Experimental observation of quantum contextuality beyond Bell nonlocality

Zheng-Hao Liu,[1,2] Hui-Xian Meng,[3,4] Zhen-Peng Xu,[5] Jie Zhou,[3] Sheng Ye,[1,2] Qiang Li,[1,2] Kai Sun,[1,2] Hong-Yi Su,[6] Adán Cabello,[7] Jing-Ling Chen,[3,*] Jin-Shi Xu,[1,2,†] Chuan-Feng Li,[1,2,‡] and Guang-Can Guo[1,2]

[1]*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, People's Republic of China*

[2]*CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, People's Republic of China*

[3]*Theoretical Physics Division, Chern Institute of Mathematics, Nankai University, Tianjin 300071, People's Republic of China*

[4]*School of Mathematics and Physics, North China Electric Power University, Beijing 102206, People's Republic of China*

[5]*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Straße 3, 57068 Siegen, Germany*

[6]*Graduate School of China Academy of Engineering Physics, Beijing 100193, People's Republic of China*

[7]*Departamento de Fisica Aplicada II, Universidad de Sevilla, E-41012 Sevilla, Spain*

(Received 2 June 2019; published 22 October 2019)

Bell nonlocality and quantum contextuality are two important resources which, however, behave very differently in many tasks where the quantum correlation has an edge over its classical counterpart. In this work, we directly compare these two behaviors, which via the Cabello-Severini-Winter graph-theoretical approach, can be associated with a same exclusivity graph. In particular, we consider the exclusivity graph that leads to I_{3322}-type inequalities in the Bell and the contextual scenarios, respectively. We find that maximal violation of Bell inequality is around 6.251, and maximal violation of the noncontextuality inequality is around 6.588 for a five-dimensional system and 6.571 for a four-dimensional system, respectively. The results predict a gap of $\Delta \approx 0.32$ between quantum contextuality and Bell nonlocality. We then present an experimental observation of the gap by employing both the maximally entangled photon pairs from the spontaneous parametric down-conversion process and the single photons encoded into qudits from an intrinsic defect in gallium nitride. Our results will further deepen the understanding of different types of quantum correlations.

DOI: 10.1103/PhysRevA.100.042118

I. INTRODUCTION

Bell nonlocality (BN) and quantum contextuality (QC) have been proved to be intrinsic properties of nature. A physical system exhibits Bell nonlocality if the outcomes of ideal measurements (i.e., those that yield the same outcome when repeated and do not disturb compatible measurements) cannot be reproduced by local-hidden-variable (LHV) models, in which measurement outcomes are independent of spacelike separated measurements performed on each distant subsystem [1]. In a similar sense, a physical system exhibits quantum contextuality if the outcomes of ideal measurements cannot be reproduced by noncontextual-hidden variable (NCHV) models in which measurement outcomes are independent of the measurement contexts (i.e., the set of jointly measurable observables that are actually measured) [2]. Both LHV and NCHV models are referred to as classical models in many situations, as they, as well as their breakdowns in quantum mechanics (QM), share several similarities indeed [3]. For instance, (i) both BN and QC reflect the confliction between quantum mechanics and some deterministic theories in which measurement outcomes are well defined even before measurements are performed; and (ii) every Bell inequality becomes a noncontextuality inequality if the constraint of spacelike separation between measurements on subsystems is removed [3,4].

As two important correlations, BN and QC have been studied widely, and in most cases separately, in the literature. They have also found significant applications in quantum information processing, such as speeding-up quantum algorithms [5], quantum key distribution [6], communication complexity [7], and randomness certification [8]. Very recently, several fundamental relations between these two quantum correlations were found. First, e.g., a fundamental monogamy relation exists between QC and BN. The theoretical result was presented in [9] and then observed in experiments [10,11]. Second, via the Cabello-Severini-Winter (CSW) graph-theoretical approach [12], the different behaviors between BN and QC can be compared in a same *exclusivity* graph [13] with which BN is proved to be tightly bounded by QC [14].

Considering an arbitrary Bell inequality written in the probability form, one easily has an exclusivity graph due to the exclusive relations among the probability events. Then, based on the same graph, one can also construct a noncontextuality inequality. Let $\Delta = I_{\mathrm{QC}} - I_{\mathrm{BN}}$ denote the gap between QC and BN, where I_{QC} and I_{BN} denote the maximal violations of the noncontextuality inequality and Bell inequality, respectively. In general, one has $\Delta \geqslant 0$. For $\Delta = 0$, it means Bell nonlocality saturates quantum contextuality, and $\Delta > 0$ means quantum contextuality beyond Bell nonlocality. The purpose of this work is to perform an experiment to observe

*chenjl@nankai.edu.cn
†jsxu@ustc.edu.cn
‡cfli@ustc.edu.cn

the gap Δ, thus demonstrating the phenomenon of quantum contextuality beyond Bell nonlocality. The point is to choose an appropriate exclusivity graph for which Δ is sufficiently large such that QC beyond BN can be observed in the experiment. The paper is organized as follows. In Sec. II, we shall study the exclusivity graph associated with the I_{3322}-type inequality, which theoretically gives $\Delta \approx 0.3$ [4]. In Sec. III, we shall then demonstrate the phenomenon of QC beyond BN by experiment. Within the experimental deviation, the experimental results coincide with the theoretical predictions. Conclusions will be drawn in the last section.

II. THEORY: QC BEYOND BN

Let us start from the symmetric Bell inequality $I_{3322} \leqslant 0$ [15], a natural generalization of the Clauser-Horne-Shimony-Holt (CHSH) inequality [16] from two measurement settings to three measurement settings. The Bell inequality can be written in the following form [4] (see Appendix A):

$$\begin{aligned} I = {} & P(0,0|0,1) + P(0,0|0,2) + P(0,0|1,0) \\ & + P(0,0|1,2) + P(0,0|2,0) + P(0,0|2,1) \\ & + P(0,1|1,1) + P(1,0|1,1) + P(1,1|1,1) \\ & + P(0,1|2,2) + P(1,0|2,2) + P(1,1|2,2) \\ & + P(1,\_|0,\_) + P(1,\_|1,\_) \\ & + P(\_,1|\_,0) + P(\_,1|\_,1) \overset{\text{LHV}}{\leqslant} 6, \end{aligned} \tag{1}$$

where I is a substitute for I_{3322}^{CSW}, $P(m,n|i,j) \equiv P(a_i = m, b_j = n) = |\langle v_{AB}|\Phi_{AB}\rangle|^2$ denotes the probability that Alice's ith measurement outcome is m and Bob's jth measurement outcome is n, $|\Phi_{AB}\rangle$ is the quantum state, $|v_{AB}\rangle = |a_i = m\rangle \otimes |b_j = n\rangle$ is the factorizable joint-measurement vector for Alice and Bob, $P(m,\_|i,\_) \equiv P(a_i = m)$ denotes the marginal probability of Alice's ith measurement result being m, and similarly for $P(\_,n|\_,j)$. For simplicity, we use $R_1, R_2, \ldots, R_{16}$ to denote the 16 terms in inequality (1) later. Inequality (1) holds for any two d-dimensional systems (i.e., two qudits). For two qubits, it is a tight inequality with the maximal quantum violation $I_{\text{BN}} = 25/4 = 6.25$, and we give the optimal observables and quantum state in Appendix A. When the dimension d increases, the maximal quantum violation slowly increases. It has been shown that the maximal violation tends to $I_{\text{BN}} \approx 6.250\,875\,38$ when $d \to \infty$ [17]. In other words, for the I_{3322}-type inequality, the value of Bell nonlocality is less than 6.251.

