ARTICLE   OPEN

From contextuality of a single photon to realism of an electromagnetic wave

Marcin Markiewicz[1], Dagomir Kaszlikowski[2,3], Paweł Kurzyński[4,2] and Antoni Wójcik[4]

Violations of Bell inequalities have been an incontestable indicator of non-classicality since the seminal paper by John Bell. However, recent claims of Bell inequalities violations with classical light have cast some doubts on their significance as hallmarks of non-classicality. Here, we challenge those claims. The crux of the problem is that such classical experiments simulate quantum probabilities with intensities of classical fields. However, fields intensities measurements are radically different from single-photon detections, which are primitives of any genuine Bell experiment. We show that this fundamental difference between field intensities measurements and single photon detections shifts the classical bound of relevant Bell inequalities to its non-signaling limit, leaving no place for their violations.

INTRODUCTION

The quantum-to-classical transition in optical interferometry can be observed either in a single-particle or multi-particle interference.[1] The multi-particle interference is commonly regarded as more fundamental, including purely quantum phenomena such as the Hong-Ou-Mandel (HOM) effect[2] and multi-photon violations of Bell inequalities.[3] The quantum-to-classical transition in these phenomena has different origins. In the HOM setting it can be attributed to the strength of particle indistinguishability[4] whereas multi-photon Bell inequality violations are tied to the coherence strength between entangled photons or, in the limit of many particles, to the vanishing ability of revealing single particle properties with multi-photon measurements.[5]

A single-photon interference scenario is much simpler to describe. To illustrate it, let us consider the simplest case—a Young double-slit experiment.[6] Here the classical limit is achieved by increasing the average number of photons prepared in a coherent state or a mixture of such states. In this limit, the interference pattern does not change. What changes is the physical meaning of a mathematical formalism used to describe the experiment—the probability amplitudes of a single photon become amplitudes of a classical electromagnetic wave. Straightforward as it seems, a relatively recent 'discovery' of Bell inequality violations with classical light, dubbed 'classical entanglement', makes the whole classical to quantum transition less obvious.

Everything started in 1996 with a paper by Patrick Suppes et al.[7] They proposed an interferomteric experiment to violate a Bell inequality with classical light. A year later Robert Spreeuw introduced the concept of classical entanglement between two different degrees of freedom of a single classical light beam.[8] Moreover, he showed that this entanglement leads to violations of the CHSH (Clauser-Horne-Shimony-Holt) inequality.[9] He subsequently generalised this idea to more degrees of freedom,

demonstrating a classical version of the GHZ (Greenberger-Horne-Zeilinger) paradox.[10]

Classical entanglement occurs between two or more properties of an individual system and as such it does not require spatial separation unlike the standard EPR scenario. Because of this, the classical entanglement was largely dismissed by other researchers as a mere curiosity, irrelevant in the context of quantum non-locality.[11] However a few years later a variety of papers appeared, discussing a similar concept of intrasystem entanglement in different physical implementations. Eberly, Qian et al. and further Aiello et al. in a series of papers[12–14] developed a theory of "bipartite" entanglement between polarisation and position degrees of freedom in stochastic light beams.[15] Later, Eberly et al. performed an experiment achieving a strong violation of the CHSH inequality with entangled states of stochastic light fields.[16] A similar violation of the CHSH inequality with classical entanglement between fields in two optical resonators was proposed by Snoke.[17] Finally Frustaglia et al.[18] derived a procedure, following earlier ideas of Cerf et al.[19] and Spreeuw,[10] which allows to reconstruct probability distributions coming from any quantum correlations tests using classical optical circuits. As an illustration of their method, the authors of[18] performed an experiment with microwave circuits showing violations of the CHSH[9] and Mermin[20] type inequalities.

Experimental demonstrations of Bell inequalities violations with classical light have profound physical implications. Snoke[17] and Qian et al.[16] hypothesised that a Bell inequality violation does not testify a presence of quantum entanglement in a given physical system—it may as well be a classical entanglement. Another hypothesis by Frustaglia et. al.[18] is that the bounds on the strength of quantum correlations (so called Tsirelson bounds) are not restricted to quantum physics, but arise naturally in classical systems which simulate quantum correlations. If these claims were

true, we would have to reconsider the role of Bell inequalities in probing quantum to classical transition.

In this paper we challenge these claims by showing that one does not observe any violation in classical regime if the Bell inequalities are properly derived. More precisely, Bell inequalities test if probability distributions of measurement results are contextual[21]—the feature commonly accepted as an indicator of non-classicality. From the mathematical perspective, Bell inequalities are based on properties of exclusivity relations between jointly measurable events.[22] A proper structure of such exclusivity implies existence of a test that can distinguish between contextual (non-classical) and non-contextual (classical) probability distributions.

We show a non-contextual physical model based on quantum-to-classical transition from single photons to classical waves. Our model proves that classical waves are not contextual and thus they can be still called classical. Moreover, we demonstrate that the proper classical bounds, i.e., the bounds respecting the correct exclusivity structure of detection events for Bell tests with classical light are equal to the non-signaling bounds on the correlations' strength. Therefore, there is no place for any classical contextuality in such systems.

RESULTS

Photon distribution in the classical limit

Consider a single photon in a polarisation state $\sqrt{p_H}|H\rangle + \sqrt{p_V}|V\rangle$, where H and V denote horizontal and vertical polarisations, respectively, and $p_H + p_V = 1$. When incident on a polarising beam splitter (PBS), the photon can either go through and become H polarised, or be reflected and become V polarised. These are two possibilities occuring randomly with probabilities p_H and p_V, if one decides to detect the photon after the PBS. Denote these two possible outcomes as {0, 1} and {1, 0}. This scenario is a physical implementation of a binary ±1 random variable X, where the outcome {1, 0} is associated with the value +1 and {0, 1} with −1.

Next, let us consider two indistinguishable photons in the above state entering the same PBS port. They are uncorrelated and therefore they scatter on the PBS independently.[23] The photons cannot be distinguished and therefore there are only three exclusive outcomes: {2, 0}, {1, 1} and {0, 2}. These outcomes cannot be interpreted as products of two single-photon outcomes becasue of indistinguishability, i.e., events {1, 0} × {1, 0}, {1, 0} × {0, 1}, {0, 1} × {1, 0}, and {0, 1} × {0, 1} are meaningless. Moreover, unlike in the single-photon case, for two photons, statements "photon is detected on the left" and "photon is detected on the right" are not exclusive because there is a chance that photons can be detected on both sides. Interestingly, the average number of photons in each output is proportional to single-photon scattering probabilities, i.e., $\overline{n}_H = 2p_H$ and $\overline{n}_V = 2p_V$.

In general, for N photons scattering on the PBS one can observe $N + 1$ exclusive outcomes: {N, 0}, {N − 1, 1}, etc. Note, that although the scenario consists of only two distinguishable modes, the exclusivity structure depends on the number of photons, since the number of exclusive events scales linearly with N. This is drastically fewer than 2^N outcomes observable for distinguishable particles. For N photons the single-photon random variable X is ill-defined because of indistinguishability. However, it is possible to define a random variable whose outcomes are given by $\mathcal{X} = (n_H - n_V)/N$, i.e., the difference between photon numbers in the output ports divided by the total number of photons. Note, that $-1 \leq \mathcal{X} \leq 1$ and $\mathcal{X} = X$ for $N = 1$. Additionally, since each photon is transformed independently and according to the same rule, the average number of photons in each polarisation mode is given by $\overline{n}_H = Np_H$ and $\overline{n}_V = Np_V$. Because of this $\langle \mathcal{X} \rangle = p_H - p_V$ does not depend on N and the most probable outcome state is $\{\overline{n}_H, \overline{n}_V\}$.

Finally, let us discuss the classical limit. In this case the total number of photons is undetermined but their average number is large ($\langle N \rangle \gg 1$). In quantum theory such situations are usually represented by a high amplitude coherent state.[24] Once we go to the classical limit, it is quite natural to treat the beam of light as a continuous object that can be split into portions in an arbitrary way. The PBS transforms a single beam with intensity I into two beams, the H polarised beam with intensity I_H and the V polarised one with I_V. This is predicted by both, classical and quantum theories. In the classical limit the average value of random variable $\mathcal{X}$ becomes $(I_H - I_V)/I$. However, since the intensities of two beams are given by $I_H = Ip_H$ and $I_V = Ip_V$, we get $\langle \mathcal{X} \rangle = p_H - p_V$, as expected. Note, that for Poissonian light the fluctuations of $\mathcal{X}$ scale as $1/\sqrt{N}$, therefore in the classical limit $\mathcal{X}$ can be treated as a deterministic variable.

