Two copies of EPR state of light lead to refutation of EPR ideas

Krzysztof Rosołek,1 Magdalena Stobińska,1, 2 Marcin Wieśniak,1 and Marek Żukowski1

1Institute of Theoretical Physics and Astrophysics, University of Gdańsk, ul. Wita Stwosza 37, 80-952 Gdańsk, Poland
2Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

We discuss the four mode squeezed vacuum produced in Type-II parametric down conversion. For proper phase relations it is a polarization singlet with an undefined number of photons. Even for a finite gain, Γ, its invariance properties, and perfect correlations (seemingly) allow to introduce EPR elements of reality, which are photon numbers. From another perspective such a state gives us two copies states which in the limit of Γ → ∞ leads the (non-normalizable) EPR state. A chained Bell inequality, built for the elements of reality, derived basing on the geometric concept of distance, is grossly violated. In the limit of infinitely many settings, we get a GHZ-type contradiction for the two copies of the EPR state: elements of reality point to a correlation, while quantum mechanics to anti-correlation. The proposed scheme is feasible in the low gain regime.

Introduction. The quantum realm is very distant from our daily experience, which shapes our view on what is reality. The phenomena are counterintuitive, and the formalism, which describes our observations of quantum reality, even more. The predictions have only a statistical nature. One cannot predict with certainty the response of individual quantum systems to all possible experimental situations. Some quantum predictions seem paradoxical.

The so-called Einstein-Podolsky-Rosen paradox was an attempt to show that the quantum description of reality cannot be considered complete. Elements of reality, properties of a system which can be established with perfect accuracy, without in any way disturbing it, were to be the missing component of the theory. EPR used perfectly correlated systems to show that such elements of reality are (seemingly) derivable from quantum predictions and the principle of locality (relativistic no action at a distance). There were some additional hidden assumptions in the reasoning of EPR, like freedom of an experimenter to choose the observable to be measured, and putting on equal footing the actual experimental situation realized for the given individual system, and a possible different (complementary) one. The second of these was directly addressed by Bohr in his response in his famous sentences “...there is essentially the question of an influence on the very conditions which define the possible types of predictions regarding the future behavior of the system […] In fact, it is the mutual exclusion of any two experimental procedures, permitting unambiguous definition of complementary physical quantities, which provides room for new physical laws the coexistence of which at first sight appear irreconcilable with the basic principles of science.” Bohr’s response discussed only the assumptions of EPR, and did not challenge their “technical” results.

However 29 years later, 30 years ago, Bell showed that there is a technical flaw in the EPR reasoning, if one applies it to the Bohm’s version of the paradox for a two-spin 1/2 singlet. Simply elements of reality are incompatible with quantum mechanics, as they must satisfy the Bell inequality, while quantum predictions violate it. The EPR program breaks down in the case of the simplest possible maximally entangled state. Even a more striking proof of that was given 25 years later by the GHZ trio (Greenberger-Horne-Zeilinger) for three spins 1/2. It can be put: if elements of reality exist then 1 = −1. All this now expanded to an “industry” which uses violations of various new and old Bell inequalities also in practical applications, like: reduction of communication complexity, randomness generation, device-independent quantum cryptography. The violations, in a device independent way, indicate the presence of entanglement.

For many years the following question was unresolved. Do we have Bell’s theorem also for the original EPR state? In the momentum representation the state is δ(p1 + p2), where pi is the momentum of i-th EPR “particle”. Such singular objects do not exist in the Hilbert space. Nevertheless, as distributions can be approximated by well behaved functions, one can study properties of proper wave functions which in some limit give us δ(p1 + p2). In Ref. Bell shows that the Wigner distribution for the EPR state is always non-negative, thus there is no chance for Bell’s theorem. The distribution gives an explicit local hidden variable model for all observables which are functions of positions and momenta.

Reid and Drummond in and showed that we can can have an optical approximation of the EPR state in the form of a radiation of non-degenerate optical parametric amplifier: a two mode squeezed vacuum. This opened the exciting prospect of actually observing approximate “original” EPR correlations in the lab. Bell’s Theorem for such approximate EPR states was finally given in and . The idea was to use different observables than the ones of EPR. For a given photon field mode the operator p is equivalent to 1√(a−a†) while the position x can be modelled by 1√(a+a†). One can use also a general pair of quadratures: Q(θ) = 1√2(eiθa+h.c.), where h.c. stands for the Hermitian conjugate of the previous expression, and Q(θ + π/2). Such pairs of conjugate observables lead always to non-negative Wigner functions for the squeezed vacuum. Instead of these Cohen...
used an approach which effectively led to an experiment describable in a Hilbert space larger than the original one (a specific interferometer, or coupling the EPR state to a pair of spin 1/2 ancillas). Both approaches did not relate directly to the properties of a squeezed vacuum. In Ref. 14 no enlargement of Hilbert (Fock) space was required, as highly non-classical observables were used to do the trick: displaced parity operators (the parity operator was defined by \((-1)^h\), where \(h = a^\dagger a\); this allows to transform the eigenvalues to the usual \(\pm 1\)).

