the value for $\lambda$ for each particular pair will be different. In pairwise measurements, partner is still compared to partner, but the mean will be reduced by the entropic penalty from cancellations arising from the stochastic distribution of values for $\lambda$ to give a $\cos^2\sigma$ curve of half-amplitude (the LR2 curves are analyzed in more detail in Part B9). With ordered populations misaligned, the standard probability approach above gives the green curves in Figs. 3E, the Malus’ law result. From the perspective of the Furry argument, one conclusion is obvious, that, contrary to the conventional view, a simple mathematical treatment fully constrained by LR limits can account for all these behaviors. What it can’t account for is the alignment of frames implicit in the Copenhagen treatment.

In all cases, the sinusoidal shape of the curve simply shows Malus’ law in action, nothing more. No natural vectorial state could generate the zigzag; no scalar state could generate a sinusoid.

(iii) Derivation of the $\sim 2$ limit from the elemental count, - the BCHSH inequality. It has been suggested in many discussions (but in particular with Tony Leggett, Richard Gill, and Jan-Åke Larsson), that the model Bell discussed here was simply an example; other models are available in early work he cited. In his preface to the 1987 edition, Bell provides a check list of his
publications that contribute to the discussion of his LR model. Any model considered would have been constrained by the statistical limits, but the only other physical model discussed was the one in this seminal paper, borrowed from the vectorial model of Bohm and Aharonov⁹ and discussed further in Part B. He returned to his zigzag model in every other discussion; he never retracted it, continued to promote it¹⁹, and amplified his case by including the CHSH argument, which provided the same limit³⁹, thus strongly supporting his alternative approach. This argument, first hinted at by Clauser et al.¹⁵, further developed by Bell¹⁹,³⁸,³⁹ and by Clauser and Horne²⁰, expressed with clarity by Leggett¹⁶,⁵⁸ and reviewed comprehensively by Shimony²² is now examined. It was based on showing from the elemental coincidence values of ±1, that the value for \( S_{LR(\text{el})} = E_{\alpha,\beta} + E_{\alpha',\beta'} + E_{\alpha',\beta} - E_{\alpha,\beta'} \) from eq. 1 is limited by the maximum of 2, and noting that this constrains the mean value from any LR population to \( S_{LR} \leq 2 \), to set a limit for all LR models. This limit was then compared to the \( S_{NL} = 2.83 \) value from the NL treatment at canonical angle differences. The maximal value of 2 defines the LR limit, the NL expectation of \( \leq 2.83 \) unambiguously exceeds that limit (both features are also demonstrated in the simulation), so the conclusion that no LR model that conforms to Bell’s constraints could match the NL expectations might seem unassailable¹⁶,²². However, although the math is correct, the conclusion is wrong. It depends on two additional mistakes: (1) the assumption that the maximal value of 2 applies only to the LR case, and (2) the assumption that it is directly comparable to the canonical value for \( S_{NL} \) of 2.83. That neither assumption is valid becomes obvious from examination of the outcome curves scaled to four units through the \( S \) parameter (Fig. 4). The same sinusoidal curve in the range \( \pm 2 \cos^2 \sigma \) can be generated from either the vLR or NL model, (disproving (1)); and, although both the maximal value of 2 and the NL limit of 2.83 belong to the same outcome curve, they describe different properties of the curve, so are not comparable (disproving (2)). These two mistakes cannot be blamed on Bell; they were first suggest by CHSH, but have been perpetrated through their acceptance by the whole community¹⁶,³⁰,⁵⁷,⁵⁸.

(a) The range \( \pm 2 \) is a consequence of the choice of elemental values of \( \pm 1 \). Since \( \cos \sigma \) varies between \( \pm 1 \), the same elemental values and the same maxima and minima also define the NL curve (Fig. 2). This is the curve discussed in all theoretical treatments, and claimed in experimental reports. As shown in the simulation, at appropriate alignment, the NL and vLR curves are the same. The maximal value and the properties of the curve therefore apply to both models. The \( S \) parameter can take any value in the range \( \pm 2 \) (from eqs. 1a, 1b):

\[-2 \leq S = E_{\alpha,\beta} + E_{\alpha',\beta'} + E_{\alpha',\beta} - E_{\alpha,\beta'} \leq 2\]
Although the count constrains the *maximal amplitude* of the outcome curve to 2, the value of a point is constrained by the ±2 limits; any difference between two points is constrained to the 4-unit range of the curve.

(b) That the LR limit of ≤2 is problematic should then be obvious. The LR limit of 2.0 comes from a *singular* point, the maximum of the curve. In contrast, the NL limit comes from *differences between two points* on the curve. At any pair of canonical values for σ, the difference between the two points is $2\sqrt{2}$, falling symmetrically about 0 *within* the scale range of ±2 (right scale pertaining to the CHSH count in Fig. 4). For example, at canonical settings of 22.5 and 67.5 shown in Fig. 4, the $S$ values are 1.414 and -1.414, with the difference of ~2.83 applying to both models. Exclusion of local realism based on the comparison between the maximal value and the difference is then obviously absurd; what conclusion could be drawn from “...the value of ≤ 2 from the maximum of the curve constrains the NL model, and the $S_{NL}$ expectation of ≤2.83 unambiguously exceeds that limit...”? (The points on Bell’s zigzag are 1.0 and -1.0, giving a limit of 2 from the first inequality, but if the model is wrong, this is of historical interest only.)

(c) For any simple count of coincidences (Fig. 4, left scale, the anti-correlation count (red points) or the Boolean coincidence count of Fig. 3A), the curve will fall naturally in the range 0 - 4, and any ordered population in which frames align will give the full-visibility sinusoidal $4\cos^2\sigma$ curve. Values at the canonical intercepts show the same difference ≤2.83, but this is unremarkable when compared to the amplitude limit of 4.

Limits of <2 as derived above provide no basis for discrimination between local and non-local models, - they reflect instead either a poor choice of LR model or a poor treatment or both. When the photon and polarizer frames are aligned, there is no difference between $v_{LR}$ and NL expectations. The failure to find Bell’s LR expectations experimentally is unremarkable since it was based on an unrealistic model.

d) *Fitting the sinusoidal curves.* When real vectors in distinct frames for photons and polarizers are used to represent values pertinent to a $v_{LR}$ model, the results differ from NL expectations only when the frames are misaligned. For aligned populations, the full-visibility sinusoid can be derived directly from the Malus’ law yield differences in each ray (right panel of Fig. 3C and legend). This can be seen in the conventional NL treatment (cf.\textsuperscript{22}), which actualizes photons with their vectors in that same alignment to generate the same yields.
When oriented LR populations are not aligned, the more complete probabilistic analysis giving the green curves (Fig. 3E) is required. In the simulation, the points are obtained (as a function of \( \sigma \)) by counting the pairwise coincidences, finding the mean by the integration of \( eq. \ 4 \), using RHE1, and calculation of \( S_{\text{LR}} \) from the sum of eight outcomes as discussed above (Part A, 5a). Analytically, the green curves are derived from the same eight outcomes but calculated from the Malus' law expectations. These treatments embody all the BCHSH realistic prescriptions for LR treatments and generate simulated or theoretical curves that fit all outcomes using oriented populations. If both outcomes of a polarization analyzer are measured, the same eight terms contribute to the mean count from experimental pairwise comparisons. The set is formally equivalent to those in play from the matrix operation of the NL treatment applied in the plane of measurement (Fig. 2).

e) **Known unknowns.** No elemental measurement can lead to complete specification. Elemental events involving photons are necessarily quantized, but classical statistics will still apply (see\(^85\), and Part B). The simulation demonstrates that the sinusoidal outcome curves are not a consequence of indeterminacy, or of inseparability, superposition, or any of the algebraic paraphernalia said to be required to deal with QM uncertainty. As long as the “uncertainties” are distributed normally about the mean, the counts of elemental pairwise coincidences will be sufficient. Attribution to the ‘entangled state’ of ‘super-correlations’\(^57\) on the basis of the sinusoidal curve is nonsensical; the shape of the curve requires nothing more than Malus’ law operating on pairs in vectorial correlation (discrete in the \( v_{\text{LR}} \) model, LR bivectors in Clifford algebra treatments\(^80,86-88\), or the pairs actualized with aligned vectors in the NL treatment). The requirement of mathematical complexities is a consequence of the superposition, and the necessity of resolving it, and superposition is a consequence of the uncertainty principle.

The \( v_{\text{LR}} \) model involves no “hidden variables”. All the information needed to account for the outcome is carried as intrinsic properties of discrete quantized entities; the information arrives with the photon. There is then no need for “conspiracies” to explain the results; the model is immune to closure of communication ‘loop-holes’ (discussed at greater length later). The only requirement is for the behavior at the discriminators to be natural. On the other hand, the full-visibility amplitude depends on alignment; the difference in the two approaches as to how that is achieved is a separate issue. In the \( v_{\text{LR}} \) case, the process is transparent, but not so in the NL case, as discussed at length in Part B.
6. Why are these conclusions important?

The literature is full of claims (cf.16,19,21,22,48,58) that no LR theory could yield a curve that departs from the limit of \( \leq 2 \). Indeed, that limit is nowadays the de facto criterion used for evaluation of the success of experiments in supporting the non-local picture, as shown by claims to “…have measured the \( S \) parameter… of Bell’s inequalities to be \( 2 < S < 2.83 \), thus violating the classical value of 2 by \( n \) standard deviations…” or similar45,54,63-66,89. In these comparisons, including in recent “loop-hole-free test of local realism” using electrons70 or photons64-66,69, the targets of 2 from the zigzag or CHSH count (or similar64) are worthless because they are irrelevant. Examples in the popular literature that justify the NL picture based on LR models that use scalar dichotomic qualities such as colored socks56, live and dead cats90, hard/soft, black/white, red/green properties29,71,91 etc., or even the polymorphic quantum cakes92, serve only to confuse. Such properties are simply inappropriate to the vectorial states involved and could never generate sinusoidal curves on analysis with polarizers. The simulation highlights a general problem with the BCHSH approach, - that Bell’s LR model is unable to represent a state with the vectorial properties needed in application of Malus’ law. The inequalities so far discussed show that i) no model that omits vectors can generate the sinusoidal curves claimed as supporting NL expectations; and that ii) the vLR model can generate the full-visibility curve within the probabilistic constraints of local realism. Justification for non-locality must be demonstrated through principles that recognize these features.

PART B. MAYBE EINSTEIN GOT IT RIGHT

A philosophical hurdle. A newcomer engaging with the entanglement community quickly learns that acceptance of the NL case is general and is based on a conviction that Einstein lost the argument with Bohr; quantum uncertainties preclude assignment of intrinsic properties, and therefore require treatments in which entities evolve in a superposition of indeterminate states, which are actualized only on measurement. The strength of the case was apparent in development of the H-atom model; as atomic spectroscopy provided energy levels for electrons, the potentials for occupancy of orbitals were mapped, completed on inclusion of the spin states in the Schrödinger wave equation, and then more widely applied to flesh-out the periodic table. The standard Copenhagen interpretation, and its ancillary orthonormal treatment of spin states became deeply embedded in the zeitgeist of quantum physics. Superposition requires a non-local framework, necessarily treated in the wavy domain, and LR treatments cannot account for the outcomes claimed9,11. Advances over the 85 years since EPR represent an academic heritage through many generations of a success that has spawned Nobel laureates galore, revolutionized physics and chemistry, and fathered many of the innovations
that drive our modern economies. Bell’s treatment was brilliantly framed within this tradition, and
the experimental validation of his theorem in the ’70s and ’80s cemented the non-local view and ex-
tended a confidence that similar treatments should also be applied in the wider spatial context.

Insofar as the conventional treatments relate to condensed systems and/or atomic scales, I
have no argument with the main conclusions from this spectacular record. However, tests based on
entangled photons require evolution of entangled partners to space-like separation before measure-
ment, and non-local effects are then consequential9. Even Bell expressed misgivings about this side
of his theorem56, and confidence must now be further eroded by the wonkiness of the stool on losing
two of its legs. The ‘failure’ of local realism arises from the fact that its constraints are real, local,
and second law compliant. Given the antecedents he invoked9,11 and the LR model he constructed,
Bell’s conclusion that all LR models could be excluded was justified, and his ideas gained traction
because that model was uncritically accepted. The zeitgeist trapped him, and apparently the commu-
nity in general, in a box that excluded Einstein’s model. Despite advances in physics, the same justi-
fications are still applied, so I will first examine the conventional NL case, and then see if any para-
dox remaining can be resolved in light of progress beyond the Copenhagen interpretation.

In the vLR model, quantum scale entities carry properties intrinsic to discrete states, - Ein-
stein’s ‘elements of reality’2,83. Einstein’s case is well known from his criticism of the Copenhagen
interpretation as incomplete1,93, but this argument had been discounted by von Neumann who proved
that “hidden variables” were not needed in the Copenhagen approach. It was from his analysis of von
Neumann’s case that Bell became interested in the entanglement debate (see Section 2 below). Ein-
stein’s model was also effectively excluded on the basis of the mathematical constraints9,11, which
showed Bell that he was right.

Bell himself had earlier18 suggested in the context of the “hidden variables” debate that “…if
states with prescribed values...could actually be prepared, quantum mechanics would be observably
inadequate....”, - perhaps a recognition that such states might become available. With PDC, the
phase-matching is determined (prescribed) by conservation laws, orientation is determined by the
pump laser, and in type II PDC for example, the two emergent beams are empirically demonstrated
to be orthogonally polarized (cf.89, and Paul Kwiat, personal communication). Paradoxically, the ma-
nipulation of these populations in state preparation is explicitly based on full knowledge of their de-
terminate properties. In Sagnac interferometric applications68,69,94 (see SI, 3 ii) a) - c)), the polarized
components of both PDC outputs (from passage clockwise or counter-clockwise through the interfer-
ometre) are separated by a beam splitter, and used without mixing to provide distinct polarized popu-
lations used in measurement. But under Copenhagen precepts use of such information is forbidden
when an indeterminable superposition is invoked for entangled states. Taking to heart Bell’s recognition that states with “prescribed values” are contrary to QM, in this section I re-examine the conventional NL treatment and its experimental support and find both wanting.