The above scenarios are schematically represented in Fig. 1. Although we considered only a single PBS, the similarity between classical intensities and probabilities generated by photonic distributions would also hold if one used an arbitrary number of linear optical devices (PBS, standard beam splitters (BS), phase shifters, etc). In this case the whole setup is equivalent to a multiport corresponding to a more complex random variable or a sequence of random variables.

To summarise, we see that the same average value $\langle \mathcal{X} \rangle$ is predicted by both, quantum and classical theories, since this value does not depend on N. Nevertheless, the underlying exclusivity structure of the outcomes of $\mathcal{X}$ strongly depends on N. This fact causes some bizarre interpretation difficulties in experimental Bell-type scenarios with classical light. We will discuss this problem in more details in the following sections.

Correlations in the classical limit

Next, we show that unlike photonic distribution, the correlations between spatially separated photons strongly depend on N and on the exclusivity structure of detection events. The problem of quantum correlations in the classical limit was discussed in details before (see for example ref. [5]) so we provide here only a simple example.

Consider a pair of photons in an entangled polarisation state $\sqrt{p_H}|HH\rangle + \sqrt{p_V}|VV\rangle$. These two photons are shared between two spatially separated parties, Alice and Bob, who measure their photons polarisations with respective PBSs. As before, we represent the two polarisation possibilities by {1, 0} and {0, 1}. Moreover, we can also use the ±1 random variables X_A and X_B, defined in the same way as X above, to represent the measurement of Alice and Bob. Alice and Bob register either {1, 0} × {1, 0} with probability p_V, or {0, 1} × {0, 1} with probability p_H. The average values and the corresponding correlations are $\langle X_A \rangle = \langle X_B \rangle = p_H - p_V$ and $\langle X_A X_B \rangle = 1$.

Next, let us consider N such photonic pairs shared between Alice and Bob who measure polarisation on all pairs at the same time. As before, the local measurements are represented by the random variables $\mathcal{X}_A = (n_{AH} - n_{AV})/N$ and $\mathcal{X}_B = (n_{BH} - n_{BV})/N$. Again, there are correlations between the photonic pairs giving $n_{AH} = n_{BH} = n_H$ and $n_{AV} = n_{BV} = n_V$, hence Alice registers the same photon distribution as Bob and $\mathcal{X}_A = \mathcal{X}_B$. However, for N entangled pairs the correlations $\langle \mathcal{X}_A \mathcal{X}_B \rangle$ are much weaker. Note, that the outcome $\{n, N - n\} \times \{n, N - n\}$ happens with probability $\frac{N!}{n!(N-n)!} p_V^n p_H^{N-n}$, thus

$$\langle \mathcal{X}_A \mathcal{X}_B \rangle = \sum_{n=0}^{N} \left(\frac{2n - N}{N}\right)^2 \frac{N!}{n!(N-n)!} p_V^n p_H^{N-n}$$

$$= \frac{N - 4p_V p_H (N-1)}{N}. \tag{1}$$

For example, for $p_H = p_V = 1/2$ one gets $\langle \mathcal{X}_A \mathcal{X}_B \rangle = 1/N$. Interestingly, $\langle \mathcal{X}_A \rangle = \langle \mathcal{X}_B \rangle = p_H - p_V$ and in the limit of the large number
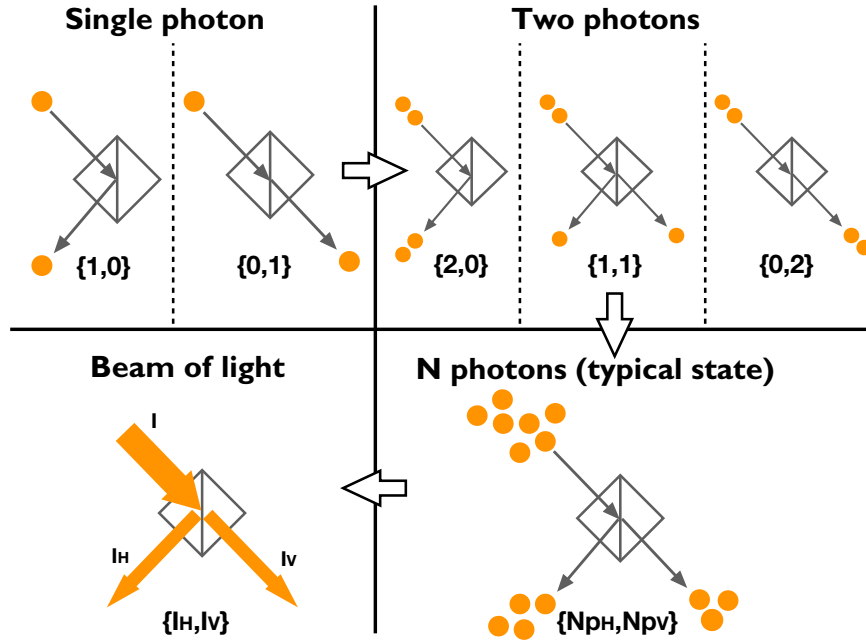
Fig. 1 Schematic representation of the classical limit in an experiment with uncorrelated photons. Single photon on a polarising beam splitter (PBS) can either go through or reflect. There are two exclusive outcomes: the photon is either registered on the left with the polarisation V or on the right with the polarisation H. The corresponding probabilities are p_V and p_H, respectively ($p_H + p_V = 1$). Two photons on the PBS can produce three exclusive outcomes: both on the left with probability p_V^2, one on the left and one on the right with probability $2p_H p_V$, and both on the right with probability p_H^2. N photons on the PBS can produce $N + 1$ exclusive outcomes, however the most probable events are those with approximately Np_V photons on the left and Np_H photons on the right. A classical beam of light of intensity I is split on the PBS into two beams. In principle there is a continuum of outcomes, however one always observes the one with the corresponding intensities I_H and I_V, where $I_H = Ip_H$ and $I_V = Ip_V$

of photons

$$\lim_{N \to \infty} \langle \mathcal{X}_A \mathcal{X}_B \rangle = 1 - 4p_V p_H = \langle \mathcal{X}_A \rangle \langle \mathcal{X}_B \rangle. \tag{2}$$

To conclude, the values $\langle \mathcal{X}_A \rangle$ and $\langle \mathcal{X}_B \rangle$ do not depend on N. However, the correlation between $\mathcal{X}_A$ and $\mathcal{X}_B$, $\langle \mathcal{X}_A \mathcal{X}_B \rangle$, does. As a consequence, in the classical limit of large $\langle N \rangle$ the two random variables get practically uncorrelated. Therefore, the classical limit of Bell-type scenarios based on correlations between many particles can always be explained by a classical theory (for more details see ref. [5] and the methods). The idea of a classical simulation of quantum correlations using classical beams uses different approach, and in the next section we focus on correlations between random variables defined for the same particle.

Bell inequalities in the clasical limit

The idea of hidden variables goes back to the father founders of quantum theory. Some of them, like Einstein, could not accept the fact that the theory is fundamentally random. A programme to explain the quantum randomness as a lack of knowledge of the system's state was developed. In particular, it was assumed that the quantum state provides only a partial information about the true state of the system. The remaining information exists, encoded in some parameters, but is inaccessible to an experimenter. These inaccessible parameters were called hidden variables. The standard hidden variable approach to quantum theory studies if it is possible to assign outcomes to all possible measurements in a consistent way, without violating commonly accepted features of nature, such as locality or non-contextuality. Bell inequalities[11] serve as the most common tool to check whether a hidden-variable description is possible for a given experimental scenario.