Here we want to take a different approach, which directly uses photon number operators. The other specific trait of our approach is that we use a four mode squeezed vacuum, which effectively gives us an approximation of the EPR state in two copies. Entangled states of light with mean number of photons of order of ten are easily accessible in laboratories 17. The standard method of their generation employs a type-II parametric down-conversion (PDC) process, which occurs in a nonlinear crystal pumped by a laser beam 11. The process is described by the Hamiltonian \(H = ig(a_H^\dagger b_V^\dagger + e^{i\phi}a_V^\dagger b_H^\dagger) + h.c.,\) where in the notation for creation operators letters \(a, b\) stand for spatial modes, and subscripts for \(H, V\) polarizations. The phase factor, \(e^{i\phi},\) is controllable, and we shall fix it to \(-1\). The output is a superposition of maximally entangled \(2N\)-photon polarization singlet states

\[
|\Psi(-)\rangle = \sum_{N=0}^{\infty} \lambda_N |\psi_{N}^{(-)}\rangle, \tag{1}
\]

where \(\lambda_N = \cosh^{-2} \Gamma \sqrt{N+1} \tanh \Gamma, \sum_{N=0}^{\infty} \lambda_N^2 = 1,\) and

\[
|\psi_{N}^{(-)}\rangle = \frac{1}{\sqrt{\lambda_N}} (a_H^\dagger b_V^\dagger - a_V^\dagger b_H^\dagger) |N\rangle \tag{2}
\]

\[
= \frac{1}{\sqrt{\lambda_N}} \sum_{n=0}^{N} (-1)^n |n_H, (N+n)V\rangle_a |(N-n)H, nV\rangle_b.
\]

The symbol \(|n_H, (N-n)V\rangle_a\) denotes \(n\) horizontally and \(N-n\) vertically polarized photons in beam \(a\). Similar notation holds for the beam \(b\). In the case of component states \(|\psi_{N}^{(-)}\rangle\) each beam contains \(N\) photons in total. Polarizations of beam \(a\) and \(b\) separately are undefined, but due to equal photon numbers in the orthogonal polarizations they are anti-correlated. The strength of the interaction \(\Gamma\) is governed by the coupling \(g\) and the interaction time \(t\). For a strong pump, that is high \(g\), the crystal output is called bright squeezed vacuum (BSV).

As the unitary transformation leading to the state is given by \(e^{\gamma H t}\) it can be factorized to \(e^{i\delta V t} e^{i\delta V t}\), where \(H_{V,H} = ig(a_H^\dagger b_V^\dagger) + h.c.\) and \(H_{V,H} = -ig(a_V^\dagger b_H^\dagger) + h.c.\) Thus as the initial state is vacuum in all modes we get two copies of an approximate EPR state in the form of two squeezed two-mode vacua. One for modes \(a_V\) and \(b_V\) and the second for modes \(a_H\) and \(b_H\). The copies differ by a phase (sign) factor in the generating Hamiltonians.

The state \(|\Psi(-)\rangle\) has some other interesting properties. Consider a Bell experiment (Fig. 3). Two spatially separated observers, Alice and Bob observe radiation of a pulse pumped source producing the four mode squeezed vacuum. They can control the orientation of their local (linear) polarizers, defined by \(\theta_A\) and \(\theta_B\), respectively, and count photons at the local outputs. Thus, the result of the local measurement for run (pulse) \(k\) is a certain number of photons counted at Alice’s side \(n^{(k)}(\theta_A)\), and at Bob’s side \(m^{(k)}(\theta_B)\). The state has the property that for \(\delta_B = \theta_A + \frac{\pi}{4}\) one always has \(n^{(k)}(\theta_A) = m^{(k)}(\theta_A + \frac{\pi}{2})\). This is a direct consequence of the fact that the generating Hamiltonian has a form invariant with respect to pairs of orthogonal polarizations in which one expresses it: \(H = ig(a_H^\dagger b_V^\dagger - a_V^\dagger b_H^\dagger) + h.c.,\) where \(\theta^\perp = \theta + \frac{\pi}{2},\) and \(\theta_A/B = 0\) stands for \(H\). As a matter of fact this holds for any pair of elliptic polarizations. To understand this fact it is enough to recall that \(\frac{1}{\sqrt{2}}(a_H^\dagger b_V^\dagger - a_V^\dagger b_H^\dagger) |0\rangle\) is a two-photon polarization singlet, which is invariant with respect to \(U \otimes U\) transformations (effectively, global rotations). Thus, the four mode squeezed vacuum shares an important property with a two-qubit singlet, which was used by Bohm to present a simple version of the EPR argument 3. In a way this is a kind of polarization super-singlet with undefined number of photons.