1. Bell’s third inequality provides a discriminating case

My vLR model is simply an extension of Einstein’s idea. In one sense it is trivial, - if a model matches the alignment expected from NL, of course it will give the same outcome. However, its consequences have apparently not previously been appreciated. Likely the mathematical constraints\(^{11}\) were seen as sufficient justification for excluding such discussion. Otherwise, it would have been obvious that the limits of \(\leq 2\) could not exclude local realism, and invocation of that limit would then represent deliberate obfuscation. Since that is anathematic, whatever has been discussed is something different that in effect reflects the constraints from the Furry limits\(^{11}\).

In Part A, I showed that two of the legs claimed as supporting this stool provide neither support nor justification for exclusion of local realism. This failure might be expected to worry the entanglement community. Since the limits of \(\leq 2\) from the zigzag and the CHSH counts cannot themselves exclude local realism, what does? In discussion with colleagues, loss of the two legs has been dismissed as uninteresting, specifically because of the antecedent cases\(^{9,11}\). Bell’s third inequality is in essence the same as the Bohm-Aharonov vectorial LR model. He framed this in the seminal paper through “…consider the result of a modified theory…in which the pure singlet state is replaced in the course of time by an isotropic mixture of product states…”, for which he suggested a correlation function, \((e^{-\frac{i}{3} \alpha \cdot \beta})\). In contrast to the model giving the zigzag, the pairs here retain vectorial properties. The factor \(\frac{1}{3}\) is not explained, but the reduction in amplitude was expected under earlier treatments\(^{9}\). The reduced amplitude is implicit in eq. 4 when the electron pairs are in stochastic orientation because of the entropic penalty incurred by the disordered state\(^{9,11,17,14,75}\). A similar reduced amplitude is shown by simulation for vLR photon pairs generated with a stochastic photon source, where the penalty gives the half-amplitude LR2 curves (Fig. 3B). Bell was clearly right here; with a stochastic population, no LR vectorial model could generate the full amplitude curve. However, in this case, the reason is thermodynamic; LR models are constrained by the second law, but the conventional NL treatment generates an ordered and aligned state from a stochastic source, so apparently is not. It is this paradox that I now want to explore.

Leggett (personal communication) has suggested a concise expression of the distinction arising from Bell’s third inequality through the following cases, in which \(\theta_1\) and \(\theta_2\) are orientations of the
fixed and variable polarizers, respectively:

- **Case 1:** For any possible choice of $\theta_1$, there exists an LR model, $T$, such that for all $\theta_2$,
  $$f_T(\theta_1, \theta_2) = f_{NL}(\theta_1, \theta_2).$$

- **Case 2:** There exists an LR model, $T$, such that for any possible choice of $\theta_1$, and for all $\theta_2$,
  $$f_T(\theta_1, \theta_2) = f_{NL}(\theta_1, \theta_2).$$

In framing these cases, Leggett perhaps recognized that the LR0 outcome of my simulation demonstrates model $T$ for Case 1; I take this as a partial validation of the conclusions in the first part of the paper. Failure of the $v$LR treatment to predict model $T$ for Case 2, the full-amplitude rotational invariance, leaves that as the remaining justification for exclusion of local realism. Despite claims from others to the contrary\textsuperscript{81,86,95} (discussed in SI, section 3 (iii)), my simulation shows, in agreement with Bell and the earlier analyses\textsuperscript{9,11}, that constraints from local realism mean that no LR model can predict the full-visibility rotational invariance expected from the NL treatment\textsuperscript{22,96}.

Leggett’s distinction omits an important consideration that will figure in further discussion, - the information about vectorial properties of the photon source that can be gleaned from the generating transition. Though excluded in the NL case, and therefore of no relevance, these properties have to be considered in any vectorial realistic model, and any comparison has to include a dissection of their fate in the NL case. For photons carrying intrinsic properties, Case 1 then becomes more highly restricted (and therefore more easily tested), limited to the situation in which the photon population is ordered, and its reference frame is aligned with $\theta_1$. For Case 2, the outcome predicted remains unconstrained, in the sense that the NL outcome is independent of the initial orientation of frames, or of whether the population is stochastic or ordered. Any population of pairs considered as initially in an indeterminate superposition with propensities in dichotomic correlation must give the same full-amplitude rotational invariance for any reference frame at the polarizers (see below). However, since in the stochastic case, the initial state is clearly disordered, and the outcome observed depends on actualization of an ordered and aligned photon state at the polarizers, to be credible the treatment would have to include a mechanism, including a work-term, to account for the ordering. I argued below that this requirement cannot be naturally satisfied.

2. **How credible is the NL case for non-locality?**

My simulation demonstrates that the restriction to the standard probability treatment from LR constraints\textsuperscript{11} does not prevent one such model, my $v$LR, from generating an outcome that matches the
QM expectation. However, my discussion also emphasizes that the match is found only with an ordered population under conditions of alignment explicitly introduced to match the alignment implicit in the orthodox NL treatment. From this, the conclusion follows that an important consideration must be of the process through which alignment occurs. In the NL case, the expectation of full-visibility rotational invariance can be framed through a few primary premises:

(i) Since Heisenberg uncertainties preclude attribution of intrinsic properties to discrete quantum entities, entangled states must be treated as in an indeterminable superposition of all possible states during evolution to the measurement context.

(ii) Correlations are represented by dichotomic spin states in the wavefunction, and the binary terms of the Pauli matrices are also dichotomic, and in principle represent the correlations. On operations of the matrices, the binary terms are assigned propensities such that the entities are actualized with real vectors in a frame aligned with the discriminator reference frame.

(iii) The matrix operations are assumed to represent a physical behavior leading, under experimental conditions, to actualization in alignment at the discriminators.

Ironically, the first premise sets up the entangled pair in a state from which vectorial information from the source is excluded. Since correlated vectorial properties, revealed on measurement at space-like separated stations, are essential to the outcome, the processes through which they become available and aligned, should have raised all those concerns implicit in Einstein’s “incompleteness” argument. Extending the irony, Bell justified his use of the term “hidden variables” in the context of that argument. However, quantum uncertainties demand a superposition, subsequent application of the orthonormal treatment provided a consistent resolution, von Neumann’s formalization showed that no “hidden variables” were needed, and the mathematical constraints excluded Einstein’s model. Although Bell in his critique of von Neumann found logical inconsistencies that allowed “hidden variables” in some contexts, he also found that they did not apply to his QM treatment, and he therefore saw no reason to consider them in that context. Instead, in a triple-irony, the inherent difficulties were transferred to LR theories. The “hidden variables” were needed there because in setting up his LR model, Bell had stripped the “entangled” state of its vectorial character.

In the Introduction, I discussed the Furry constraints showing that no LR model limited by standard probabilistic constraints can account for expectations from non-local treatments. Resolution of real entities from the superposition entails a measurement context involving conjoined probabilities at separate stations, so the mathematics must cope both with representation of the starting state,
and with the process through which actualization as separate entities is implemented. The conventional conclusion is that local realism should be excluded. In Part A I showed that my vLR model also demonstrates that no LR model can explain the NL expectations (Furry’s conclusion). If this is settled, then further argument along these lines is futile. I now want to reframe the discussion by suggesting a different approach, - that we agree that the two models are irreconcilable and ask how to test each of them on their own merits. As an experimentalist, I would approach this task by asking how to discriminate between the two models. Ideally, this would require experimental protocols that conform physically to the two different physical states on which the models are based. Protocols based on PDC sources are ideal for testing Einstein’s view because the determinate nature of the process ensures that the photon source conforms to the LR model tested. Unfortunately, this determinate nature is then in direct contradiction with the superposition, the physical state claimed to have been tested in all recent reports.

In fact, the Copenhagen interpretation is the problem. There are two main points. 1) The superposition of states cannot be represented by any protocol using PDC as the source of photon pairs. The properties of photons in the cones output from the crystal are prescribed by conservation laws and the properties of pump laser. The state is fully determined and is shown experimentally to be so. There is no sense in which the output could be claimed as in an indeterminate superposition. 2) The treatment generates an outcome inconsistent with the second law. This is most obvious when an atomic cascade provided the source; the atomic beam is stochastic and the flash that populates the excited state is stochastic, so the population generated must also be stochastic. Even when lasers were used in activation, the orientation was chosen so as to preserve the stochastic nature. By definition, these sources generate a disordered state, which is thereby indeterminate. However, the conventional NL treatment leads to expectation of the same full-visibility outcome with all orientations of the photons. This is a consequence of the alignment of the photon frame with the polarizer frame implemented in the matrix operations. This ordered outcome came from a disorder state, and the ordering is achieved without an input of work. Even when applied in experiments using PDC sources, a similar ordering is needed to get the full-visibility rotational invariance when starting from a misaligned state. The same outcome is claimed for all orientations of the photon frame; ordered and aligned, ordered and misaligned, or stochastic.

In the context of experimental test, the vLR model now introduces a scenario comparing local and non-local perspectives in which both are vectorial and compliant with the fundamental QM tenet. In my vLR interpretation, all energy exchanges are necessarily local and quantized, and probabilisti-
cally constrained. From this perspective, when properties are intrinsic and explicit, the vectorial information is carried by the photon, and therefore cannot be “hidden”, and the simulation shows that no conspiracies are needed to explain the outcomes simulated. On the other hand, the above premises necessarily leave the conventional NL model still as “incomplete” as it was when EPR challenged Bohr\cite{1,14,93}.

In the next sections I examine the conventional NL case in greater detail in order to understand its justification, and in particular how the contravention of the second law is explained. I then explore how the experimental tests are claimed to discriminate between the two models.

3. **Indeterminacy and superposition of the entangled state**

The fundamental tenet of quantum mechanics is that all energy exchanges at that level are quantized. From a research career in photosynthetic mechanisms, it is obvious that at the molecular level interactions of photons with electrons in molecular orbitals involve local exchanges in which energy is conserved, the charge of the electron is inviolate, and photons are neutral and interact through transfer of momentum. This is in contrast with the electromagnetic nature of light inherent in Maxwell’s equations and treatments incorporating them, from which the interaction of light with harmonic oscillators is through the fields of the wave. This distinction is discussed in the closing sections, but the neutrality of photons makes mechanistic involvement through electromagnetic properties untenable, and I will therefore adopt the former perspective as my starting point.

*The uncertainty principle and its application are not appropriate to entangled pairs*

Despite the extensive literature, I can see no reason to believe that quantum uncertainties should exclude treatments in which discrete quantum entities have intrinsic properties. In the conventional treatment, superposition is called for because uncertainties preclude assignment of definite properties to quantum entities. The treatment dates back to the period when the electronic orbital occupancies of the H-atom model were being sorted out, and the question of what information could be included led to recognition of the need for a probabilistic approach. The minimal uncertainty derived for electrons by Heisenberg ($\sigma_r \sigma_p \geq \frac{3\hbar}{2}$ in the 3-D case) precluded assignment of definite properties, but, the argument applied was in a scenario of simultaneous measurement of conjugate variables for position and momentum on a single electron. In the context of occupancy of electron orbitals, the conclusion was appropriate. Can this approach be applied to photon pairs? Although a similar minimal uncertainty has been suggested for photons\cite{98}, the *logic* of this approach cannot simply apply.
Since detection consumes the photon, it can occur only once; there’s no way to detect a photon twice, so the equivalent experiment could never be attempted. But that is not a problem in the entanglement context because we are dealing with two different photons, correlated in orientation, interrogated through refractive interactions (which return $h\nu$ without loss), separately detected at different stations. A similar argument applies to electron pairs.

Measurement on quantum-scale entities involves two distinct components, - interrogation and detection. From the experimentalist’s perspective, although detection of a photon is a one-off event, it does allow one certainty; a recording of the time and place of its arrival. This here-and-now information is all that is directly available from detection, but that leaves as a separate issue how to determine other properties. To access those, experiments have necessarily examined populations. In principle, the properties of the photons in a population can be determined when the source and pathway of evolution are known; a measurement can then be interpreted in terms of information available from the generating transition, and from interrogation during evolution to detection. Interrogation is of interest only when it does not consume the photon. The processes can be selective (transmission through a filter, prism or monochromator), reflective when mirrors are used, or refractive in, for example, use of lenses, HWPs, or polarizers. A single photon will experience multiple refractive interactions in its path to detection, but refractive events are loss-less, so this is not a problem. Useful information can be gleaned from analysis, because, without changing the energy of the selected photons, the path is perturbed in time and/or in space. At a particular frequency (determining the refractive index), refractive behavior probes the remaining property of the photon, the polarization vector. At the elemental level, to engage, the vector of the photon must match a vector for electronic displacement, - along the polarizer axis in a polarizer. The probability that a photon will excite a displacement will then be given by Malus’ law. The uncertainty here is in that classical probability term. Since elemental measurements are made on populations, such statistical uncertainties are ironed out in the mean. Epistemologically, the essential point is that because in the entanglement context, two separate photons are detected separately, and the refractive events are loss-less, quantized interrogation through refraction entails only the statistical uncertainty inherent in probabilities of Malus’ law at the elemental level.