Let us consider the CHSH scenario,[9] which is the simplest Bell test involving four ± 1 binary random variables A_0, A_1, B_0, and B_1. In a classical theory these four random variables are jointly distributed which implies that the following inequality must be satisfied

$$-2 \leq \langle A_1 B_1 \rangle + \langle A_0 B_1 \rangle + \langle A_1 B_0 \rangle - \langle A_0 B_0 \rangle \leq 2, \tag{3}$$

where

$$\langle A_i B_j \rangle = \sum_{a_i, b_j = \pm 1} a_i b_j p(A_i = a_i, B_j = b_j). \tag{4}$$

In quantum theory it is possible to find a set of binary observables represented by Hermitian matrices, such that $[A_i, B_j] \equiv A_i B_j - B_j A_i = 0$ (for $i, j = 0, 1$), but $[A_0, A_1] \neq 0$ and $[B_0, B_1] \neq 0$. This means that A_i and B_j can be jointly measured, but it is not possible to jointly measure A_0 and A_1 or B_0 and B_1. Interestingly, for quantum correlations $\langle A_i B_j \rangle$ the inequality (3) can be violated up to $\pm 2\sqrt{2}$ for an optimal choice of the state and observables. The violation implies that the measured correlations cannot be described by classical theories.

The simplest quantum system where such a scenario is possible has four levels. In the original Bell-type scenario we have two spatially separated systems, e.g., two polarisation entangled photons discussed in the previous section, see also Fig. 2a. In this case A_0 and A_1 correspond to the polarisation properties of the first photon, whereas B_0 and B_1 correspond to the polarisation properties of the second photon. However, using the arguments from the previous section, a large number of indistinguishable entangled pairs would produce $\langle A_i B_j \rangle \approx \langle A_i \rangle \langle B_j \rangle$ in the classical limit. Thus, the CHSH inequality (3) would not be violated.

Let us now discuss another implementation of the CHSH scenario. This time the four-level system is made of a single photon which can occupy four modes, e.g., two polarisation modes (H and V) and two spatial modes (a and b). As a result the
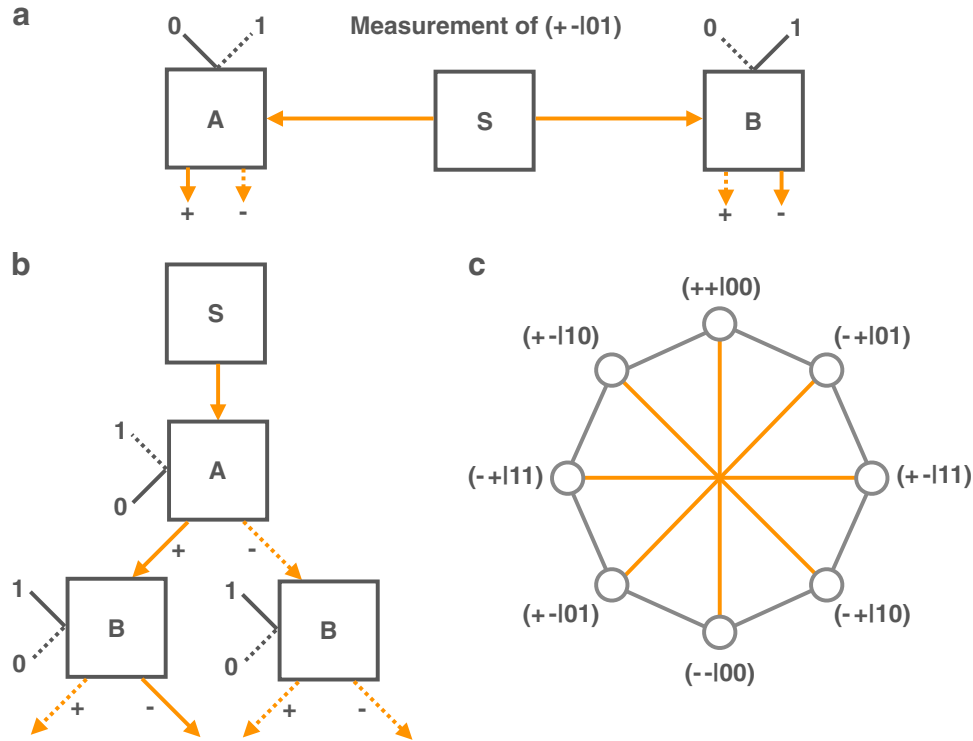
Fig. 2 Clauser-Horne-Shimony-Holt (CHSH) Bell-type scenario. **a** Nonlocal setting in which a source S emits two correlated particles flying to Alice and Bob. Each of them performs one of the two measurements (denoted by 0 or 1). Each measurement produces a binary outcome (±1). Here, we present an instance in which Alice chooses to measure 0 and Bob chooses to measure 1. Alice's outcome is + and Bob's is −, hence they jointly register an event (+ − |01). **b** The same instance, but in a local scenario. Classical entanglement can only be tested in such scenarios. The measurement of A is performed before B. It is generally assumed that both properties are compatible (they commute in QM sense), therefore the order of measurement is irrelevant. **c** The exclusivity graph for the CHSH scenario. The vertices correspond to measurement events and the edges represent the exclusivity relations. Orange edges correspond to exclusivity of measurement outcomes for the same settings, e.g., $(+ + |ij)$ and $(- - |ij)$. Grey edges correspond to exclusivity of measurement outcomes in which the second measurement has different settings, e.g., $(+ + |i0)$ and $(- - |i1)$. This exclusivity can be tested in the setting represented in **b**, by choosing B_0 for the second left measuring device and B_1 for the second right measuring device—detailed discussion in the text

photon can be in one of four possible states a_H, a_V, b_H and b_V, or in an arbitrary superposition of them. The properties A_0 and A_1 can be associated with spatial modes, whereas B_0 and B_1 can be associated with polarisation. For example, A_0 can assign +1 to mode a and −1 to b. On the other hand, A_1 can assign ±1 to orthogonal superpositions of modes, like $|a\rangle \pm |b\rangle$. Similarly, B_0 can assign +1 to polarisation H and −1 to V, whereas B_1 can assign +1 to the right-handed circular polarisation and −1 to the left-handed one.

The Hilbert space of the system is a tensor product of two Hilbert spaces: the one corresponding to spatial modes and another one to polarisation. However, this time the system cannot be divided into parts that can be separated from each other. Still, it is possible to speak of entanglement between these two degrees of freedom, but this entanglement has nothing to do with nonlocality. Nevertheless, violation of the CHSH inequality with A_i and B_j confirms the presence of entanglement between spatial modes and polarisation. This entanglement gives non-classical correlations that can be attributed to contextuality rather than to nonlocality.

The properties A_i and B_j can be measured sequentially, as in the ref. [18] and the measurement of one property does not disturb the measurement of the other. More precisely, such a measurement can be implemented in a setup in which the system goes through the measuring device corresponding to A_i and then through one of the two measuring devices corresponding to B_j. The schematic representation of this setup is shown in Fig. 2b. Because A_i and B_j commute, the results of the measurements do not depend on their order, i.e., B_j can be measured before A_i. The measurements

lead to four possible outcomes that we denote by $(+ + |ij)$, $(+ − |ij)$, $(− + |ij)$ and $(− − |ij)$. The result $(+ − |ij)$ corresponds to $A_i = +1$ and $B_j = −1$.

A single run of the experiment makes one of the four detectors, placed after the outputs, click. The probabilities of these clicks are $p(+ + |ij)$, $p(+ − |ij)$, $p(− + |ij)$ and $p(− − |ij)$. They can be estimated after many experimental runs and used to evaluate correlations $\langle A_i B_j \rangle = p(+ + |ij) − p(+ − |ij) − p(− + |ij) + p(− − |ij)$. One can observe violation of the CHSH inequality if in each experimental run the photon is prepared in the same special state and the measurements A_i and B_j are properly chosen. Although the setup is interpreted as a measurement of two random variables, it can also be viewed as a measurement of a single degenerate random variable X_{ij} whose outcomes are products of the outcomes of A_i and B_j. Therefore, $\langle A_i B_j \rangle = \langle X_{ij} \rangle$.

What would happen if in a single experimental run one used many identical photons or a classical beam of light? From our initial discussion we know that the intensities at the outputs would be proportional to $Np(+ + |ij)$, $Np(+ − |ij)$, $Np(− + |ij)$ and $Np(− − |ij)$, where N is the number of photons. In the classical limit one would deal with a beam of light whose intensities would be $I(+ + |ij)$, $I(+ − |ij)$, $I(− + |ij)$ and $I(− − |ij)$. Moreover, $I(+ + |ij)/I = p(+ + |ij)$, etc., where I is the input intensity.