This feature of \(|\Psi(-)\rangle\) allows one to perform an EPR-like reasoning, with different observables than the ones considered by EPR or Bohm. Using the property \(n^{(k)}(\theta_A) = m^{(k)}(\theta_A + \frac{\pi}{2})\) one can fix the Bob’s value for the \(k\)-th run, and his possible setting \(\theta_B = \theta_A + \frac{\pi}{2}\), without actually measuring it, by a distant physically non-disturbing measurement at Alice’s side, and vice versa. The experiment can have a configuration in which the acts of measurement of Alice and Bob are spatially separated (in the relativistic meaning). Thus, \(n^{(k)}(\theta_A)\) and \(m^{(k)}(\theta_B)\) are elements of reality. Note that this holds for any \(\theta_A\) and thus for any \(\theta_B\). To show that is unacceptable in quantum mechanics, one should seek for Bell inequalities satisfied by the elements of reality, and violated by quantum predictions.

We shall show that such inequalities exist. The double EPR-like-supersinglet leads to predictions which disagree with the ideas of EPR. The Bell inequalities are based on the concept of distance, and the fact that any properly defined distance satisfies polygon inequalities.

Take two stochastic variables \(V(\lambda)\) and \(W(\lambda),\) governed by a joint probability \(\rho(\lambda)\). Their “separation” can be measured for example by \(D(V, W) = \int |V(\lambda) - W(\lambda)| \rho(\lambda)\). This function of two variables satisfies all defining properties of a distance: \(D(V, V) = 0,\) \(D(V, W) = D(W, V) \geq 0\) and the triangle inequality \(D(V, Z) \leq D(V, W) + D(W, Z)\). The last property is due to the fact that for any three real (or complex) numbers \(a, b, c\) one has: \(|a - c| \leq |a - b| + |b - c|\).

**Distance Bell inequality.** Allow Alice and Bob to choose freely between several local settings of their polarizers, \(\theta_A,\) and \(\theta_B,\) respectively. For a more concise notation we put \(n_i^{(k)}(\theta_A)\) and \(n_j^{(k)}(\theta_B)\).

A triangle inequality implies polygon inequalities of an
arbitrary length (see Fig. 1). Put \( i, j = 1, \ldots, L \). One has

\[
\sum_{i=1}^{L} |m_i^{(k)} - n_i^{(k)}| + \sum_{i=1}^{L-1} |m_i^{(k)} - n_i^{(k)}| \geq |m_1^{(k)} - n_L^{(k)}|,
\]

(3)
a polygon inequality for numbers. For averages, \( \langle |m_i - n_j| \rangle = \frac{1}{R} \sum_{k=1}^{R} |m_i^{(k)} - n_j^{(k)}| \), where \( R \) is the number of runs, we get an inequality for elements of reality:

\[
\sum_{i=1}^{L} \langle |m_i - n_i| \rangle + \sum_{i=1}^{L-1} \langle |m_{i+1} - n_i| \rangle \geq \langle |m_1 - n_L| \rangle.
\]

(4)

This Bell inequality is in a form of a chain (the concept of chained inequalities see [15]): if the number of vertices grows, the sums on the left-hand side gain extra terms but the inequality as a whole does not change its form. In a geometric interpretation presented in Fig. 1, this corresponds to adding new red segments.

One arrives at the same inequality, if in the derivation one uses local hidden variable (LHV) models. Assume that \( m_i \) and \( n_i \) depend on some hidden parameters (causes, variables) \( \lambda \). The ‘distance’ can be put as

\[
\langle |m_i - n_j| \rangle = \int d\lambda \rho_{hv}(\lambda) \langle |m_i(\lambda) - n_j(\lambda)| \rangle,
\]

(5)

where \( \rho_{hv}(\lambda) \) is a probability distribution of LHVs. Obviously it satisfies (4).

Below we show that inequality (4) is violated by the EPR supersinglet \( \ket{\psi^{(-)}} \), as well as by its component states, the 2N-photon singlets \( \ket{\psi_N^{(-)}} \).

