Rapid Communications

Experimental entropic test of state-independent contextuality via single photons

Dengke Qu,[1,2] Paweł Kurzyński,[3] Dagomir Kaszlikowski,[4,5] Sadegh Raeisi,[6] Lei Xiao,[1] Kunkun Wang,[1] Xiang Zhan,[7] and Peng Xue[1,*]

[1]*Beijing Computational Science Research Center, Beijing 100084, China*
[2]*Department of Physics, Southeast University, Nanjing 211189, China*
[3]*Faculty of Physics, Adam Mickiewicz University, Umultowska 85, 61-614 Poznań, Poland*
[4]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore*
[5]*Department of Physics, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore*
[6]*Department of Physics, Sharif University of Technology, Tehran, Iran*
[7]*School of Science, Nanjing University of Science and Technology, Nanjing 210094, China*

(Received 13 December 2019; accepted 18 May 2020; published 1 June 2020)

Recently, an inequality satisfied by noncontextual hidden-variable models and violated by quantum mechanics for all states of a four-level system has been derived based on an information-theoretic distance approach to nonclassical correlations. In this Rapid Communication, we experimentally demonstrate the violation of this inequality with single photons. Our experiment offers a method to study a distinction between quantum and classical correlations from an information-theoretic perspective.

DOI: 10.1103/PhysRevA.101.060101

Quantum theory [1–6] outperforms classical ones for certain communication [7] and computational tasks [8]. Classical and quantum information processing scenarios differ on a fundamental level [9]. If one performs a test X with outcomes $\{x\}$, the information content of the test's statistics can be quantified via Shannon entropy $H(X) = -\sum_x P(X = x)\log_2 P(X = x)$, regardless of whether the tested system was classical or quantum. In order to detect nonclassicality in such tests, one should understand what the typical classical properties, revealed by Shannon entropies, are and how quantum ones differ. It is known that quantum correlations are stronger than classical ones, but when it comes to entropies of quantum tests, the differences are much more sophisticated. For example, nonclassical features of entropies coming from quantum experiments can be amplified if one postprocesses measured data, e.g., mixes nonclassical and classical probability distributions [10].

Contextuality is one of the major differences between quantum and classical physics [3]: It states that measurement results of some physical property may depend on how this property is measured. Tests of contextuality focused mainly on the probability distribution of measurement results. Recently, the entropic tests of quantum contextuality were introduced [11–13] and further investigated experimentally [14]. However, these tests are state dependent, i.e., a departure from classical behavior can be detected only if the system is prepared in some special state. This paradigm was changed in Ref. [15] where the entropic approach to the state-independent contextuality was proposed, allowing nonclassical features to be observed for any state of the system. This is done with the help of a recently discovered multipartite information-theoretic distance for binary measurements. This distance measure uses Shannon entropy and yields a state-independent, multipartite noncontextual inequality that resembles the correlation-based inequality [5,16–18]. In this Rapid Communication, we demonstrate its violations with single photons.

The advantage of using entropies instead of linear probabilistic inequalities is the fundamental interpretation of what violations mean. If you violate a linear Bell-like inequality you conclude that there are no noncontextual hidden variables. Violations of entropic inequalities also tell you that there are no noncontextual hidden variables but they also indicate that there is an information deficit between contextual theories and noncontextual ones. One can tell from our experimental data that contextual theories such as quantum mechanics contain "negative" information. Of course at this stage of our research we do not know what this negative information really means, but this is the stepping stone towards an experimental data analysis using classical algorithmic entropies that do not use probability theory (Shannon entropies are upper bounds on algorithmic entropies). Classical algorithmic entropy is one of the most fundamental descriptions of information and any test of quantumness that uses them is the most primitive test there is.

