Editors' Suggestion

Nonclassicality of bright Greenberger-Horne-Zeilinger–like radiation of an optical parametric source

Konrad Schlichtholz, Bianka Woloncewicz, and Marek Żukowski
International Centre for Theory of Quantum Technologies, University of Gdansk, 80-308 Gdansk, Poland

(Received 26 January 2021; accepted 5 April 2021; published 30 April 2021)

With the emerging possibility to obtain emissions of triples of Greenberger-Horne-Zeilinger–entangled photons via direct parametric generation, we study here bright emissions of this kind which involve higher-order emissions of two triples, three triples, etc. Such states would constitute a natural generalization of the four-mode (two beams plus polarization) squeezed vacuum. We show how to avoid technical difficulties related to straight-ahead generalization of the usual approximate description of parametric down-conversion which uses a classical pump field. Using Padé approximation, we turn the first terms of the nonconverging perturbation expansion into elements of a converging series. This allows us to study nonclassicality of the new bright generalized Greenberger-Horne-Zeilinger states. We present both violations of local realism by such a radiation and simple entanglement indicators tailored for this case.

DOI: 10.1103/PhysRevA.103.042226

I. INTRODUCTION

Multiphoton interferometry is extensively studied and used in the context of revealing nonclassical phenomena [1]. Parametric down-conversion (PDC) is a robust source of entangled photon pairs (for early trailblazing steps see [2]) and with the use of special techniques can be used to observe correlations characteristic of three- and four-photon entangled states [3–6] and cluster and Dicke states [7–9]. These states find use in demonstrating paradoxical predictions of quantum mechanics and are observable in laboratories and in experimental applications of quantum information theory [10], quantum metrology [11], cryptography, communication protocols [12,13], imaging [14], and related research. Thus, nowadays, PDC is considered a versatile tool that can be employed to demonstrate nonclassicality or quantum communication protocols, etc., with quantum optics.

All that we mentioned above is related to techniques and observations related to phenomena due to emissions of a suitable fixed numbers of pairs of PDC photons. Still "bright" states of undefined photon number of pairs, e.g., bright squeezed states of light [15], which can be produced via PDC for higher pump powers [16] and involve induced emission of multiple pairs, also exhibit nonclassical properties (see, e.g., [17–19]) and can be used to demonstrate, e.g., Einstein-Podolsky-Rosen-Bell nonclassicality.

An emblematic example of nonclassical light of undefined photon number is a 2×2 mode bright squeezed vacuum generated via type-II parametric down-conversion which exhibits Einstein-Podolsky-Rosen–like (EPR-like) anticorrelations of Stokes observables for the two beams [20]. Its singletlike form gives the invariance of polarization effects under any pair of identical unitary transformations of polarization performed on both subsystems. Thus, it is commonly considered a generalization of the Bell singlet state, as it shares a lot of its properties [21].

The following question emerges. As the 2×2 mode squeezed vacuum can be a generalization of the singlet Bell state, can we have similar analogs for Greenberger-Horne-Zeilinger (GHZ) states, i.e., states that demonstrate quantum features of the GHZ state for qubits and simultaneously have an undefined photon number? Can they be obtained via a suitable PDC process? A parametric process which produces photon correlations in three beams via emissions of triples is well defined quantum optically [22] and hence can be achievable in the laboratory. Several experimental attempts to obtain three-photon down-conversion were performed successfully [23–25]. Yet nonlinear crystals are not the only possible tool to obtain such states. In [26] the authors report an observation of three-photon parametric down-conversion in a superconducting parametric cavity.

As the process is becoming experimentally feasible, it is high time to give it an effective theoretical description and see what types of nonclassicality we can expect. It was shown in [27] that a straightforward generalization of the usual approximate description of PDC processes to three-photon emissions is impossible. The source of the problem is the parametric approximation in which emitted photons are treated in a quantum optical way, while the pump field is approximated by the classical wave. This approximation works perfectly fine for two-photon down-conversion, but its generalization to three- (or more) photon processes is impossible. One must describe the pump field as a quantum one [28]; for a specific application of this method see [29]. One may be tempted to think that the description of the pump in the form of a coherent state is the one avoiding approximations and that it takes into account its depletion due to the emission of the photon triples. But that is not so. The description is an approximation because we take into account only emissions into three phase-matched directions which we collect. In reality, on the one hand, we have emissions into other phase-matched directions, and the depletion is much higher. Also other loss processes contribute. On the other hand, to observe a reasonable count of the triples the pump must be very strong, and its depletion due to emissions of the triples would be negligible with respect to its intensity.

We present an approximate method of how to avoid mathematical difficulties emerging for higher-order parametric Hamiltonians which still allows a classical approximation of the pump field. This is done using a form of Padé approximation: we turn the first terms of the nonconverging perturbation expansion into elements of a converging series.

The nature of the cubic nonlinearity of the crystal polarization is such that only a three-beam emission process is feasible, i.e., one pump photon splitting into three down-converted photons. Thus, we show specific results and figures only for the three-beam case. However, we discuss also currently infeasible higher-order processes (requiring even higher nonlinearities and thus most likely out of experimental reach). The aim of this exercise is to test our approximation method in even more demanding situations.

We construct entanglement indicators and Mermin-GHZ Bell inequalities which allow us to reveal the GHZ-like properties of the radiation and give our estimates of critical (maximal) values of the pump parameters, which allow us to see entanglement or violate the inequalities.

II. BRIGHT GHZ STATES

In the famous EPR paper [30], the authors described a thought experiment which, in their opinion, pointed out the incompleteness of quantum mechanics. In Bohm's version of the EPR experiment, a particle of spin 0 decays into two $\frac{1}{2}$ spin particles in the singlet state that are sent in opposite directions [31]. The particles are correlated in such a way that after performing a spin-component measurement on the first particle one can predict with certainty the result of a measurement of the same spin component of the second particle; that is, we are able to predict the result of a remote identical measurement without performing the actual measurement on the other particle. Thus, following EPR, such a result must be an "element of physical reality." As "elements of reality" are not present in the quantum-mechanical description, EPR concluded that quantum theory is not complete [30]. Such was the birth of local realism.

However, in 1964 Bell showed that it is impossible to construct a local realistic [local hidden variable (LHV)] model that would explain all possible correlations between two such spins and would agree with statistical predictions of quantum mechanics for measurements of arbitrary pairs of spin components. In 1989 Greenberger, Horne, and Zeilinger showed that for three or four spins one could directly show that the concept of "elements of reality" is at odds with quantum predictions [32].

With emerging bright parametric sources of three-beam entanglement, one should check to what extent the highly nonclassical properties of GHZ states are also shared with their "bright" versions.

A. Pitfalls of the parametric approximation (classical pump)

Here, we shall study the technicalities concerning the theoretical description of the multiphoton GHZ-like state (bright GHZ) of n beams of light which is a kind of n-beam generalization of the 2×2 mode squeezed vacuum. We assume that each beam has two orthogonal polarization modes, but equivalently, one can imagine that there are n pairs of beams, with each pair directed to a different observer who is equipped with a Mach-Zehnder interferometer, into which the local beams enter (each via a different entry port). Such an interferometer is capable of performing any U(2) transformation of the pair of modes, and thus, it is endowed with the power to show the same type of interference effects as a universal polarization beam splitter. Thus, we shall use the "polarization picture" throughout just for the simplicity of presentation. However, we do not suggest here that the polarization version of the experiment would be more feasible (as a matter of fact, it seems less feasible due to the complicated phase matching required for such a case, whereas the two-beams-per-observer situation is much clearer).

Let us introduce the following notation:

$$\hat{A}_n^\dagger = \prod_{X=1}^{n} \hat{a}_X^\dagger, \tag{1}$$

$$\hat{B}_n^\dagger = \prod_{X=1}^{n} \hat{b}_X^\dagger, \tag{2}$$

where $a_X^\dagger$ and $b_X^\dagger$ are creation operators for two orthogonal polarization modes of the Xth party's beam.

The parametric approximation in which the "pump" is described as a classical field [5,18,33] leads to the following Hamiltonian:

$$\hat{H}_n = \gamma(\hat{A}_n^\dagger + \hat{B}_n^\dagger) + \text{H.c.}, \tag{3}$$

where γ is an effective coupling with the classical pumping field. Parametric approximation is simple, very intuitive, and widely used in the description of a quantum system interacting with intense electromagnetic field [5,15,18]. For $n = 2$ the unitary transformation with Hamiltonian (3) acting on a vacuum state produces a 2×2 mode squeezed vacuum state, with perfect correlations for Stokes observables, which is an analog of the two-qubit Bell state $|\Phi^+\rangle$ [20]. Still, like any other approximation, this one also has a range of applicability that requires investigation in every considered case. The approximation causes no mathematical problems in the case of two-photon down-conversion. Still, it is known that for $n > 2$ an expression of the form $\exp[it\gamma H_n]$ is not a well-defined unitary transformation (the expansion series does not converge) [27,33]. Thus, a straightforward generalization for $n > 2$ is impossible. Still, one can show that for the pump treated as a (coherent) quantum field, the fully quantum Hamiltonian leads to a well-defined evolution. This approach is way more demanding than the parametric one. Nevertheless, this is not the only option. The Hamiltonian (3) can be used with a suitable approximation that allows the convergence of a perturbation series (for example, see, [22]).