We will test inequality (4) within the quantum theory for the photon number resolving measurements. The observables which represent these measurements are the photon number operators \( a_i^\dagger a_i \) (Alice) and \( b_i^\dagger b_i \) (Bob). The polarization rotations performed by Alice and Bob result in the change of the measurement basis according to the transformation \( a_i = \cos \theta_i^A a_i + \sin \theta_i^A a_i^\dagger \), with \( a_i^\dagger \) denoting annihilation operator for the horizontal (vertical) polarization. The notation for the beam \( b \) is different. Because of the fact that the perfect correlations for the considered state occur for orthogonal polarizations, we put \( b_i = -\sin \theta_i^B b_H + \cos \theta_i^B b_V \). The distance Bell inequality (4) requires the following

\[
\text{LHS} = \sum_{i=1}^{L} \langle a_i^\dagger a_i, b_i^\dagger b_i \rangle + \sum_{i=1}^{L-1} \langle a_i^\dagger a_{i+1}, b_i^\dagger b_i \rangle \\
\geq \langle a_1^\dagger a_1, b_1^\dagger b_L \rangle = \text{RHS}.
\]

(6)

Violation of the distance inequality by the 2N photon singlets, \( \ket{\psi_N^{(-)}} \). The measurements which we suggest for Alice and Bob are shown in Fig. 2. The relative angle between the polarization settings by Alice \( A_i \) and Bob \( B_i \) is constant and reads \( \theta = \frac{\theta_i}{\pi} \). Each subsequent setting of Alice changes by 2\( \theta \) thus, the angle between \( A_{i+1} \) and \( B_i \) is also \( \theta \). The angle between the first Alice’s setting \( A_1 \) and the last Bob’s setting \( B_L \) is \( \theta' = \frac{(2L-1)\pi}{4L} \).

The singlets (2) also have the invariance with respect to \( U \otimes U \) discussed above. Thus, the setting rotations \( a_i \) and \( b_i \) will appear in inequality (4) only in the form of the relative angle, \( \theta \) or \( \theta' \).

If \( \theta = 0 \), Alice and Bob have the same settings and the perfect correlations (2) between the orthogonally polarized components in beams \( a \) and \( b \) are observed. Even for small \( \theta \neq 0 \) the correlations deteriorate. For the settings \( \theta_1^A = 0 \) and \( \theta_1^B = \theta \) the probability to register \( n \) photons in Alice’s channel \( H \) and \( m \) in Bob’s channel \( V_{\theta} \), denoted below as \( V + \theta \) (the other one: \( H + \theta \)) reads

\[
p_{Q}^N(n,m | \theta) = \langle \psi_N^{(-)} | (|h_{n}, (N-n)_{V_{\theta}} \rangle a_i |(N-m)_{H+\theta}, m_{V+\theta} \rangle a_i) |^2.
\]

The distance inequality (4) takes the form

\[
\text{LHS} = \sum_{n=0}^{N} |m-n| \langle 2L-1 \rangle p_{Q}^N(n,m | \theta) \\
\leq \sum_{n=0}^{N} |m-n| p_{Q}^N(n,m | \theta') = \text{RHS}.
\]

(8)

Let us estimate the RHS for a large number of settings (parametrized by \( L \)). With \( L \) very large \( \theta' \approx \frac{\theta}{\pi} \) and \( \frac{\pi}{2} \) is the limit for \( L \to \infty \). This means that Bob’s measurements are effectively in the same basis as Alice’s, however the polarizer’s outputs are flipped. Thus there is no \( n = m \) correlation characteristic for singlets for outputs \( H \) on one side and \( V \) on the other (and other way
round). As Bob’s \( H \) is now \( V \), etc. we have a perfect anti-correlation, \( m = N - n \). The formula (2) tell us, that \( p_Q^N(n, N - n | \pi/2) = 1/N + 1 \). The RHS reaches its maximum possible value (see Supplementary Material).

When estimating the LHS, notice that

\[
|n_H, (N - n)V\rangle_a |(N - m)H + \theta, m_V + \theta\rangle_b \tag{9}
\]

is proportional to \( a_v^m a_V^{(N-n)} b_H^{m+\theta} b_V^{n+\theta} |0\rangle \). As \( b_H^{m+\theta} = b_H^{\theta} \cos \theta + b_V^{\theta} \sin \theta \) and \( b_V^{n+\theta} = b_V^{\theta} \cos \theta - b_H^{\theta} \sin \theta \), we see the following. For \( \theta = 0 \) we recover the formula for perfect singlet correlations: formula (7) tells us that

\[
|\langle \psi^{(-)}_N| |n_H, (N - n)V\rangle_a |(N - m)H, m_V\rangle_b |^2 \tag{10}
\]

is non-zero only for \( n = m \). However for non-zero \( \theta \) we see that all ‘new’ terms in the amplitude in formula (7), new with respect to (10), are proportional to some power of \( \sin \theta \). The lonely term proportional to \( \cos 2N \theta \) does not contribute to \( p_Q^N(n \neq m|\theta) \). Thus the difference between \( p_Q^N(n \neq m|0) \) = 0, and \( p_Q^N(n \neq m|\theta) = 0 \) is a polynomial in \( \sin \theta \) with the lowest power equal to 2. As \( \theta = \pi/(4L) \), the lowest order terms LHS of (9) behave as \( (2L - 1) \sin^2 \frac{\pi}{2N} \). With \( L \rightarrow \infty \), LHS approaches zero.