Now we present an entropic version of the state-independent contextuality proof commonly known as the Peres-Mermin square [4,19,20]. It is derived for a four-level system that can be represented as a composition of two qubits that are in the same place, or even encoded on the same system. Therefore, nonlocality is of no importance in this case as it is not in a nonlocal Bell scenario. There are nine binary ± 1 observables that can be measured on this system. Based on the compatibility relations, we perform these different measurements in the following triples,

$$\{A, a, \alpha\}, \{B, b, \beta\}, \{C, c, \gamma\}, \{A, B, C\}, \{a, b, c\}, \{\alpha, \beta, \gamma\}, \tag{1}$$

*gnep.eux@gmail.com

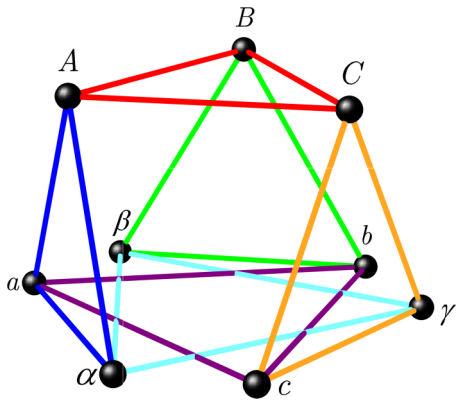

FIG. 1. Illustration of compatibility relations among the nine binary-outcome observables. Compatible observables are connected by links of the same color.

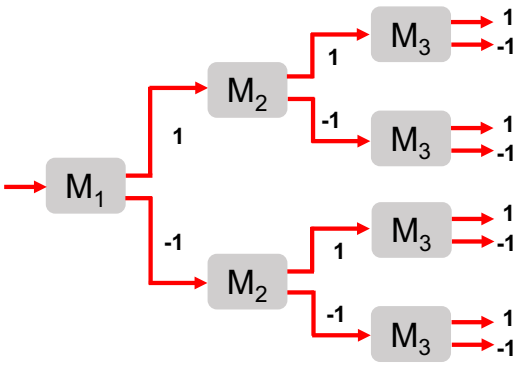

FIG. 2. Devices for measuring six sets of nine observables to test inequality (4).

as shown in Fig. 1. The classical reasoning based on the noncontextuality hypothesis implies that we have the distribution $\prod_{i=1}^{6} q_i = 1$ for measured products

$$q_1 = Aa\alpha, \quad q_2 = Bb\beta, \quad q_3 = Cc\gamma,$$
$$q_4 = ABC, \quad q_5 = abc, \quad q_6 = \alpha\beta\gamma. \tag{2}$$

In quantum theory, we take

$$A = X \otimes \mathbb{1}, \quad a = \mathbb{1} \otimes X, \quad \alpha = X \otimes X,$$
$$B = \mathbb{1} \otimes Y, \quad b = Y \otimes \mathbb{1}, \quad \beta = Y \otimes Y,$$
$$C = X \otimes Y, \quad c = Y \otimes X, \quad \gamma = Z \otimes Z, \tag{3}$$

where X, Y, and Z are Pauli operators. In quantum theory $\prod_{i=1}^{6} q_i = -1$ as $q_1 = \cdots = q_5 = 1$ and $q_6 = -1$ for any quantum state. On the other hand, in any noncontextual realistic theory (NRT), i.e., a theory where the outcomes of A, B, etc., are predetermined (realism) and do not depend on the context in which they are measured (noncontextuality), one has $\prod_{i=1}^{6} q_i = 1$. This is because each such an outcome appears exactly twice in two different products q_i and q_j. Interestingly, in both NRT and quantum theory, we have $H(q_i) = 0$, since the products are well defined. In NRT, we have $P(q_i = 1) = 1$ for $i = 1, \ldots, 6$, whereas, in quantum theory, we have $P(q_i = 1) = 1$ for $i = 1, \ldots, 5$ and $P(q_6 = -1) = 1$. The products q_i are well defined in Eq. (2). Thus in both cases the entropy $H(q_i) = 0$ is satisfied.