B. Convergence via Padé method

Our approach is based on two steps. First, we expand $e^{iH_n t}$ acting on the vacuum state. As noted earlier, there are problems with convergence. To address this problem, we apply Padé approximants. By combining the two approximations, we get a convergent formula.

Consider the following Hamiltonian:

$$H_n^A = \gamma \hat{A}_n^\dagger + \text{H.c.} \tag{4}$$

For $n = 1$ the unitary transformation with Hamiltonian (4) produces a coherent state, and for $n = 2$ it produces a two-mode squeezed vacuum. For $n = 3$ the Hamiltonian (4) corresponds to the Hamiltonian presented in [26]. Thus, after time t we seem to have

$$|\Sigma^n\rangle = e^{iH_n^A t}|\Omega\rangle = \sum_{k=0}^{\infty} \frac{(i\Gamma)^k}{k!}(\hat{A}_n^\dagger + \hat{A}_n)^k|\Omega\rangle, \tag{5}$$

where $\Gamma = \gamma t$ is the amplification gain. However, formula (5) for more than two parties is meaningless! The series in (5) does not converge; that is, the sum of probabilities tends to infinity instead of 1. Thus, (5) is not a well-defined state. The vacuum state is not an analytical vector for unitarylike expression $e^{iH_n^A t}$ based on Hamiltonian (4) for $n > 2$ [27]. However, the expansion of (5) is only a formal description, which is an approximate form of the considered state. We shall introduce an additional, compensatory approximation that allows for its convergence.

Note that (5) can be set as follows:

$$|\Sigma^n\rangle = \sum_{k=0}^{\infty} C_k^n (\hat{A}_n^\dagger)^k|\Omega\rangle, \tag{6}$$

where C_k^n are coefficients. We show in the Appendix that C_k^n can be expanded as follows:

$$C_k^n = \sum_{l=0}^{\infty} \frac{(i\Gamma)^{n+2l}}{(k+2l)!} P_{k+2l}^{k,n}, \tag{7}$$

where $P_l^{k,n}$ obey the recurrence relation: $P_l^{k,n} = P_{l-1}^{k-1,n} + (k+1)^n P_{l-1}^{k+1,n}$. Still, the series (7), just like (5), does not converge. Its infinite sequence of the partial sums does not have a finite limit. To impose convergence, we use a numerical method of Padé approximants [34].

Characteristics of an n-mode squeezedlike state with Padé approximants and convergence of the photon number

Even if a power series does not converge, we can still derive an alternative convergent approximation. Note that the reason for nonconvergence is the fact that we treat the pumping field as classical. Thus, we do not have overall energy conservation in the case of emitted photons. Still, it is obvious that only the first dozen or so of the expansion terms matter because the process of emission has a very low probability. Hence, we shall seek an approximation that is suitable for such a case. There are many methods that allow one to extract information from a power series outside of its convergence radius. One such method, which is extensively used in numerical calculations, is Padé approximants.

TABLE I. Probability $p(k)$ of observing k single photons emitted in a coherent state ($n = 1$), photons pairs from two mode-squeezed vacuum ($n = 2$), and triples of photons from three-beam radiation ($n = 3$). The range of k is 1–10, and the calculation is performed for a constant value of amplification gain $\Gamma = 0.8$. For $n = 2$ obtained values are consistent with theoretical results. Note that the probability of vacuum increases with n. This is due to the fact that higher-order states are generated in processes with a higher degree of nonlinearity. As k increases, the probability starts to increase with n. The probability of observing k triples of photons for three-beam radiation is higher than the probability of getting k pairs of photons for $n = 2$ for the same Γ starting from $k = 4$.

		$p(k)$	
k	$n = 3$	$n = 2$	$n = 1$
0	0.60	0.55	0.53
1	0.16	0.24	0.34
2	0.074	0.11	0.11
3	0.040	0.048	0.023
4	0.024	0.021	0.0037
5	0.016	0.0093	0.00047
6	0.011	0.0041	5×10^{-5}
7	0.0087	0.0018	4.6×10^{-6}
8	0.0066	0.0008	3.7×10^{-7}
9	0.0052	0.00035	2.6×10^{-8}
10	0.0042	0.00016	1.7×10^{-09}

Padé approximants are based on the idea of reformulating power series $\sum c_n x^n$ into a limit of a sequence of the ratio of polynomials. Elements of this sequence have the following form:

$$Q_M^N(x) = \frac{\sum_{n=0}^{N} X_n x^n}{\sum_{m=0}^{M} Y_m x^m}, \tag{8}$$

where X_n and Y_m are such that the first $(N + M + 1)$ terms of the Taylor series expansion of $Q_M^N(x)$ match the first $(N + M + 1)$ terms of $\sum c_n x^n$.

We denote by $[N/M]$ the respective $Q_M^N(x)$. We use a diagonal series of approximants, i.e., $[N/N]$, and the highest degree of approximants is $[40/40]$ in order to avoid machine epsilon and other numerical errors.

Still, we must remember that convergence of coefficients in the Fock space is not sufficient itself. We must also ensure convergence of the average photon number of the superposition (6). That convergence will determine the range of applicability of Padé approximants for our expansion. To test our method we reconstructed the expansion coefficients for the $n = 1$ case, i.e., those for a coherent state and the coefficients for $n = 2$ that are for a two-mode squeezed vacuum (generated by PDC; see Table I).

Let us consider the problem of the convergence of the total photon number. The problem was pointed out in [28]. We shall define the range of amplification gain for which the expectation value of the photon number converges. If $p(k)$ converges faster than $\sum_{k=1}^{\infty} \frac{1}{k^2}$ from some k, the average photon number is always finite. For the realistic case of $n = 3$, the critical applicable amplification gain is around $\Gamma = 0.9$. Thus, here, our approximation breaks down, and therefore, the results for Γ approaching 0.9 most likely do not describe the real

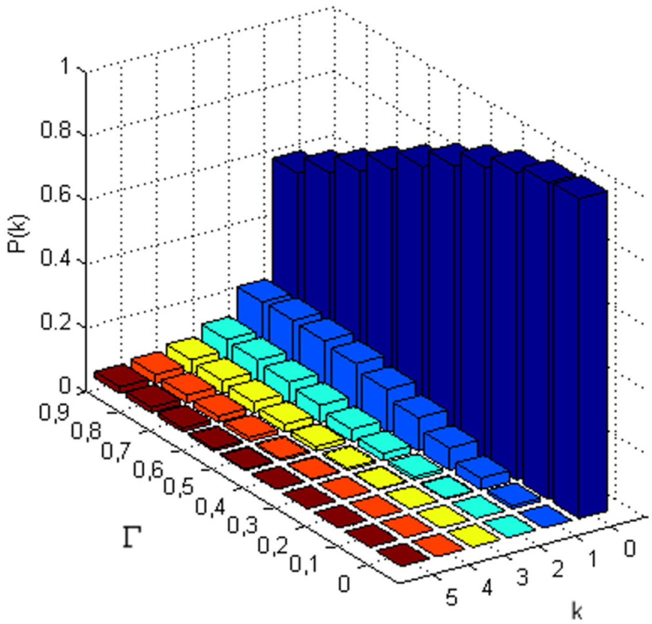

FIG. 1. Probability $p(k)$ of observing k triplets of photons for $k = 0, \dots, 5$ as a function of the amplification gain Γ. The probability of a vacuum event decreases when Γ increases. Also, all probabilities approach zero when the number of triples goes to infinity: $p(k \to \infty) \to 0$. This tendency is typical also for a two-mode squeezed vacuum.

situation. Figure 1 shows the probabilities $p(k)$ of observing k triples of photons as a function of the amplification gain Γ.