Based on the exclusive relations among the 16 probability events in the inequality (1), one can immediately plot an exclusivity graph G as shown in Fig. 1 (see also Ref. [4]). The vertices of graph G represent probability events, and two connected vertices imply that two events are exclusive with one another, i.e., the two corresponding projectors or joint-measurement vectors are mutually orthogonal. Accordingly, from graph G one can have the noncontextuality inequality as follows:

$$I = \sum_{i \in V} P_i \overset{\text{NCHV}}{\leqslant} \beta \overset{\text{QM}}{\leqslant} \vartheta(G), \tag{2}$$

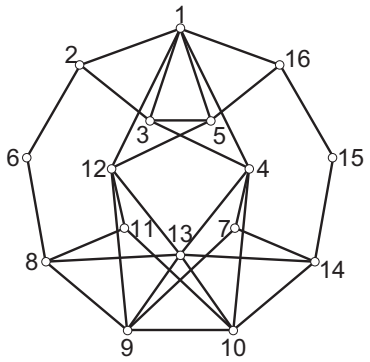

FIG. 1. Exclusivity graph associated with I. For Bell nonlocality, the 16 events in inequality (1) correspond to 1: 1, 1|2, 2, 2: 0, 0|2, 0, 3: 1, 0|2, 2, 4: 0, 0|2, 1, 5: 0, 1|2, 2, 6: _, 1|_, 0, 7: _, 1|_, 1, 8: 0, 0|1, 0, 9: 1, 0|1, 1, 10: 0, 1|1, 1, 11: 1, _|1, _, 12: 0, 0|1, 2, 13: 1, 1|1, 1, 14: 0, 0|0, 1, 15: 1, _|0, _, and 16: 0, 0|0, 2, and the spacelike measurements between Alice and Bob guarantee the mutual orthogonality. For quantum contextuality, 16 vertices correspond to 16 vectors $|A_i\rangle$'s, and $|A_i\rangle \perp |A_j\rangle$ when vertex i and vertex j are connected.

where V denotes the vertex of graph G, and $P_i \equiv P_{|\psi\rangle}(A_i = 1)$ denotes the probability of obtaining result 1 when the observable $|A_i\rangle\langle A_i|$ is measured on the state $|\psi\rangle$, namely, $P_{|\psi\rangle}(A_i = 1) = |\langle A_i|\psi\rangle|^2$. Here $\beta = 6$ is the classical bound of the NCHV models, and $\vartheta(G)$ is the maximal violation of the noncontextuality inequality and is equal to the Lovász number [18] describing the Shannon capacity of G.

The Lovász number is computed by $\vartheta(G) = \max_{\{|\psi\rangle,|A_i\rangle\}} \sum_{i \in V} |\langle A_i|\psi\rangle|^2$. In Ref. [4], the authors computed $\vartheta(G) = I_{\text{QC}} \approx 6.588\,412\,87$ without giving $|A_i\rangle$ and $|\psi\rangle$ that yield the Lovász number. Here we show that the above $\vartheta(G)$ can be achieved in a five-dimensional system, and we also list the corresponding optimal measurement settings and optimal state (see Table I in Appendix A). For the I_{3322}-type inequalities, the gap between QC and BN is given by $\Delta = I_{\text{QC}} - I_{\text{BN}} \approx 0.337$, which is sufficiently large, hence the I_{3322}-type inequality is a very good candidate to experimentally demonstrate QC beyond BN.

Remark 1. The difference between QC and BN can be understood in this way: For QC, in order to evaluate the quantity $\mathcal{Q} = \sum_{i \in V} |\langle A_i|\psi\rangle|^2$, one has to numerate all possible $|A_i\rangle$'s and $|\psi\rangle$ in the Hilbert space of arbitrary dimensions d, with all orthogonal relations of $|A_i\rangle$'s being satisfied, such that the maximal value $\mathcal{Q}^{\max}$ equals the Lovász number $\vartheta(G)$. For the graph in Fig. 1, the minimum dimension of obtaining the Lovász number is 5. Nevertheless, for testing the Bell inequality, it is additionally required that $|A_i\rangle$'s have to be factorizable, i.e., $|A_i\rangle = |v_i\rangle_{d_1} \otimes |u_i\rangle_{d_2}$, and so the dimension ought to be factorized to $d = d_1 \times d_2$ accordingly. For instance, let us take $d_1 = d_2 = 2$. By running over all possible $|A_i\rangle$'s and $|\psi\rangle$ under the factorizability requirement, one can obtain the maximal value $\mathcal{Q}^{\max} = 6.25$, which exactly is the maximal violation of inequality (1) for two qubits in the Bell scenario.

Remark 2. For the vector $|A_i\rangle$'s, the minimal dimension to fulfill the exclusive relations shown in Fig. 1 is 4. In a four-dimensional system, the maximal violation of

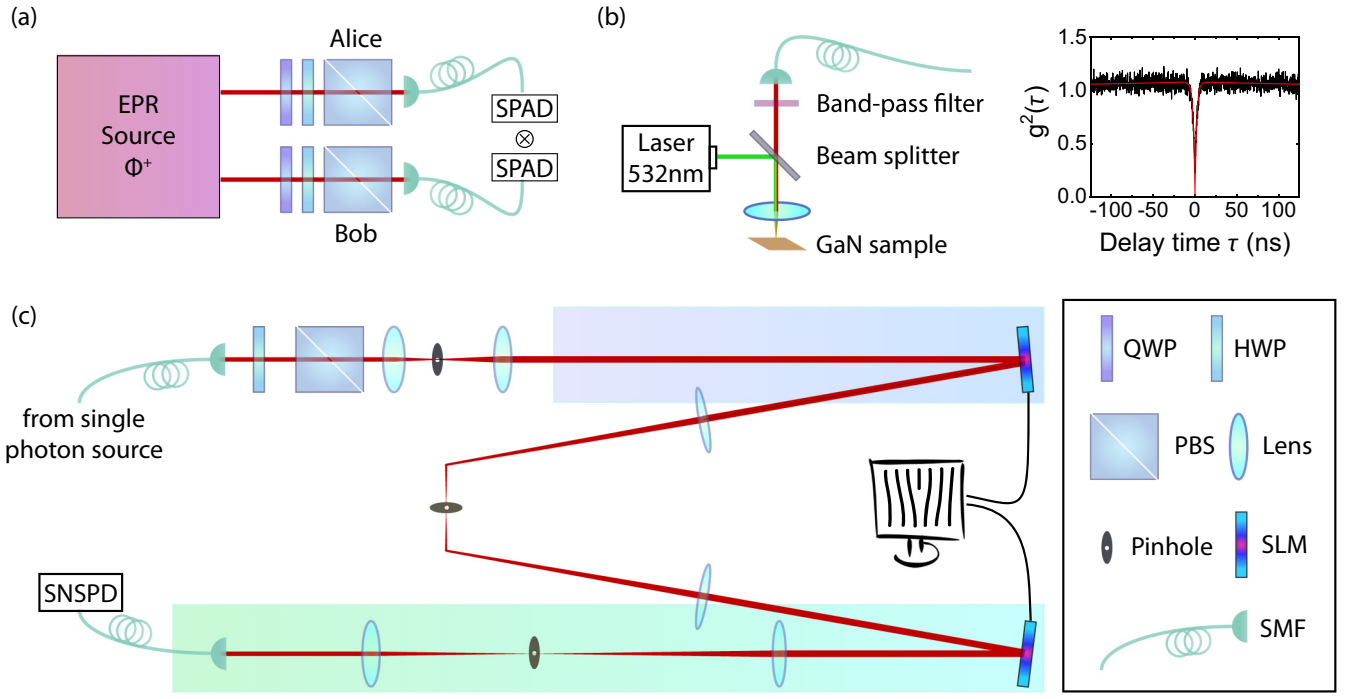

FIG. 2. Experimental setup. (a) The two-particle setup to check BN. Two maximally polarization-entangled photons were separately sent to Alice and Bob, going through their polarization discrimination system and recorded by single-photon avalanche detectors (SPADs). (b) The single-photon source, prepared by exciting an intrinsic defect in a bulk GaN sample. The second-order photon correlation function at zero delay without background correction was 0.225, clearly exhibiting the signature of photon antibunching and confirming the character of single-photon emission. Single photons were further filtered by a bandpass filter with a central wavelength of 780 nm and a bandwidth of 25 nm, and sent to the contextuality test setup. (c) The single-body contextuality test setup. The single photons from (b), after passing through the beam expanding system with two lenses and a pinhole between them, were sent to a spatial light modulator (SLM) controlled by a computer to prepare the high-dimensional states, which is denoted by a violet box. The measurement setup consists of another SLM and a single-mode fiber (SMF), and the counting apparatus here was a superconducting nanowire single-photon detector (SNSPD).

noncontextuality inequality (2) is $I_{\mathrm{QC}} \approx 6.571\,52$, which is slightly smaller than the Lovász number $\vartheta(G)$. The optimal settings (i.e., $|A_i\rangle, i = 1, 2, \ldots, 16$) and the state ($|\psi\rangle$) are given in Table IV of Appendix A. The gap is around $\Delta \approx 0.320$. Therefore, in this work we shall choose a four-dimensional system instead of a five-dimensional system to demonstrate QC beyond BN.