In addition, one could consider a random variable

$$\mathcal{X}_{ij} = \frac{n(+ + |ij) − n(+ − |ij) − n(− + |ij) + n(− − |ij)}{N}, \tag{5}$$

where $n(+ + |ij)$ is the number of photons in the output $(+ + |ij)$, etc. For a single photon $\langle A_i B_j \rangle = \langle X_{ij} \rangle = \langle \mathcal{X}_{ij} \rangle$. For $N > 1$ it is

impossible to assign definite values to A_i, B_j and to assign X_{ij} to individual photons. However, $\langle X_{ij} \rangle$ can be evaluated and in the classical limit one gets

$$\langle X_{ij} \rangle = \frac{I(++|ij) - I(+-|ij) - I(-+|ij) + I(--|ij)}{I}. \tag{6}$$

Thus, it is possible to prepare a classical state of light such that

$$\langle X_{11} \rangle + \langle X_{01} \rangle + \langle X_{10} \rangle - \langle X_{00} \rangle = \pm 2\sqrt{2}. \tag{7}$$

The above may lead to a discussion whether the classical light has some nonclassical properties.[7,8,10,12–14,16–18] In the following sections we show that for more than one photon the classical bound is different than ±2. One needs to remember that although $\langle X_{ij} \rangle$ does not depend on N, the random variable X_{ij} and the corresponding exclusivity structure of events strongly depends on N, therefore in order to understand what is really going on it is better to examine the CHSH scenario from the point of view of events, not averages.

Exclusivity and classical bounds

The contextuality and non-locality scenarios can be studied within the graph theoretical model developed by Cabello, Severini, and Winter.[22] This model is based on the exclusivity structure of measurable events and offers three different approaches to deriving bounds on sums of probabilities of measurable events. The first approach assigns logical truth/false values to each event and the resulting bound on the sum of probabilities, obtained by optimising over all possible assignments, is known as the classical bound. The second approach is equivalent to the quantum mechanical way of assigning probabilities to the events, which is based on the Born rule. The corresponding quantum bound is known as Lovasz theta function and is equal or greater than the classical bound. Finally, the third approach uses the minimal assumption, namely that the sum of probabilities of mutually pairwise exclusive events is bounded by one. This assumption is known as the E-principle[25] and the corresponding bound is known as the non-signaling bound, since general non-signaling theories obey the E-principle. The non-signaling bound is equal or grater than both, classical and quantum bounds. It is worth to mention that there is a growing interest in finding ways to derive quantum bounds from non-signaling bounds using additional physical assumptions.[25–27] In the following we discuss the CHSH scenario from the point of view of its exclusivity structure.

The CHSH inequality can be rewritten with probabilities of detection events. Since $\langle A_i B_j \rangle = 1 - 2p(+-|ij) - 2p(-+|ij) = 2p(++|ij) + 2p(--|ij) - 1$, the inequality (3) becomes

$$\begin{aligned} &p(+-|11) + p(-+|11) + p(+-|01)+ \\ &p(-+|01) + p(+-|10) + p(-+|10)+ \\ &\quad p(++|00) + p(--|00) \le 3. \end{aligned} \tag{8}$$

This inequality can be derived in a completely different way. The upper bound equal to three comes from the exclusivity structure of events. Firstly, the events $(++|ij)$, $(--|ij)$, $(+-|ij)$, and $(-+|ij)$ are pairwise exclusive. This is because they correspond to different outcomes of the same measurements. For example, $(++|00)$ cannot happen together with $(--|00)$. In addition, two events are exclusive if they share the same measurement settings and the corresponding outcomes are different. This means that $(+\#|ij)$ is exclusive to $(-\#|ik)$ and $(\#+|ij)$ is exclusive to $(\#-|kj)$; Here $\#$ denotes an arbitrary outcome. For example, $(+-|10)$ is exclusive to $(-+|11)$ and $(--|00)$ is exclusive to $(-+|10)$. Such example can be realised in quantum theory by events corresponding to projections onto states $|0\rangle \otimes |0\rangle$ and $|1\rangle \otimes (\alpha|0\rangle + \beta|1\rangle)$. Although $|0\rangle$ and $\alpha|0\rangle + \beta|1\rangle$ are in general nonorthogonal states, the exclusivity is provided by the orthogonality of $|0\rangle$ and $|1\rangle$ in the first Hilbert space. Verification of this type of exclusivity can be implemented in the sequential

scenario represented in Fig. 2b in which the second left measuring device is set to B_0 and the second right to B_1.

The exclusivity structure of the eight events can be represented with the exclusivity graph[22] whose vertices correspond to events and edges to exclusivity between two events, see Fig. 2c. The upper bound of (8) is derived under assumption that the eight events are jointly distributed.[28] The joined probability distribution (JPD) is constructed over all possible assignments of 1/0 (truth/false) values to these events. In principle there are 2^8 possible assignments, however the value 1 cannot be simultaneously assigned to two exclusive events. This significantly reduces the number of possible assignments. The maximum value of the sum of the eight probabilities is given by the maximal number of events that can be assigned the value 1. The problem of finding the maximal number of events that can be assigned 1 is equivalent to the graph theoretical problem known as maximum independent set.[22] An independent set of a graph is a set of disconnected vertices. We are looking for a set with the largest possible number of vertices. In general, it is an NP-hard problem but it is solvable for our graph. Note, that the set of events that are assigned 1 must correspond to the independent set of the exclusivity graph, since two events from such set cannot be exclusive. It is easy to find that the maximum independent set of the graph from Fig. 2c contains three vertices. Therefore, the sum of the eight probabilities cannot be larger than three if these probabilities originate from some JPD. Not surprisingly, quantum theory can go as high as $2 + \sqrt{2}$ and it cannot be modelled with any JPD.

(Non-)contextuality of many indistinguishable particles and a proper classical bound

Hidden variable theories that aim to explain standard photonic experiments associate outcomes with detector clicks. However, here we show that this approach cannot be applied to experiments where more than one indistinguishable photon arrives at the detector. Instead, we propose an alternative hidden variable model, which provides a classical explanation of the Bell-like experiments done on electromagnetic waves. Within the entire discussion we adapt an operational approach in which a photon, or in general an indistinguishable particle, is identified with a detector click. The 1/0 assignment corresponds to a deterministic non-contextual (NC) model. The photon is assigned at most one event from each set of pairwise exclusive events. Such a set makes a measurement context—a set of events that can be jointly measured. If a context is complete, i.e., it consists of all possible measurement outcomes, the photon is assigned exactly one event. However, in the scenario considered here all contexts are not complete and contain exactly two events.

To properly discuss the problem of non-classicality of correlations in Bell-type scenarios for classical light, we need to redefine the introduced exclusivity graph model so that a transition from a single photon to a macroscopic electromagnetic wave is transparent. Instead of assigning events to a photon one should tie a photon to an event. This is a subtle difference but it leads to fundamental consequences once we deal with more than one photon. More precisely, a photon is assigned to at most one single event in each measurement context, where 1 corresponds to a photon and 0 to no photon event. In this new picture the events can be considered as modes and 1/0 as occupation numbers. The NC model assigns a well defined occupation number to each mode. The exclusivity of modes leads to conservation of the particle number—since there is a single photon in the system there could be at most a single photon in each context. If two exclusive events (modes) were assigned one, then there would exist a context containing two photons, which would contradict conservation of the particle number. The above interpretation was proposed for the first time in the ref.[29] This approach is discussed
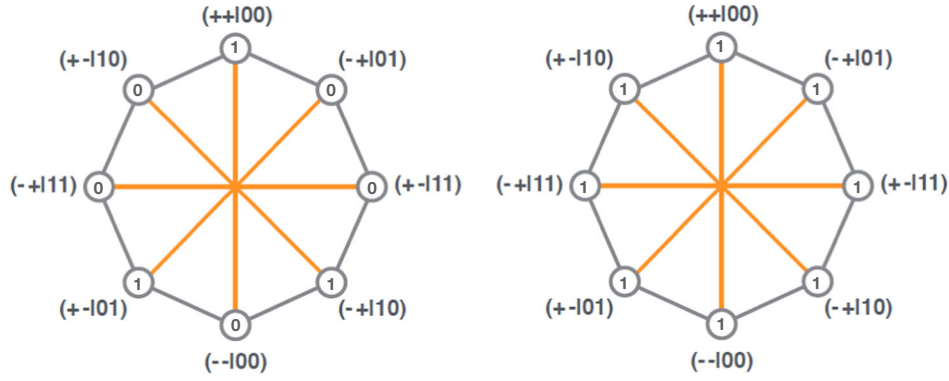
Fig. 3 Examples of particle-assignments to events in the CHSH exclusivity graph. Left—one particle, right—two particles

in details in the Methods section. It should be emphasised that the introduced model is very general and describes the single-photon-to-classical-wave transition in Bell-type scenarios irrespective of the direct physical implementation, which may be introduced in many different scenarios.[7,10,14,16–19]