III. NONCLASSICAL PROPERTIES OF THREE-PARTY SIX-MODE BRIGHT GHZ

Applying the results from previous sections, we construct a GHZ-like state $|\mathrm{BGHZ}\rangle$ which can be generated with the use of Hamiltonian (3) for $n = 3$. Since operators $\hat{A}_3$ and $\hat{B}_3$ commute, we have

$$e^{i\Gamma(\hat{A}_3^\dagger+\hat{A}_3+\hat{B}_3^\dagger+\hat{B}_3)} = e^{i\Gamma(\hat{A}_3^\dagger+\hat{A}_3)}e^{i\Gamma(\hat{B}_3^\dagger+\hat{B}_3)}. \tag{9}$$

Thus, we can put the state into the following form:

$$|\mathrm{BGHZ}\rangle = \sum_{k=0}^{\infty}\sum_{m=0}^{k} C_{k-m}^3 C_m^3 (\hat{A}_3^\dagger)^{k-m}(\hat{B}_3^\dagger)^m|\Omega\rangle, \tag{10}$$

where C_Q^3 for $Q = k$ or $Q = k - m$ can be obtained with the Padé approximants described in previous sections. From (10) we can see that the state $|\mathrm{BGHZ}\rangle$ is symmetric under a change in indices $(k - m \leftrightarrow m)$; that is, the amplitudes of probability for states obtained with the action of operators $(\hat{A}_3^\dagger)^q(\hat{B}_3^\dagger)^p$ and $(\hat{A}_3^\dagger)^p(\hat{B}_3^\dagger)^q$ on the vacuum are equal.

We denote by $\hat{T}$ the correlation tensor, the elements of which are given by $T_{ijk} = \langle \hat{S}_i^1\hat{S}_j^2\hat{S}_k^3\rangle$, where $\hat{S}_q^X$ is the qth normalized Stokes operator for the Xth party introduced in [35] and rediscovered in [36], the form which we use here. For a general theory of such quantum Stokes operators, see [37]. These operators for the Xth party can be represented with photon number operators for the respective modes as follows:

$$\langle \hat{S}_q^X\rangle = \left\langle \hat{\Pi}^X \frac{(\hat{n}_j^X - \hat{n}_{j_\perp}^X)}{(\hat{n}_j^X + \hat{n}_{j_\perp}^X)}\hat{\Pi}^X\right\rangle. \tag{11}$$

In the formula $j, j_\perp$ denote a pair of orthogonal polarizations of one of three mutually unbiased polarization bases $j = 1, 2, 3$. Further down, we assign index 1 for polarizations $\{45^\circ, -45^\circ\}$, index 2 for $\{R, L\}$ (circular), and index 3 for the $\{H, V\}$ basis. The projector $\hat{\Pi}^X = 1 - |\Omega^X\rangle\langle\Omega^X|$, where $|\Omega^X\rangle$ is the vacuum state of the Xth party, makes the formula well defined, as it does not allow zero eigenvalues for the denominator. The zeroth operator is $\langle\hat{S}_0^X\rangle = \langle\hat{\Pi}^X\rangle$. The normalized quantum optical Stokes operators (11) allow one to straightforwardly introduce Bell inequalities for optical fields based on photon number observables [36,37].

One can show (see the Appendix) that nonvanishing elements of $\hat{T}$ are $T_{111} = t$ and $T_{122} = T_{212} = T_{221} = -t$, where t is given by

$$t = \sum_{k=1}^{\infty}\sum_{m=0}^{k}\Bigg(\left(C_{k-m-1}^3\right)^*\left(C_{m+1}^3\right)^*\frac{[(k-m)!(m+1)!]^3}{k^3} + \left(C_{m-1}^3\right)^*\left(C_{k-m+1}^3\right)^*\frac{[m!(k-m+1)!]^3}{k^3}\Bigg)C_m^3C_{k-m}^3. \tag{12}$$

Thus, we have the same set of nonvanishing elements of the correlation tensor with the same relative signs as for the three-qubit (spin-$\frac{1}{2}$) $|\mathrm{GHZ}\rangle$ state. This observation shows that, indeed, $|\mathrm{BGHZ}\rangle$ has the same type of correlations as a three-qubit state which is a mixture of "white noise" and a $|\mathrm{GHZ}\rangle$ state.

A. Mermin-GHZ-like Bell inequality violation by $|\mathrm{BGHZ}\rangle$

We are going to derive the Mermin-like Bell inequality [38] for three-beam optical fields and local measurements of (normalized) Stokes parameters. As shown in [36], thus far, we do not have a Bell inequality which involves standard Stokes parameters.

Observer X measures the intensity of a light beam using an analyzer of jth polarization. The outcomes given as intensities in the presence of local hidden variables λ can be written as $I_j^X(\lambda)$ and $I_{j_\perp}^X(\lambda)$. As we want to model the quantum Stokes parameters, their values are natural numbers [they must agree with the eigenvalue spectrum of the number operators used in (11)]. The distribution of λ is denoted $\rho(\lambda)$.

With the above model for intensities, we introduce local hidden variables $S_j^X(\lambda)$, which represent the predetermined values of stokes parameters (11): for $I_j^X(\lambda) + I_{j_\perp}^X(\lambda) \neq 0$,

$$S_j^X(\lambda) = \frac{I_j^X(\lambda) - I_{j_\perp}^X(\lambda)}{I_j^X(\lambda) + I_{j_\perp}^X(\lambda)}, \tag{13}$$

and if $I_j^X(\lambda) + I_{j_\perp}^X(\lambda) = 0$, then $S_j^X(\lambda) = 0$ see [36].

Consider the following expression:

$$\begin{aligned} &S_1^1(\lambda)S_1^2(\lambda)S_1^3(\lambda) - S_1^1(\lambda)S_2^2(\lambda)S_2^3(\lambda) \\ &-S_2^1(\lambda)S_1^2(\lambda)S_2^3 - S_2^1(\lambda)S_2^2(\lambda)S_1^3(\lambda). \end{aligned} \tag{14}$$

The values of $S_j^X(\lambda)$ are bounded by ± 1. As (14) is linear with respect to all $S_j^X(\lambda)$, we can find extremal values of (14) considering only border values, i.e., for which $|S_j^X(\lambda)| = 1$. With that, we get the bound of (14) equal to 2. The LHV averages for terms of (14) are given by

$$\langle S_i^1(\lambda)S_j^2(\lambda)S_k^3(\lambda)\rangle_{\mathrm{LHV}} = \int d\lambda\rho(\lambda)S_i^1(\lambda)S_j^2(\lambda)S_k^3(\lambda). \quad (15)$$

Thus, the following generalization of the Mermin inequality holds:

$$\big|\langle S_1^1(\lambda)S_1^2(\lambda)S_1^3(\lambda) - S_1^1(\lambda)S_2^2(\lambda)S_2^3(\lambda) - S_2^1(\lambda)S_1^2(\lambda)S_2^3(\lambda) - S_2^1(\lambda)S_2^2(\lambda)S_1^3(\lambda)\rangle_{\mathrm{LHV}}\big| \leqslant 2. \quad (16)$$

In the quantum case, if we straightforwardly calculate inequality (16) for $|\mathrm{BGHZ}\rangle$, we see that it is not violated due to the high probability of vacuum events in $|\mathrm{BGHZ}\rangle$. To bypass this problem, we shall modify inequality (16). This can be done by reformulating observables (13) in such a way that we can assign the value -1 for the case in which no light detection occurs (this concept was first introduced in [36]). The ideas of [39] were our inspiration. For the modified hidden values, if $I_j^X(\lambda) + I_{j_\perp}^X(\lambda) \neq 0$, we have the same approach as for (13), $S_j^X(\lambda) \to S_j^{X'}(\lambda) = S_j^X(\lambda)$, but when $I_j^X(\lambda) + I_{j_\perp}^X(\lambda) = 0$, we assign $S_j^{X'}(\lambda) = -1$. Still, the bound for the reformulated Bell inequality remains the same because we have $-1 \leqslant S_j^{X'}(\lambda) \leqslant 1$. Hence, the modified Mermin-like inequality has the same form as inequality (16):

$$|\langle S_1^{1'}(\lambda)S_1^{2'}(\lambda)S_1^{3'}(\lambda) - S_1^{1'}(\lambda)S_2^{2'}(\lambda)S_2^{3'}(\lambda) - S_2^{1'}(\lambda)S_1^{2'}(\lambda)S_2^{3'}(\lambda) - S_2^{1'}(\lambda)S_2^{2'}(\lambda)S_1^{3'}(\lambda)\rangle_{\mathrm{LHV}}| \leqslant 2. \quad (17)$$

In the quantum case, we reformulate the normalized Stokes operators in the following way:

$$\hat{S}_j^X \to \hat{S}_j^{X'} = \hat{S}_j^X - |\Omega^X\rangle\langle\Omega^X|. \quad (18)$$

For $|\mathrm{BGHZ}\rangle$ we get,

$$\langle \hat{S}_i^{1'}\hat{S}_j^{2'}\hat{S}_k^{3'}\rangle_{\mathrm{BGHZ}} = \langle \hat{S}_i^1\hat{S}_j^2\hat{S}_k^3\rangle_{\mathrm{BGHZ}} - |\langle\Omega|\mathrm{BGHZ}\rangle|^2, \quad (19)$$

because the expectation values of combinations of two Stokes operators and one projector into vacuum vanish, e.g., $\langle \hat{S}_i^1\hat{S}_j^2|\Omega^3\rangle\langle\Omega^3|\rangle_{\mathrm{BGHZ}} = 0$ and $\langle \hat{S}_i^1|\Omega^2\rangle|\Omega^3\rangle\langle\Omega^2|\langle\Omega^3|\rangle_{\mathrm{BGHZ}} = 0$.