If the superposition is a pure state of zero entropy, the uncertainty principle might well apply; then actualization must involve a conjoined operation to reveal two real separate entities, demanding the fancy math. But the argument works only if a superposition is required. In the previous sections I show that superposition is not needed for generation of a full visibility curve under aligned
conditions. In that case, the Furry constraints\textsuperscript{11}, which discriminate only if the whole caboodle is invoked, would lose their force. The critical issue would then be the question of how alignment is implemented.

Bell saw the distinction between the views of Bohr and Einstein as between “…wavy quantum states on the one hand, and in Bohr's ‘classical terms’ on the other…” (see Preface\textsuperscript{56}, and \textsuperscript{99}). This view likely reflects the probabilistic constraints\textsuperscript{9,11}. From the correspondence principle, either a particulate or a wavy treatment could be used; choice of one must then have an equivalent expression in the other which gives the same result. From the fundamental tenet, the particulate choice must always be primary; at the elemental level photons are quanta. Following Dirac\textsuperscript{10,37,100}, advances over eight decades in quantum field theory (QFT) have dealt with particle/wave duality by coopting Maxwell’s electromagnetic wave treatment; different interpretations have led to many different models\textsuperscript{101}, all of them non-local, justified by mathematical treatments of increasing sophistication, and a massive literature demanding a technical expertise beyond my skill-set. Since the BCHSH argument did not explicitly invoke Maxwell’s approach, I defer further discussion to the concluding section. Notwithstanding that issue, all NL treatments in superposition depend the math required to deal with that, and lead to the probabilistic exclusion of the LR case\textsuperscript{11}. To set against this, the non-local nature of all treatments has its problems. Apart from the claim to have eliminated Einstein’s model, no experimental test has been suggested that can eliminate any of the others. This obviously raises the question of whether the problem lies in the common non-local feature of their treatments.

For entangled photon pairs, the complexities inherent in representation in Hilbert space have perhaps also been overhyped. The spatial complexity is restricted because the action vector is orthogonal to the \(z\)-axis of propagation. This limits measurement of vectorial parameters to the \(x\), \(y\) plane, which simplifies consideration to the unit circle (Fig. 2 and legend). The photon vectors defining the unit circle projections are already aligned, and could be seen as electric or momentum vectors, and either as coming from the complex plane, or as the intrinsic vectors of the \(v\)LR treatment. The projections give the roots of the Malus’ law \(\cos^2 \sigma\) observables. Values returned on taking squares have the same Malus’ law values in either treatment. The only component of the treatment that discriminates is that implementing alignment.

Practice elsewhere in physics deals in photon populations without any prerequisite for such complexities. For example, in determining the history of our universe, modern cosmology invokes three simple notions. (i) Measurements on a uniform population can reveal common properties of the discrete photons making it up. (ii) Those properties are intrinsic and determined by the transitions in
which the photons were generated (the spectra inform us of the distant chemistry). The properties may also be modulated by local fields, for example to generate polarization. (iii) Properties are in principle conserved on travel over space-like distances. Although frequencies are modified during their journey by relativistic effects, and gravitational refraction can distort the path, these features allow interrogations that provide, in a measurement context, information about the source and its environment, and the path. Without these assumptions, we could not construct a history of the cosmos. Similar principles underlie all spectroscopic applications, and my simulation of a vLR photon populations is based on these same assumptions. If we recognize with EPR that each photon of the pair carries its own vectorial property, the idea of an ontic indeterminacy has meaning only in the context of an indeterminate superposition.

I can see no simple model of the photon other than Einstein’s that is compatible with the quantized nature of its interactions, the time and length scales given by its frequency, and with relativistic constraints. Uncertainties are statistical, - an epistemological challenge. A photon can only be detected once, but in the entanglement context we have two of them, and refractive interrogation is loss-less, so the logic of the uncertainties from measurement on a single entity cannot be taken to necessitate a treatment starting with states in superposition.

4. What is known from the generating transition?

Under all experimental protocols reported, useful information has been available about the transition generating the initial state. The correspondence principle\textsuperscript{8} would require that such information be considered under any QM treatment. Dirac\textsuperscript{100} suggests that indeterminacy and superposition necessitate a probabilistic treatment. However, Planck\textsuperscript{102} derived $k$ and $h$ by applying the classical probabilistic treatment of Boltzmann to a quantized distribution. In vectorial treatments, probabilities are given by Malus’ law yields, so the result of a set of pairwise measurements of correlations should simply generate the statistical outcome expected from such properties, as it does in my simulation. Problems arise not in the vLR case, which, by recognizing discrete properties, allows a natural treatment, but in the NL case, where, contrary to the correspondence principle, vectorial properties are excluded by fiat. In the superposition then required, correlations have to be represented in a dichotomic Bell-state through the spin topologies implicit in the spin quantum numbers. But these are not vectorial. A mechanism generating specific vectorial properties in the process of actualization must then be invoked to account for the behavior at the discriminators.

(a) Cascade sources. With cascade sources, frequencies are known from spectral lines, experimental limits for vectorial correlations from conservation of angular momentum are well known;
these determinate properties are used in design of protocols. When photons from arc-discharges are used to excite the atomic beam, since orientation of both the atom beam and the photon source are stochastic, orientation of the pairs after excitation must also be stochastic. Bell was right to recognize that their stochastic nature has real consequences, but wrong in how he applied those insights (Part A, section 5).

Excitation of the atomic beam by a laser would lead to photoselection\textsuperscript{40,41}. This was demonstrated by polarization detected along the orthogonal $y$-axis when the laser was polarization along the $z$ axis of propagation\textsuperscript{40}. Photo-selection along the $z$ axis would excite a population of atoms with vectors in that direction, but it would be isotropic in the $x$, $y$ plane of measurement. Experimentally, the singles-counts measured in that plane would then follow the stochastic pattern\textsuperscript{40}, as expected from this analysis.

(b) \textit{Sources generated by parametric down conversion}. The stochastic model has no relevance in the case of PDC sources. Laser excitation of PDC provides well-oriented photon pairs correlated through conservation of energy and angular momentum (phase-matching\textsuperscript{103}); all properties are defined with respect to those of the pump laser\textsuperscript{63,89,103,104}. The photon pairs separate into two populations, orthogonal in orientation in Type II PDC, with determinate vectorial properties ($H$ and $V$ are real orientations in the pump reference frame); they are, in Bell’s usage\textsuperscript{18}, prescribed. The behavior of one population is independent of measurement in the other\textsuperscript{49,89}. The PDC output is clearly a state the “with prescribed values” of Bell’s caveat\textsuperscript{18} (see introduction to Part B). Should we not take Bell’s conclusion seriously, and worry about the adequacy of the NL approach in that context?

PDC sources have been used in all recent photon-based experiments claimed to support the non-local perspective (cf.\textsuperscript{34,45,48,54,89,104}), and knowledge of these determinate properties is exploited in design of experimental protocols. There is a degree of schizophrenia in play here; the engineering side requires acknowledgement of the determinate nature, while, since all test are premised on the non-local treatment, the theoretical side has to shun it. For example, in a seminal paper\textsuperscript{45} entangled pairs were generated in complementary cones by PDC in a type-II phase-matching BBO crystal excited at 351 nm. The output from PDC was two overlapping cones, one with $H$ and the other with $V$ photons. For any pair, the correlated partner was in the other cone. Photons at 702 nm were selected from the two intersection points of the cones. The Bell state was $HV/VH$, so the photons at either intersection were not pairwise correlated, but by selecting the opposing intersection points, the entangle partners of each pair were separated and sent to the different stations. After further tweaking, for example by insertion of a half-wave plate in one channel to change the Bell-state to $HH/VV$ (‘state-
preparation'), the mixed populations from the intersections were sent to separate stations for polariza-

In another application\textsuperscript{63}, compensating crystals were used to maximize the output of entan-
gled photons from a double-crystal BBO type-I source, - a pretty exercise in optical design.

In a type-II configuration using a different crystal for PDC (periodically poled KTiOPO\textsubscript{4})\textsuperscript{89},
and pumped at \textasciitilde{}400 nm, the co-linear cones at \textasciitilde{}800 nm overlapped completely, but the contribu-
tions of the two orthogonal orientations could be distinguished by polarizer at orthogonal rotations\textsuperscript{89}. Entangled partners at the chosen frequency were in opposite halves of the overlap, so could be sepa-
ated using mirrors and irises, allowing a much larger fraction of the population to be tested.

A recent variant of this approach\textsuperscript{94} used the same crystal for PDC, but configured in a Sagnac
interferometer. This configuration, discussed in detail in the \textit{SI} (section 6 Bb, and Fig. \textit{SI}._1A), was
also used in entanglement experiments with the source located in a satellite to test communications
loop-holes in\textsuperscript{68}, and in other recent spectacular over-kills in the context of such loopholes\textsuperscript{65,66,69}. As
discussed in the \textit{SI}, some features of the Sagnac configuration are far from conventional. In particu-
lar, the signal and idler beams were fully polarized when projected to measurement stations, and the
\textit{H} and \textit{V} photons at each station came from the two separate PDC processes, so neither station sees
the mixture in ontic dichotomy implicit in the conventional NL treatment.

The general point I want to emphasize is that both in cascade experiments and PDC-based
protocols, detailed information is available from knowledge of the source. Manipulation of the “en-
tangled” populations in state preparation is clearly based on exploitation of that information, - on
specific determinate properties of the source, and on classical behavior at refractive elements. Uncer-
tainty is introduced from a mixing of two separate populations for which properties are determinate.
With cascade sources, the stochastic nature introduces classical uncertainty. With well-determined
PDC sources, the dichotomic state generated is fully determined, and quite different from the dichot-
omy \textit{by predicate} of the orthonormal application. The appearance of indeterminacy from mixing does
not mean that intrinsic properties and their correlations are lost. Photon properties, when intrinsic,
would be retained, and photons would then behave at polarizers according to Malus’ law\textsuperscript{89} to gener-
ate my \textit{vLR} outcome. Given that the initial state is determinate, two \textit{ontic changes} would be required
to generate NL expectations; one, on generation of the pair, to a state in superposition in which vec-
tors are \textit{lost}, and the other, the notoriously vague “reduction of the wave packet”\textsuperscript{38}, to the determina-
ble states revealed on measurement, with vectors aligned with the polarizer frame. Absent a mecha-
nism, this is plain silly.
5. How is full-amplitude rotational invariance implemented under NL predicates?

Two overlapping scenarios need to be examined, the first a mathematical model, and the second an attempt to relate the model to processes in the real world.

The model. Bell adapted the conventional QM approach\textsuperscript{17,19} for entangled electron pairs, summarized in the matrix operation of eq. 2. In the BCHSH consensus, essentially the same approach was applied to photon pairs\textsuperscript{9,21}. As shown by eq. 2, the matrix operations are vectorial, with photons (represented as binary values in the matrices) and polarizers at both stations, and an implicit simultaneity of action. As discussed below, the spin states are potentialities and correlation between them gives the phase difference, but these are then taken to also indicate propensities. In operation of the Pauli matrices, the propensities in effect determine that the photon pairs become aligned with the reference polarizer. If the photon was $H$, it will emerge in the ordinary ray, if $V$ in the extraordinary ray, and the partner photon at station 2 will be actualized with the orthogonal vector implicit in the phase difference. The constraints of the matrices are then resolved as observables, - the Malus’ law yields.

Shimony provides a useful summary\textsuperscript{22} of the above behavior; in effect, the alignment of the photon frame with the fixed polarizer implemented in the matrix operation is described “…by substituting the transmission axis of analyzer I for $x$ and the direction perpendicular to both $z$ and this transmission axis for $y$”, where $x$ and $y$ are $H$ and $V$ in the wavefunction equation.

To attain the outcome claimed, the primary premises (Section 3) require auxiliary assumptions.

1) Application of Malus’ law requires that at both stations, photons and polarizers have real vectors, so actualization would have to occur before the photons reach the polarizers. Since vectorial information for the photons is exclude in the superposition, vectors have to be provided in the process of actualization. The fixed polarizer is the only reference frame, and the difference from the variable polarizer is referred to that zero. The potentialities of the spin states are represented by the binary matrix elements. In effect, the propensities take on a vectorial agency in the matrix operations, which then implement an alignment constrained so as follow Shimony’s prescription above. Only then do they account for the outcomes claimed (see Section 6, 3) and 4) below).

2) In recent applications using PDC sources, both polarizers are fixed, and their discriminator function is implemented by using HWPs to rotate the beams before they arrive at the polarizers. This requires a modified treatment to cover the HWP function, - actualization has to occur at or before the discriminating HWPs. However, since similar refractive components are used upstream in state preparation, the question of why actualization does not occur there becomes problematic.
3) The NL outcome depends only on the angle difference between polarizers, \( \sigma \). Operation of the matrices implies simultaneous actualization at both stations; in the math this is no problem, because both polarizers are engaged, but in the real world, the two stations are space-like separated, and information about polarizer settings is only available locally. The question of how simultaneity is achieved then becomes problematic.

6. Relating the model to the real world

The math may be elegant but fitting it to the physics is not. Several features are missing, mainly because the spatial element can be ignored in the math but cannot be avoided in the physics. Firstly, when starting from a stochastic state, or with a misaligned ordered state, an input of work is required to generate the aligned state needed to explain full visibility. Secondly, if the polarizers are to retain their natural function, photons must have real properties before they arrive there. Actualization in the ordered state would have to precede interaction with the polarizer. If HWPs are used to provide the discriminator function, actualization must be before them. This becomes complicated because, thirdly, the stations are space-like separated; a value for the vector for the photon at the reference station can be known only at the instant of actualization, but that information is needed simultaneously at the other station to allow actualization of the second photon in appropriate correlation. Fourthly, in the mechanism suggested the spin states are assigned propensities with a causal role, but there is no obvious physical basis for this. Analysis of these anomalies reveals a common problem: the premises do not match the properties of the system under study.