In order to make our discussion as general as possible and not restricted to the case of classical and quantum optics let us consider the above described model for arbitrary indistinguishable particles. Our generalised model allows for arbitrary schemes of assigning particles to modes, therefore no physical constraints on the statistics (bosonic or fermionic behaviour), correlations and possible interactions are made. The only physical assumption we impose, is that there exists a macroscopic limit of a strong beam of the particles possessing two properties. The first one states that the occupation number ratios tend to intensity ratios of the macroscopic beam. The second one assumes that the ratio of the standard deviation of the particle number in a given mode to the average particle number tends to zero:

$$\frac{\sigma_n}{\langle n \rangle} \to 0. \tag{9}$$

The last assumption guarantees that the ratios of intensities of the macroscopic beam are fully deterministic. Note that the physical example of single-photon-to-classical-wave transition satisfies both conditions, since macroscopic EM wave can be treated as a limit of a strong beam of photons in a coherent state with particle number fluctuations following Poissonian statistics with standard deviation σ_n of the form $\sqrt{\langle n \rangle}$. In the introduced model the notion of exclusivity changes its implementation. Namely let us assume that at most N photons can be assigned to a single event. Then each experimental context, that is any set of mutually exclusive events, can be also filled with at most N photons.

Let us now discuss the bound for a CHSH inequality within the above model. One can rewrite the inequality (8) as

$$\langle n(+-|11)\rangle + \langle n(-+|11)\rangle + \langle n(+-|01)\rangle +$$
$$\langle n(-+|01)\rangle + \langle n(+-|10)\rangle + \langle n(-+|10)\rangle + \tag{10}$$
$$\langle n(++|00)\rangle + \langle n(--|00)\rangle \leq \mathcal{C},$$

where $n(++|ij)$, etc., are occupation numbers of the corresponding events and $\mathcal{C}$ is the NC bound on the sum of these numbers, which in the case of a single particle equals to three. A single particle violates this bound.

Now, consider the same CHSH scenario, but this time inject two indistinguishable particles to the system. The exclusivity and particle number conservation imply that there could be at most two particles per context. The possible occupation numbers are 0, 1, or 2. Since each context consists of only two events, one can assign a single particle to each event, see Fig. 3. Therefore, for two particles $\mathcal{C} = 8$, which is the maximal possible sum of non-

contextually assigned occupation numbers over all events. We see that the bound depends on the number of particles. We divide (10) by N and rewrite it as

$$\frac{\langle n(+-|11)\rangle}{N} + \frac{\langle n(-+|11)\rangle}{N} + \frac{\langle n(+-|01)\rangle}{N} +$$
$$\frac{\langle n(-+|01)\rangle}{N} + \frac{\langle n(+-|10)\rangle}{N} + \frac{\langle n(-+|10)\rangle}{N} + \tag{11}$$
$$\frac{\langle n(++|00)\rangle}{N} + \frac{\langle n(--|00)\rangle}{N} \leq \mathcal{P}_C(N),$$

where $\mathcal{P}_C(N) = C/N$. For a single particle $\mathcal{P}_C(1) = 3$, whereas for two particles $\mathcal{P}_C(2) = 4$. Therefore, for $N = 2$ there is no violation and the measurements can be described by NC occupation number assignments.

As far as we know, the value $\mathcal{P}_C(2) = 4$ cannot be reached in any experimental setup, although it is allowed in our model. This is becasue the maximal quantum value of $2 + \sqrt{2}$ (attainable in the CHSH scenario) does not depend on the physical implementation of the experiment. In particular it does not depend on the dimension of the state space of the physical system, which in our case translates to independence on the particle number N.

Finally, let us consider the classical limit $\langle N \rangle \gg 1$. This time the system is described by a classical beam of intensity I for which the inequality (11) reads

$$\frac{I(+-|11)}{I} + \frac{I(-+|11)}{I} + \frac{I(+-|01)}{I} +$$
$$\frac{I(-+|01)}{I} + \frac{I(+-|10)}{I} + \frac{I(-+|10)}{I} + \tag{12}$$
$$\frac{I(++|00)}{I} + \frac{I(--|00)}{I} \leq \mathcal{P}_{cl},$$

where N in the denominator of (11) was replaced by $\langle N \rangle$ due to the particle number uncertainty. The maximal experimentally attainable value of the left-hand side for an EM wave is still $2 + \sqrt{2}$ because the classical light beam in any linear optical setup behaves in the same way as a single-photon probability amplitude. The right-hand side can be evaluated in two ways. Firstly, in the classical limit the total intensity$I$ that is distributed between the events can be treated as a continuous property. Therefore, in the NC model one can assign $I/2$ to each event and as a result $\mathcal{P}_{cl} = 4$. This is the main result in this section: The corresponding CHSH inequality (12) cannot be violated by any macroscopic beam following assumptions of our model, like for example the classical light.

The other approach, which also confirms the above result, does not assume that the intensity is a continuous property. We consider two cases. First, let us take even N. The number of particles per context cannot be greater than N and since each context contains two events, one simply assigns $N/2$ particles per context. This leads to $\mathcal{P}_C(N) = 4$. Next, we consider odd N. In this case it is easy to show that one can assign $(N-1)/2$ particles to five events and $(N+1)/2$ particles to three events, such that there are at most N particles per each context, see Fig. 4. As a result one gets $\mathcal{P}_C(N) = 4 - \frac{1}{N}$. In the classical limit N is undetermined,
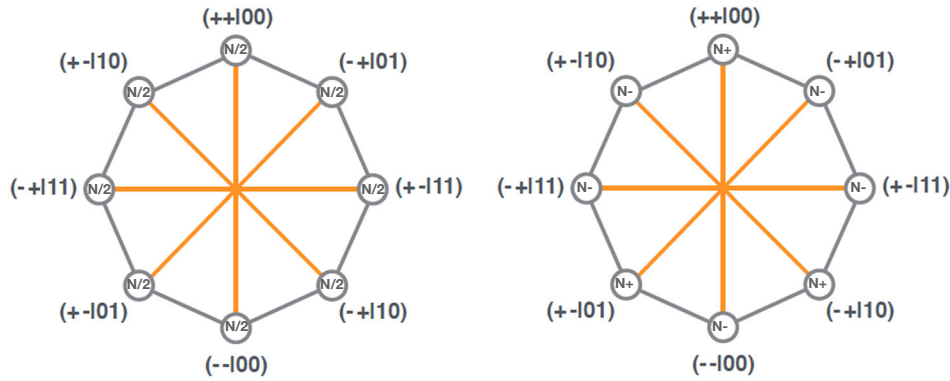
Fig. 4 Assignment of particles to events in the exclusivity graph of the CHSH scenario. Left—even number of particles, right—odd number of particles, where $N_+ = (N+1)/2$ and $N_- = (N-1)/2$

therefore $\mathcal{P}_{cl} = \langle \mathcal{P}_C(N) \rangle$. However, since $\langle N \rangle \gg 1$ the dominating terms in $\langle \mathcal{P}_C(N) \rangle$ correspond to large values of N and hence $\mathcal{P}_{cl} \approx 4$.

To conclude, we see that in the experiments discussed above light beams, as expected, do not exhibit any quantum behaviour. Quantum behaviour is only possible for a single photon and already for $N > 1$ one observes noncontextual (classical) behaviour. This is because for $N \geq 2$ the non-contextual bound $\mathcal{P}_C(N)$ is either 4 or $4 - \frac{1}{N}$ and is always greater than the physically attainable value of $2 + \sqrt{2} \approx 3.41$. We have only considered the CHSH scenario, however in the Methods section we show that for an arbitrary contextuality scenario the bound $\mathcal{P}_{cl}$ is always greater or equal than what can be achieved in the classical limit and therefore classical systems are always noncontextual and can never violate any Bell-type inequality.

DISCUSSION

It is commonly accepted that the violation of some Bell inequality is an indicator of non-classical behaviour. It turns out that these violations are strongly related to the wave-particle duality. In quantum theory the same object can manifest a wave-like behaviour in one experiment and a particle-like behaviour in some other experiment. The wave-particle duality does not occur in classical theory, i.e., classical objects are either waves or particles, but never both. No one ever doubts that classical particles are localised and discrete whereas classical waves are delocalised and continuous.