Figure 2 shows the left-hand side of inequality (19) as a function of the amplification gain. Note that the range of Γ for which inequality (17) is violated covers almost the whole range for Γ for which the Padé approximation works.

B. Violation of the Mernin-like inequality by |BGHZ⟩ for the case of imperfect detection efficiency

We study here the resistance of the above results with respect to photon losses. We assume a model of the experimental setup in which all photon losses are modeled as inefficient detectors. The standard quantum optical model for that is as follows. The lossy detector is defined as a perfect detector (with efficiency $\eta = 1$) with a beam splitter of transitivity $\sqrt{\eta}$ in front of it. Assume that k_j^X photons reach the

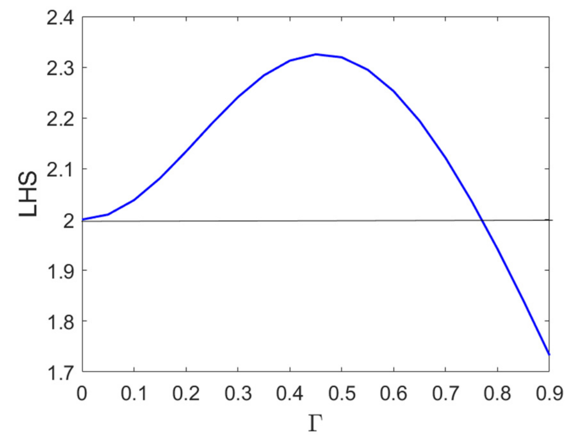

FIG. 2. Left-hand side of inequality (19) as a function of amplification gain Γ (Γ_t). The threshold value of Γ, such that for all $\Gamma < \Gamma_t$ inequality (19) is violated, is $\Gamma_t = 0.77$.

detector of observer X which collects photons of polarization j. However, due to the losses, only κ_j^X counts are registered ($\kappa_j^X \leqslant k_j^X$). The probability of registering κ_j^X photons for imperfect efficiency η is then given by the binomial distribution:

$$p(\kappa_j^X|k_j^X) = \binom{k_j^X}{\kappa_j^X}\eta^{\kappa_j^X}(1-\eta)^{k_j^X-\kappa_j^X}. \quad (20)$$

For simplicity let us consider just the expectation value $\langle \hat{S}_3'^1\hat{S}_3'^2\hat{S}_3'^3\rangle$, where the lower index 3 stands for the measurement in basis $\{H, V\}$ [see (11)]; then, e.g., for $|\phi\rangle = |k_H^1, k_V^1, k_H^2, k_V^2, k_H^3, k_V^3\rangle$ in the presence of losses we get

$$\langle \hat{S}_3'^1\hat{S}_3'^2\hat{S}_3'^3\rangle_\phi = \lim_{\epsilon\to 0}\prod_{X=1}^{3}\sum_{\kappa_H^X=0}^{k_H^X}\sum_{\kappa_V^X=0}^{k_V^X} p(\kappa_H^X|k_H^X) \times p(\kappa_V^X|k_V^X)\left(\frac{\kappa_H^X - \kappa_V^X}{\kappa_H^X + \kappa_V^X + \epsilon} - \delta_{0,\kappa_H^X+\kappa_V^X}\right), \quad (21)$$

where δ_{pq} describes the Kronecker delta.

In order to calculate other elements of inequality (17) it is enough to apply a unitary transformation that links Stokes operators (see Appendix A5).

Obviously, the value of the threshold efficiency η_t such that for the $\eta < \eta_t$ inequality (17) is not violated varies depending on the amplification gain Γ. We calculated η_t for the range of Γ for which (17) is violated (see Fig. 3).

IV. DETECTION OF ENTANGLEMENT OF |BGHZ⟩

We discuss two entanglement indicators for the $|\mathrm{BGHZ}\rangle$ state. The first one is based on an entanglement indicator for $|\mathrm{GHZ}\rangle$ for qubits presented in Ref. [40]. The second one is derived from the above Mermin-like Bell inequality.

The entanglement indicator for three qubits given in Ref. [40] reads

$$\hat{w} = \tfrac{3}{2}\mathbb{1} - \hat{\sigma}_1^1\hat{\sigma}_1^2\hat{\sigma}_1^3 - \tfrac{1}{2}\left(\hat{\sigma}_3^1\hat{\sigma}_3^2 + \hat{\sigma}_3^2\hat{\sigma}_3^3 + \hat{\sigma}_3^1\hat{\sigma}_3^3\right), \quad (22)$$

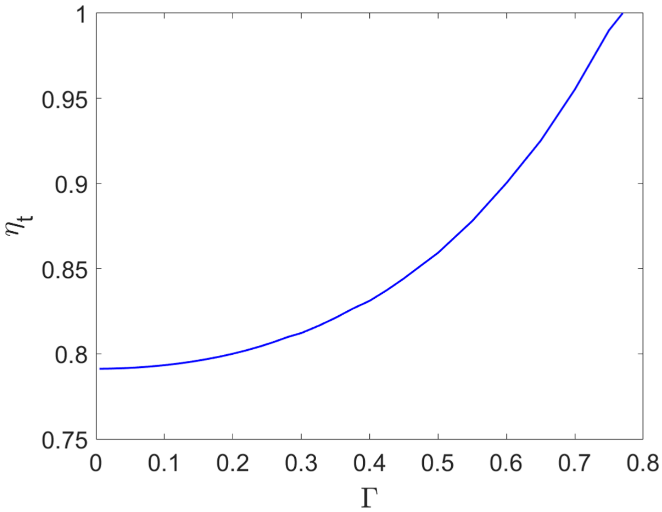

FIG. 3. Threshold efficiency η_t as a function of amplification gain Γ. As expected, the value of η_t increases with Γ. Note that for small values of amplification gain ($\Gamma \to 0$) threshold efficiency $\eta_t = 0.79$, and it conforms with η_t for qubits given in [3]. That is because for $\Gamma \to 0$ our $|\text{GBHZ}\rangle$ effectively becomes a superposition of $|\text{GHZ}\rangle$ and vacuum.

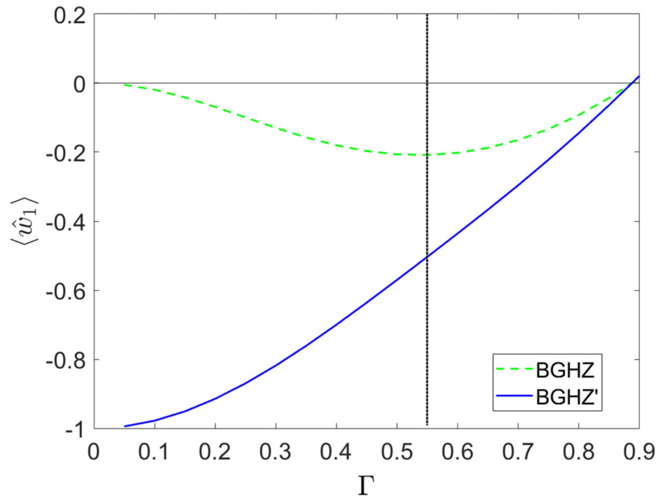

FIG. 4. Expectation value of entanglement indicator (23) for $|\text{BGHZ}\rangle$ as a function of amplification gain Γ. The lower blue line is for state ρ', from which we removed the vacuum contribution and renormalized it, (24). This procedure allows one to see better violations for low Γ, as in the original state in such a case the vacuum term is dominant. The value -1 points to the theoretical value for the three-qubit entangled GHZ state and the original entanglement witness (22). This is because, for a very small Γ, the state $|\text{BGHZ}\rangle$ is effectively just a superposition of vacuum and a single three-photon GHZ emission. Removal of vacuum and renormalization leaves just the GHZ state.

where σ_k^X denotes the kth Pauli matrix related to measurement performed by the Xth party. Using the isomorphism between Pauli matrices and normalized Stokes operators given, e.g., in Ref. [37], we straightforwardly obtain an entanglement indicator for $|\text{BGHZ}\rangle$:

$$\hat{w}_1 = \tfrac{3}{2}\hat{S}_0^1\hat{S}_0^2\hat{S}_0^3 - \hat{S}_1^1\hat{S}_1^2\hat{S}_1^3 - \tfrac{1}{2}\big(\hat{S}_3^1\hat{S}_3^2\hat{S}_0^3 + \hat{S}_0^1\hat{S}_3^2\hat{S}_3^3 + \hat{S}_3^1\hat{S}_0^2\hat{S}_3^3\big). \tag{23}$$

The isomorphism is simply a replacement of Pauli operators by normalized Stokes operators, namely, $\hat{\sigma}_\nu^X \to \hat{S}_\nu^X$, where $\nu = 0, 1, 2, 3$.