1) There is no evidence for any ontic vectorial dichotomy in populations of electrons or photons. The conventional formalism requires an intrinsic dichotomy of spin states. Interpretations involving such states date back to the seminal work in which a beam of silver atoms, each with an unpaired 5s electron, was analyzed using Stern-Gerlach magnets. The atoms were aligned by the magnets into two well-defined populations, interpreted as showing that the partition of the atoms was determined by the spins of their electrons, - that electrons come in two different spin states, \( \uparrow \) (up) and \( \downarrow \) (down) with an intrinsic dichotomy. The partitioning into well-defined populations was then further interpreted as showing propensities for partitioning. The spin states became aligned with the field of the Stern-Gerlach magnet as determined by their propensities; the up electrons were expected to emerge in the upper population, and the down electrons in the lower population. The conversation is further complicated by what happens when the ordered populations are test with S-G magnets at orthogonal orientation. Bell has a lucid discussion of
these features, and an honest evaluation of the epistemological difficulties, but no simple explanation. The debate about propensities has a history going back to Dirac and von Neumann, recently discussed as the projection postulate by Graft\textsuperscript{106}, which is revisited later. However, the dichotomy implicit in this assignment cannot be taken as ontic, because, as fundamental particles, all electrons must be the same (cf.\textsuperscript{107}). Expectation of such a dichotomy came from a misunderstanding, - the dichotomy of the discriminator function was interpreted as demonstrating an intrinsic dichotomy of the particles. This dichotomy has become deeply embedded in the standard orthonormal treatment, and the potentialities implicit in the spin quantum numbers have been invested with the vectorial agency of propensities.

Although spin may be an intrinsic property of the electron, spin quantum numbers are not. Like all the electronic quantum numbers of the hydrogen atom, they are topological designators. In their atomic context, the topologies of orbital occupancy are defined by the lower quantum numbers, and the electrons in completed orbitals are paired. The spin quantum numbers, $s = \pm \frac{1}{2}$, determine the relative orientation of the spin states, with potentialities (the phase difference) given by $\pi/2s$; the spin axes of a pair then have opposite vectors (\uparrow and \downarrow) so that their magnetic effects cancel in their attraction. This makes sense in the context of electron pairing; they really are entangled through the work term from attraction that is expressed in the additional bond stability. In that case, the orthonormal operations give the ensemble results implicit in modern quantum chemistry. But such topological consequences of spin cannot be in play with unpaired electrons.

When unpaired, nuclear and electron spins can be explored experimentally using NMR or EPR (electron paramagnetic resonance, not to be confused with the authors). When placed in a magnetic field, they are then found to distribute into separate populations following a Boltzmann distribution, \begin{equation*}
\frac{n_{\text{up}}}{n_{\text{down}}} = \exp \left( - \frac{\hbar \nu}{kT} \right),
\end{equation*}
reflecting the energy difference between up and down states. The fractional excess of the more stable $n_{\text{down}}$ population, which becomes more obvious as $T$ is lowered, leads to a net absorption when the populations are flipped on illumination by radio or microwaves at resonance. The Boltzmann distribution in the magnetic field does not show intrinsic dichotomy. If the spins were in an ontic dichotomy, they would distribute evenly, and no NMR or EPR signal would be detectable.

If electrons are not intrinsically dichotomic, the sorting in the Gerlach-Stern experiment\textsuperscript{105} has to be explained by an alternative mechanism, and a trivial one is available. The separation by
the magnets into the two ordered spin populations reflects two different but overlapping processes in measurement. With atoms (and their 5s electrons) initially at stochastic orientation, spins would have partitioned with equal probability to either the ↑ or ↓ channel on encountering the magnetic field, based on simple projections. They are ordered into distinct sets through an overlapping “active” role of the magnetic field, which provides a force to pull the spins (and their atoms) into two well-defined populations. (There is a subtle difference from the effects of the field in the EPR magnet. In a biological context, the spins are from unpaired electrons in molecular orbits, often locked into polymeric structures, and unbudgeable, so EPR explores the alignment and flipping of a population of spins, initially in a stochastic mix of orientations, and in spin echo applications, of their relaxation.) The observed distribution into well-defined populations does not reflect an intrinsic dichotomy of propensities in the incoming population, but its stochastic nature, the initial partitioning, and the active alignment.

2) The alignment of photons at a polarizer does not require dichotomy. The assumption of dichotomy embedded in the treatment of electrons led to the standard orthonormal treatment, and was generically applied to the treatment of photons, but it is even more obviously inappropriate in that case. The observable properties, the polarization vector and the frequency, are both determined by the generating transition. A polarization analyzer has refractive elements in orthogonal orientations, partitioning photons into H or V polarized populations, but, although this might suggest an intrinsic dichotomy in the photon population, 200 years of experimental application (cf.108,109) shows no justification for any such property. In those cases where coupled transitions yield pairs in dichotomic spin states, the correlation is determinate, and well-explained by conservation laws. In entanglement experiments, the polarization vector of each partner is the critical property tested through refractive interactions. The alignment observed can be explained economically by the quantized nature of each elemental interaction. When refraction occurs, the full action (hv) is expended in the displacement of an electron (the initial quantized event) over a timescale represented by 1/v. The lossless recoil returns a photon with the same action and with the vector of the reverse displacement. The displacement vectors available in the polarizer are aligned along the polarization axis, at angle φ in the ordinary path (or orthogonal in the extraordinary path). With the photon similarly aligned, the displacement and its reverse have the same vector. If not aligned, the behavior will reflect the Malus’ law probability (cos²θ) that a photon with its momentum vector at λinit can induce the transition along the vector φ, with θ = (φ - λinit). For the fraction given by this probability (the remaining fraction goes to the orthogonal path), the
photon is returned with a vector reflecting the recoil, with $\lambda_{\text{final}} = \phi$. Subsequent refractive engagements would involve this same vector for electron displacement, to give the polarized outcome determined on measurement. With an ordered population, at any particular polarizer orientation, all photons will contribute to a mean to give the Malus’ law yield; with a stochastic population the distribution will follow a $0.5\cos^2\theta$ curve, centered on the polarizer vector (see Part A 5, (iii)). This explanation is local, realistic, and completely natural, is consistent with 200 years of experimental validation, and requires no intrinsic dichotomy.

3) *The matrix operations do not describe a natural process leading to alignment.* In protocols using photons to test Bell’s theorem, the three frames involved (that of the correlated photon pair, and those of the polarizers at two separate stations), are defined by space-like separated actions, generation for the pair, and measurement at separate stations. Evolution of the partner photons to measurement necessarily involves independent space-time histories, as is well-understood in the entanglement community. Orientation for a particular state should entail reference to its independent frame, because the frames can be (and, in many protocols, are) rotated independently. For the photon pair, the frame is initially provided by the transitions generating the partners, but this information is discarded in the NL treatment. It may be mathematically elegant and convenient that the matrix operations actualize the vectors of the photon pair in the common frame of the fixed polarizer, but this is a *mathematical device* that includes no work term or physical mechanism for alignment, and no recognition of complications associated with the distribution of a conjoined state over space-like separated locations, and the processes leading to their subsequent separate actualization.

4) *Quantum ‘magic’; identifying the Maxwellian demon.* In operation of the Pauli matrices, the only vectors known are those of the polarizers. The spin states represent correlations, not vectors. In the math, the binary matrix elements represent potential spin states for photons of a pair, the matrix operation involves both polarizers, and all processes occur together. However, the real actions are space-like separated, so neither station has access to all the information needed for the full operation. The expectation of rotational invariance depends on alignment of the photon frame with the fixed polarizer. Even if we accept the convention of an ontic dichotomy, the spin quantum numbers, $s = \pm 1$, carry no vectorial information, neither do the spin states, nominally $H$ and $V$, derived from them. Only the spin states are represented in the wavefunction. The correlation through the phase difference may appear to be vectorial but could become so only by reference to a real vectorial frame. The only frame defined is that of the polarizers. Nevertheless, the matrix operations lead to actualization of photons in alignment with that frame. As briefly noted above,
Graft has recently raised a similar worry in discussion of what he calls the projection postulate problem, which was an important topic of debate first introduced by Dirac in the first edition of his book in 1930, and shortly after endorsed by von Neumann. As Graft puts it, the “...Dirac-von Neumann projection transforms the pre-measurement density matrix to the post-measurement density matrix...”. With this pedigree, it is not surprising that when the topic was revisited in 1951, the postulate was widely accepted. This ordering of the state would, in the real world, require a work term, but none is available. In effect a Maxwellian demon is then needed to create order out of disorder. From the above, the action of the demon comes from the assignment of a vectorial agency to the spin states. In the math, this works because the operation is vectorial, but the matrix elements are binary; at station 1, either an H photon represented by a 0 element in the matrix becomes a vector aligned with the reference polarizer at 0, or a V photon is actualized orthogonal. In either case, the partner photon at station 2 is actualized in the complementary orientation. Then on rotation of the variable polarizer, the outcome is as shown in Fig. 2. But this alignment is not natural; in effect, the propensities pre-determine the alignment through the projection postulate.

5) The vectors required for projections at the variable polarizer (station 2) depend on instantaneous transfer of information about the partner photon actualized at station 1. As noted above, the matrix operation implements all processes in the frame of the reference polarizer, but without any recognition of problems associated with spatial separation. In real experiments, no information is held in common by the experimentalists at the two stations. The polarizer vectors are known, but only locally. The matrix operation leads to simultaneous actualization of photon vectors, both in the frame of the fixed polarizer, correlated through their phase difference, and aligned through their propensities. However, the partner photons are necessarily at the two separate stations. If the reference polarizer is at station 1, the vectorial information becomes available there, but only at the instant of actualization. Actualization of the photon at station 2 with vectorial propensity expected from the event at station 1 could then not happen simultaneously, because the information needed at station 2 could not arrive in time. The absurdity would disappear if the superposition consisted of discrete photons in pairs carrying propensities as real properties, - essentially as vectors in a frame pre-aligned with the reference polarizer, - but such properties are not allowed in the superposition. Only in the vLR model do photons have real vectors as intrinsic properties. Partners are correlated, and can be in alignment with a known polarizer frame if explicitly arranged. Then, nothing else is needed, because all the information arrives with the photon.
Even if quantum entities were always to arrive in a superposition with a dichotomic distribution of states correlated as suggested by the spin quantum numbers, the NL expectations would still require a real mechanism for alignment. There is no obvious justification either for a superposition, or for such a dichotomy, or for simultaneity, or the work done on alignment. Since the demon to implement the alignment by is now exposed as such, the treatment looks to be inadequate on all counts.

7. **What do the experimental protocols tell us about the processes needed to implement the NL expectations?**

In the experimental context, the entangled state must evolve in a spatiotemporal framework from its site of generation to measurement at space-like separated stations where real properties are actualized. As discussed above, the treatment is missing several important causal mechanistic processes connecting events. Rather than belabor these shortcomings, consider a more general question, - that of the ‘…notoriously vague “reduction of the wave packet”…’ that Bell worried about (cf.38) but left unresolved. Actualization can only happen once, and if the alignment implemented in the matrix operations is to be credible, would have to occur simultaneously in both channels. For each refractive event, the photon has to be there, fully represented by discrete properties. With PDC sources, HWPs are used both in “state preparation”, and also to provide the discriminator function. However, this compounds the vagueness (cf.45,89,94) because if physics is the same throughout, actualization in any process upstream from the polarizers would lead to actualization of both photons with real vectors. Unless physics changes on the fly, all HWPs must behave the same. In the standard interpretation, actualization occurs “on measurement”. For the purist this means at the detector; if this is taken to mean “immediately before or at the discriminator”, that would be a HWP at one of the remote stations. For example, in a seminal paper45 (Fig. 1), the discriminator function was through HWPs P1 an P2, used to rotate the frame in front of fixed polarizers before the detectors. However, in that paper HWP1 was used in channel 1 to change the Bell-state from \( HV/VH \) to \( HH/VV \). That refractive engagement would have to rotate the frame in channel 1 without effecting the ‘entangled’ partners in channel 2. It would have to be different from the one leading to actualization of a photon at P1 and simultaneous actualization in channel 2 in appropriate alignment. Refractive engagement also occurs even further upstream where HWP0 across both cones and additional BBO crystals were used to correct time lags. In another example from a Sagnac interferometric configuration94 (see Fig. SI_1A), each photon passes through six or more refractive interactions between the pump laser and the discriminating HWPs before the detectors, assumed to be the site of actualization. Each of those refractive interactions involves a specific local function critical to state preparation. Unless physics does
change on the fly, this brings up the question of why actualization didn’t occur at an earlier refractive element. Unless it occurred at the first element encountered, the failure to actualize the pair on any subsequent engagement has to be explained away by a different specific mechanism and auxiliary hypothesis for each. Actualization at the first element would solve this dilemma, but if that happened, the lifetime of the state in superposition would be limited to a short (ps) span between the pump laser and the first HWP encountered, where nothing else happens. The transient superposition would then make no difference to the outcome, because the “state preparation” processes subsequent to the first encounter would all be determinate. Superposition is for practical purposes redundant, and any quantum ‘magic’ a fantasy.

It will be obvious from the above that, as already noted, the disparity between the determinate population generated by PDC and the superposition required by the conventional QM treatment entails a degree of schizophrenia from the purists.