Remark 3. To experimentally test quantum contextuality, we need to recast inequality (2) to the following noncontextuality inequality (by adding the second term) based on theorem 1 of Ref. [19], i.e.,

$$I = \sum_{i\in V} P_{|\psi\rangle}(A_i = 1) - \sum_{(i,j)\in E} P_{|\psi\rangle}(A_i = 1, A_j = 1) \overset{\mathrm{NCHV}}{\leqslant} \beta \overset{\mathrm{QM}}{\leqslant} \vartheta(G), \tag{3}$$

where E denotes the vertex and edge sets of the graph G, and $P_{|\psi\rangle}(A_i = 1, A_j = 1) \equiv P_{|\psi\rangle}(A_i = 1)P_{|A_i\rangle}(A_j = 1)$ represents the experimental imprecision of $|A_i\rangle\perp|A_j\rangle$ when they are measured successively. Introducing the second term in (3) is crucial and necessary, as it is used to evaluate the compatibility requirement in any contextuality test in a practical experiment [19]. Certainly, the second term becomes zero in the ideal case and can thus be neglected. Crucially, the graph G needs to be extended with extra projectors that complete the full sets of compatible observables in order that every projector is measured as part of a complete basis [20]. In the work, we make measurements in the following six complete bases: $\{|A_1\rangle, |A_2\rangle, |A_3\rangle, |A_{17}\rangle\}$, $\{|A_1\rangle, |A_4\rangle, |A_7\rangle, |A_{18}\rangle\}$, $\{|A_1\rangle, |A_5\rangle, |A_{16}\rangle, |A_{19}\rangle\}$, $\{|A_1\rangle, |A_9\rangle, |A_{12}\rangle, |A_{20}\rangle\}$, $\{|A_1\rangle, |A_6\rangle, |A_8\rangle, |A_{21}\rangle\}$, and $\{|A_1\rangle, |A_{14}\rangle, |A_{15}\rangle, |A_{22}\rangle\}$. Here, $|A_{17}\rangle, |A_{18}\rangle, \ldots, |A_{22}\rangle$ are uniquely determined by the other three vectors in a complete basis.

Thus, the detection probability of each observable, say, $|A_2\rangle\langle A_2|$, can be obtained as $P_{|\psi\rangle}(A_2 = 1) = N_{|\psi\rangle}(A_2)/[N_{|\psi\rangle}(A_1) + N_{|\psi\rangle}(A_2) + N_{|\psi\rangle}(A_3) + N_{|\psi\rangle}(A_{17})]$, where $N_{|\psi\rangle}(A_i)$ is the number of counts of obtaining result 1 when $|A_i\rangle\langle A_i|$ is measured on the state $|\psi\rangle$, and similarly for the measurements of the other observables.

III. EXPERIMENTAL SETUP AND RESULTS

We are now ready to experimentally test the two distinct quantum values of I separately, namely, $I_{\mathrm{BN}} \leqslant 6.25$ and $I_{\mathrm{QC}} \leqslant 6.57$, to explicitly demonstrate QC beyond BN. The experimental setup is illustrated in Fig. 2. We start from I_{BN}. Based on the maximally entangled state $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$ of two photonic polarization qubits, we checked the maximal violation of inequality (1). A type-II β-barium borate (β-BBO) crystal was pumped by a frequency-doubled

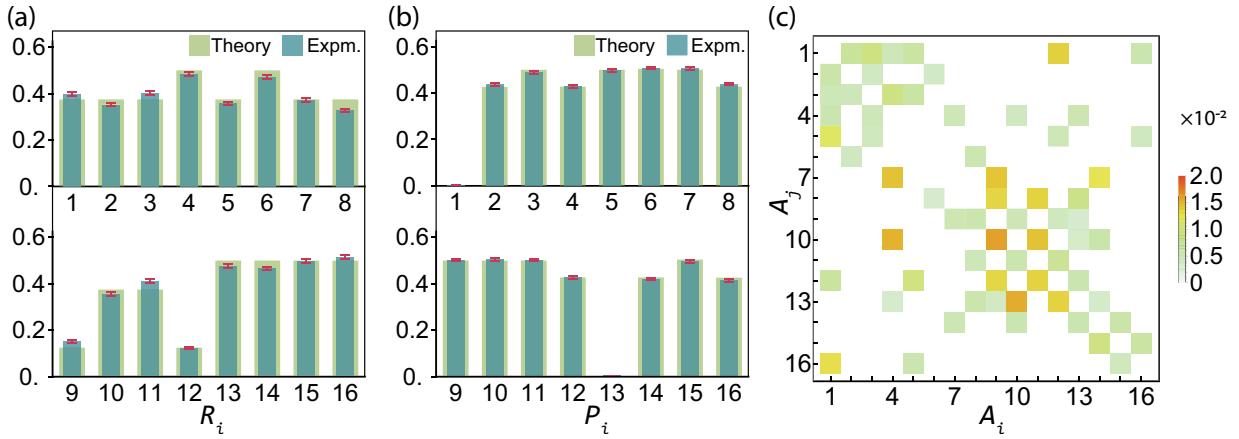

FIG. 3. Experimental results of detection probabilities for computing I_{BN} (a) and I_{QC} (b and c). (a) Coincidence counting probabilities of Alice and Bob. The settings are chosen to produce the 16 terms, namely, $R_1 \sim R_{16}$ in (1). (b) Detection probabilities of the 16 optimal projectors, representing the 16 vertices of the corresponding exclusivity graph, with the input state $|\psi\rangle$. The probability is equivalent to $P_1 \sim P_{16}$ in (2). (c) Detection probabilities of nonzero $P_{|A_i\rangle}(A_j = 1)$, where $|A_i\rangle$ and $|A_j\rangle$ are connected by 32 edges in the exclusivity graph.

femtosecond laser to generate a polarization-entangled biphoton, which was further sent to Alice and Bob, where a polarization beam splitter (PBS), proceeded by a quarter-wave plate (QWP) and a half-wave plate (HWP), was utilized to construct a polarization discrimination system. Two single-photon avalanche detectors (SPADs) recorded photon counting rate after the polarization discrimination system, and their coincidence counts were proportional to the probability of detecting two photons with a certain setting.

For each set of measurements, the total photon count was about 8000, and results are shown in Fig. 3(a). Substituting the measurement results into inequality (1) gives a result of $I_{\mathrm{BN}} = 6.165 \pm 0.012$. Here, the standard deviation was deduced from the Poisson distribution. The result surpasses prediction of any possible LHV model but, due to imperfections in the experiment, falls sightly below the Tsirelson bound, 6.25.

We now turn to investigate I_{QC}. In a four-dimensional space, the single-body quantum violation of inequality (3) was checked. To construct this synthetic space, we exploit the orbital angular momentum (OAM) of light. The manipulation of states was accomplished by modulating the phase of the photon's wave function.

By exciting an intrinsic defect in a bulk GaN sample, we obtained an ultrabright, photostable single-photon source at room temperature [21]. A bandpass filter with a central wavelength of 780 nm and a width of 25 nm was used to filter the fluorescence of the defects and match the operation wavelength of the contextuality test setup. The total photon counting rate past the filter was about 0.94 Mcps.

The single photon was directed to the setup by a single-mode fiber (SMF) and spatial filtered into fundamental Gaussian mode. Using a spatial light modulator (SLM), a coordinate-related wave function, originating from displayed holograms on the SLM, was bestowed to photons [22]. The photons would then carry a wave function resembling superpositions of Laguerre-Gaussian modes LG_p^l with OAM number of l [23] and be prepared to desired initial four-dimensional states $|\psi\rangle$ after being diffracted by the first SLM. The four eigenvectors are set to be $|i_1\rangle = |1\rangle$, $|i_2\rangle = |3\rangle$, $|i_3\rangle = |-1\rangle$, and $|i_4\rangle = |-3\rangle$. A $4f$ system mapped the first and the second SLM at its input and output plane, with an aperture inserted at the focal plane to filter off the unwanted zero- and higher-order diffraction terms. The second SLM converted specific OAM states $|\alpha\rangle$ back to the Gaussian mode $|0\rangle$, which was subsequently postselected by a SMF, as photons carrying nonzero OAM cannot be focused to a spot and collected by a SMF. In addition, two telescope lenses were inserted between the second SLM and SMF to adjust beam size and optimize fiber coupling efficiency. Photons were then sent to a superconducting nanowire single-photon detector (SNSPD), and the counting rate was proportional to the probability of successful postselection, i.e., $|\langle\alpha|\psi\rangle|^2$.

The input state was prepared to be $|\psi\rangle = (|i_1\rangle + |i_2\rangle)/\sqrt{2}$, and the 16 corresponding projectors $|A_j\rangle\langle A_j|$ ($j = 1, 2, \ldots, 16$) were then measured on six sets of complete bases. To test the inequality (3), we also need to check the no-signaling condition, which requires the local marginal probabilities of successive measurements to be irrelevant to the other settings being chosen, either forward or backward in time, as suggested in [19]. Our experimental results confirm the no-signaling condition in which the detailed information is shown in Appendix B.