In the limit of macroscopic population \( (N \rightarrow \infty) \) we obtain a striking contradiction: \( 0 \geq \infty \).

**Inefficient detection.** One can introduce a Bell parameter in the form \( B_q^N = \text{LHS} - \text{RHS} \), which according to LHV theories is always positive \( B_q^N \geq 0 \). We have investigated the fragility of violations of \( B_q^N \geq 0 \) and violation in the case of imperfect detection efficiency \( \eta \), for various numbers of settings \( L \) and photon population \( 2N \). The exact dependence of the Bell parameter on \( \eta \) is given in the Supplementary Material (Section II). In Fig. 3c, we show the simplest case \( 2N = 2 \). Here the state \( |\psi_{N=1}^{(-)}\rangle \) is a two-photon polarization singlet state. As expected, violation of the local bound is possible only above the usual minimal value of efficiency of detectors \( \eta > \frac{1 + \sqrt{2}}{1 + \sqrt{2}} \approx 83\% \).

The violation takes its maximal value \(-0.4\) in the limit of perfect detection, \( \eta = 1 \). The violation is possible for all \( L \)’s but the most pronounced one is obtained for the highest (considered) number of settings \( L = 10 \) and reaches \(-0.8\). For larger photon numbers, in order to observe any violation, the minimal efficiency as well as the minimal number of settings increase, see Fig. 3c. For \( 2N = 8 \) \( L \geq 4 \) and for \( 2N = 12 \) \( L \geq 6 \). Interestingly, the minimal efficiency required for violation can be smaller for larger number of settings, compare the green and yellow curves in Fig. 3.

Violation by the EPR-Bohm supersinglet. The states \( |\psi^{(-)}_N\rangle \) can be produced postselectively from \( |\Psi^{(-)}\rangle \) via photon number measurement (which is inherent in the measuring scheme), the Bell inequality 1 can be also directly tested with the (bright) four mode squeezed vacuum state. Fig. 4 depicts a simple experimental scheme for this test. The quantum value of the Bell parameter reads \( B_q = \sum_{N=0}^{\infty} \lambda^2 N B_q^N \). The exact form of the Bell parameter for large \( L \) given in the Supplementary Material (Section III) and reads \( B_q^N = -\frac{1}{N^2 + N + 1} \) for an odd \( N \) and \( B_q^N = -\frac{1}{N^2 + N + 1} \) for even \( N \). Fig. 5 shows \( B_q^N \) as a function of number of settings \( L \) for different photon numbers \( 2N \). Independent on \( N \), for sufficiently large \( L \), \( B_q^N \) always tends to a certain negative value. Therefore, in this limit \( B_q < 0 \) also holds, which proves violation for (11). This result is independent of the statistics \( \lambda \) in \( |\Psi^{(-)}\rangle \). Direct calculation of the inequality (8) shows that \( LHS \rightarrow 0 \) and \( RHS \rightarrow e^{2\gamma} \). If we take a very large \( \Gamma \), we approach \( 0 \geq \infty \).

**Conclusions.** A four-mode bright squeezed vacuum (BSV) can be thought of as two copies of an approximate EPR state, or a super-singlet. Even without going to the unphysical limit the perfect correlations allow one formulate an EPR argument for elements of reality, however for different variables than the original considerations of EPR [2]. Still, the variables are “natural” – mere photon numbers. However, the distance-like Bell inequality invalidates the EPR reasoning. In the limit of an infinitely long chained polygon inequality, based on the EPR concepts, for specific settings the LHS predicts a perfect photon-number correlation for orthogonal polarizations, while the RHS gives is the maximum possible “distance” for the photon number observables if measured for BSV. We have a situation similar to the GHZ paradox, because actually the for parallel polarizations photon-numbers show perfect correlation! Elements of reality point to opposite correlation than the actual one. One can expect various applications of the presented process in quantum information. For small values of \( \Gamma \) the experiment is feasible. One can use a modification of the techniques of Ref. [14]. However the analysis would necessarily involve a fair sampling assumption.

**ACKNOWLEDGMENTS**

The work is a part of EU project BRISQ2. MS and KR were supported by the EU 7FP Marie Curie Career Integration Grant No. 322150 “QCAT”, NCN grant No. 2012/04/M/ST2/00789, FNP Homing Plus project No. HOMING PLUS/2012-5/12 and MNiSW co-financed international project No. 25867/PR2012/2. MZ is supported by TEAM project of FNP.

[1] Y. Aharonov and D. Rohrlich, Quantum Paradoxes: Quantum Theory for the Perplexed, Wiley-VCH.

[2] A. Einstein, B Podolsky, N. Rosen, Phys. Rev. 47, 777 (1935).
FIG. 3. Bell parameter $B_N^\eta$ as a function of efficiency of detectors $\eta$ evaluated for different number of settings $L$: $L = 2$ (red line), $L = 4$ (blue line), $L = 6$ (green line), $L = 8$ (orange line), $L = 10$ (yellow line) for (a) $2N = 2$, (b) $2N = 8$ (c) $2N = 12$.