A “QM simulation” mode can be called in the program, in which pairs are in Bell-states $VH/HV$ or $HH/VV$, photons are labeled as $H$ or $V$, and kets are assigned at random to stations but with equal distribution. The measurement subroutine than uses a few “If…then…else” statements (see code) to implement Shimony’s prescription for the outcome at station 1: “…by substituting the transmission axis of analyzer I for $x$ and the direction perpendicular to both $z$ and this transmission axis for $y$”, where $x$ and $y$ are $H$ and $V$ in the wavefunction equation. This is supplemented by appropriate auxiliary assumptions for the behavior at the second polarizer. (It is assumed that polarizer 1 defines the frame for actualization.) With any population of pairs in a dichotomy of states, the program then generates curves showing the peculiar invariant behavior (blue symbols in the left frame of Figs. 2, 5). This is an interesting outcome because it shows that Shimony’s summary is all that is needed; any model in the dichotomic correlation leading to that alignment would suffice. However, any suggestion that this describes a realistic process would be greeted with well-deserved derision by the wider physics community. Any such process would have to invoke a physically coherent mechanism, including a work term that could reorder the incoming flux, and synchronization of timing and differential function for the polarizers at the two stations. Since no appropriate terms are available, the routine would have to invoke a Maxwellian demon to implement the invariant behavior. However, as noted above, this is also necessary in the NL treatment. The matrix operations of the NL treatment generate the same outcome as the demon of the simulation, in effect by implementing the same alignment, also without identifying appropriate work terms. Any such process would then necessarily be encumbered by the ‘tensions’ with relativity and the second law$^{22}$. They exist because the processes postulated are unnatural.
8. **In summary, the NL predicates are surprisingly flimsy**

(i) Although quantized events must limit the precision of elemental measurements on a single quantum entity, there is no reason to suppose that this excludes intrinsic properties. The conventional justification for superposition in terms of simultaneous measurements of conjugate properties on a single entity does not apply in pairwise measurements on separate partners in a population of photon pairs. No photon can be detected twice, but that does not matter because we have two of them, and only need to detect each once. Refractive interrogations return $E$ in full, so no uncertainty invoking a causal role for $h\nu$ can be justified.

(ii) When determinate and separate $H$ and $V$ populations are generated by PDC, they cannot be in a superposition and are not indeterminate. This is particularly obvious in configurations using a Sagnac interferometer (see SI, Section 3 ii) b), c)). In more conventional applications, the mixing of determinate populations from separate cones in state preparation modifies neither the discrete properties of each photon of a pair, nor the correlation between partners directed to different separate stations.

(iii) The matrix operations through which alignment of the photon and the polarizer frames is implemented represent a mathematical device, not a real process. The alignment of frames depends on assignment to spin states of vectorial agency (propensities), and their representation as binary elements in the matrices. Operation with reference of the fixed polarizer at zero then brings about alignment of frames. With a stochastic or misaligned source, the alignment involves an ordering for which no physical justification is provided. The attribution of propensities dates back to the Gerlach and Stern experiments\textsuperscript{105}, and is based on a misunderstanding of the mechanism leading to partitioning of the outcome.

(iv) The matrix operations require a dichotomic photon population, but there is no reason to suppose such dichotomy is ontic. Even with dichotomy, there is no reason to believe that spin states (potentialities) can cause differential partitioning. With PDC sources, the dichotomy is determinate, reflecting real vectors that do cause differential partitioning, but that is fully expected under classical predicates.

(v) The mathematical paraphernalia of the NL treatment is needed only because vectors inherent in the generating transition are excluded in the superposition.
(vi) The ordering of a disordered photon population occurs without a work term, - the characteristic of a Maxwellian demon, whose machinations are now explained. The assumption that the propensities align the spin states implies a causal role which has no clear basis in physics. Even if the math allowed it, the process would still lack a work-term for alignment and a mechanism for simultaneous actualization.

None of these features, - superposition, ontic dichotomic properties, alignment steered by propensities, - is justified. They do not correspond in any obvious way to the physical state tested in recent experiments with PDC sources. All are necessary for prediction of the NL expectation of fully visibility rotational invariances. None are necessary to generate the full-visibility curve under aligned conditions.

9. Experimental outcomes

In comparing simulated with experimental outcomes, it is worth noting that the simulation is ideal: pairs are generated simultaneously, all photons are counted, and the coincidence window is vanishingly narrow (there are no “accidentals”, and corrections are unnecessary). This simplicity is useful but may mask features of consequence in interpretation of important functions. In particular, refractive behaviors are modeled in terms of linear polarization and empirical behavior, without any attempt to treat the complexities in the wavy realm implicit in phase delays or elliptical elements introduced by HWP/QWPs. With this caveat, the simulation is useful in addressing the question of how well the two hypotheses survive Popper’s test\textsuperscript{112}. Does the behavior in the real-world match that expected from NL or vLR models?

i) Experiments with cascade sources

The early experiments with stochastic pairs from atomic cascades provide the most obvious challenge to local realism. The orthodox NL treatment was developed in this context and led to expectation of the full-visibility rotational invariance, which, with a stochastic state, cannot be predicted by LR models.

What would we expect from a naïve perspective? The cascade occurs on decay from an excited state (which can be generated through several different protocols) in two sequential transitions, distinguishable by their different energies\textsuperscript{23,40-42,44}. Conservation laws determine that the photons fly off in opposite directions. In all reports, coincidences on pairwise measurement at different stations were maximal with aligned polarizers, demonstrating that partners always had approximately the
same orientation in the \(x, y\) plane. From knowledge of the source, the correlated kets might be represented as \(|\theta_1^a \theta_1^b\rangle\) or \(|\theta_2^b \theta_2^a\rangle\) \(\ldots\) \(|\theta_{n-1}^a \theta_{n-1}^b\rangle\) or \(|\theta_n^a \theta_n^b\rangle\), where each partner in a pair has the same orientation, and (for example, in a Ca-cascade) is either blue (4227 Å) or green (5513 Å). Experimental protocols are designed to take advantage of all these properties, through use of different filters at the two stations is at the expense of a loss of half the coincidences. In a stochastic population, every ket must be at a different orientation, readily dealt with by simulation (as here) or analysis\(^9,^72,^83\) to give the LH2 curve. On the other hand, to match NL expectations, an alignment of each pair with the fixed polarizer would be needed to get full visibility. If this happened, all pairs would become aligned with the reference polarizer, and the outcome would be polarized. No experimental result has shown this polarized behavior. To explain this, two conditions are needed, the ontic dichotomy of pairs in orthogonal orientation (nominally \(HH\) or \(VV\)), and actualization of each pair in alignment with the polarizer frame, independent of \(\theta\), but different for \(HH\) or \(VV\), as mimicked in QM simulation mode.

Can a detailed analysis of the experimental and analytical protocols distinguish these two possibilities? To avoid disrupting the flow of the argument, details have been relegated to the SI (see Section 6A), but the conclusions can be summarized as follows:

(a) Expectations with cascade experiments have been represented by an equation suggested by Freedman and Clauser\(^23\), and presented as the QM case:

\[
\frac{R(\sigma)}{R_0} = 0.25(\epsilon_M^1 + \epsilon_m^1)(\epsilon_M^2 + \epsilon_m^2) + 0.25(\epsilon_M^1 - \epsilon_m^1)(\epsilon_M^2 - \epsilon_m^2)F_0 \cos 2\sigma \quad (eq. 6).
\]

The result claimed (after adjusting for accidentals) was a rotationally invariant sinusoidal curve approaching the full amplitude (0.5 in this accounting). However, this is not expected from \(eq. 6\). The first right-hand term \(0.25(\epsilon_M^1 + \epsilon_m^1)(\epsilon_M^2 + \epsilon_m^2)\), RHT1) is independent of \(\sigma\), so could not contribute to the sinusoidal curve. The \(\sigma\)-dependent RHT2, which alone could generate the sinusoidal curve expected, would return a maximal value of 0.25. However, it is not obvious what the two terms represent. If both represent the same process, the equation would give a sinusoid of half-amplitude, offset by a constant value for RHT1. This is similar to the half-amplitude LR2 curve, and the analytic expectation of Thompson\(^72,^83\).

(b) Since the source is undoubtedly initially stochastic, to generate the NL expectation, each pair would have to become alignment with the reference polarizer, thus ordering the population. Cancellation of accidentals would make RHT1 redundant, and a rescaled equation,

\[
\frac{R(\sigma)}{R_0} = 0.5(\epsilon_M^1 - \epsilon_m^1)(\epsilon_M^2 - \epsilon_m^2)F_0 \cos 2\sigma \quad (eq. 7),
\]
would then be needed to account for the full amplitude claimed. This is, in effect, what the matrix operations promise. However, this would entail all the problems of the conventional NL treatment dealt with above, including missing work terms and lack of mechanism. Two differences from the NL expectation were previously noted by Thompson\(^\text{72,83}\); in the amplitude (0.5 claimed, 0.25 expected), and in the lack of an offset in the experimental curves. The offset is expected from RHT1, and the loss could be accounted for by cancellation on subtraction of accidentals, but there’s no accounting for the former, and there are several other troubling features, as discussed in the SI (Section 3, i) a).

(c) In her analysis, Caroline Thompson\(^\text{72,83}\) had noted that the stochastic cascade sources would be expected to show the half-amplitude outcome (rather than the full-amplitude claimed), and she analyzed the data from Aspect’s thesis to demonstrate that this results could be naturally obtained by applying a different subtraction of accidentals. She also noted that a simple integration of pairwise expectations would result in a constant offset of the curve, not seen in results reported. The LR2 curve of the simulation (Fig. 6) demonstrates the same result as her analysis. It is centered at the middle of the amplitude scale, and consequently is offset from the minimum and maximal values by 0.25 of the scale. Since in the simulation there are no “accidentals”, these offsets must be explained solely through the Malus’ law behavior (see SI, Section 3 i)). The offsets are expected from the stochastic distribution of pair orientations (Section 6, 2); with HV/VH populations, practically all pairs will generate some coincidences even when the polarizers are nearly aligned, and some will fail to generate coincidences when they are nearly orthogonal. Integration over a population will then reflect the cancellations associated with these losses. Between these extremes, the simulation follows the half-amplitude \(\cos^2 \sigma\) curve of eq. 6. The green curves in Fig. 6, calculated from orientations of a random pair sampled from each population, map out an envelope of possible Malus’ law outcome curves, which covers the full scale. The LR2 curve (which fits the simulated points) passes through the center of the envelope. The envelope is always symmetrical about the LR2 curve, indicating that the cancellations distribute symmetrically as expected from the isotropic source. Any vectorial approach compliant with second law and relativistic constraints would generate the LR2 outcome, including the offsets, if the stochastic nature of the source was considered.

(d) Fry and Thompson\(^\text{40}\) excited a \(^{200}\)Hg cascade using a laser polarized along the axis of propagation (the ±z-axes). The singles counts measured in the x,y-plane orthogonal to that axis where independent of polarizer orientation, - the rotationally independent behavior expected at each station from the stochastic state or from a dichotomic distribution. However, strong polarization in
counts along an axis (the -y-axis) orthogonal to the laser polarization was detected in the $x,z$-plane. If the populations were intrinsically dichotomic, the photons emitted along the orthogonal axis would not be polarized.

The program includes a simulation of the Freedman and Clauser (FC) count. As with the other counting protocols, this produces the LR2 outcome simply by accumulating coincidences from pairwise counts reflecting Malus’ law behavior. Fig. 7 compares the curves for an $HH/VV$ population, either with pairs both ordered and in frame with the fixed polarizer (circles), or at random orientation (diamonds), generated either by simulation of the FC count (cyan), or by a conventional coincidence count scaled to $\cos^2\sigma$ (red). The curves generated using the FC count are all half the amplitude of those from coincidence counts, reflecting the normalization protocol above, but otherwise show the same features.

It is perhaps a little droll that, had the Freedman and Clauser result followed the expectation of their equation, the outcome would have shown the half-amplitude LR2 curve, as expected from the Bohm-Aharonov treatment. This would have given a limit of $S_{QM}\leq 1.414$ when scaled to the conventional ±2. Since the LR model they expected was Bell’s zigzag showing a full amplitude curve and $S_{LR}\leq 2$ at canonical angle difference, this would have engendered an entirely different discussion of the “inequalities”.

ii) Experiments with PDC sources.