The first term of I_{QC} inequality was evaluated by directly referring to detect probabilities of the 16 corresponding projectors on six full sets of compatible observables. In order to obtain all probabilities needed for the exclusivity graph, there were in total 19 independent projectors to be measured. The measurement settings of these projectors were shown in Appendix B. The counting rate of each projector was then recorded and the detection probability was calculated. The integration time for each projection measurement was 10 s. Shown in Fig. 3(b), the results were fairly close to prediction. The experimental result of the first term of I_{QC} was 6.548 ± 0.020, an error estimated by assuming Poisson distribution.

The second term of I_{QC} inequality represents experimental imperfections of orthogonality between states denoted by connected vertices in the exclusivity graph. Similar to the way the

first term was calculated, the second term was also evaluated by adding up mutual projection probabilities of the states mentioned above, and their fluctuation was also estimated by Poisson distribution. The measurement result is displayed in Fig. 3(c), yielding a value of -0.060 ± 0.005 without correcting background noise from SNSPD and stray light.

By adding the two terms together, the final result of quantum contextuality is $I_{\mathrm{QC}} = 6.488 \pm 0.025$, exceeding the classical bound by about 19 standard deviations and the Tsirelson's bound derived from Bell nonlocality by over 9 standard deviations. We thus successfully observed quantum contextuality beyond Bell nonlocality.

IV. CONCLUSION

We have studied the different behaviors between Bell nonlocality and quantum contextuality associated with the same exclusivity graph by comparing quantum violations of the I_{3322}-type Bell inequality and the noncontextuality inequality. Quantum contextuality has been investigated in the five- and four-dimensional systems, and we have found that the gap between QC and BN is $\Delta \approx 0.3$, which is experimentally observable. We have performed an experimental observation of the gap in a four-dimensional system to demonstrate the phenomenon of QC beyond BN. Within experimental deviation, the results are in good agreement with the theoretical predictions. Our results indicate a deep-rooted connection between graph theory and quantum correlations, and will further deepen the understanding of different types of quantum correlations.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (Grant No. 2016YFA0302700), the National Natural Science Foundation of China (Grants No. 61725504, No. 61490711, No. 11774335, and No. 11821404), the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS) (Grant No. QYZDY-SSW-SLH003), Science Foundation of the CAS (Grant No. ZDRW-XH-2019-1), Anhui Initiative in Quantum Information Technologies (Grants No. AHY060300 and No. AHY020100), and the Fundamental Research Funds for the Central Universities (Grants No. WK2030380017 and No. WK2470000026). H.-X.M. was supported by a project funded by the National Natural Science Foundation of China (Grant No. 11901317). H.-Y.S. was supported by the National Natural Science Foundation of China (Grant No. 11905209) and the China Postdoctoral Science Foundation (Grant No. 2018M630063). J.-L.C. was supported by the National Natural Science Foundations of China (Grant No. 11875167) and the Fundamental Research Funds for the Central Universities, Nankai University (Grant No. 63191507). J.Z. was supported by the Nankai Zhide Foundation. The processing of GaN samples was partially carried out at the USTC Center for Micro and Nanoscale Research and Fabrication.

Z.-H.L., H.-X.M., and Z.-P.X. contributed equally to this work.

APPENDIX A: THE THEORETICAL PART

1. The Bell inequality and noncontextuality inequality associated with I_{3322}^{CSW}

The symmetric version of the inequality $I_{3322} \leqslant 0$ proposed in Ref. [15] is

$$I_{3322} \overset{\mathrm{LHV}}{\leqslant} 0, \tag{A1}$$

with

$$\begin{aligned} I_{3322} = {} & P(0,0|0,1) + P(0,0|0,2) + P(0,0|1,0) \\ & + P(0,0|1,2) + P(0,0|2,0) + P(0,0|2,1) \\ & - P(0,0|1,1) - P(0,0|2,2) - P(0,\_|0,\_) \\ & - P(0,\_|1,\_) - P(\_,0|\_,0) - P(\_,0|\_,1). \end{aligned} \tag{A2}$$

Using the equation

$$-P(A_x = a, B_x = b) = -1 + \sum_{(a',b') \neq (a,b)} P(A_x = a', B_y = b') \tag{A3}$$

to replace probabilities with minus signs by the corresponding positive probabilities, we obtain

$$I_{3322} = I - 6, \tag{A4}$$

where

$$\begin{aligned} I \equiv I_{3322}^{\mathrm{CSW}} = {} & P(0,0|0,1) + P(0,0|0,2) + P(0,0|1,0) \\ & + P(0,0|1,2) + P(0,0|2,0) + P(0,0|2,1) \\ & + P(0,1|1,1) + P(1,0|1,1) + P(1,1|1,1) \\ & + P(0,1|2,2) + P(1,0|2,2) + P(1,1|2,2) \\ & + P(1,\_|0,\_) + P(1,\_|1,\_) + P(\_,1|\_,0) \\ & + P(\_,1|\_,1). \end{aligned} \tag{A5}$$

TABLE I. Numerically optimal settings and optimal state in a five-dimensional system for $I_{\mathrm{QC}} \approx 6.588\,412\,87$.

$\lvert A_i\rangle$	i_1	i_2	i_3	i_4	i_5
$\lvert A_1\rangle$	0	0	0	0	1
$\lvert A_2\rangle$	0.538648	−0.458664	0.0322917	−0.706005	0
$\lvert A_3\rangle$	0.672155	−0.188775	−0.359722	0.619009	0
$\lvert A_4\rangle$	0	0.95651	0	0.2917	0
$\lvert A_5\rangle$	0.188775	−0.672155	−0.359722	−0.619009	0
$\lvert A_6\rangle$	0.192831	−0.671551	−0.44115	0.563224	0
$\lvert A_7\rangle$	0.912455	0	−0.409176	0	0
$\lvert A_8\rangle$	−0.946087	0	0	0.323911	0
$\lvert A_9\rangle$	0	1	0	0	0
$\lvert A_{10}\rangle$	1	0	0	0	0
$\lvert A_{11}\rangle$	0	0.912455	0.409176	0	0
$\lvert A_{12}\rangle$	0.95651	0	0	0.2917	0
$\lvert A_{13}\rangle$	0	0	1	0	0
$\lvert A_{14}\rangle$	0	−0.946087	0	0.323911	0
$\lvert A_{15}\rangle$	−0.671551	0.192831	0.44115	0.563224	0
$\lvert A_{16}\rangle$	−0.458664	0.538648	−0.0322917	−0.706005	0
$\lvert\psi\rangle$	−0.674033	0.674033	0.302255	0	0

Therefore, the inequality $I_{3322} \leqslant 0$ can be written as

$$I \overset{\text{LHV}}{\leqslant} 6. \tag{A6}$$

The maximal Bell nonlocality (BN) of inequality (A6) for two qubits is exactly $I_{\text{BN}} = 6.25$, which can be obtained in the following case:

$$\begin{cases} A_0 = B_0 = \begin{pmatrix} \frac{\sqrt{2-\sqrt{3}}}{2} & \frac{\sqrt{2+\sqrt{3}}}{2} \\ \frac{\sqrt{2+\sqrt{3}}}{2} & \frac{\sqrt{2}-\sqrt{6}}{4} \end{pmatrix}, A_1 = B_2 = \begin{pmatrix} \frac{\sqrt{2+\sqrt{3}}}{2} & \frac{\sqrt{2-\sqrt{3}}}{2} \\ \frac{\sqrt{2-\sqrt{3}}}{2} & -\frac{\sqrt{2+\sqrt{3}}}{2} \end{pmatrix}, A_2 = B_1 = \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \\ \rho = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} = |\Phi^+\rangle\langle\Phi^+|, \quad |\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\ P(a,b|i,j) = \text{Tr}\left[\frac{I_2+(-1)^a A_i}{2} \otimes \frac{I_2+(-1)^b B_j}{2}\rho\right]. \end{cases} \tag{A7}$$

For the exclusivity graph G (see Fig. 1) associated with I, a noncontextuality (NC) inequality is constructed with the form

$$I = \sum_{i\in V} P_{|\psi\rangle}(A_i = 1) - \sum_{(i,j)\in E} P_{|\psi\rangle}(A_i = 1, A_j = 1) \overset{\text{NCHV}}{\leqslant} 6, \tag{A8}$$

where V and E respectively denote the vertex and edge sets of graph G, $P_{|\psi\rangle}(A_i = 1)$ denotes the probability of obtaining result 1 when the observable $|A_i\rangle\langle A_i|$ is measured on the state $|\psi\rangle$, i.e., $P_{|\psi\rangle}(A_i = 1) = |\langle A_i|\psi\rangle|^2$, $P_{|\psi\rangle}(A_i = 1, A_j = 1)$, i.e., $P_{|\psi\rangle}(A_i = 1)P_{|A_i\rangle}(A_j = 1)$, represents the experimental imprecision of $|A_i\rangle \perp |A_j\rangle$ when vertex i and vertex j are connected in the graph.