FIG. 4. Setup for testing the distance-Bell inequality with entangled bright squeezed vacuum state. PDC - parametric down conversion crystal, PBS - polarizing beam splitter. The detectors measure photon numbers. $\theta$ and $\theta^\perp$ denote measurements in a basis rotated with respect to the initial one H,V (denoted as $\theta = 0$.)

FIG. 5. The Bell parameter $B_N^L$ as a function of number of settings $L$ evaluated for different number of photons $2N$: $2N = 2$ (red line), $2N = 4$ (blue line), $2N = 6$ (black line), $2N = 8$ (green line), $2N = 10$ (orange line), $2N = 12$ (yellow line).

[3] N. Bohr, Phys. Rev. 48, 696 (1935).
[4] D. Bohm, Quantum Theory (Prentice-Hall, Englewood Cliffs, NJ, 1951).
[5] J. S. Bell, Physics 1, 195 (1964).
[6] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Bell’s Theorem, Quantum Theory, and Conceptions of the Universe, M. Kafatos (Ed.), Kluwer, Dordrecht, 69-72 (1989).
[7] C. Brukner, M. Zukowski, J.-W. Pan, and A. Zeilinger, Phys. Rev. Lett. 92, 127901 (2004).
[8] S. Pironio et al., Nature 464, 1021 (2010).
[9] A. Acin et al., Phys. Rev. Lett. 98, 230501 (2007).
[10] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. 81, 865 (2009).
[11] J.-W. Pan et al., Rev. Mod. Phys. 84, 777 (2012).
[12] J. S. Bell, in: New Techniques and Ideas in Quantum Measurement Theory, ed. D. M. Greenberger (The New York Academy of Sciences, New York, 1986) p. 263, see also U. Leonhardt, Phys. Lett. A 182, 195 (1993), L. M. Johansen, Phys. Lett. A 236, 173 (1997).
[13] M. D. Reid and P.D. Drummond, Phys. Rev. Lett. 60, 2731 (1988).
[14] M. D. Reid, Phys. Rev. A 40, 913 (1989).
[15] O. Cohen, Phys. Rev. A 56, 3484 (1997).
[16] K. Banaszek, K. Wodkiewicz, Phys. Rev. A 58, 4345 (1998).
[17] T. Sh. Iskhakov, M. V. Chekova, G. O. Rytikov, and G. Leuchs, Phys. Rev. Lett. 106, 113602 (2011).
[18] P. M. Pearle, Phys. Rev. D 2 1418 (1970); S. L. Braunstein, C. M. Caves, Annals of Physics 202, 22-56 (1990); J. Pykacz and E. Santos, J. Math. Phys. 32, 1287 (1991); A. Cabello, Phys. Rev. A 58, 1687 (1998). L. Hardy, Phys. Lett. A 161, 21-25 (1991); J. Barrett, A. Kent, and S. Pironio, Phys. Rev. Lett. 97, 170409 (2006). P. Kurzyński, D. Kuszkowski, Phys. Rev. A 89, 012103 (2014); M. Zukowski, A. Dutta, Phys. Rev. A 90, 012106 (2014).
[19] M. Radmark, M. Zukowski, M. Bourennane, Phys. Rev. Lett., 103, 150501 (2009).
SUPPLEMENTARY MATERIAL

I. PROOF OF THE INEQUALITY (3)

Inequality (3) will be proved by induction with respect to L. Let us denote by $S_L$ the right side of inequality (3). We will show that $S_L \geq 0$. For $L = 1$ (we are not interested in that case physically but it does not matter in the proof) $S_L = |m_1 - n_1| - |m_1 - n_1| = 0$ and inequality (3) is satisfied. Let us assume that $S_{L-1} \geq 0$ holds. We have that $S_L = S_{L-1} + R$, where $R = |m_1 - n_{L-1}| + |m_L - n_L| + |m_L - n_{L-1}| - |m_1 - n_L|$. Let us notice that after substitution $L - 1 \rightarrow 1$ and $L \rightarrow 2$ $R = |m_1 - n_1| + |m_2 - n_2| + |m_2 - n_1| - |m_1 - n_2|$ what is equal to $S_L$ for $L = 2$. Thus $R \geq 0$ from induction hypothesis for $L = 2$ and finally $S_L \geq 0$ which completes the proof.