Using (23), we can detect entanglement of $|\text{BGHZ}\rangle$. Still, the indicator (23) performs quite weakly; for example, for small Γ it does not detect entanglement. This is due to a large number of vacuum events in $|\text{BGHZ}\rangle$. To improve detection of entanglement, we use the approach presented in [41]. Let us denote the density matrix for $|\text{BGHZ}\rangle$ as $\hat{\rho}$. We remove the vacuum contribution from $\hat{\rho}$ in the following way:

$$\hat{\rho} \to \hat{\rho}' = \frac{1}{\text{Tr}(\Pi\hat{\rho}\hat{\Pi})}\hat{\Pi}\hat{\rho}\hat{\Pi}, \tag{24}$$

where $\hat{\Pi} = \hat{\Pi}^1\hat{\Pi}^2\hat{\Pi}^3$. Note that as this is a product of local operations, it does not create new entanglement, and thus, the procedure is admissible. Figure 4 shows the violation of condition (23) for $\hat{\rho}$ and $\hat{\rho}'$ as a function of amplification gain Γ.

Note that the procedure (24) is definitely commonplace in the laboratory, as the replacement of ρ by ρ' amounts to cases in which, despite the source being switched on (here, it is good to imagine that the pump field is pulsed), for the given run at each of the measurement stations we receive no counts. Such lack-of-counts events are usually by default ignored.

Let us now move to the other entanglement indicator, which is inspired by the Mermin-like inequality. It is well known that Bell inequalities are entanglement indicators. Still, one can improve their functioning in that role by taking the Bell operators linked with them and calculating, for specific settings, the maximal value for a separable state. This may lead to a lower bound than for local hidden variables (as a separable state can be viewed as a specific local-hidden-variable model). This allows one to create an entanglement indicator (witness) which is more efficient than the initial Bell inequality.

Consider the following operator:

$$\langle\hat{M}\rangle = \big\langle\hat{S}_1^1\hat{S}_2^2\hat{S}_2^3 + \hat{S}_2^1\hat{S}_1^2\hat{S}_2^3 + \hat{S}_2^1\hat{S}_2^2\hat{S}_1^3 - \hat{S}_1^1\hat{S}_1^2\hat{S}_1^3\big\rangle. \tag{25}$$

Let us search for its highest value for fully separable states, i.e.,

$$\rho^{1,2,3} = \sum_\lambda p_\lambda|\psi(\lambda)\rangle\langle\psi(\lambda)|, \tag{26}$$

where

$$|\psi(\lambda)\rangle = f_\lambda^1(\hat{a}^\dagger)f_\lambda^2(\hat{b}^\dagger)f_\lambda^3(\hat{c}^\dagger)|\Omega\rangle \tag{27}$$

and $f_\lambda^1(\hat{a}^\dagger)$, $f_\lambda^2(\hat{b}^\dagger)$, and $f_\lambda^3(\hat{c}^\dagger)$ are functions of powers of creation operators for polarization modes corresponding to optical beams 1, 2, and 3 of the property that $f_\lambda^1(\hat{a}^\dagger)|\Omega\rangle$ is a proper state in the Fock space.

For every λ we have

$$\langle\hat{S}_1^1\hat{S}_1^2\hat{S}_1^3\rangle_{\lambda_{\text{sep}}} = \langle\hat{S}_1^1\rangle_{\lambda_{\text{sep}}}\langle\hat{S}_1^2\rangle_{\lambda_{\text{sep}}}\langle\hat{S}_1^3\rangle_{\lambda_{\text{sep}}}, \tag{28}$$

where $\langle\hat{O}\rangle_{\lambda_{\text{sep}}} = \langle\psi(\lambda)|\hat{O}|\psi(\lambda)\rangle$ for an operator $\hat{O}$. In order to find the bound, it is enough to use the following property of the normalized Stokes operators (see Ref. [41]). If

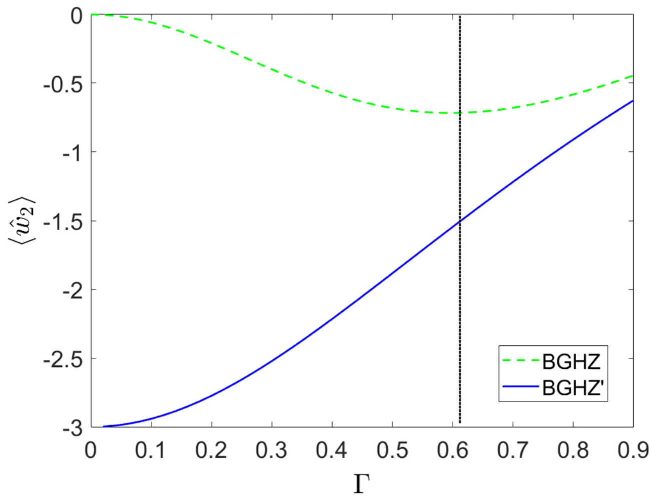

FIG. 5. Comparison of the violation of (32) for $\hat{\rho}$ and $\hat{\rho}'$ as a function of amplification gain Γ. See also the caption of Fig. 4. Note that in the case of indicator (32) we seem to have slightly more robust violations of the separability threshold than for (23).

one constructs a three-dimensional "Stokes" vector $\langle \vec{\hat{S}}^X \rangle = (\langle S_1^X \rangle, \langle S_2^X \rangle, \langle S_3^X \rangle)$, one has $||\langle \vec{\hat{S}}^X \rangle|| \leqslant 1$, that is, $\langle \hat{S}_1 \rangle^2 + \langle \hat{S}_2 \rangle^2 + \langle \hat{S}_3 \rangle^2 \leqslant 1$. Thus, $\langle \vec{\hat{S}}^X \rangle$ is a vector in the Bloch ball.

Note that (25) involves only two different local measurements, so it is enough to consider only $\langle \hat{S}_1 \rangle^2 + \langle \hat{S}_2 \rangle^2 \leqslant 1$. Hence, the relevant Stokes vector of maximal length can have the following representation:

$$\langle \overrightarrow{\hat{S}^X} \rangle = (\cos \alpha_X, \sin \alpha_X), \tag{29}$$

where α is a certain angle. As the value of operator $\hat{M}$ for pure separable states is proportional to the product of the lengths of the Stokes vectors for each of the beams, using the above representation, we search for the maximum of (25) by bounding from above the following expression:

$$\begin{aligned} &\cos \alpha_1 \sin \alpha_2 \sin \alpha_3 + \sin \alpha_1 \cos \alpha_2 \sin \alpha_3 \\ &\quad + \sin \alpha_1 \sin \alpha_2 \cos \alpha_3 - \cos \alpha_1 \cos \alpha_2 \cos \alpha_3 \\ &\quad\quad = \cos(\alpha_1 + \alpha_2 + \alpha_3) \leqslant 1. \end{aligned} \tag{30}$$

Thus, we get

$$-1 \leqslant \langle \hat{M} \rangle_{\text{sep}} \leqslant 1. \tag{31}$$

This bound is two times smaller than the one for local-hidden-variable models.

Still, inequality (31) may be an inefficient entanglement indicator in the case when the state contains a significant vacuum component or admixture. The trick of considering only the nonvacuum part of the state, given by (24), leads to a new entanglement indicator (witness),

$$0 \leqslant \langle \hat{M} + \hat{\Pi}^1 \hat{\Pi}^2 \hat{\Pi}^3 \rangle_{\text{sep}} = \langle \hat{w}_2 \rangle_{\text{sep}}. \tag{32}$$

For $|\text{BGHZ}\rangle$ we get

$$\langle \hat{w}_2 \rangle_{\text{BGHZ}} = -4t + 1 - |C_0^3|^4, \tag{33}$$

where $|C_0^3|^4$ is the probability of vacuum events for all beams. Figure 5 shows the violation of separability conditions (25) and (32) as a function of amplification gain Γ.