The optimal quantum contextuality (QC) of inequality (3) is just the Lovász number of G, i.e., $I_{\text{QC}} = \vartheta(G) \approx 6.588\,412\,87$, which is obtained in the five-dimensional quantum system. In Table I, we list the optimal settings and the optimal state numerically.

Remark 4. If we take the following analytical settings and state in Table II , then the maximal QC is

$$I = \sum_{i=1}^{16} |\langle i|\psi\rangle|^2 - 0 = [355 + 209\cos(2\tau) + 112\sin(2\tau)]/90, \tag{A9}$$

whose maximal point is $\tau = \arctan[(-209 + 5\sqrt{2249})/112]$ with the maximum $(71 + \sqrt{2249})/18 \approx 6.579\,09$. And if we choose $\tau = 2\pi/25$, then $\sum_{i=1}^{16} |\langle i|\psi\rangle|^2 \approx 6.578\,94$. It is easy to see that this numerical contextuality is close to the optimal one with an error of 0.0014.

2. QC of I in a four-dimensional system

In the four-dimensional quantum system, we take

$$\begin{cases} |\psi\rangle = (\cos[\theta], \sin[\theta]\cos[\phi], \sin[\theta]\sin[\phi]\cos[\tau], \sin[\theta]\sin[\phi]\sin[\tau]) \\ |A_i\rangle = (\cos[\theta_i], \sin[\theta_i]\cos[\phi_i], \sin[\theta_i]\sin[\phi_i]\cos[\tau_i], \sin[\theta_i]\sin[\phi_i]\sin[\tau_i]), \end{cases} \tag{A10}$$

with $i = 1, 2, \ldots, 16$, then we have

$$\begin{aligned} P_{|\psi\rangle}(A_i = 1) &= |\langle A_i|\psi\rangle|^2 \\ &= (\cos[\theta_i]\cos[\theta] + \sin[\theta_i]\sin[\theta](\cos[\phi_i]\cos[\phi] + \sin[\phi_i]\sin[\phi](\cos[\tau_i]\cos[\tau] + \sin[\tau_i]\sin[\tau])))^2 \end{aligned} \tag{A11}$$

and

$$\max\{I : \theta, \phi, \tau, \theta_i, \phi_i, \tau_i\} \approx 6.571\,52. \tag{A12}$$

To obtain Eq. (A12), it is sufficient to assume the settings and state have the form in Table III, and the angles are as follows:

$$\begin{cases} \phi_2 \approx 1.015\,602\,970\,716\,640\,5, \phi_4 \approx 1.180\,253\,821\,474\,319\,2, \\ \phi_3 = \arctan\frac{2\cos\phi_4}{\sin^2\phi_4}, \phi_1 = -\arctan\frac{\cos\phi_2\cos\phi_3 - \sin\phi_2\sin\phi_3\sin\phi_4}{\cos\phi_4\sin\phi_3}, \phi_5 = -\arctan\frac{\cos\phi_1\cos\phi_2 + \phi_4}{\sin\phi_1}. \end{cases} \tag{A13}$$

TABLE II. Analytical settings and state in a five-dimensional system for I.

$\|A_i\rangle$	i_1	i_2	i_3	i_4	i_5
$\|A_1\rangle$	0	0	0	0	1
$\|A_2\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{\sqrt{2}}$	0
$\|A_3\rangle$	$\frac{3}{2\sqrt{5}}$	$\frac{1}{2\sqrt{5}}$	$\frac{1}{\sqrt{10}}$	$\sqrt{\frac{2}{5}}$	0
$\|A_4\rangle$	0	$\frac{2\sqrt{2}}{3}$	0	$-\frac{1}{3}$	0
$\|A_5\rangle$	$\frac{1}{2\sqrt{5}}$	$\frac{3}{2\sqrt{5}}$	$\frac{1}{\sqrt{10}}$	$-\sqrt{\frac{2}{5}}$	0
$\|A_6\rangle$	$\frac{1}{2\sqrt{5}}$	$\frac{3}{2\sqrt{5}}$	$\frac{1}{\sqrt{10}}$	$\sqrt{\frac{2}{5}}$	0
$\|A_7\rangle$	$\frac{2\sqrt{2}}{3}$	0	$\frac{1}{3}$	0	0
$\|A_8\rangle$	$\frac{2\sqrt{2}}{3}$	0	0	$-\frac{1}{3}$	0
$\|A_9\rangle$	0	1	0	0	0
$\|A_{10}\rangle$	1	0	0	0	0
$\|A_{11}\rangle$	0	$\frac{2\sqrt{2}}{3}$	$\frac{1}{3}$	0	0
$\|A_{12}\rangle$	$\frac{2\sqrt{2}}{3}$	0	0	$\frac{1}{3}$	0
$\|A_{13}\rangle$	0	0	1	0	0
$\|A_{14}\rangle$	0	$\frac{2\sqrt{2}}{3}$	0	$\frac{1}{3}$	0
$\|A_{15}\rangle$	$\frac{3}{2\sqrt{5}}$	$\frac{1}{2\sqrt{5}}$	$\frac{1}{\sqrt{10}}$	$-\sqrt{\frac{2}{5}}$	0
$\|A_{16}\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{\sqrt{2}}$	0
$\|\psi\rangle$	$\frac{\cos\tau}{\sqrt{2}}$	$\frac{\cos\tau}{\sqrt{2}}$	$\sin\tau$	0	0

In this case, Table III becomes the following Table IV.

In experiment, one can choose $\phi_2 = 1.015, \phi_4 = 1.180$, and then the orthogonal relationship can be satisfied with errors 10^{-32}, i.e.,

$$\sum_{(i,j)\in E} P_{|\psi\rangle}(A_i = 1, A_j = 1) \approx 9.629\,65 \times 10^{-33} < 10^{-32}, \tag{A14}$$

and $\max I \approx 6.571\,52$.

TABLE III. Settings and state in a four-dimensional system for I.

$\|A_i\rangle$	i_1	i_2	i_3	i_4
$\|A_1\rangle$	0	0	1	0
$\|A_2\rangle$	$\cos\phi_1\cos\phi_2$	$\sin\phi_1$	0	$-\cos\phi_1\sin\phi_2$
$\|A_3\rangle$	$\cos\phi_3$	$\cos\phi_4\sin\phi_3$	0	$\sin\phi_3\sin\phi_4$
$\|A_4\rangle$	0	$\sin\phi_4$	0	$-\cos\phi_4$
$\|A_5\rangle$	$\cos\phi_4\sin\phi_3$	$\cos\phi_3$	0	$-\sin\phi_3\sin\phi_4$
$\|A_6\rangle$	$\cos\phi_4\cos\phi_5$	$\sin\phi_5$	0	$\cos\phi_5\sin\phi_4$
$\|A_7\rangle$	1	0	0	0
$\|A_8\rangle$	$\sin\phi_4$	0	0	$-\cos\phi_4$
$\|A_9\rangle$	0	1	0	0
$\|A_{10}\rangle$	1	0	0	0
$\|A_{11}\rangle$	0	1	0	0
$\|A_{12}\rangle$	$\sin\phi_4$	0	0	$\cos\phi_4$
$\|A_{13}\rangle$	0	0	1	0
$\|A_{14}\rangle$	0	$\sin\phi_4$	0	$\cos\phi_4$
$\|A_{15}\rangle$	$\sin\phi_5$	$\cos\phi_4\cos\phi_5$	0	$-\cos\phi_5\sin\phi_4$
$\|A_{16}\rangle$	$\sin\phi_1$	$\cos\phi_1\cos\phi_2$	0	$\cos\phi_1\sin\phi_2$
$\|\psi\rangle$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0	0

TABLE IV. Numerically optimal measurement settings and state in a four-dimensional system for $I_{\mathrm{QC}} \approx 6.571\,52$.