II. DERIVATION OF PROBABILITY IN LOSSLESS DETECTION CASE

Our goal is to calculate the following probability

$$p_{Q}^{N}(n, m \mid \theta) = |\langle \Psi(-) | n, N - n, (N - m)_{H+\theta}, m_{V+\theta} \rangle|^2.$$ 

In order to do that we need to express the state $|(N - m)_{H+\theta}, m_{V+\theta}\rangle$ in terms of $\{H, V\}$ basis. The two bases, i.e. $\{H, V\}$ and $\{H + \theta, V + \theta\}$ are linked by the following unitary transformation (it is just a rotation in operators space)

$$\begin{pmatrix} b_{H+\theta}^j \\ b_{V+\theta}^j \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} b_{H}^j \\ b_{V}^j \end{pmatrix}.$$ (11)

Applying Eq. (11) we obtain

$$\langle (N - m)_{H+\theta}, m_{V+\theta} \rangle = \frac{(b_{H+\theta}^{N-m}(b_{V+\theta}^m)^m)}{\sqrt{(N-m)!m!}} |0\rangle$$

$$= \frac{1}{\sqrt{(N-m)!m!}} \sum_{p=0}^{N-m} \sum_{q=0}^{m} \left( \begin{array}{c} N - m \\ p \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) (-1)^q \sin^N m - p + q (\cos \theta)^{m-q+p} (b_{H}^p b_{V}^q)^{N-(p+q)} |0\rangle$$

$$= \frac{1}{\sqrt{(N-m)!m!}} \sum_{p=0}^{N-m} \sum_{q=0}^{m} \left( \begin{array}{c} N - m \\ p \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) (-1)^q \sin^N m - p + q (\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))! |p+q, N-(p+q)\rangle.$$ 

Now we are ready to compute $p_{Q}^{N}(n, m \mid \theta)$. The probability amplitude is given by

$$\langle \Psi(-) | n, N - n, (N - m)_{H+\theta}, m_{V+\theta} \rangle = \frac{1}{\sqrt{(N+1)(N-m)!m!}} \sum_{k=0}^{N} \sum_{p=0}^{N-m} \sum_{q=0}^{m} \left( \begin{array}{c} N - m \\ p \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) (-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!$$

$$(k, N - k, N - k, k|p+q, N-(p+q)\rangle)$$

$$= \frac{1}{\sqrt{(N+1)(N-m)!m!}} \sum_{k=0}^{N} \sum_{p=0}^{N-m} \sum_{q=0}^{m} \left( \begin{array}{c} N - m \\ p \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) (-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!$$

$$\delta_{n,k}\delta_{p+q,N-k} = \frac{(-1)^n}{\sqrt{(N+1)(N-m)!m!}} \sum_{p=0}^{N-m} \sum_{q=0}^{m} \left( \begin{array}{c} N - m \\ p \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) (-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!$$

Now we get rid of the Kronecker delta $\delta_{p+q,N-n}$, we notice that $p \in (0, N-m)$ and $0 \leq q = N-n-p \geq m-n$ what implies the summation from $q = i = \max\{0, m-n\}$ to $j = \min\{N-n, m\}$. Hence we obtain the probability amplitude

$$\langle \Psi(-) | n, N - n, (N - m)_{H+\theta}, m_{V+\theta} \rangle = \frac{(-1)^n \sqrt{\xi_{nm}^{(N)}} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) \frac{(-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!}{\sqrt{(N+1)(N-m)!m!}}}{\sqrt{(N+1)(N-m)!m!}} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right) \frac{(-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!}{\sqrt{(N+1)(N-m)!m!}} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right)$$

$$= (-1)^n \sqrt{\xi_{nm}^{(N)}} (\cos \theta)^{N-(2q+n-m)} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right)$$

$$= (-1)^n \sqrt{\xi_{nm}^{(N)}} (\cos \theta)^{N-(2q+n-m)} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right)$$

where $\xi_{nm}^{(N)} = \frac{(N-n)!m!}{(N+1)(N-m)!m!}$. Now we square [12] and get probability $p_{Q}^{N}(n, m \mid \theta)$

$$p_{Q}^{N}(n, m \mid \theta) = \xi_{nm}^{(N)} (\cos \theta)^{2N} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right)$$

$$\frac{(-1)^q (\sin \theta)^{N-m-p+q}(\cos \theta)^{m-q+p} \sqrt{(p+q)!}(N-(p+q))!}{\sqrt{(N+1)(N-m)!m!}} \sum_{q=i}^{j} \left( \begin{array}{c} N - m \\ q \end{array} \right) \left( \begin{array}{c} m \\ q \end{array} \right)$$

Now we are ready to compute $p_{Q}^{N}(n, m \mid \theta)$.