When comparing these two entanglement conditions (23) and (32), it seems that (32), derived from the Mermin-like inequality, is more efficient than (23). Comparing Figs. 4 and 5, we see that (32) is violated for a broader range of Γ, and thus, it is more robust. For example, 50% of the negative value for $\Gamma \to 0$ is reached in the case of $\hat{w}_1$ for $\Gamma \approx 0.55$, whereas for $\hat{w}_2$, $\Gamma \approx 0.61$. When analyzing Fig. 5, one must have in mind that our approximation breaks down around $\Gamma \approx 0.9$.

V. FINAL REMARKS

We have shown that a version of the Padé approximation is a candidate for an effective description of the bright GHZ states, which cures, to some extent, the pitfalls of the usual parametric approximation, of the kind that works well for two-beam PDC. As such states are interesting and are on the verge of experimental feasibility, the results presented here may be a useful tool for further investigations and for estimating the influence of the higher-order emissions on, for example, some quantum information protocols. An open question is finding better entanglement indicators, which would be more efficient for higher Γ.

ACKNOWLEDGMENTS

This work is a part of the "International Centre for Theory of Quantum Technologies" project (Contract No. 2018/MAB/5), which is carried out within the International Research Agendas Programme (IRAP) of the Foundation for Polish Science (FNP) cofinanced by the European Union from the funds of the Smart Growth Operational Programme, axis IV: Increasing the research potential (Measure 4.3).

APPENDIX

1. Equivalence of representations of an *n*-beam multiphoton state

Here, we show that formulas (5) and (6) are equivalent, that is, that

$$|\Sigma_n\rangle = \sum_{k=0}^{\infty} \frac{(i\Gamma)^k}{k!} (\hat{A}_n^\dagger + \hat{A}_n)^k |\Omega\rangle = \sum_{k=0}^{\infty} C_k^{(n)} (\hat{A}_n^\dagger)^k |\Omega\rangle.$$

First, note that $\hat{A}_n|\Omega\rangle = 0$, so every term with an operator A_n on the very left is equal to zero. We now show that

$$\hat{A}_n^l (\hat{A}_n^\dagger)^p |\Omega\rangle = \left[\prod_{j=1}^{l} (p - j + 1)^n \right] (\hat{A}_n^\dagger)^{p-l} |\Omega\rangle. \tag{A1}$$

Let us analyze the action of operator $\hat{A}\hat{A}^\dagger$ on the n-beam k-photon state $|k_1 \cdots k_n\rangle$:

$$\hat{A}\hat{A}_n^\dagger |k_1 \cdots k_n\rangle = \left[\prod_{i=1}^{k} (k_i + 1) \right] |k_1 \cdots k_n\rangle \tag{A2}$$

because $\hat{a}_X\hat{a}_X^\dagger|k_1\cdots k_n\rangle = (\hat{a}_X^\dagger\hat{a}_X+1)|k_1\cdots k_n\rangle = (\hat{n}_X+1)|k_1\cdots k_n\rangle = (k_X+1)|k_1\cdots k_n\rangle$. Hence,

$$\begin{aligned}\hat{A}_n^l(\hat{A}_n^\dagger)^p|k_1\cdots k_n\rangle &= \prod_{i=1}^n \frac{\sqrt{k_i+p}!}{\sqrt{k_i!}}\hat{A}_n^{l-1}\hat{A}_n(\hat{A}_n^\dagger)\\ &\quad\times|(k_1+p-1)\cdots(k_n+p-1)\rangle\\ &=\left[\prod_{i=1}^n(k_i+p)\right]\hat{A}_n^{l-1}(\hat{A}_n^\dagger)^{p-1}|k_1\cdots k_n\rangle.\end{aligned}\tag{A3}$$

Iterating such operation l times, we obtain

$$\begin{aligned}&\hat{A}_n^l(\hat{A}_n^\dagger)^p|k_1\cdots k_n\rangle\\ &=\prod_{j=1}^l\left[\prod_{i=1}^n(k_i+p-j+1)\right](\hat{A}_n^\dagger)^{p-l}|k_1\cdots k_n\rangle.\end{aligned}\tag{A4}$$

In the case of $|k_1\cdots k_n\rangle = |\Omega\rangle$ relation (A4) takes the form of (A1). Thus, for $p = l$ the action of $\hat{A}_n^l(\hat{A}_n^\dagger)^l$ boils down to only a real coefficient. Therefore, all terms in (5) can be set as $(\hat{A}_n^\dagger)^l|\Omega\rangle$ for some integer l multiplied by some coefficient. From that, we conclude that the formula (A1) is correct.

2. Derivation of coefficients C_k^n

First, let us notice that each coefficient C_k^n is composed of an infinite sum of coefficients $P_l^{k,n}$ that stay by $(\hat{A}_n^\dagger)^k$ operators linked with the action of the lth power of the Hamiltonian (4) on $|\Omega\rangle$. Using the structure of (4) and state (6), we can propose the following form of C_k^n:

$$C_k^n = \sum_{l=0}^\infty \frac{(i\Gamma)^l}{l!}P_l^{k,n}.\tag{A5}$$

Now let us define boundary conditions for $P_l^{k,n}$. Note that $P_l^{k,n} = 0$ if $k<0$ or $l<0$. Also, analyzing the first elements (6), we conclude that $P_0^{k,n} = \delta_{0k}$. In the next step, we realize that there are only two ways to obtain a term proportional to $(\hat{A}_n^\dagger)^k|\Omega\rangle$ with Hamiltonian (4) acting on a state of the form $\sum_i p^i(A_n^\dagger)^i|\Omega\rangle$ (p^i is some coefficient):

$$\hat{A}_n(\hat{A}_n^\dagger)^{k+1}|\Omega\rangle = (k+1)^n(\hat{A}_n^\dagger)^k|\Omega\rangle,\tag{A6}$$

$$\hat{A}_n^\dagger(\hat{A}_n^\dagger)^{k-1}|\Omega\rangle = (\hat{A}_n^\dagger)^k|\Omega\rangle,\tag{A7}$$

where we used identity (A1).

Finally, combining (A6) and (A7), after noticing that $H_n^l|\Omega\rangle = H_n(H_n^{l-1}|\Omega\rangle)$, we get the following recurrence pattern for $P_l^{k,n}$:

$$P_l^{k,n} = P_{l-1}^{k-1,n} + (k+1)^n P_{l-1}^{k+1,n}.\tag{A8}$$

Let us now make some remarks on when we have $P_l^{k,n} = 0$:

(i) Since $(\hat{A}_n^\dagger + \hat{A}_n)^k$ acts on $|\Omega\rangle$ [see (5)], we have $P_l^{k,n} = 0$ if $k > l$. Also, $P_k^{k,n} = 1$.

(ii) We have $P_l^{k,n} = 0$ if $(l \mod 2) \neq (k \mod 2)$. The proof of this statement uses mathematical induction. One can check easily that $P_1^{0,n} = 0$. Let us assume that $P_{k+1}^{k,n} = 0$. Using (A8), we construct $P_{k+2}^{k+1,n}$

$$P_{k+2}^{k+1,n} = P_{k+1}^{k,n} + (k+2)^n P_{k+1}^{k+2,n} = 0,\tag{A9}$$

TABLE II. First few coefficients of $P_l^{k,n}$.

	k			
l	0	1	2	3
0	1	0	0	0
1	0	1	0	0
2	1	0	1	0
3	0	$1+2^d$	0	1
4	$1+2^d$	0	$1+2^d+3^d$	0

which proves that the assumption is correct for every k. Now let us make another induction assumption that $\forall_{k\in N}\, P_{k+1+2l}^{k,n} = 0$ for some $l \in N$. Now, we have to check if this also holds for $l+1$. We calculate $P_{1+2(l+1)}^{0,n} = P_{2+2l}^{1,n} = 0$ due to the assumption. We go through an analogous reasoning to that for the case of $P_{k+2}^{k+1,n}$ and as a result we conclude that the given thesis is correct and each second element of an infinite sum (A5) vanishes.

Consequently, sum (A5) turns into its simplified form (7).

3. The explicit form of $P_l^{k,n}$ coefficients

Here, we give the explicit form of nonvanishing coefficients $P_l^{k,n}$ for $l > k$:

$$P_l^{k,n} = \sum_{i=1}^{k+1} i^n \sum_{j=1}^{i+1} j^n \cdots \sum_{y=1}^{x+1} y^n,\tag{A10}$$

where we have $\frac{l-k}{2}$ sums. The proof is based on mathematical induction and goes as follows. First, we calculate the first few $P_l^{k,n}$ coefficients that are given in Table II. We observe that they obey relation (A10).