Bell inequalities are derived under assumption that measured properties, such as position, are well defined and the phenomenon of superposition does not occur. This is a typical particle-like approach which can be partially justified in quantum regime by the fact that at the end of each experiment one registers a single click of some detector. However, the violation of the corresponding inequality implies that this assumption needs to be reconsidered. This is because the measurements in the Bell tests exploit the phenomenon of superposition and in this sense the violation is a manifestation of the wave-like behaviour.

The situation is different in the classical theory in which there is no wave-particle duality. In particular, the classical light behaves as a wave all the time and instead of clicks one registers continuous intensities. Therefore, there is no justification to apply the standard Bell inequality to such system. In this case Bell inequalities need to be rederived using properly chosen assumptions. This is what we show in this work.

Our results have the following implications on the previously discussed Bell-type tests with classical light. It was stressed by Snoke[17] and Qian et al.[16] that violation of some Bell inequality by classical light is in fact a confirmation of the presence of some kind of entanglement. Here we show that the value of $2 + \sqrt{2}$

obtainable by classical light can indicate some form of strong correlations, but these correlations are fully describable by a noncontextual model. Therefore, classical entanglement does not give rise to contextual behaviour, it does not have any features of quantum entanglement apart from mathematical analogy on the level of complex vector spaces.

Next, Frustaglia et al.[18] argued that the bounds on the strength of quantum correlations (so called Tsirelson bounds) are not restricted to quantum theory but arise naturally in classical systems that simulate quantum correlations. There are attempts to derive these bounds using only the exclusivity structure of detection events.[25,30] However, we showed that in the CHSH scenario discussed above the classical bound resulting from the exclusivity structure of detection events is $\mathcal{P}_{cl} = 4$. Therefore, the value of $2 + \sqrt{2}$ in the classical regime must originate from some physical constraints. Our model for assigning particles to modes allows for an arbitrary distribution of particles across the modes, since no concrete mechanism for such a distribution is specified. Adding appropriate constraints on the possibility of interaction between the particles and the way they can be correlated may lead to a recovery of the Tsirelson bound.

It is interesting to understand why quantum systems and classical light lead to the same value. The reason for this is that quantum probability amplitudes and classical electromagnetic amplitudes transform in the same way (apart from the non-deterministic quantum collapse induced by the measurement). This is why classical wave theory can simulate some aspects of quantum theory. Nevertheless, the interpretation of both amplitudes is completely different. The value of $2 + \sqrt{2}$ would be much more fundamental if it resulted from both, the exclusivity structure of events and from the transformations allowed by the theory.

Finally, we would like to note that our work leads to an open problem. Is it possible to find a classical system, other than a classical wave, which would provide a value larger than $2 + \sqrt{2}$? Perhaps there are some additional constraints preventing this from happening.

METHODS

Nonlocality and contextuality

Original Bell inequalities are statistical tests that verify whether the statistics of measurements' outcomes performed on spatially separated systems fulfil the conditions of locality and realism.[9] Locality means that any action executed on one system does not affect the other one. Realism means that each measurable property has a well defined outcome, irrespective of whether this property is measured or not. Initially Bell-type tests were derived from the properties of probability distributions for correlations of measurement outcomes.[9] It was early recognised by Fine[28] that any Bell inequality is equivalent to the existence of a joint probability distribution (JPD) for all measurable properties.

Violation of Bell inequalities by quantum systems, typically referred to as quantum nonlocality, is an example of a broader class of phenomena—quantum contextuality.[21] Simply speaking, contextuality is a property of a physical system where the outcome of some property A may depend on whether it is measured with B or with C. In typical contextuality scenario one considers a system with a number of measurable properties. The goal is to find a noncontextual (NC) assignment of outcomes to all measurements, i.e., to assign an outcome to each property in a way that does not depend on what is assigned to other properties. Just like in the case of nonlocality, contextuality is equivalent to a lack of JPD for all measurable properties.[28] Finally, note that noncontextual scenarios do not require space-like separated measurements, therefore Bell-type scenarios constitute a subclass of contextual scenarios, in which the commeasurable observables can be identified with spatially separated physical systems.[31] This is the reason why we focus on a more general phenomenon of contextuality.

Contextuality scenarios underline the role of the exclusivity structure of events behind Bell-type inequalities, which is not explicit in the standard correlation-based approach. In a series of papers[22,25,32,33] Cabello et. al. show that all contextuality tests can be derived from the exclusivity structure of measurement events. Assume that the set of all events v_i, the precise meaning of which is defined separately in the physical scenario, can be decomposed into subsets $\mathcal{C}_k = \{v_r\}$ called measurement contexts, such that all the events in a single context correspond to outcomes of some experiment. Therefore, the events in a single context can be jointly measured. Because of various constraints, some of the events cannot be simultaneously true, the property which is known as exclusivity in probability theory. We point out that the notion of exclusivity is a fundamental property of Kolmogorovian probability theory, since elementary events, which constitute a sample space in any statistical model, must be mutually exclusive.

The exclusivity structure of the set of events for a contextuality test can be represented in the exclusivity graph[22] in which adjacent vertices represent exclusive events. An example of such a graph is presented in Fig. 2c. For commeasurable observables A_i with outcomes a_i, each vertex represents an event, which is a conjunction of all single detection events for a fixed experimental context. Namely $v_k = (a_1, \ldots, a_n | A_1, \ldots, A_n)$, where (with a slight abuse of notation) the set $\{a_i\}$ represents actual measurement outcomes, obtained for fixed properties $\{A_i\}$ that constitute a measurement context.

Two events $v_a = (a_1, \ldots, a_n | A_1, \ldots, A_n)$ and $v_b = (b_1, \ldots, b_n | B_1, \ldots, B_n)$ in some contextuality scenario are exclusive if and only if:

$$(A_1 = B_1 \,\&\, a_1 \neq b_1)\text{OR}(A_2 = B_2 \,\&\, a_2 \neq b_2)\text{OR}\ldots\text{OR}(A_n = B_n \,\&\, a_n \neq b_n). \tag{13}$$

Having defined the exclusivity structure for a given scenario, one may attach different kinds of probabilistic structures to the set of events. We stress that the phrase probabilistic structure does not refer to a Kolmogorovian model, instead it can be understood within the context of generalised probabilistic theories.[34] The models can be ordered with respect to the strength of allowed correlations and this strength can be measured by the maximal value of $\mathcal{P} = \sum_i p(v_i)$ allowed within the model. Note, that although here we consider the inequalities $\sum_i c_i p(v_i) \leq \mathcal{P}_C$ with all coefficients $c_i = 1$, it is important to observe that any inequality with rational coefficients (including standard correlation inequalities) can be transformed into one with all coefficients equal to one—for details see ref. [35]

We discuss three classes of models. The most restrictive one is the classical model. It assumes that all events can be attributed to a single sample space (not necessarily as atomic events, but compound events as well), and that a joint probability distribution exists for all v_i. Then the maximal value of $\mathcal{P}$, denoted by $\mathcal{P}_C$, known as noncontextual local hidden variable (NCHV) or simply classical bound, is given by the vertex independence number of the exclusivity graph—the maximal number of non-adjacent vertices. To rephrase, the classical bound $\mathcal{P}_C$ corresponds to the maximal number of events that can be assigned truth value 1. This mathematical model represents a physical situation where all observables are commeasurable.

The second model, involving stronger correlations, is the quantum probabilistic model.[36] Here each event corresponds to a projector $P(v_i)$. Each projector is assigned a probability $p(v_i) = \text{Tr}(\rho_{Pi}(v_i))$, where ρ represents a quantum state. For the context $\mathcal{C}_k$ consisting of mutually exclusive events the corresponding projectors make a mutually orthogonal

set and therefore:

$$\sum_{v_i \in \mathcal{C}_k} P(v_i) \leq 1. \tag{14}$$

This means that either (equality) the projectors within a context represent a von Neumann measurement or (sharp inequality) a von Neumann measurement extension is possible. The maximum value $\mathcal{P}_Q$ of $\mathcal{P}$ for a quantum model, known as the Tsirelson's bound, is bounded by the Lovász number of the exclusivity graph.[22] This bound is often tight.[32] The quantum model is contextual if and only if $\mathcal{P}_C < \sum_i p(v_i) \leq \mathcal{P}_Q$. This means that an assignment of outcome probabilities can be done within each context separately and that the model cannot be extended to a single Kolmogorovian probabilistic model.