III. CONTRADICTION WITH QUANTUM MECHANICAL PREDICTIONS

We show that considering inequality (8) one obtains contradiction with quantum mechanics. Inequality (8) can be written as
\[
\sum_{N=0}^{\infty} \lambda_N^2 \sum_{n,m=0}^{N} |m-n| L \Lambda(n,m | N, \theta) \geq 0 \quad (13)
\]

where

\[
L = (2L-1)p_Q(n,m | \theta) \quad (14)
\]

Here the notation \(LHS\) and \(RHS\) stand for the left-hand side and the right-hand side of the inequality (13). We easily see that for \(L \to \infty, \theta = \pi/4L \to 0\) and \(\theta' = (2L-1)\pi/4L \to \pi/2\). At first we consider right-hand side of the inequality (13). Please notice that

\[
\sum_{n,m=0}^{N} |m-n| p_Q^{N}(n,m | \theta') =
\sum_{n+m<N}^{N} |m-n| p_Q^{N}(n,m | \theta') +
\sum_{n+m>N}^{N} |m-n| p_Q^{N}(n,m | \theta') +
\sum_{n=m=N}^{N} |m-n| p_Q^{N}(n,m | \theta').
\]

In Eq.(12) we have that \(i = q_{\text{min}} = \max\{0, m-n\}\) and \(j = q_{\text{max}} = \min\{N-n, m\}\). For simplicity we introduce notation \(K(q) = 2q + n - m\). We have to consider three cases. 1° If \(n + m < N\) then \(q_{\text{max}} = m\) and \(K(q_{\text{max}}) = n + m < N\). Hence obviously \(K(q) < N\) and \((\cos \theta')^N (\tan \theta')^{K(q)} \to 0\) if \(L \to \infty (\theta \to 0)\). 2° Similarly if \(n + m < N\) then \(q_{\text{max}} = N - n\) and \(K(q_{\text{max}}) = 2N - (n + m) < N\). Hence again \(K(q) < N\) and \((\cos \theta')^N (\tan \theta')^{K(q)} \to 0\). From 1° and 2° we conclude that the first two sums in (15) vanish for \(L \to \infty\). 3° If \(n + m = N\) we obtain that \(q_{\text{max}} = m = N - n\) and \(K(q_{\text{max}}) = n + m = N\). Since \((\cos \theta')^N (\tan \theta')^{K(q)} \to 1\) and \((\cos \theta')^N (\tan \theta')^{K(q)} \to 0\) for \(K(q) < K(q_{\text{max}})\), thus only the last component of the third sum in (15) contributes. Since for \(n + m = N\) and \(q = m\), \(\xi_{nm}^{(N)} = \frac{1}{N+1}\) and \((N-m) \binom{m}{N-n-q} = 1\) we obtain that \(p_Q^{N}(n,m | \theta') \to \epsilon_{nm}^{(N)} = \frac{1}{N+1}\) which is a number substantially greater than 0. Hence the expression (15) for \(L \to \infty\) takes the following form

\[
\sum_{n,m=0}^{N} |m-n| p_Q^{N}(n,m | \theta') =
\frac{1}{N+1} \sum_{n,m=0,n+m=N}^{N} |m-n| =
\begin{cases}
\frac{2}{N+1} \sum_{n=0}^{N-1} (N-2n), & \text{for } N \text{- odd} \\
\frac{2}{N+1} \sum_{n=0}^{N-2} (N-2n), & \text{for } N \text{- even}
\end{cases}
\]

(16)

Using Eq. (16) and the fact that \(\lambda_N^2 = \cosh^{-4}(g)(N+1) \tanh^{2N}(\Gamma)\) we are able to calculate that \(RHS\) equals

\[
RHS = \frac{\sinh^3(2\Gamma)}{\sinh(4\Gamma)} > 0
\]

(17)

Now we show that parameter \(\Lambda(n,m | N, \theta) \to 0\) for \(\theta \to 0\) \((L \to \infty)\). Using expression for \(p_Q^{N}(n,m | \theta)\) parameter \(\Lambda\) can be written as

\[
\Lambda(n,m | N, \theta) = \xi_{nm}^{(N)} (\sqrt{2L-1} (\cos \theta)^N \sum_{q=0}^{q-1} (N-m-q) (N-n-q) (\tan \theta)^{2q+n-m})
\]

(18)

Please notice that for \(m \neq n\) (if \(m = n, |m-n|=0\)) and \(q \geq \max\{0, m-n\}\) we have that \(K(q) = 2q + n - m > 0\). Using de L’Hospital rule is is easy to calculate that \(\sqrt{2L-1} \tan \theta \to 0\) for \(L \to \infty\). From this fact and using that \((\cos \theta)^N \to 1\) and \((\tan \theta)^l \to 0\) \((l\) is arbitrary positive number) for \(L \to \infty\) we obtain that \(\sqrt{2L-1} (\cos \theta)^N (\tan \theta)^{K(q)} \to 0\) for \(L \to \infty\). Hence \(\Lambda(n,m | N, \theta) \to 0\). This implies that \(LHS \to 0\). From the above considerations we also notice that in the limit \(L \to \infty\) parameter \(C_q\) is negative

\[
\lim_{L \to \infty} C_q < 0.
\]