$\|A_i\rangle$	i_1	i_2	i_3	i_4
$\|A_1\rangle$	0	0	1	0
$\|A_2\rangle$	0.469723	0.453743	0	0.757283
$\|A_3\rangle$	0.746839	0.253161	0	0.614931
$\|A_4\rangle$	0	0.924703	0	−0.38069
$\|A_5\rangle$	0.253161	0.746839	0	−0.614931
$\|A_6\rangle$	0.249899	0.754381	0	0.607009
$\|A_7\rangle$	1	0	0	0
$\|A_8\rangle$	0.924703	0	0	−0.38069
$\|A_9\rangle$	0	1	0	0
$\|A_{10}\rangle$	1	0	0	0
$\|A_{11}\rangle$	0	1	0	0
$\|A_{12}\rangle$	0.924703	0	0	0.38069
$\|A_{13}\rangle$	0	0	1	0
$\|A_{14}\rangle$	0	0.924703	0	0.38069
$\|A_{15}\rangle$	0.754381	0.249899	0	−0.607009
$\|A_{16}\rangle$	0.453743	0.469723	0	0.757283
$\|\psi\rangle$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	0	0

Remark 5. In fact, for $\phi_2 = \pi/3, \phi_4 = 7\pi/18$, we can obtain that the orthogonal relationship is strictly satisfied $[\sum_{(i,j)\in E} P_{|\psi\rangle}(A_i = 1, A_j = 1) = 0]$ and $\max I \approx 6.562\,33$.

APPENDIX B: THE EXPERIMENTAL PART

1. Testing experimental reliability of orbital angular momentum qudit

The two experimental tests rely on a photonic polarization qubit and an OAM qudit, and for the latter one it is vital to confirm that the measurements performed in experiment ideally represent the terms in (3). Due to the complexity of a four-dimensional system, we opted out of a complete state tomography and, instead, chose to test the orthonormality of the six sets of complete bases used in our experiment.

The Born rule declares that for an observable corresponding to a self-adjoint operator A, the probability of obtaining result λ_i on state $|\psi\rangle$ is $\langle\psi|\Pi_i|\psi\rangle$, where Π_i is the projection operator onto the eigenspace of A corresponding to the eigenvalue of λ_i. It follows immediately from the Born rule that when we prepare an eigenstate of Π_i, the probability of getting λ_i in the measurement A is 1, while the probabilities of getting other $\lambda_{j\neq i}$'s are 0. For an ideal measurement, the complete set of bases should then be orthonormal, regardless of the exact bases settings we choose. Figure 4 shows the result of testing orthonormality of the six sets of complete bases. When we prepared the state corresponding to one specific outcome and then measured the corresponding projector, we (almost) always obtained that certain outcome. Since the projectors to orthogonal bases commute with each other, the null counting rates between orthogonal states and projectors can also serve as proof of nondisturbance of compatible measurements. We thus infer that the measurements performed in the experiment are overall approximately ideal, and the data is suitable to be analyzed in the CSW framework using (3).

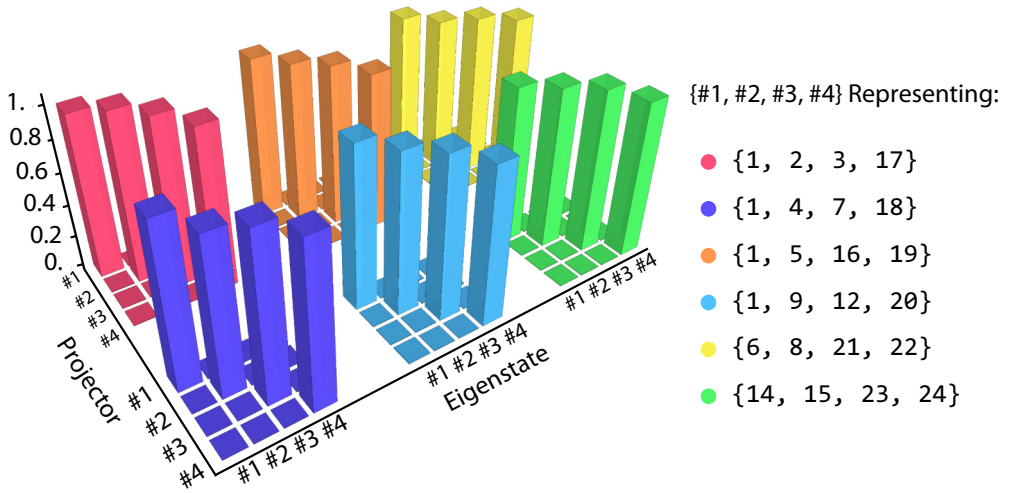

FIG. 4. Testing the orthonormality of the six sets of complete measurements used in the experiment by enumerating the detection probability between bases within each set. For example, the (#2, #3) slot in the darkest blue histogram shows the measured detection probability for $|\langle A_7|A_4\rangle|^2$.

2. Details about the single-photon source

The sample used in this experiment was a 2-μm-thick magnesium (Mg)-doped gallium nitride (GaN) film on a 2-mm undoped GaN layer grown on sapphire substrate. This kind of sample was reported in [21], and we followed their practice to pretreat the sample with acetone and isopropanol (1:1) solution and clean in ultrasound bath. The defects distributed over the sample surface, and a confocal microscopy system was exploited to excite the sample and collect emitted single photons.

In our experiment, a 532-nm continuous-wave laser was used to excite the GaN sample. The laser was reflected by a dichroic mirror focused onto the sample by an objective lens with high numeric aperture of 1.25. The fluorescence was collected by the same objective and filtered by the dichroic mirror and the bandpass filter.

To examine attributes of the photoluminescence (PL) signal, it was coupled to a single-mode fiber and either guided to a SPAD for photon counting and spectrum measurement (the bandpass filter was removed for this measurement) or a Hanbury Brown-Twiss (HBT) interferometer setup for single-photon verification.

(1) The second-order autocorrelation was measured to confirm the single-photon attribute. The experimental value of $g^2(\tau)$ at $\tau = 0$ was 0.225. As the dip fell below the threshold of 0.5, we confirmed that the PL emitter was a genuine single-photon source.

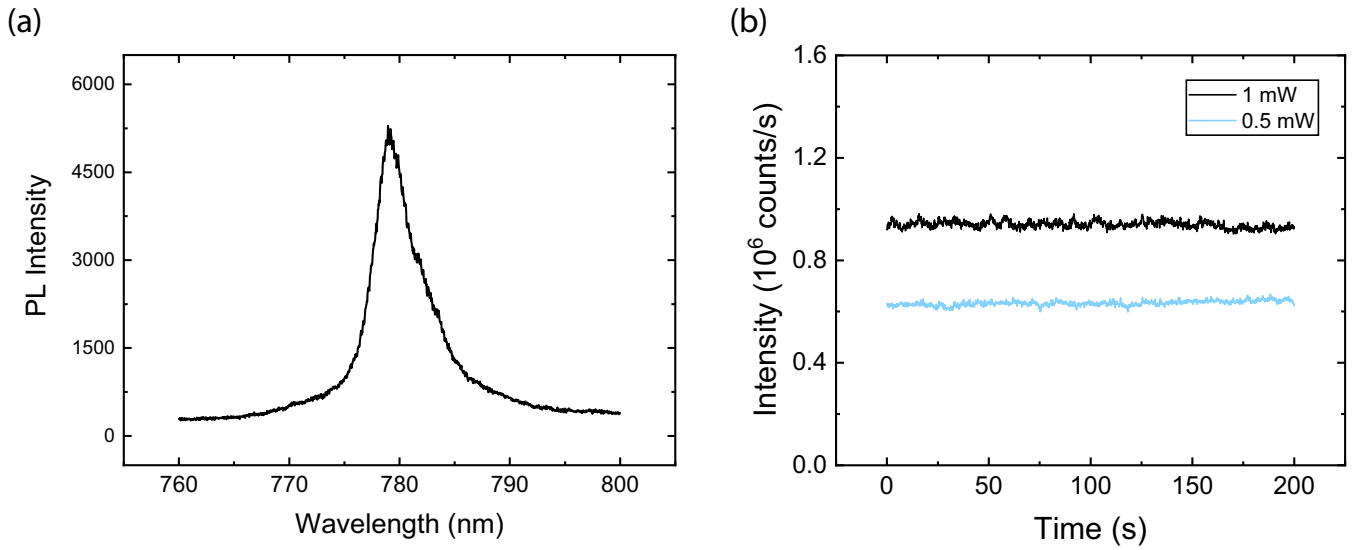

FIG. 5. The properties of the single-photon source used in the contextuality experiment. (a) The spectrum. The central wavelength of the emitter was 779.61 nm and the FWHM was 4.97 nm. (b) Photostability stats with respect to two different excitation powers, 0.5 mW and 1 mW.