The third model, leading to the strongest correlations, is defined by a sole demand that its probabilistic structure fulfils the exclusivity (E) principle in its final version[25] (whose other variant, known as local orthogonality, was independently developed in[30]). The E principle states that the sum of probabilities corresponding to a subset of pairwise exclusive events is bounded by 1. This principle is a conjunction of two facts: (1) Cabello's principle called the Specker's principle,[33] which states that the set of pairwise exclusive events is jointly exclusive, and (2) the property of a Kolmogorovian model that the sum of probabilities of jointly exclusive events is at most 1. Therefore, the probabilistic model obeying the E principle can be defined as the set of numbers $0 \leq p(v_i) \leq 1$, fulfilling the normalisation condition $\sum_{v_i \in \mathcal{C}_k} p(v_i) \leq 1$ for each context.

Note that both, classical model and the quantum model, obey the E principle. This does not mean that a model obeying the E principle is classical or quantum. There are models obeying the E principle that are more contextual than quantum theory. This is because the maximal possible value $\mathcal{P}_E$ of $\mathcal{P}$ for the model obeying the E principle can be greater than $\mathcal{P}_Q$. One can derive it using the properties of an exclusivity graph. The detailed derivation is given in the appendix of[25] and states that the maximal possible value is given by the fractional packing number of the graph, defined as:

$$\max \sum_i \omega_i, \text{ s.t. } \left(\forall_i 0 \leq \omega_i \leq 1 \text{ and } \sum_{i \in \mathcal{S}} \omega_i \leq 1 \right), \tag{15}$$

The maximum is taken over all cliques $\mathcal{S}$ of the graph. A clique is a subset of mutually adjacent vertices (pairwise exclusive events). Therefore the cliques of an exclusivity graph correspond to all the contexts $\mathcal{C}_k$. The Lovász number is bounded by the fractional packing number. However in many cases the fractional packing number is larger than the Lovász number.

To conclude, the bounds for the three models obey the following relation

$$\mathcal{P}_C \leq \mathcal{P}_Q \leq \mathcal{P}_E. \tag{16}$$

For example, in the case of the exclusivity graph presented in Fig. 2c one has $\mathcal{P}_C = 3$, $\mathcal{P}_Q = 2 + \sqrt{2}$ and $\mathcal{P}_E = 4$. In general, if the sum of probabilities of events in the exclusivity graph is bounded by $\mathcal{P}_C$ then we say that the model is noncontextual. Otherwise it is contextual. It is worth mentioning that in the context of Bell inequalities the bound $\mathcal{P}_E$ coincides with the non-signaling bound for the corresponding inequality.

Contextuality and indistinguishability

Now we discuss events that occur in scenarios in which instead of a single particle one uses N indistinguishable particles. If the number of particles N in the system increases so does the number of possible detection events. This affects the structure of the exclusivity graph and that of bounds $\mathcal{P}_C$, $\mathcal{P}_Q$ and $\mathcal{P}_E$. Instead of changing the graph, one can keep the exclusivity graph in the same form and change the range of values that are assigned to each event, which warrants a slight redefinition of the entire scenario defined precisely in section II.E. Here we focus on the classical model because we want to find a criterion for which the system of N indistinguishable particles is noncontextual. We will show that within the classical model the bound $\mathcal{P}_C$ depends on N, i.e., $\mathcal{P}_C = \mathcal{P}_C(N)$. Finally, we show that in the classical limit $\mathcal{P}_{cl} = \mathcal{P}_E$, therefore classical light, for which the ratio of output intensities to the initial intensity is the same as the probabilities generated by a single photon, is always noncontextual.

For $N = 1$ one assigns truth values 1/0 to each event and tries to maximise the number of events that are assigned 1 under the exclusivity constraint. The value 1 corresponds to events that will be observed in an experiment, provided a proper measurement is performed, whereas 0

corresponds to events that will not be observed. However, 1 can be interpreted as assignment of a particle to an event and 0 as an assignment of the absence of a particle. One can generalise this approach to the case $N > 1$. This time each event v_i is assigned a value $n_i = 0, 1, \ldots, N$. The values n_i can be interpreted as occupation numbers. This approach was suggested in the ref.[29] The exclusivity of v_i and v_j translates to $n_i + n_j \leq N$. In general, the sum of occupation numbers in each context $\sum_{v_i \in C_k} n_i \leq N$. Therefore, the exclusivity of events implies that the number of particles per context cannot be larger than the total number of particles. If the context consists of all measurement outcome events then the number of particles in it must be equal to N. This corresponds to the particle-number conservation. In full analogy to the original model, which assigns probabilities to the events, the strength of correlations for assigning particles to the events is represented by the expression $\mathcal{P}(N) = \frac{1}{N} \sum_i n_i$, where the sum is taken over all events in the exclusivity graph. The classical model in the original scenario translates to an occupation-based model with fixed global assignment of occupation numbers. In the redefined scenario we assume that the ratios of occupation numbers tend to intensities of a macroscopic beam. Moreover we assume, that the fluctuations of the particle number described by the standard deviation σ_n follow the assumption $\sigma_n/\langle n \rangle \to 0$, which guarantees that the intensities are deterministic quantities.

We are looking for the maximal value $\mathcal{P}_C(N)$ of the expression $\mathcal{P}(N)$ optimised over all allowed NC global assignments. For $N = 1$ the NC occupation-based model has the same constraints as the deterministic model with assigned probabilities, thus $\mathcal{P}_C(1) = \mathcal{P}_C$. However, for $N > 1$ the constraints are different and $\mathcal{P}_C(N) \neq \mathcal{P}_C$. In particular, we are interested in the classical limit $\langle N \rangle \gg 1$. The system, corresponding to a classical beam, is modelled by a collection of particles whose total number is undetermined, but its average number is large. We are looking for $\mathcal{P}_{cl}$ which is the maximum of $\frac{1}{\langle N \rangle} \sum_i \langle n_i \rangle$. We write $l_i \equiv \langle n_i \rangle$. We call l_i the intensity assigned to the event v_i. There is a fundamental difference between intensities and occupation numbers. In general $\langle n_i \rangle \neq n_i$, however due to our assumptions the classical limit is deterministic and therefore $\langle l_i \rangle = l_i$. Moreover, unlike occupation numbers, intensities can be treated as continuous variables $l_i \in [0, I]$, where I is the total intensity. In this case we are looking for $\mathcal{P}_{cl}$ which is a maximum of

$$\sum_i \frac{l_i}{I}, \tag{17}$$

where $I = \langle N \rangle$ is the total intensity. The ratios $0 \leq l_i/I \leq 1$ are continuous numbers and the only constraint on them is that within each context: $\sum_{v_i \in C_k} \frac{l_i}{I} \leq 1$.

By taking $p_i = l_i/I$ the above model is equivalent to the third model of the original scenario, which only needs to obey the E principle. Therefore the noncontextuality bound in the classical limit is $\mathcal{P}_{cl} = \mathcal{P}_E$. To sum up, any classical beam which can be described by deterministic intensity ratios defined as the limits of particle's occupation number ratios, fulfils the following relation:

$$\sum_i \frac{l_i}{I} \leq \mathcal{P}_E = \mathcal{P}_{cl}, \tag{18}$$

which proves that it is always noncontextual.

Non-applicability of the methods used to derive the quantum bound $\mathcal{P}_Q$ with the E principle

In[25] a general method of derivation of the maximal quantum bound for a given Bell-type scenario was introduced. This method is based on the idea, that the E principle can be applied not only to a single copy of a Bell-type test represented by some exclusivity graph G, but also to k independent runs of such identical Bell tests carried on different physical systems. Such a compound experiment is described by an exclusivity graph that is represented by the so called disjunctive product G^k of k copies of the graph G.[37] It is assumed that the E principle should be applicable also to any clique (representing measurement context) in the graph G^k. It might seem that such a procedure should not lead to any new conclusions, since the carried k experiments are completely independent. However it turns out that application of the E principle to a clique in G^k restricts the possible assignments in much stronger way than in the case when it is applied to a single copy of G. This is because each vertex V of G^k represents a joint event, which consists of a conjunction of k independent events corresponding to some k vertices $\{v_i\}$ of G. Therefore the assigned

probability factorises into $p(V) = p(v_1) \cdot \ldots \cdot p(v_k)$. It can be easily shown that application of E principle to some clique in G^k consisting of some number of vertices V places stronger restriction on the possible values of $p(v_i)$.