The entropic noncontextual inequality derived in Ref. [15] reads

$$H(\alpha\beta\gamma) \leqslant H(Aa\alpha) + H(Bb\beta) + H(ABC) + H(abc) + H(Cc\gamma), \tag{4}$$

where

$$H(X_iX_jX_k) = -P(x_ix_jx_k = -1)\log_2 P(x_ix_jx_k = -1) - P(x_ix_jx_k = 1)\log_2 P(x_ix_jx_k = 1) \tag{5}$$

is the Shannon entropy of the probability distribution associated with the measurements X_i, X_j, and X_k and the corresponding outcomes x_i, x_j, and x_k. The distributions of outcomes in both quantum and NRT do not violate the inequality. However, equal mixing of the quantum and NRT distributions maximally violates the inequality for *any* quantum states.

We now demonstrate the experimental test of the entropic state-independent noncontextuality inequality. The purpose of the experiment is to test different quantum states of a single-particle system. In the experiment we use a single photon that simulates two qubits encoded in different degrees of freedom—polarization and propagation modes. The basis for our two qubits is encoded as $\{|0\rangle = |UH\rangle, |1\rangle = |UV\rangle, |2\rangle = |DH\rangle, |3\rangle = |DV\rangle\}$, where U (D) denotes the upper (lower) spatial mode of a single photon and H (V) denotes their horizontal (vertical) polarizations. The photon pairs are generated in the spontaneous parametric down-conversion process, where one of the photons is a *trigger* heralding an arrival of the signal photon that we use to test the inequality. A polarization beam splitter (PBS), a beam displacer (BD), and half-wave plates (HWPs) are used to prepare the photonic four-level system in 27 different quantum states ready for testing.

Sequential measurements of three compatible observables on the same photon are shown in Fig. 2. M_i $(i = 1, 2, 3)$ describes the setup for measuring one of the nine observables. After the preparation stage, the photons enter the device M_1 through the input and yield one of the two possible outcomes. Next, the photon enters devices M_2, then M_3, and finally it is detected at one of the eight outputs.

The experimental setup is shown in Fig. 3. Observables $a = \mathbb{1} \otimes X$ and $B = \mathbb{1} \otimes Y$ are simply rotations on the polarizations of photons keeping the spatial mode unchanged. Observables X and Y can be written as $M = \sum_{i=H,V} m_i |m_i\rangle\langle m_i|$, where $|m_i\rangle$ is an eigenstate of M and m_i is the corresponding eigenvalue. A polarization rotation is defined $U_M = |H\rangle\langle m_H| + |V\rangle\langle m_V|$ and is applied on the polarization of the photons, which can be implemented by wave plates at certain setting angles following by a polarizing beam splitter (PBS). The overlap between the initial state and $|m_i\rangle$ is equal to the probability of the photons being measured in the basis state $|i\rangle \in \{|H\rangle, |V\rangle\}$.

Measurements $A = X \otimes \mathbb{1}$ and $b = Y \otimes \mathbb{1}$ are performed only on the spatial modes. First, we use beam displacers (BDs) to split and then combine the photons with certain polarizations into the same spatial mode, which amounts to a basis transformation between spatial and polarization modes. Polarization rotations are done via wave plates followed by