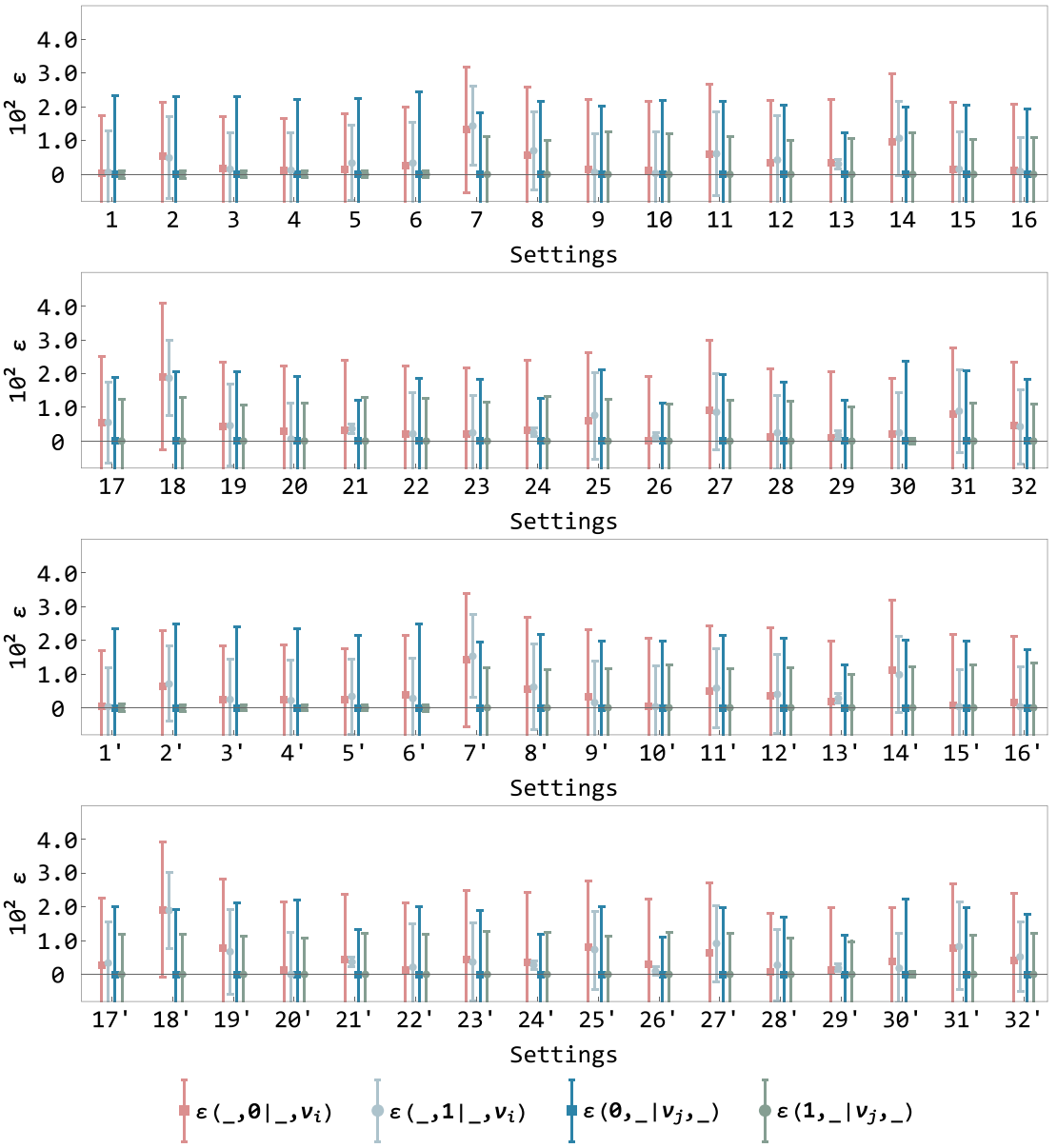

FIG. 6. Eliminating the possibility of signaling between sequential measurements. By exhibiting null results for the four signaling factors $\varepsilon(0,\_|v_i,\_)$, $\varepsilon(1,\_|v_i,\_)$, $\varepsilon(\_,0|\_,v_j)$, and $\varepsilon(\_,1|\_,v_j)$, the nonsignaling condition required in [19] is preserved. In each plot, the horizontal coordinate corresponds to experimental settings, i.e., different choices of states and projectors represented by edges of the exclusivity graph. The settings with primed notation swapped eigenstates and projectors. The four signaling factors are shown in four different colors.

(2) Spectrum. Figure 5(a) shows the spectrum of the emitter we used in the contextuality experiment with the background subtracted. It had a central wavelength of 779.61 nm and a FWHM of about 4.97 nm.

(3) Photostability. The PL intensity versus time under continuous-wave laser excitation was recorded and plotted in Fig. 5(b). The standard deviation of photon counting during a span of 200 s was below 1.5%. As a result, the photon source was narrow linewidth, bright, and stable. The photon source was then connected to the contextuality experiment setup by SMF.

3. Testing nonsignaling condition of sequential measurements

The nonsignaling condition, indicating that local marginal probabilities of Alice are independent of Bob's measurement setting, and vice versa [24], was also checked for testing contextuality in the CSW approach, as suggested in [19,20]. From [19], the statistics of the second measurements v_j affected by the first measurements v_i can be calculated as

$$\varepsilon(\_,0|\_,v_j) = |P_{|\psi\rangle}(v_j=0) - P_{|\psi\rangle}(v_i=0,v_j=0) - P_{|\psi\rangle}(v_i=1,v_j=0)|, \tag{B1a}$$

$$\varepsilon(\_,1|\_,v_j) = |P_{|\psi\rangle}(v_j=1) - P_{|\psi\rangle}(v_i=0,v_j=1) - P_{|\psi\rangle}(v_i=1,v_j=1)|, \tag{B1b}$$

$$\varepsilon(0,\_|v_i,\_) = |P_{|\psi\rangle}(v_i=0) - P_{|\psi\rangle}(v_i=0,v_j=0) - P_{|\psi\rangle}(v_i=0,v_j=1)|, \tag{B1c}$$

$$\varepsilon(1,\_|v_i,\_) = |P_{|\psi\rangle}(v_i=1) - P_{|\psi\rangle}(v_i=1,v_j=0) - P_{|\psi\rangle}(v_i=1,v_j=1)|. \tag{B1d}$$

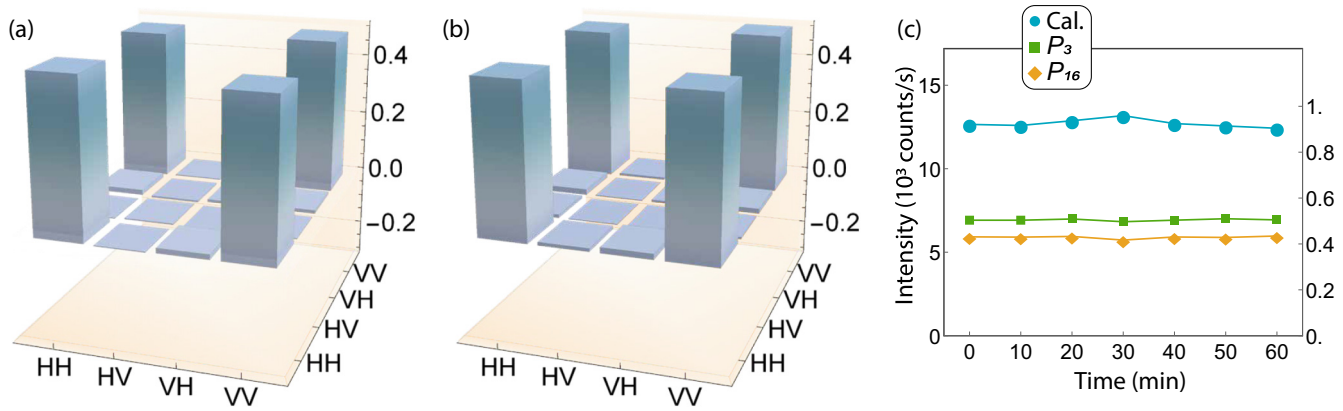

FIG. 7. Confirming the system's stability during the experiment. (a, b) The density matrices deduced from tomography of the biphoton state at the beginning and the end of the experiment. The imaginary part of density matrices are negligible. The two deduced density matrices have a mutual fidelity of over 99.1% and fidelity with $|\Phi^+\rangle$ of at least 97.5%. (c) Counting rates and probability recorded at integer minutes in multiples of 10. Integration time 10 s. Cerulean: projecting on $|\psi\rangle = (|i_1\rangle + |i_2\rangle)/\sqrt{2}$, corresponding to unit probability. Shamrock: projecting on $|A_3\rangle$. Marigold: projecting on $|A_{16}\rangle$.