Let us assume that relation (A10) remains valid for $P_{l-1}^{k,n}$. We apply the recursive formula (A8) for the lth step. We get

$$P_l^{k,n} = P_{l-1}^{k-1,n} + (k+1)^n P_{l-1}^{k+1,n}.\tag{A11}$$

Note that the element $P_{l-1}^{k-1,n}$ is composed of $(l-k)/2$ sums, and $P_{l-1}^{k+1,n}$ has $(l-k-2)/2$ sums:

$$\begin{aligned}P_l^{k,n} &= \sum_{i=1}^{(k+1)-1} i^n \sum_{j=1}^{i+1} j^n \cdots \sum_{y=1}^{x+1} y^n\\ &\quad+(k+1)^n \sum_{j=1}^{(k+1)+1} j^n \cdots \sum_{y=1}^{x+1} y^n.\end{aligned}\tag{A12}$$

The second term extends the first sum $\sum_{i=1}^{(k+1)-1}$ by the element corresponding to $i = k+1$. Thus, by merging two terms of (A12), we get (A10).

4. Derivation of elements of the correlation tensor for $|\mathrm{BGHZ}_3\rangle$

Let us consider state $|\mathrm{BGHZ}_3\rangle$ of the following form:

$$|\mathrm{BGHZ}_3\rangle = \sum_{k=0}^\infty\sum_{m=0}^k C_{k-m}^3 C_m^3 (\hat{A}_3^\dagger)^{k-m}(\hat{B}_3^\dagger)^m|\Omega\rangle.\tag{A13}$$

Standard Stokes operators for the Xth beam of light can be written as follows [17]:

$$\begin{aligned}\hat{\Theta}_1^X &= \hat{b}_X^\dagger \hat{a}_X + \hat{a}_X^\dagger \hat{b}_X,\\ \hat{\Theta}_2^X &= i(\hat{b}_X^\dagger \hat{a}_X - \hat{a}_X^\dagger \hat{b}_X),\\ \hat{\Theta}_3^X &= \hat{a}_X^\dagger \hat{a}_X - \hat{b}_X^\dagger \hat{b}_X.\end{aligned} \tag{A14}$$

Using formula (A14), we recall the structure of normalized Stokes operators:

$$\hat{S}_j^X = \hat{\Pi}^X \frac{\hat{\Theta}_j^X}{\hat{a}_X^\dagger \hat{a}_X + \hat{b}_X^\dagger \hat{b}_X} \hat{\Pi}^X, \tag{A15}$$

where $\hat{\Pi}^X = \hat{I}^X - |\Omega^X\rangle\langle\Omega^X|$. We are going to calculate elements of the correlation tensor for $|\mathrm{BGHZ}_3\rangle$. Let us start with $\langle S_3^1 S_3^2 S_3^3\rangle$. For simplicity, we use standard Stokes operators $\hat{\Theta}_j^X$, but our conclusions remain valid for normalized ones. Let us calculate how standard Stokes operators act on the components of $|\mathrm{BGHZ}_3\rangle$ of the following type: $|\phi_{k-m,m}\rangle = |k-m, m\rangle_1 |k-m, m\rangle_2 |k-m, m\rangle_3$ and $|\phi_{m,k-m}\rangle$. We get

$$\hat{\Theta}_3^1 \hat{\Theta}_3^2 \hat{\Theta}_3^3 |\phi_{k,k-m}\rangle = (k-2m)^3 |\phi_{k-m,m}\rangle \tag{A16}$$

and

$$\hat{\Theta}_3^1 \hat{\Theta}_3^2 \hat{\Theta}_3^3 |\phi_{k-m,k}\rangle = (2m-k)^3 |\phi_{m,k-m}\rangle. \tag{A17}$$

Due to the symmetry of probability amplitudes in $|\mathrm{BGHZ}_3\rangle$ we have the same amplitudes of probability for $|\phi_{m,k-m}\rangle$ and $|\phi_{k-m,n}\rangle$. Hence, after applying $\langle \mathrm{BGHZ}_3|$ to (A16) and (A17), these two terms cancel out. Obviously, also $S_3^1 S_3^2 S_3^3 |\phi_{k,k}\rangle = 0$. It follows that $\langle \hat{S}_3^1 \hat{S}_3^2 \hat{S}_3^3\rangle = 0$.

Now, let us make two observations. First, note that $\hat{a}_X^\dagger \hat{b}_X$ and $\hat{b}_X^\dagger a_X$ flip one photon between modes without changing the total number of photons at the Xth party, whereas operators $\hat{a}_X^\dagger \hat{a}_X$ and $\hat{b}_X^\dagger \hat{b}_X$ do not change the number of photons between modes. Also, let us observe that we have to take into consideration the terms of $\langle \hat{S}_i^1 \hat{S}_j^2 \hat{S}_k^3\rangle$ which are products of operators $\hat{A}_3, \hat{A}_3^\dagger, \hat{B}_3$, and $\hat{B}_3^\dagger$ because $|\mathrm{BGHZ}_3\rangle$ is a superposition of states of the form $(\hat{A}_3^\dagger)^k (\hat{B}_3^\dagger)^l |\Omega\rangle$ and so, due to the first observation, other terms acting on $|\mathrm{BGHZ}_3\rangle$ will flip photons only in some parties, leaving the rest of them with unchanged numbers of photons in modes, making the state orthogonal to $|\mathrm{BGHZ}_3\rangle$.

Therefore, we can easily conclude that all elements of $\hat{T}$ that contain one or two $\hat{S}_3$ operators are equal to zero.

Next, we consider element $\langle \hat{S}_1^1 \hat{S}_1^2 \hat{S}_1^3\rangle$. Note that operator $\hat{\Theta}_1^1 \hat{\Theta}_1^2 \hat{\Theta}_1^3$ due to the second observation effectively takes the following form: $\hat{B}_3^\dagger \hat{A}_3 + \hat{A}_3^\dagger \hat{B}_3$. Hence, we have

$$\begin{aligned}&(\hat{B}_3^\dagger \hat{A}_3 + \hat{A}_3^\dagger \hat{B}_3)|\phi_{k-m,m}\rangle\\ &\quad = [(k-m)(m+1)]^{3/2} |\phi_{k-m-1,m+1}\rangle\\ &\quad + [(m)(k-m+1)]^{3/2} |\phi_{k-m+1,m-1}\rangle\end{aligned} \tag{A18}$$

and

$$\begin{aligned}&(\hat{B}_3^\dagger \hat{A}_3 + \hat{A}_3^\dagger \hat{B}_3)|\phi_{m,k-m}\rangle\\ &\quad = [(m)(k-m+1)]^{3/2} |\phi_{m-1,k-m+1}\rangle\\ &\quad + [(k-m)(m+1)]^{3/2} |\phi_{m+1,k-m-1}\rangle.\end{aligned} \tag{A19}$$

After taking into account probability amplitudes and taking into account the full structure of normalized Stokes operators (with the total photon number in the denominator), we get

$$\begin{aligned}&m^{3/2}(k-m)^{3/2} C_m^3 C_{k-m}^3 \langle \mathrm{BGHZ}_3 | \hat{S}_1^1 \hat{S}_1^2 \hat{S}_1^3 | \phi_{k,k-m}\rangle\\ &\quad = \Bigg((C_{k-m-1}^3)^* (C_{m+1}^3)^* \frac{[(k-m)!(m+1)!]^3}{k^3}\\ &\quad + (C_{m-1}^3)^* (C_{k-m+1}^3)^* \frac{[m!(k-m+1)!]^3}{k^3} \Bigg) C_m^3 C_{k-m}^3\end{aligned} \tag{A20}$$

and

$$\begin{aligned}&m^{3/2}(k-m)^{3/2} C_m^3 C_{k-m}^3 \langle \mathrm{BGHZ}_3 | \hat{S}_1^1 \hat{S}_1^2 \hat{S}_1^3 | \phi_{k-m,k}\rangle\\ &\quad = \Bigg((C_{k-m-1}^3)^* (C_{m+1}^3)^* \frac{[(k-m)!(m+1)!]^3}{k^3}\\ &\quad + (C_{m-1}^3)^* (C_{k-m+1}^3)^* \frac{[m!(k-m+1)!]^3}{k^3} \Bigg) C_m^3 C_{k-m}^3.\end{aligned} \tag{A21}$$

Note that formulas (A20) and (A21) are equal. Finally, we obtain

$$\begin{aligned}&\langle \hat{S}_1^1 \hat{S}_1^2 \hat{S}_1^3\rangle\\ &\quad = \sum_{k=1}^{\infty} \sum_{m=0}^{k} \Bigg((C_{k-m-1}^3)^* (C_{m+1}^3)^* \frac{[(k-m)!(m+1)!]^3}{k^3}\\ &\quad + (C_{m-1}^3)^* (C_{k-m+1}^3)^* \frac{[m!(k-m+1)!]^3}{k^3} \Bigg) C_m^3 C_{k-m}^3.\end{aligned} \tag{A22}$$

Let us now consider elements which contain one operator $\hat{S}_1$ and two operators $\hat{S}_2$. Analogous to the reasoning above, the product of such operators turns out to be $i^2(\hat{B}_3^\dagger \hat{A}_3 + (-1)^2 \hat{A}_3^\dagger \hat{B}_3) = -(\hat{B}_3^\dagger \hat{A}_3 + \hat{A}_3^\dagger \hat{B}_3)$. Thus, these elements have the same absolute value as (A22).