The easiest possible case is given by the contextuality test of Klychko-Can-Binicioglu-Schumovsky (KCBS),[38] in which the original exclusivity graph G is a 5-cycle, and therefore the E principle allows for assignment of the probability at most $\frac{1}{2}$ to each vertex leading to the bound $\mathcal{P}_E = \frac{5}{2}$. On the other hand the graph G^2 representing the exclusivity structure for two independent KCBS experiments has a 5-vertex clique K, and the E principle applied to K allows for assignment of a probability at most $\frac{1}{5}$ to each vertex of K. Note that a probability assigned to each vertex V_i of K is a product of probabilities corresponding to different subsets of vertices of G: $p(V_i) = p(v_i)p(u_i)$. Now, K can be chosen such that the vertices in K, which are of the form $\{v_i \times u_j\}$, contain all the vertices from the elementary graphs G. Hence assignments $p(v_i)$ and $p(u_j)$ are also assignments to the entire graph G of a single KCBS test. Since $\sum_{i=1}^5 p(v_i)p(u_i) \leq 1$, and the set $\{p(v_i)\}$ is any permutation of the set $\{p(u_j)\}$, we obtain the restriction $\sum_i p(v_i) \leq \sqrt{5}$, which exactly reproduces the quantum bound $\mathcal{P}_Q$.

It can be easily seen that the above derivation cannot be applied to the case of classical waves, described by the NC occupation-based model defined in the previous section. This is because in the limiting case of the modified model the vertices of the graph are assigned relative intensities of light instead of probabilities of events (17). Now taking two such independent experiments the intensity of two independent waves is not a product of their intensities but instead it is their sum (assuming they propagate in a linear medium). On the other hand, the joint probability of two independent events is a product of their probabilities. This implies that the entire derivation of the bound cannot be performed in the case of correlations for classical waves, since it is based on the assumption that the relative quantities which describe single events translate to their products in the product graph for two independent experiments. This property holds for probabilities, but not for macroscopic intensities, which shows, that the exclusivity model for two independent experiments based on disjunctive product of exclusivity graphs has no well-defined macroscopic limit. Nevertheless classical waves still obey the bound $\mathcal{P}_Q$. As we showed, this cannot be directly derived from the exclusivity structure of the experiment. It holds because classical waves follow the same evolution rules as quantum probability amplitudes.

AUTHOR CONTRIBUTIONS

All authors researched, collated, and wrote this paper.

ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

REFERENCES

1. Pan, J.-W. et al. Multiphoton entanglement and interferometry. *Rev. Mod. Phys.* **84**, 777–838 (2012).
2. Hong, C. K., Ou, Z. Y. & Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**, 2044–2046 (1987).
3. Lo, H.-P. et al. Experimental violation of bell inequalities for multi-dimensional systems. *Sci. Rep.* **6**, 22088 (2016).
4. Jin, R.-B. et al. Monotonic quantum-to-classical transition enabled by positively correlated biphotons. *Phys. Rev. A* **95**, 062341 (2017).
5. Ramanathan, R., Paterek, T., Kay, A., Kurzynski, P. & Kaszlikowski, D. Local realism of macroscopic correlations. *Phys. Rev. Lett.* **107**, 060405 (2011).

6. Aspden, R. S., Padgett, M. J. & Spalding, G. C. Video recording true single-photon double-slit interference. *Am. J. Phys.* **84**, 671–677 (2016).

7. Suppes, P., Acacio de Barros, J. & Sant'Anna, A. S. A proposed experiment showing that classical fields can violate bell's inequalities. https://arxiv.org/abs/quant-ph/9606019 (1996).

8. Spreeuw, R. J. C. A classical analogy of entanglement. *Found. Phys.* **28**, 361 (1998).

9. Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969).

10. Spreeuw, R. J. C. Classical wave-optics analogy of quantum-information processing. *Phys. Rev. A* **63**, 062302 (2001).

11. Brunner, N., Cavalcanti, D., Pironio, S., Scarani, V. & Wehner, S. Bell nonlocality. *Rev. Mod. Phys.* **86**, 419–478 (2014).

12. Qian, X.-F. & Eberly, J. H. Entangled states of light in classical polarization theory. *Opt. Lett.* **36**, 4110 (2011).

13. Qian, X.-F. & Eberly, J. H. Entanglement is sometimes enough. https://arxiv.org/abs/1307.3772v1 (2013).

14. Aiello, A., Töppel, F., Marquardt, Ch, Giacobino, E. & Leuchs, G. Quantum-like nonseparable structures in optical beams. *New J. Phys.* **17**, 043024 (2015).

15. Wolf, E. *Introduction to the Theory of Coherence and Polarization of Light* (Cambridge University Press, Cambridge, 2007).

16. Qian, X. F., Little, B., Howell, J. C. & Eberly, J. H. Shifting the quantum-classical boundary: theory and experiment for statistically classical optical fields. *Optica* **2**, 611 (2015).

17. Snoke, D. W. A macroscopic classical system with entanglement. https://arxiv.org/abs/1406.7023 (2014).

18. Frustaglia, D. et al. Classical physics and the bounds of quantum correlations. *Phys. Rev. Lett.* **116**, 250404 (2016).

19. Cerf, N. J., Adami, C. & Kwiat, P. G. Optical simulation of quantum logic. *Phys. Rev. A.* **57**, R1477–R1480 (1998).

20. Mermin, N. D. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Phys. Rev. Lett.* **65**, 1838–1840 (1990).

21. Kochen, S. B. & Specker, E. The problem of hidden variables in quantum mechanics. *J. Math. Mech.* **17**, 59 (1967).

22. Cabello, A., Severini, S. & Winter, A. (Non-)contextuality of physical theories as an axiom. https://arxiv.org/abs/1010.2163 (2010).

23. Loudon, R. Fermion and boson beam-splitter statistics. *Phys. Rev. A* **58**, 4904 (1998).

24. Scully, M. O. & Zubairy, M. S. *Quantum Optics* (Cambridge University Press, Cambridge, 1997).

25. Cabello, A. Simple explanation of the quantum violation of a fundamental inequality. *Phys. Rev. Lett.* **110**, 060402 (2013).

26. Amaral, B., Cunha, M. T. & Cabello, A. Exclusivity principle forbids sets of correlations larger than the quantum set. *Phys. Rev. A* **89**, 030101 (2014).

27. Henson, J. Bounding quantum contextuality with lack of third-order interference. *Phys. Rev. Lett.* **114**, 220403 (2015).

28. Fine, A. Hidden variables, joint probability, and the bell inequalities. *Phys. Rev. Lett.* **48**, 291–295 (1982).

29. Kurzynski, P. Contextuality of identical particles. *Phys. Rev. A* **95**, 012133 (2017).

30. Fritz, T. et al. Local orthogonality as a multipartite principle for quantum correlations. *Nat. Commun.* **4**, 2263 (2013).

31. Markiewicz, M. et al. Unified approach to contextuality, nonlocality, and temporal correlations. *Phys. Rev. A* **89**, 042109 (2014).

32. Cabello, A., Danielsen, L. E., López-Tarrida, A. J. & Portillo, J. R. Basic exclusivity graphs in quantum correlations. *Phys. Rev. A* **88**, 032104 (2013).

33. Cabello, A. Specker's fundamental principle of quantum mechanics. https://arxiv.org/abs/1212.1756 (2012).

34. Short, A. J. & Barrett, J. Strong nonlocality: a trade-off between states and measurements. *New J. Phys.* **12**, 033034 (2010).

35. Cabello, A. Simple method for experimentally testing any form of quantum contextuality. *Phys. Rev. A* **93**, 032102 (2016).

36. Streater, R. F. Classical and quantum probability. *J. Math. Phys.* **41**, 3556 (2000).

37. Hammack, R., Imrich, W. & Klavzar, S. *Handbook of Product Graphs*. (CRC Press, Boca Raton, FL, 2011).

38. Klyachko, A. A., Can, M. A., Binicioglu, S. & Shumovsky, A. S. Simple test for hidden variables in spin-1 systems. *Phys. Rev. Lett.* **101**, 020403 (2008).