Populations generated by PDC have correlated pairs in two separate orthogonally polarized output cones, $H$ and $V$ with respect to the pump laser as reference. In the vLR model, these properties would be sufficient to account for all results except the full visibility of rotational invariance. In view of the inadequacies already covered, some skepticism might be appropriate. Have full visibility and rotational invariance both been consistently observed under all protocols? An alternative explanation for full visibility is that the alignments needed to generate NL expectations were established by refractive manipulations during state preparation. There are ambiguities in most descriptions of protocols that make it uncertain which explanation applies. This is an incomplete list but seems representative. Outcomes that match those reported can be readily generated in my simulation; some examples are shown in Fig. 8 and others are discussed at greater length in the SI (see SI, Section 3 ii) and Figs. SI-2A,B). The analysis there suggests the following conclusions.
a) The photons reaching the measurement stations are not in a superposition of states. The ever-increasing technical ingenuity of the experimentalists in design of optical configurations for testing the entangled states requires specific and detailed information about the determinate properties of photon pairs generated by PDC, and explicit processes for refractive interactions used in state preparation. These populations are not in the indeterminate superposition essential to the premises of the NL treatment. Especially in the Sagnac interferometer-based systems, where all photons detected arrive at the measurement station in fully polarized beams, the notion of an indeterminate superposition is simply untenable except as a philosophical abstraction (see SI, Section 6B, (iii)).

b) Simultaneity of actualization is not needed to get the NL result. Whatever process is involved, it is not adequately described by the matrix operations. In recent tests, the communication loophole has been “closed” by testing entanglement at stations well-separated from each other and from the source. In several of these, the distances between source and the two stations differ significantly. As a consequence, entangled photons arrived at their measurement stations at different times, and in several cases the difference in travel times was very substantially greater than the coincidence window width in the ns range. Reconciliation of this difference could be achieved only by time-tagging photon arrival at each station (using coordinated atom clocks), and looking for the coincidence peak on sliding the data set at one station with respect to the other. For example, in Ursin et al., the source and one measurement station were located on La Palma, and the other measuring station was on Tenerife, and the time difference between detection at the two stations was 487 μs. When data sets were centered at that time-difference, the coincidence window was 0.8 ns. In Yin et al., using a Sagnac-configured source in a satellite, the travel distance to measuring stations separated by 1200 km varied between 500 and 2000 km, and experiments required a sophisticated “…high-precision acquiring, pointing, and tracking (APT) technology…” to coordinate the space-time differences between different stations and the satellite source. This was achieved by transmission of additional information about orientation (via polarization of a separate supplementary laser) and about distances from the source (via a separate pulsed laser and GPS-synchronized atom clocks). This additional information was necessarily full determinate. As in, time-tagged data sets were aligned by sliding to find the coincidence point. In Handsteiner et al., also using the Sagnac configuration, the buildings housing the receiver stations were at significantly different distances (~500 m) from the source, and coincidences were again found by sliding time-tagged data. Clearly, in these experiments there is no question of the simultaneity of actualization implicit in the matrix operations.
c) **Closure of the communication loophole does not exclude local realism.** In experiments by Weihs et al.\(^4^8\), by Handsteiner et al.\(^6^9\), and by Giustina et al.\(^6^6\), closure of this loophole has been used to test for conspiracy in setting discriminators. Switching of discriminator settings between fixed values exploring canonic angles differences, was triggered by a random number generators in\(^4^8,\(^6^6\), or by the random arrival of photons from separate cosmic light sources in\(^6^9\). In\(^4^8\), there was little ambiguity in description of the protocol for the data shown, and simulation (see Fig. 8B) showed that the outcome could be readily explained within vLR constraints, and thus supports local realism (SI, section 3 ii), e)); in\(^6^9\), the settings themselves were only briefly described, but with ambiguities similar to those discussed elsewhere in this paper. Unless the ambiguities are clarified, these experiments do not exclude local realism.

d) **The NL outcome does not depend on each station receiving a state in ontic dichotomy.** In the Sagnac interferometric configuration (Fig. SI_1A), all beams reaching the measurement station are polarized, and the results reported can be accounted for in the vLR framework simply on that basis (see SI 3 ii), b), c) and Fig. SI_1B-D). For each PDC process, the signal and idler beams are separated by a PBS. The beams arriving at each station come from different PDC processes, and the separate ‘entangled’ beams from each PDC go to the two different stations, but are not in a superposition of states. The mean singles-counts at each detector were 0.5, which might suggest a dichotomic superposition of states, but what the simulation shows is that it is not a consequence interference, but of simple subtraction of polarized beams in a well-determined but out-of-phase correlation (Fig. SI_1B-D). Since the \(H\) and \(V\) beams arriving at each station involve different PDC processes and the entangled beams are separate, there is no realistic way to describe the behavior through the conventional NL approach.

e) **No feature of the NL treatment is needed to explain the behavior reported.** The alignment of frames in the matrix operation depends on actualization with reference to the fixed polarizer. With PDC sources, the special rotational invariance could then be trivially demonstrated: - for any protocol demonstrating the full amplitude sinusoid, rotate the reference polarizer and repeat the experiment, without changing other settings. If the outcome is the same at any setting of the reference polarizer, the case is made. In all experiments reported based on PDC sources, a much more complex procedure, inspired by the canonical settings, has been used. In all of these protocols, the polarizers have been fixed, and discrimination has been implemented through beam rotations using HWPs or EOMs. The problem of where actualization occurs in such protocols has been addressed in Section 7 above. Since Malus’ law behavior is, in all cases, invoked at the po-
larizers, photons must be actualized with appropriate orientations before they get to the discrimi-
nating elements. However, this does not fix where actualization occurs because HWPs and
EOMs both implement beam rotations through simple refractive processes; if actualization is im-
plemented there, it would, unless a different physics pertains, also be implemented in any refrac-
tive event further upstream. Despite the vagueness of interpretation, the manipulations in ‘state-
preparation’ are for the most part quite detailed and specific, and clearly depend on knowledge of
real properties. If all protocols start with the determinate $H$ and $V$ orientations of the output from
PDC, and subsequent manipulations involve well-defined refractive behaviors, they will give the
same outcome whether or not entangled state had been in superposition before the first refractive
element. Actualization is needed only if the pairs were initially in superposition; if real states are
involved, then none of the algebraic paraphernalia of the standard NL approach is needed. What
is always needed for full visibility rotational invariance is the alignment of frames, implicit in the
paraphernalia in the NL approach, but always explicit in vLR treatments.

f) How believable are the outcomes claimed to exclude LR interpretations? As an example of am-
biguity in description of protocols, in recent reports from the Wong lab $^{89,94}$ either using a con-
ventional PDC protocol, or one designed around a Sagnac configuration $^{68,94}$, settings for the po-
larization analyzer in channel B were reported as $0^\circ$, $45^\circ$, $90^\circ$, or $135^\circ$. Implicitly, these represent
beam rotations implemented by the HWP. Then, values for Bell’s $S_{NL}$ parameter consistent with
full visibility at all settings were reported when the dependence on angle difference, $\sigma$, was ex-
plored on rotation of the discriminating HWP in channel A. Since the polarizer in channel A was
set to $0^\circ$, it might be assumed that the polarizer at B was also set to $0^\circ$. However, the $45^\circ$ or $135^\circ$
rotation of the beam in channel B would then result in distribution of photons with equal proba-
bility to $H$ and $V$ channels of the analyzer. This would generate a zero-visibility outcome curve
(open squares in Fig. 2, and similar in other Figs) because in pairwise comparison the $H$ or $V$
character would be randomly assigned. The full-visibility result claimed would then contradict
expectations from standard physics. The paradox could be resolved by recognizing an ambiguity
in the description; the published account refers to settings for the polarization analyzer without
explicit recognition of the two separate components. If the polarizer in the reference channel (B)
is not set at $0^\circ$ (as implied), but at the same angle as that to which the beam is rotated, the out-
come is vLR compliant. This interpretation is not in contradiction with the protocol as reported,
and simulation of this configuration then gives the same set of phase shifted LR0 curves as reported (Fig. 8A, curves 2, 4, and Fig. SI_2). Similar configurations involving an $HV/VH$ beam rotated by 45° which falls on a polarizer implicitly at 0° are the most common type of ambiguity.

An alternative arrangement using the same experimental setup could have been used to implement a slightly different configuration, as in Weihs et al., in which there is little ambiguity, at least in the description for the Figs. reported. For the data shown, both the frame shift and switching between rotations to explore changes in polarizer difference were made at the Alice’s discriminator while Bob’s was held fixed at 0°. This configuration (see Fig. 8B) leads, on simulation under vLR constraints, to the full-visibility phase-displaced outcome curves shown. Similar configurations could have been set up both in conventional protocols based on overlapping of cones, and to Sagnac interferometric implementations (see Fig. 8B and Fig. SI-2). It is important to emphasize that the outcome from the Weihs et al. protocol supports an LR model.

In conclusion, for every case considered, an alternative configuration involving trivial adjustment to the protocol allows the same set up on implementation in my simulation to generate an LR outcome that accounts for the behavior observed. The take-home message from this section is that, within the limits of ambiguity, the vLR model could perhaps account naturally for all outcomes claiming to exclude local realism, and without contravention of standard laws. NL expectations of full visibility when photons at ± 45° fall on a polarizer implicitly at 0° (and similar ambiguities in45) require a different physics than Malus’ law. If physics changes on the fly some mechanism that does not invoke a demon is needed to explain how.

10. **Closure of communication loophole does not exclude local realism.**

Several protocols, including some of the above, have been set up to eliminate loopholes involving communication between stations, generally by excluding subluminal transfer of information (see Sections 3, 4, and 6 of SI). Many of these involved impressive engineering feats, including measurements using astronomical telescopes at stations separated geographically by large distances (144 km in113), or with the PDC source in a satellite, with ground-based measurement stations separated by 1,600 km in68, or by rapid switching between specific analyzer settings or by beam rotations, with photons in flight47,53,65,113,115. Demonstration that closure of communication loopholes did not affect the outcome have been claimed to exclude local but allow non-local models.
It needs to be re-emphasized that, although these experiments would eliminate models depending on subluminal communication between stations, they are of no interest with respect to Einstein’s model as represented in the vLR simulation. When photons carry intrinsic properties, vectorial information arrives with the photon. Since correlation between partners is conserved in all models, no other communication is needed, and a natural behavior at the polarizers can account for the observations reported. On the other hand, superposition, indeterminacy, and actualization in alignment on collapse of the wavefunction, require assumptions that both conflicted with the second law, and need superluminal communication between stations. These inconsistencies cannot be shrugged away. If the speed of light is the limit, closure of communication loopholes would, under conventional NL predicates, eliminate all such non-local hypotheses from serious consideration. The assertion that transfer of information is unconstrained by laws pertaining to the rest of physics might be thought to rescue the NL approach, but this lands one clearly on the metaphysical side of a useful demarcation (cf. and Section 3 of the SI). The vLR model is the only one for which expectations are impervious to closure of communication loopholes.

11. Conclusions

The Bell’s theorem community has fostered a model that, by excluding consideration of vectors intrinsic to the initial state, by using spin quantum numbers to justify dichotomic states appropriate to orthonormal assumptions, by assigning causal agency to those states so that, on implementing a mathematical device, they generate vectors in the alignment needed, then predicts outcomes in contradiction with standard constraints that govern the rest of science. If outcomes consistent with this model had been convincingly demonstrated, science would indeed be on an exciting path. However, for each of the reports examined above, the simulation demonstrates a simple explanation that is better aligned with fundamental QM precepts, compliant with LR constraints and classical refractive behavior, and is, within ambiguities, consistent with the published account. The predicates of the conventional NL view are essential to expectation of results showing non-local behavior, but it is not obvious either that they are justified, or that the results have been demonstrated, so re-examination is surely necessary.

The response from my colleagues has been varied, but can be paraphrased as follows (cf.):

“The many different interpretations do suggest that no QM explanation is entirely satisfactory, but they all have as a common feature a non-local perspective. Within the Copenhagen interpretation, the
probabilistic arguments preclude LR explanations, and in that context, carefully designed experiments routinely show outcomes consistent with non-locality. The results are well understood, all the loopholes have been addressed, and there is no satisfactory mechanism consistent with local realism that accounts for them. That’s where the adventure in physics now lies; things really are spooky, but the paradoxes point the path to a future understanding”.

Before such a path can be found, some agreement is needed on where the paradoxes lie. Unfortunately, the literature belies the assurance of my colleagues that the problems are well understood. The exclusion of local realism based on probabilistic constraints is applicable only if the Copenhagen interpretation is well justified. Since there are real problems with that, perhaps it has been given undue weight. The ad hoc features of the predicates, and the determinacy issues are ignored, and the conflicts with relativity and the second law are either glossed over, glorified, or cloaked in metaphysics (cf. 118-123). The appearance of obfuscation is compounded by invocation of Bell’s limit of ≤2 as excluding local realism, - the only justification in many recent reports 34,51,54,65,68,69,113. The vLR model, though an obvious alternative, does not seem to have been seriously considered, likely because it would have been excluded by the Furry constraints 11. This leaves the community open to the accusation that it is acquiescing in an essentially metaphysical view of physics and philosophy 24,119,120,124,125. Until the resulting confusion is cleared up, the “kind but bewildered listener”92 might legitimately retort that the line between sophistication and sophistry seems thinly stretched.

I do not claim my vLR model excludes alternatives; the results obtained under protocols cited herein were all interpreted as showing non-locality. However, my simulation shows that all protocols examined that use PDC sources could have been configured so as to demonstrate a natural LR outcome. My criticisms of the published accounts involve ambiguities in the description of protocols, which on an alternative reading would allow a minor change in protocol that could also lead to an LR explanation. In most cases, the non-local interpretation depends on an arrangement in which a beam with photons implicitly in H and V orientation is rotated through 45° by a HWP or EOM to give photons oriented at ±45° so as to fall on a polarization analyzer at 0° (or equivalently, photons implicitly in H and V fall on an analyzer at 45°). In these configurations, all photons would have an equal probability of emerging in either the ordinary or extraordinary ray, and no information about the initial configuration would be available on measurement. This gives the flat outcome curve in my simulation expected from analysis. The outcome claimed would then require a different physics than that which was applied in design of the rest of the protocol.

With clarification, the information already available would no doubt be sufficient to distinguish between these perspectives; perhaps physics can change on the fly, the world is non-local, and
this essay will never be published. However, for each published report examined, the alternative vLR explanations are consistent with LR constraints and (within ambiguities) with the protocol described. They can be trivially generated and could be used to test hypotheses. In general, and mindful of Einstein’s caveat (“…but not simpler”), Occam’s razor favors the vLR model, because, despite its naivety, it accounts for more with greater economy than the conventional NL one.

The conventional rejection of LR models has depended heavily on the Furry constraints\textsuperscript{11}. Those apply only if the superposition of the Copenhagen interpretation is required, along with the sophisticated math to deal with its resolution. If that treatment sets up a demon, it is likely unsoundly based. I am open to any demonstration that an alternative NL treatment can without ambiguity fully account for the experimental observations while excluding local realism, so long as it does not create a demon.

Theoretical physics has obviously advanced since the Bohr-Einstein controversy and its resurrection in Bell’s argument; indeed, Wikipedia currently has some 14 different interpretations of the nature of quantum reality, the central question in the entanglement debate. From the arguments above, there are now two clear conclusions. Firstly, that the one model that has been tested, Einstein’s, should not have been excluded; secondly, as noted earlier, that the plethora of interpretations reflects the fact that no experimental test has yet been devised to exclude any of the others. This state of play is obviously unsatisfactory from a philosophical perspective\textsuperscript{112}.