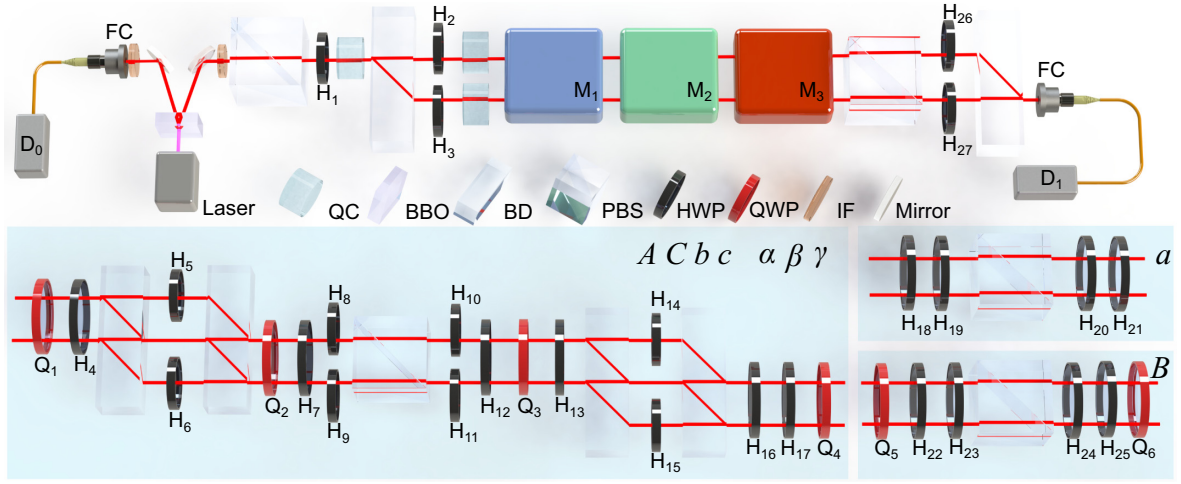

FIG. 3. Experimental setup. The heralded single photons are created via type-I spontaneous parametric down-conversion in a β-barium-borate (BBO) nonlinear crystal and are injected into the optical network (see figure for acronyms). The first polarizing beam splitter (PBS), half-wave plates (HWPs), and beam displacer (BD) are used to generate the pure qudit states. To prepare mixed states [21,22], we use a quartz crystal (QC) with a thickness (longer than coherence length of photons) placed in front of the first BD to reduce the coherence of the photons in different spatial modes. After the first BD, the photons are separated into the upper and lower path depending on their polarizations. Then a QC is inserted into each path and reduces the coherence of the photons with different polarizations. The coherence length of the photons is $L_c \approx \lambda^2/\Delta\lambda$, where λ is the central wavelength of the source and $\Delta\lambda$ is the spectral width of the source. Hence the thickness of the QC should be at least 23.97 mm. In our experiment, it is about 28.77 mm. The measurements are realized by wave plates and BDs. The photons are detected by APDs. The measurements A, C, b, c, α, β, and γ can be realized by the setup involving a PBS, four BDs, and several wave plates, whereas the setups for realizing a and B can be simplified.

a PBS as mentioned above. The measurements of the other observables α, β, C, c, and γ, which are the products of two Pauli operators, are implemented by a polarization rotation, a basis transformation between spatial and polarization modes, and another polarization rotation followed by a projective measurement in the $\{|H\rangle, |V\rangle\}$ basis.

We need to apply sequential measurements of three compatible observables on the same photon. Before the next measurement is done, we need to bring back the eigenstates of the previously measured observable. For six sets of measurements shown in (1), each set has eight different outcome distributions $\{1,1,1\}$, $\{1,1,-1\}$, $\{1,-1,1\}$, $\{1,-1,-1\}$, $\{-1,1,1\}$, $\{-1,1,-1\}$, $\{-1,-1,1\}$ and $\{-1,-1,-1\}$. With a proper choice of devices and the wave plates' angles (please find the angles in the Supplemental Material [23]), we can implement six sets of measurements with eight different outcomes. Finally, photons are detected by single-photon avalanche photodiodes (APDs). We only register coincidences between APD (D_1) and the trigger APD (D_0). For each outcome distribution of each measurement, we recorded clicks for 2 s, and registered about 20 000 single photons. The probability for more than one photon pair is less than 10^{-4}