Next, we investigate element $\langle \hat{S}_2^1 \hat{S}_2^2 \hat{S}_2^3\rangle$. Again, the operators there can be transformed into $i^3[\hat{B}_3^\dagger \hat{A}_3 + (-1)^3 \hat{A}_3^\dagger \hat{B}_3] = -i(\hat{B}_3^\dagger \hat{A}_3 - \hat{A}_3^\dagger \hat{B}_3)$. It is also easy to notice that the expectation value of this operator is zero (the respective terms cancel out).

Also, for elements with two $\hat{S}_1$ operators and one $\hat{S}_2$, one can transform the product of the operators into $i(\hat{B}_3^\dagger \hat{A}_3 - \hat{A}_3^\dagger \hat{B}_3)$. Thus, this element is equal to $-\langle \hat{S}_2^1 \hat{S}_2^2 \hat{S}_2^3\rangle = 0$.

5. Unitary transformations that link Stokes operators in the context of a mutually unbiased basis

According to formula (11), one can express Stokes operators as the ratio of the difference and the sum of photon number operators related orthogonal polarizations of a given polarization basis. The three measurement bases, $j = 1, 2, 3$, are mutually unbiased.

By this we mean the following. When we expand annihilation operators a_j and $a_{j_\perp}$ in terms of $a_{j'}$ and $a_{j'_\perp}$ of a different basis of the three, $j' \neq j$, the squares of the moduli of all expansion coefficients are $1/2$.

In other words, if we put $\hat{a}_j = \hat{a}_1(j)$ and $\hat{a}_{j_\perp} = \hat{a}_2(j)$, we have

$$\hat{a}_r^\dagger(j') = \sum_{s=1}^{2} U_{rs}(j \to j')\hat{a}_s^\dagger(j), \tag{A23}$$

where $U_{rs}(j \to j')$ is a unitary matrix linking two mutually unbiased bases in a Hilbert space of dimension 2.

The transformation (A23) gives the links between the three normalized Stokes operators, which in the notation used above read $\hat{S}_j = \frac{\sum_{s=1}^{2}(-1)^{(s-1)}\hat{a}_s^\dagger(j)\hat{a}_s(j)}{\sum_{s=1}^{2}\hat{a}_s^\dagger(j)\hat{a}_s(j)}$ (see, e.g. [37]). Applying this reasoning to formula (21), we can recover all terms from (19).

[1] J.-W. Pan, Z. B. Chen, C. Y. Lu, H. Weinfurter, A. Zeilinger, and M. Żukowski, Rev. Mod. Phys. **84**, 777 (2012).

[2] C. K. Hong and L. Mandel, Phys. Rev. A **31**, 2409 (1985); Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988); Y. H. Shih and C. O. Alley, *ibid.* **61**, 2921 (1988); M. A. Horne, A. Shimony, and A. Zeilinger, *ibid.* **62**, 2209 (1989); L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995); P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih, Phys. Rev. Lett. **75**, 4337 (1995).

[3] A. Cabello, D. Rodríguez, and I. Villanueva, Phys. Rev. Lett. **101**, 120402 (2008).

[4] D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger, Nature (London) **403**, 515 (2000).

[5] H. Weinfurter and M. Żukowski, Phys. Rev. A **64**, 010102(R) (2001).

[6] D. Ding, Y. He, F. Yan, and T. Gao, Quantum Inf. Process. **17**, 243 (2018).

[7] N. Kiesel, C. Schmid, U. Weber, G. Tóth, O. Gühne, R. Ursin, and H. Weinfurter, Phys. Rev. Lett. **95**, 210502 (2005).

[8] N. Kiesel, C. Schmid, G. Tóth, E. Solano, and H. Weinfurter, Phys. Rev. Lett. **98**, 063604 (2007).

[9] W. Wieczorek, R. Krischek, N. Kiesel, P. Michelberger, G. Tóth, and H. Weinfurter, Phys. Rev. Lett. **103**, 020504 (2009).

[10] D. Bouwmeester, J.-W. Pan, K. Mattle, M. Eibl, H. Weinfurter, and A. Zeilinger, Nature (London) **390**, 575 (1997).

[11] C. E. Vollmer, C. Baune, A. Samblowski, T. Eberle, V. Händchen, J. Fiurášek, and R. Schnabel, Phys. Rev. Lett. **112**, 073602 (2014).

[12] T. Jennewein, Ch. Simon, G. Weihs, H. Weinfurter, and A. Zeilinger, Phys. Rev. Lett. **84**, 4729 (2000).

[13] D. S. Naik, C. G. Peterson, A. G. White, A. J. Berglund, and P. G. Kwiat, Phys. Rev. Lett. **84**, 4733 (2000).

[14] S. P. Walborn, C. H. Monken, S. Pádua, and P. H. Souto Ribeiro, Phys. Rep. **495**, 87 (2010).

[15] U. L. Andersen, T. Gehring, C. Marquardt, and G. Leuchs, Phys. Scr. **91**, 053001 (2016).

[16] T. S. Iskhakov, I. N. Agafonov, M. V. Chekhova, and G. Leuchs, Phys. Rev. Lett. **109**, 150502 (2012).

[17] C. Simon and D. Bouwmeester, Phys. Rev. Lett. **91**, 053601 (2003).

[18] M. V. Chekhova, G. Leuchs, and M. Żukowski, Opt. Commun. **337**, 27 (2015).

[19] P. R. Sharapova, G. Frascella, M. Riabinin, A. M. Pérez, O. V. Tikhonova, S. Lemieux, R. W. Boyd, G. Leuchs, and M. V. Chekhova, Phys. Rev. Research **2**, 013371 (2020).

[20] P. R. Sharapova, O. V. Tikhonova, S. Lemieux, R. W. Boyd, and M. V. Chekhova, Phys. Rev. A **97**, 053827 (2018).

[21] K. Rosołek, M. Stobińska, M. Wieśniak, and M. Żukowski, Phys. Rev. Lett. **114**, 100402 (2015).

[22] S. L. Braunstein and R. I. McLachlan, Phys. Rev. A **35**, 1659 (1987).

[23] C. Okoth, A. Cavanna, N. Y. Joly, and M. V. Chekhova, Phys. Rev. A **99**, 043809 (2019).

[24] N. A. Borshchevskaya, K. G. Katamadze, S. P. Kulik, and M. V. Fedorov, Laser Phys. Lett. **12**, 115404 (2015).

[25] F. A. Domínguez-Serna, A. B. U'Ren, and K. Garay-Palmett, Phys. Rev. A **101**, 033813 (2020).

[26] C. W. Sandbo Chang, C. Sabín, P. Forn-Díaz, F. Quijandría, A. M. Vadiraj, I. Nsanzineza, G. Johansson, and C. M. Wilson, Phys. Rev. X **10**, 011011 (2020).

[27] R. A. Fisher, M. M. Nieto, and V. D. Sandberg, Phys. Rev. D **29**, 1107 (1984).

[28] M. Hillery, Phys. Rev. A. **42**, 498 (1990).

[29] K. Banaszek and P. L. Knight, Phys. Rev. A **55**, 2368 (1997).

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

[31] D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951).

[32] D. M. Greenberger, M. A. Horne, and A. Zeilinger, *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, 1989), pp. 69–72.

[33] D. Mogilevtsev, Opt. Commun. **118**, 565 (1995).

[34] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

[35] Q. Y. He, M. D. Reid, T. G. Vaughan, C. Gross, M. Oberthaler, and P. D. Drummond, Phys. Rev. Lett. **106**, 120405 (2011).

[36] M. Żukowski, M. Wieśniak, and W. Laskowski, Phys. Rev. A **94**, 020102(R) (2016).

[37] J. Ryu, B. Woloncewicz, M. Marciniak, M. Wieśniak, and M. Żukowski, Phys. Rev. Research **1**, 032041(R) (2019).

[38] N. D. Mermin, Pshys. Rev. Lett. **65**, 1838 (1990).

[39] A. Garg and N. D. Mermin, Phys. Rev. D **35**, 3831 (1987).

[40] G. Tóth and O. Gühne, Phys. Rev. Lett. **94**, 060501 (2005).

[41] M. Żukowski, W. Laskowski, and M. Wieśniak, Phys. Rev. A **95**, 042113 (2017).