Following the approach of [20], the four terms of ε were calculated by experimentally observable probabilities:

$$P_{|\psi\rangle}(v_i = 0) = 1 - P_{|\psi\rangle}(v_i = 1), \tag{B2a}$$

$$P_{|\psi\rangle}(v_i = 1, v_j = 1) = P_{|\psi\rangle}(v_i = 1) \quad P_{|v_i\rangle}(v_j = 1), \tag{B2b}$$

$$P_{|\psi\rangle}(v_i = 0, v_j = 1) = P_{|\psi\rangle}(v_i = 0) \quad P_{|v_i^\perp\rangle}(v_j = 1), \tag{B2c}$$

$$P_{|\psi\rangle}(v_i = 1, v_j = 0) = P_{|\psi\rangle}(v_i = 1) - P_{|\psi\rangle}(v_i = 1, v_j = 1), \tag{B2d}$$

$$P_{|\psi\rangle}(v_i = 0, v_j = 0) = P_{|\psi\rangle}(v_i = 0) - P_{|\psi\rangle}(v_i = 0, v_j = 1), \tag{B2e}$$

and we need to check that

$$\varepsilon(\_, 0|\_, v_j) \approx \varepsilon(0, \_|v_i, \_) = 0, \tag{B3a}$$

$$\varepsilon(\_, 1|\_, v_j) \approx \varepsilon(1, \_|v_i, \_) = 0. \tag{B3b}$$

The key point to evaluate (B2a)–(B2e) is to obtain $P_{|\psi\rangle}(v_i = 0, v_j = 1)$, which requires preparation of $|v_i^\perp\rangle$, namely, the state obtained after a projective measurement of v_i with outcome 0 on a system initially prepared in state $|\psi\rangle$. For a path-encoded photonic qudit, $|v_i^\perp\rangle$ can be prepared by the blocking method, see [20], for example. This method would not work for an OAM qudit. On the contrary, the task of preparing a known state is not harder than scripting a hologram for SLM, so in our experiment we prepare *ab initio* the states needed in (B2c).

The Gram-Schmidt orthonormalization gives

$$|v_i^\perp\rangle = \frac{|\widetilde{v_i^\perp}\rangle}{||\widetilde{v_i^\perp}\rangle||} = \frac{|\psi\rangle - \langle v_i|\psi\rangle|v_i\rangle}{||\psi\rangle - \langle v_i|\psi\rangle|v_i\rangle||} = \frac{|\psi\rangle - \langle v_i|\psi\rangle|v_i\rangle}{\sqrt{1 - |\langle v_i|\psi\rangle|^2}}, \tag{B4}$$

and here we use the notation $|\widetilde{v_i^\perp}\rangle = |\psi\rangle - \langle v_i|\psi\rangle|v_i\rangle$. Essentially,

$$\begin{aligned} P_{|\psi\rangle}(v_i = 0, v_j = 1) &= P_{|\psi\rangle}(v_i = 0) \quad P_{|v_i^\perp\rangle}(v_j = 1) \\ &= (1 - |\langle v_i|\psi\rangle|^2) \left| \langle v_j| \frac{|\psi\rangle - \langle v_i|\psi\rangle|v_i\rangle}{\sqrt{1 - |\langle v_i|\psi\rangle|^2}} \right|^2 \\ &= |\langle v_j|\widetilde{v_i^\perp}|^2 = P_{|\widetilde{v_i^\perp}\rangle}(v_j = 1). \end{aligned} \tag{B5}$$

Notice that the values of $\langle v_i|\psi\rangle$ were already measured, as they had once occurred in the first term of I_{QC}. We can thereby script the holograms to prepare $|\widetilde{v_i^\perp}\rangle$ by virtue of these experiment data instead of deducing $\langle v_i|\psi\rangle$ from the raw settings. Measurement of other expectation values in (B2a)–(B2e) are straightforward.

From (B2d), (B2e), $\varepsilon(0, \_|v_i, \_)$ and $\varepsilon(1, \_|v_i, \_)$ automatically vanish and only the $\varepsilon(\_, 0|\_, v_j)$ and $\varepsilon(\_, 1|\_, v_j)$ terms survive. Nonetheless, the four statistical indices are unanimously calculated. The result of nonsignaling condition verification is shown in Fig. 6. Not all $\varepsilon(\_, 0|\_, v_j)$ and $(\_, 1|\_, v_j)$ perfectly vanish, and a few error bars, which were calculated from Poisson distribution, were not guaranteed to hit zero. Still, the overall small absolute value of εs suggested that the error was due to imperfections in state preparation and measurement and the no-signaling condition held.

4. Testing system stability

The stability of the setup plays an important role in the experiment discussed in this work. As the system may be subject to drifts, one is obliged to confirm the manifesting nonlocality and contextuality origin from the same system rather than being the result of an average over time. For the nonlocality experiment, we check the biphoton density matrices at the beginning and the end of the experiment, to check for the consistency of the prepared biphoton state,

and for the contextuality we sample over some detecting probabilities to see if they stay the same during the whole period of the experiment.

Stability of entanglement source. Figures 7(a) and 7(b) show the real part of the density matrices, measured at the beginning and the end of the Bell nonlocality experiment. They are deduced from 16-settings state tomography, and in both results the reconstructed imaginary parts are negligible. The density matrices reasonably resemble $|\Phi^+\rangle$, with a fidelity of 0.980 and 0.984, respectively. The fidelity between the two density matrices is as high as 0.991, suggesting two almost identical input states.

Stability of contextuality setup. Figure 7(c) provides the counting rates and detection probabilities of three characteristic measurements, namely, onto the original state $|\psi\rangle$ and the projecting bases $|A_3\rangle$ and $|A_{13}\rangle$. The counting rates are recorded every 10 min for 7 times in total to cover the time span of the whole experiment and reflect the stability of the setup. The counting rates vary in a small range of $\leqslant 3\%$, without exhibiting a noticeable trend of drifting. In conclusion, the measured system stays the same during the experiment period and the observed phenomena cannot be considered a time-averaging effect but instead are compelling evidence of nonlocality and contextuality.

[1] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).

[2] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).

[3] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).

[4] R. Rabelo, C. Duarte, A. J. López-Tarrida, M. T. Cunha and A. Cabello, J. Phys. A: Math. Theor. **47**, 424021 (2014).

[5] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).

[6] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[7] H. Buhrman, R. Cleve, S. Massar, and R. de Wolf, Rev. Mod. Phys. **82**, 665 (2010).

[8] S. Pironio, A. Acín, S. Massar, A. Boyer de la Giroday, D. N. Matsukevich, P. Maunz, S. Olmschenk, D. Hayes, L. Luo, T. A. Manning, and C. Monroe, Nature (London) **464**, 1021 (2010).

[9] P. Kurzyński, A. Cabello, and D. Kaszlikowski, Phys. Rev. Lett. **112**, 100401 (2014).

[10] X. Zhan, X. Zhang, J. Li, Y. S. Zhang, B. C. Sanders, and P. Xue, Phys. Rev. Lett. **116**, 090401 (2016).

[11] T. Li, X. Zhang, Q. Zeng, B. Wang, and X. D. Zhang, Opt. Express **26**, 11959 (2018).

[12] A. Cabello, S. Severini, and A. Winter, Phys. Rev. Lett. **112**, 040401 (2014).

[13] Two events are *mutually exclusive* if there is a measurement M such that each event corresponds to a different outcome of M. The relations of exclusivity between events can be represented in a graph in which a vertex represents an event and an edge indicates mutual exclusivity. This graph is a *graph of exclusivity* [12]. The relations of compatibility between the measurements in a Bell or Kochen-Specker (KS) scenario S can be represented in a graph in which a vertex represents a measurement and an edge indicates compatibility. This graph is called the *graph of compatibility* of S. A *context* in a Bell or KS scenario is a subset of compatible and mutually nondisturbing measurements in S. Contexts correspond to cliques in the graph of compatibility of S. A clique is a set of vertices such that every two distinct vertices are adjacent.

[14] B. Yan, Phys. Rev. Lett. **110**, 260406 (2013).

[15] N. Brunner and N. Gisin, Phys. Lett. A **372**, 3162 (2008).

[16] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969). The CHSH inequality (i.e., the I_{2222}-type inequality) is the simplest Bell inequality; however, one will obtain $\Delta = 0$ by a similar computation. Thus, the CHSH inequality is not a suitable candidate to demonstrate QC beyond BN.

[17] K. F. Pál and T. Vértesi, Phys. Rev. A **82**, 022116 (2010).

[18] L. Lovász, IEEE Trans. Inf. Theory **25**, 1 (1979).

[19] A. Cabello, Phys. Rev. A **93**, 032102 (2016).

[20] Y. Xiao *et al.*, Opt. Express **26**, 32 (2018).

[21] Q. Li *et al.*, Optica **6**, 67 (2019).

[22] E. Bolduc, N. Bent, E. Santamato, E. Karimi, and R. W. Boyd, Opt. Lett. **38**, 3546 (2013).

[23] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, Phys. Rev. A **45**, 8185 (1992).

[24] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).