In developments following Dirac, Maxwell’s view of light as consisting “…of a coupled oscillating electric field and magnetic field …”\textsuperscript{126} has been coopted as appropriate for treatment of photons. As the wave propagates, the fields interact with “harmonic oscillators”, mostly contributed by electrons in the medium through which the light passes. A colleague has suggested that “action at a distance” is implicit in invocation of Maxwell’s fields in the wavy realm and can explain the simultaneous actualization at space-like separated locations. Does this help the argument?

That the neutrality of the photon is problematic has already been noted, - they cannot interact through coulombic forces. When first presented, Maxwell’s treatment provided a spectacular synthesis, - a thing of beauty uniting all current theories of the behavior of light through representation of their wavy nature, but it required transmission through a medium in which the fields were engaged. The first blow to this edifice was the Michelson-Morley experiment that showed no effect on the speed of light measured when the orientation of the apparatus was changed with respect to earth’s trajectory through space, interpreted as showing that no effect of
such medium could be detected. Then Planck developed quantum theory requiring quantized interaction, and somewhat inconveniently, this was extended by Einstein in explaining the photoelectric effect to demonstrate a particulate perspective. These impediments have been brushed aside by invoking particle/wave duality.

In dealing with interactions between photons and electrons in quantum electrodynamics (QED), the probability expressions are essentially the solutions of Dirac’s equation\textsuperscript{10,37} for the behavior of the electron's probability amplitude, and Maxwell's equations for the behavior of the photon's probability amplitude. In interpretations through Feynman’s path integral approach, probability amplitudes are represented in Feynman diagrams. Although Feynman elegantly simplified the mathematical underpinnings\textsuperscript{127}, the formal equations are highly technical, and, in view of my limited math skills, I must leave any depth in analysis to the professionals. But, from Feynman’s account, pathways are quantized, and to approximate the effective path, the integrals must include probabilities for all paths. Interference effects then come into play because phase delays from differences in the length and refractive index of the paths, and from any different frequency for the photons\textsuperscript{127}, then tend to cancel contributions from paths more convoluted. Feynman showed that the path integral approach was equivalent to the Schrödinger equation applied to a stream of photons propagating in a wave\textsuperscript{107,128, and see A. Zee in Introduction to }\textsuperscript{127}. In emphasizing that the treatment must embrace the particulate nature of quantized events, and that expectations must therefore be expressed probabilistically, Feynman found a brilliant way to make Maxwell’s electromagnetic treatment work, but the hybrid approach is ultimately futile because it cannot be adapted to processes at the elemental level that necessary involve neutral photons.

In the wavy domain, the medium needed to support the propagation of a light wave is nowadays supplied through interpretations of the vacuum state, which is “… not truly empty but instead contains fleeting electromagnetic waves and particles that pop into and out of existence…”\textsuperscript{129}, allowing particle/antiparticle pairs to fleetingly spring into reality on interaction with the fields of the passing photon wave. However, in light of the Michelson-Morley result, these interactions cannot involve any reaction time. Since quantized interactions involve engagements of their action that are frequency dependent, and necessarily linked to time, something else has to be invoked to make this realistic. Mechanisms involving electron-positron annihilation generating 2 gamma rays photons, or a high-energy photon generating an electron-positron pair, are well
known, but these don’t fit the bill; whatever happens has to be independent of frequency. Perhaps the electron-positron pairs of vacuum bubbles are involved\textsuperscript{130}: “…Such charged pairs act as an electric dipole. In the presence of an electric field, e.g., the electromagnetic field around an electron, these particle–antiparticle pairs reposition themselves, thus partially counteracting the field (a partial screening effect, a dielectric effect). …This reorientation of the short-lived particle-antiparticle pairs is referred to as \textit{vacuum polarization}…” Replace “around an electron” with “of the light wave”, and we have a candidate medium. However, such interactions are said to “…have no measurable impact on any process…”\textsuperscript{130}; if this is so, the polarization through repositioning can involve no process engaging the known properties of photons, and is therefore immune to measurement. This would then be just another hypothesis that cannot be tested.

A more natural approach to dealing with photon neutrality would have to start by recognizing that the electromagnetic perspective is necessarily \textit{inappropriate}. An alternative is to invoke quantized transfer of momentum at the particulate level, where electromagnetic properties are not needed. The general approach should be familiar from its use in explaining Compton scattering (cf.\textsuperscript{131}), but momentum transfer is often invoked in other treatments where appropriate. In considering refraction at the elemental level, each path is defined by specific quantized refractive events, but all outcomes still reflect classical parameters (cf.\textsuperscript{132}). Bialynicki-Birula \textsuperscript{133} suggested a photon wave function in momentum representation which allows quantized interactions of neutral photons with electrons to be represented at that level. A more recent paper shows that the outcome is quite compatible with those of QFT approaches via Maxwell’s equations \textsuperscript{134}. In on-line discussion, “Maxwell”, an anonymous commentator\textsuperscript{135}, finds a hybrid treatment in which a wavy expression is developed following a form similar to those of QED, and a particulate expression in which the total momentum of the photon is applied at the quantum level along the polarization vector.

Photons appear to be electromagnetic not because of any electrical or magnetic properties, but because they are generated by, and their energy is used in, displacements of mass. In EPR and H-NMR, the difference in energies of the photons flipping the spins reflect the \(~1,836\) difference in mass. In the UV/VIS/NIR range, the most easily displaced mass, the electron, is charged, and the appearance of electromagnetic behavior of photons then simply reflects this fact. In generation, the transition of an electron to a lower energy involves a spatial displacement that yields a photon with energy, \(E = h\nu\), equal to the energy lost by the electron. The photon has
a momentum vector along which action can be applied (the electric vector in the classical description), which is determined by the vector of the displacement generating it. The photon can be thought of as a wave packet occupying $\lambda = \nu/c$ of space-time, which propagates through a medium at $c/n$ with overall momentum $p = E(n/c)$ (where $n$ is the index of refraction and $c$ the speed of light *in vacuo*). When a photon transfers its energy to an electron, the momentum vector must match an available displacement vector for the electron, the energy involved in displacement must match $hv$, and the timescale of displacement must match $1/\nu$. Events can be absorptive or refractive. If a higher-energy orbital (atomic or molecular) is unoccupied and the photon energy matches the energy difference, displacement will lead to absorption. On the other hand, if the electron has nowhere to go, it has little option but to recoil in a loss-less refractive event returning the energy as a new photon. Whether justified in terms of momentum vector or the electric field vector, projections leading to a Malus’ law outcome will be the same.

A neutral photon does not require any medium to propagate through space. Passage of a photon through a refractive medium will involve sequential refractive events each requiring transfer of momentum and its return, calculable from refractive index, path-length, frequency, etc. Many potential paths will be explored on passage of a population. The question as to how phase-differences are conserved if a process involves many sequential events and paths, was solved through the path integral approach in Feynman’s treatment, an approach that would work just as well in momentum transfer. Nothing in this description should be controversial. In contrast, in the wavy treatment, the observer-dependent non-local reality is weakly founded, in conflict with relativity and conservation laws, is open to metaphysical exploitation, and could be abandoned for any of those failings.

My initial interest in the entanglement debate arose from a concern that invocation of such a blemished philosophy undermines the scientific enterprise; if relativity and the second law are disposable, what can be held dear? Perhaps what is needed is not new science, but a fresh look at the old science. Maybe a reformulation of wavy explanations to comply with locality constraints applicable to the particulate nature of the neutral photon and quantized nature of its interactions could be productive. From the wider perspective of entanglement, a demonstration that Einstein’s local realism can account for most behaviors previously taken as justifying non-locality should be welcome, and its simplicity makes it a more attractive point of departure. In the narrative above, and in the SI (Section 4), I explore a number of cases in which the perspective of photons as momentum carriers leads to simple particulate explanations for problems, in particular refractive behaviors and interference.
effects, usually considered in the wavy domain. In an admittedly speculative discussion, I extend this viewpoint to cosmological controversies.

The simple perspective coming from Einstein’s ideas might remove tensions, clean up the obfuscation, eliminate justification for metaphysical speculation, and perhaps even put natural philosophy back on a coherent track. If nothing else, my critique could focus the target better.

Acknowledgements. The biologist claiming to understand entanglement is a standing joke among quantum physicists, so I first wish to thank Mike Weisman, Martin Gruebele, Greg Engel, John Miller, and Paul Kwiat for their patience in helping me to understand some of the issues in early discussions. I also thank Thomas Angelidis for an early correspondence relating to his LR model. I acknowledge with gratitude a useful correspondence with Sean Barrett before his tragic death in 2012, and ongoing correspondence with Arto Annila which has help me to appreciate quantum field theories, and with Richard Gill and Jan-Åke Larsson who have provided a Bell-friendly perspective from experts in debunking criticisms of Bell’s approach. I am particularly grateful to Tony Leggett for his permission to cite the distinction noted in his cases 1 and 2, and for continuing discussion, correspondence, and criticisms, always delivered with sympathy, that have corrected misconceptions and deepened my appreciation of the subtleties of the Bell’s theorem debate.
Figure Legends

Figure 1. Schematic of a typical experimental set-up for an entanglement experiment\textsuperscript{12,19}. A laser polarized at 0° (H) excites a BBO crystal, generating a population of pairs of orthogonally correlated photons through parametric down conversion. A filter at 702 nm (half the frequency of the 351 nm excitation wavelength) selects pairs of photons of the same color. These are emitted in ordinary and extraordinary cones, one H polarized, and the other V polarized. The cones overlap at two intersections where H and V photons are mixed. For each pair, the ‘entangled’ partner is in different overlaps so that correlated partners emerge in separate beams sampled at the overlaps\textsuperscript{45,49}, with equal but random allocation of HV and VH pairs to the two paths. Each beam then has a random but equal mix of H and V photons, with complementary HV or VH pairs in the two beams. The photons are channeled, for example through optical fibers, to space-like separated measurement stations where polarization analyzers sort each photon by its polarization, with probabilities according to Malus’ law. In the experimental situation, four detectors, one each in the ordinary and extraordinary rays of polarization analyzers at the two stations (elemental outcomes Q, R, S, and T at the four output rays are equivalent to elemental probabilities: $p_1(\lambda, \alpha) \equiv Q$; $p_1(\lambda, \alpha') \equiv R$; $p_2(\lambda, \beta) \equiv S$; $p_2(\lambda, \beta') \equiv T$), are used to measure the photons. Entangled partners are identified through the temporal coincidence. Photons are analyzed pair by pair at particular orientations of the polarizers, and the mean is taken over a significant population. The counting is handled by suitable electronics, involving timers, coincidence counters, etc., or equivalent recording procedures. Additional refractive elements (HWP0, HWP1, HWP2) ‘prepare’ the entangled pair before measurement (others not shown may also be included, see text). In most recent reports, discrimination is implemented by rotation of the photon frame using HWP1 and P2, while the polarizers are static. The simulation can model this set-up by using HWP1 for P1 at station 1, or HWP2 for P2 at station 2, with the polarizer fixed. However, to keep the program simple, the default mode is to use the polarizers themselves for discrimination, with the variable polarizer at station two, and function implemented by rotation of the polarizer. HWP1 and HWP2 can then be used for “state preparation”. The simulation assumes classical properties for refractive components, and perfect optical elements so that no photon is lost to imperfection.

Figure 2. Typical outcomes from vLR or QM simulation, and a unit circle used to illustrate vector projections for aligned vLR or QM model expectations. Left: A Run3 cycle generates points at canonical polarizer settings. A vLR simulation using an ordered HV/VH population aligned with polarizer 1 at 0° (LR0 curve, ●); a stochastic population at the same settings (LR2 curve, △); or an
ordered $HV/VH$ population, but with polarizer 1 at $45^\circ$ (zero visibility curve, □). Curves generated in “QM simulation” mode with a stochastic population (□), or with $HV/VH$ photons, but with polarizer 1 at $45^\circ$ (△), or with ordered pairs at any orientation, fall on the QM curve (blue). Right: the unit circle representation. All vector operations can be represented without algebraic sophistication. The vectors of the polarization analyzer at station 1 are at $0^\circ$ (for the ordinary ray) and $90^\circ$ (for the extraordinary ray). Under vLR premises, with an aligned population the vectors for an orthogonal $HV$ (or $VH$) photon pair overlay these, with orientation of photon 1 at $0^\circ$ if $H$ (yellow), or at $90^\circ$ (purple) if $V$. Under NL treatments, the dichotomic spin propensities determine that the matrix operations align the photon frame with the polarizer frame to give this same alignment. The black diagonal lines are the unit vectors for 25 angle settings of the variable polarizer 2 ($\theta_2$) spaced at 7.5° intervals, restricted to one hemisphere between $\pm 90^\circ$ for clarity. These settings were varied with respect to polarizer 1 oriented at $\theta_1=0^\circ$, so that $\sigma = \theta_2 - \theta_1 = \theta_2$. At station 1, orientation of the $H$ photon always aligns with polarizer 1, so that projection of polarizer 1 onto the two photon vectors would give 1 and 0 as values for $\cos \sigma$ and $\sin \sigma$ in the ordinary ray, with orthogonal projections in the extraordinary ray. At station 2, the red and magenta perpendiculars are projections of the polarizer vector onto the two photon vectors using the 25 values for $\sigma$, given by the intercepts at the circumference. (The projections for the orthogonal polarizations in the extraordinary ray are not shown; yields are implicit in the remainders from yields given by the ordinary ray projections.). The text boxes show (top) values returned on clicking at an intercept (here at $22.66^\circ$ (accuracy is limited by pixilation of the pointer position), for $\cos \sigma$ (left) and $\sin \sigma$ (right) (the potentialities or probability amplitudes), and (bottom) $\cos^2 \sigma$ (left) and $\sin^2 \sigma$ (right) (the Malus’ law yields, - observables or probability densities). For each polarizer 2 vector, $\theta_2 = \sigma$, the intercept at the circumference gives projections whose values on taking squares give Malus’ law yields. The relation between these observables ($\cos^2 \sigma$ values) and the potentialities ($\cos \sigma$) derives from representation in the complex plane, where the possible results from vector projections are $\pm \cos \sigma$, falling in different quadrants, and potentialities in all quadrants are explored in Hilbert space. In effect, since the only ‘moving part’ is the variable polarizer, the projections at particular settings then represent the values from Bell’s analysis, $(\vec{a}_1 \cdot \vec{a} \cdot \vec{a}_2 \cdot \vec{b}) = -\vec{a} \cdot \vec{b} = -\cos \sigma$. The total probability for detection of a photon, $\cos^2 \sigma + \sin^2 \sigma = 1.0$, and the difference $\cos^2 \sigma - \sin^2 \sigma = \cos 2\sigma$, is shown for $\sigma = 22.66^\circ$. The present implementation is limited to representation of the aligned situation. Under QM expectations, the dichotomic propensities are taken as determining that actualization is in alignment with the polarizer frame (see Part B, Section 6 for discussion of the validity of the assumption). For a vLR population, the aligned configuration shown here applies only to
conditions giving the LR0 outcome curve, when the projections (and probabilities) are the same for LR and QM models.

**Figure 3. The program window, showing typical outcome curves with different parameter settings.** Settings for controls were as follows:

**A.** The $HV/VH$ photon source of 1,000 orthogonally correlated pairs at a fixed orientation ($H$ is $0^\circ$). Malus’ law is implemented (Malus LR mode), polarizer 1 is fixed (at $0^\circ$), and polarizer 2 rotated at 50 random angles. Three coincidence count outcomes are shown (left panel), all of which follow the curves predicted from QM (LR0 curve): the default anti-correlations count ($●$); coincidence count ($□$); CHSH count ($△$, right scale). The right panel shows the single counts at all four detectors, and their mean. The Gadgets (top section of window) show values from the last set; these were examined by clicking on the Sample # slider (values for the 30th pair shown).

**B.** Photon source as in A, but with 5,000 pairs, each at random orientation. With polarizers in Malus LR mode, two successive Run1 cycles were made, one with polarizer 1 in ‘Fixed’ mode ($●$, anti-correlation; ○, CHSH count), and the second with both polarizer angles set randomly at 90 values ($□$, anti-correlation; ■, CHSH count). The correlations follow a sinusoidal curve, but with half the amplitude of QM expectations (LR2 curve). The right panel shows mean singles-counts as in A.

**C.** $HV/VH$ photon source rotated to $\lambda = 45^\circ$, with polarizer1 also rotated to $\lambda = 45^\circ$ to give a common rotational frame. In a Run3 sequence, polarizer 2 was initially set to $-45^\circ$ to give an angle difference of $-90^\circ$, and the angle was then rotated by $7.5^\circ$ between successive runs, covering a range for angle difference from $-90^\circ$ to $90^\circ$. The outcome (the mean from 5 runs with 1,000 photon pairs at each angle difference; left panel; $●$, anti-correlation; ○, CHSH count, right scale) follows the LR0 curve. The same outcome would have been found for any value of $\lambda$, as long as both photon source and static polarizer were rotate to the same value in a common rotational frame. The Analog implementation (right panel) shows calculated points and theoretical curves for Malus’ law yield differences in the ordinary rays of the polarizers for the two orthogonal photons of a pair: $\Delta I_{HV}^O = \cos^2\sigma$ (dark blue points and curve), and $\Delta I_{VH}^O = -\cos^2\sigma$, (light blue); and complementary curves $\Delta I_{HV}^E = \sin^2\sigma$ and $\Delta I_{VH}^E = -\sin^2\sigma$, (red and yellow) for the extraordinary rays. With the Malus’ law yields normalized, application to photon pair gives, for particular setting, $\alpha$, of the fixed polarizer $E_{\alpha,\beta}^{LR} = 0.5\{(\Delta I_{HV}^O - $
\( \Delta I_{VH}^O - (\Delta I_{HV}^E - \Delta I_{VH}^E) \) = \cos^2\sigma - \sin^2\sigma = \cos 2\sigma, \) etc., the same outcome as QM. Since \( \sin^2 x = 1 - \cos^2 x \), the curve is scaled to \( 2\cos^2 \sigma \), offset by -1, as with conventional scoring.

D. Random photon source, with discriminators set in Bell binary mode, and successive Run1 simulations (1,000 photon pairs) taken with polarizer1 angle fixed at 0° (closed symbols), or with both polarizers at random (open symbols). Points are shown for HV/VH source (● or ○, anti-correlation; ■ or □, CHSH count), and for HH/VV source (♦ or ◊, CHSH count and Δ, anti-correlation). All fall on the diagonals expected from Bell’s zigzag (LR1 curve). The right-panel shows the singles-counts.

E. HV/VH source, anti-correlation counts under the following rotations of polarizer 1 (P1) and photon frames: ●, P1 0°, photon 0°; ■, P1 45°, photon 45°; □, P1 45°, photon 0°; ○, P1 0°, photon 45°; ◊, P1 30°, photon 0°; ♦, P1 0°, photon 30°; Δ, P1 30°, photon 30°. The right panel shows analog points and curves for the ordinary rays (see C) that illustrate why amplitude is lost on rotation of frames.

**Figure 4. The Bell-type inequalities examined.** LR0 and LR1 curves from anti-correlation (●, □, left scale, red) and CHSH counts (●, □, right scale, gray) for populations of 1,000 pairs of HV/VH photons, simulated through Run3 cycles, with photon source static and ‘Malus LR’ polarizer mode, (LR0, closed symbols), or photon source random and ‘Bell binary’ polarizer mode (LR1, open symbols). The magenta points and symbols show the difference between the LR0 and LR1 curves (in the spirit of the \( \delta \)-function). The dashed lines show the intercepts at canonical values used to calculate the \( S_{\text{BCHSH}} \) expectations. See text for discussion.

**Figure 5. Simulation of LR and quantum mechanical outcome curves.** Successive Run1 simulations using HV/VH photon populations were performed as follows: ●, aligned frames at 0° (LR0 outcome); ○, polarizer 1 (P1) at -45°; Δ, P1 and photon frame both at -45° (LR0 outcome); Δ, P1 at -45°, QM mode; ■, stochastic photons, P1 at -45° (LR2 outcome); □, stochastic photons, QM mode. Note that at all these settings, the single-counts of the right panel showed the same pattern of behavior (yield of 0.5 at each detector).

**Figure 6. Entropic penalty for stochastic orientation in an LR population.** The sample envelope of the Malus’ law expectation curves in A, or the distribution of Malus’ law yield differences revealed through the Analog option in B, provide a visualization of the penalty. The pho-
ton population was stochastic, in $HV/VH$ state, and coincidences were detected using the anti-correlation count. A Run3 protocol using an average of 5, with polarizer 1 at $0^\circ$, generated both panels.

A: Experimental points (●), and red curve show the LH2 outcome when the photon population is stochastic. The green lines are 125 theoretical expectation curves, calculated using 360 values of $\sigma$ over the full circle, but normalized to the $\pm90^\circ$ range shown. Each point was calculated from 8 cross-products of Malus’ law yields at detectors $Q, R, S, T$ ($QS, RS, RT$, and $QT$, for $HV$ and $VH$ configurations, appropriately weighted), generated using the random angle of orientation for the first pair of each population used at the particular setting of polarizer 1 (each line showing the curve expected if a population of photon pairs with this orientation was determined at each value for $\sigma$). The curves map out a sample envelope of values, equally distributed in phase space around the mean curve, so that the mean from the stochastic population of photon pairs is extracted from within such a range of values.

B: The Malus’ law yield for each photon (from the population of 5,000 in the average of 5) in the ordinary and extraordinary rays of the two polarizers was calculated at each angle difference. The differences in yield for a correlated pair measured in the ordinary (dark and light blue) and extraordinary (red, yellow) rays were then plotted for each hemisphere. Every one of the photon pairs gave a different yield, to give a range of values in a vertical bar at each angle difference, which fall either side of the theoretical curves. These follow a $cos^2\varphi$ (Malus’ law) distribution of values, where $\varphi = \theta - \lambda$, $\theta$ is the orientation of the variable polarizer (with the fixed polarizer at $0^\circ$, $\theta = \sigma$), and $\lambda$ is the vector of the photon. The curve is centered at $\theta$, and varies with $\lambda$. For a photon population at fixed orientation, all points calculated at a particular angle difference would overlap as a single point and follow the theoretical curves $\pm cos^2\sigma$ (dark and light blue) or $\pm sin^2\sigma$ (red, yellow). The points following the LR2 curve are the mean values from coincidences, derived with appropriate sign from elemental cross-products. In the anti-correlation count used here this is implicit, but, for example from the CHSH count $(\sum_{i=1}^{n} (QS + RS + RT - QT)) / n$ (see Program Notes) it is explicit. All algorithms (with account taken of the Bell-state) follow the half-visibility curve, with envelopes to match.

Note that all photons contribute to the outcome curve (left panel), but that cancellations lead to the analytical outcome, which segregates two equal contributions, both of which depend
on \( \sigma \). This is shown in the left panel by the fractional amplitudes inside and outside the envelope, each of which at any value for \( \sigma \), sum to half the total amplitude expected from the LR0 curve.

**Figure 7. Comparison of a simulation of the Freedman-Clauser treatment (blue symbols) with the standard vLR simulation.** The treatment underlying the \( \delta \)-function of Freedman and Clauser\(^{23} \) leads to an LR outcome that matches the experimental result. Left panel: Symbols show means from measurement of 10,000 photon pairs when coincidences were counted at 100 angle differences using the FC \( \delta \)-count protocol with different sources. Points show: static \( HH/VV \) source (●, the LR0 curve at the half-scale expected from normalization and an equal mix of \( H \) and \( V \)); random \( HH/VV \) source (♦, the LR2 curve at half-scale, which shows half the amplitude of the LR0 curve, as expected from an isotropic source); and PDC EO source with \( V \) photons in channel 2 rotated to \( H \) by HWP2 at 45° (■, this generates a fully polarized source with both channels in \( H \) orientation). This last source would represent the NL outcome expect after alignment of a stochastic source with polarizer 1 set at 0° (see text). The red symbols show the outcome from the standard coincidence count at the same scale (●, static \( HH/VV \) source; ♦, random source). Right panel: mean singles-counts show the \( HH/VV \) counts, all distributed with values close to 0.5 as expected from an isotropic source. However, with the PDC EO source, where both signal (ordinary) and idler (extraordinary) channels are polarized. Then, counts with values either at 1 and 0 (at the station with the fixed polarizer), or as complementary cosine and sine curves (the polarization detected at the other station). For each value of \( \sigma \), the four mean counts at any angle difference sum to give the two photons of a pair.

**Figure 8. Simulation of early PDC-based experiments claimed as supporting non-locality\(^{45,48,89} \).**

A. Curve 1 shows successive runs as follows: ●, LR0 curve (cf. Figs. 1, 2, no ‘state preparation’); ○, HWP0 set at 45° (same curve). Curves 2-4 show the effects of setting “the analyzer in beam 1 at 45°”: Curve 2 (□) was obtained by setting HWP1 at ±45° in front of static polarizer 1 at 0° (one literal interpretation of the text in Kwiat et al.\(^{45} \) ); for curves 3, successive runs were taken first with HWP1 rotated by 22.5° with polarizer 1 fixed at 0° (△), then with the HWP1 at 0°, and polarizer 1 at 45° (■, which is equivalent, both showing zero visibility). Curve 4 (●) shows the outcome on rotation both of HWP1 by 22.5° and of polarizer 1 by 45°. Curves 2 and 4 show the same full visibility as curve 1, but curve 2 is phase-shifted by 90°, and curve 4 by 45°. This is the behavior shown in Fig. 7 of Fiorentino et al.\(^{89} \), and interpreted as demonstrating the QM expectation.
B. With polarizer 1 fixed at 0°, HWP2 was introduced in beam 2, in which the setting of polarizer 2 was separately varied over the full range. Successive curves show the outcome as HWP2 was rotated by increments of 15° over the range $\theta = \pm 45^\circ$. All curves show full visibility, but are phase displaced by $2\theta$, the angle through which the beam was rotated; the behavior is fully LR compatible. This outcome is, in effect, the same as that reported by Weihs et al.\textsuperscript{48} using EOMs to rotate the beams. The outcome is also similar to that found experimentally in\textsuperscript{89,94}, and claimed there to show the full rotational invariance expected from NL predicates. Although the outcome here has the fixed polarizer at 0°, and has both the offset and analysis rotations in the variable channel, it demonstrates that the configuration (which matches in essentials those in\textsuperscript{89,94}) can be used to generate outcome curves that appear to show full-visibility rotational invariance on rotation of one polarizer.
Figures

Parametric down conversion, filtration at 702 nm, and selection from the two cone overlaps gives two beams, with $H$ and $V$ photons of an ‘entangled’ pair in different beams.

Figure 1

Polarization analysis is achieved by rotating P1 or P2 (HWPs) with fixed polarizers.
Figure 2
Figure 3A
Figure 3B
Figure 3C
Figure 3D
Figure 3E
Figure 4
Fig. 5

**Coincidence counts v. Angle difference**

**Single counts v. Angle difference**
Fig. 6
Fig. 7
Fig. 8
Bibliography