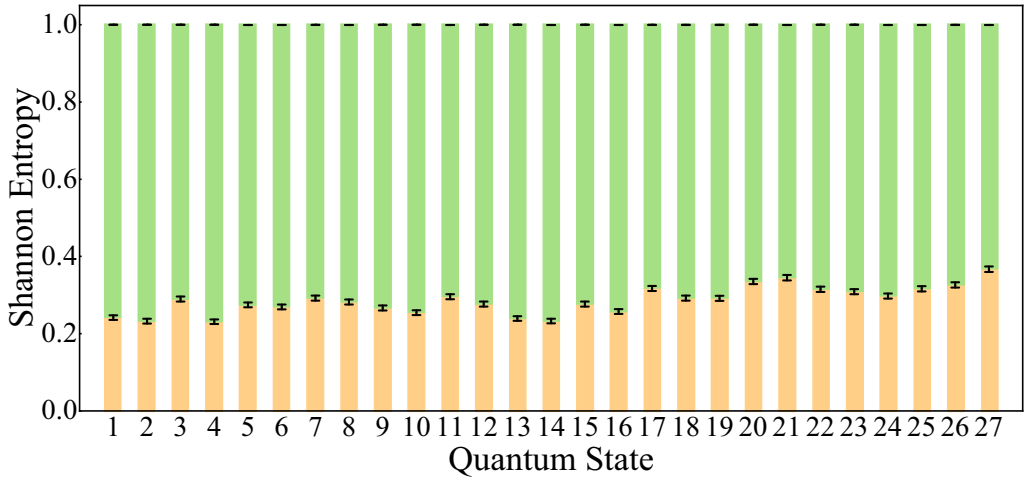

FIG. 4. State-independent violation of the entropic inequality (4). The inequality is tested for 27 different quantum states. The left-hand and right-hand sides of the inequality are shown in orange (short) and green (tall) bars, respectively.. Error bars indicate the statical uncertainty which is obtained based on assuming Poissonian statistics.

TABLE I. Experimental results of each term of the entropic inequality for 27 different states being tested. Here, SD is an abbreviation for standard deviation. Error bars indicate the statistical uncertainty which is obtained based on assuming Poissonian statistics. The last column shows the distance $\sum_i^9 (p_i - p_i')^2$, where p_i is the measured probability of the observable measured in one context and p_i' is the measured probability of the same observable measured in the other context.

State	$H(Aa\alpha)$	$H(Bb\beta)$	$H(Cc\gamma)$	$H(ABC)$	$H(abc)$	$H(\alpha\beta\gamma)$	SD	$\sum_{i=1}^9 (p_i - p_i')^2$
$\|\Psi_1\rangle = \|0\rangle$	0.03441(243)	0.04503(267)	0.05584(290)	0.04615(271)	0.06022(298)	0.99988(1)	124	0.00023(15)
$\|\Psi_2\rangle = \|1\rangle$	0.04155(259)	0.04921(277)	0.04537(269)	0.03879(253)	0.05758(292)	0.99993(1)	127	0.00058(24)
$\|\Psi_3\rangle = \|2\rangle$	0.05467(287)	0.06471(304)	0.04739(272)	0.06575(307)	0.05731(292)	0.99991(1)	108	0.00043(19)
$\|\Psi_4\rangle = \|3\rangle$	0.03834(253)	0.05590(291)	0.04805(272)	0.05295(285)	0.03568(244)	0.99990(1)	128	0.00035(18)
$\|\Psi_5\rangle = (\|0\rangle + \|1\rangle)/\sqrt{2}$	0.05146(281)	0.06297(303)	0.05843(295)	0.04875(275)	0.05298(285)	0.99977(2)	113	0.00209(46)
$\|\Psi_6\rangle = (\|0\rangle + \|2\rangle)/\sqrt{2}$	0.03811(252)	0.06311(305)	0.05173(282)	0.06159(299)	0.05473(286)	0.99974(3)	115	0.00230(48)
$\|\Psi_7\rangle = (\|0\rangle + i\|1\rangle)/\sqrt{2}$	0.05566(286)	0.05792(289)	0.05925(297)	0.05766(295)	0.06177(300)	0.99983(2)	108	0.00447(66)
$\|\Psi_8\rangle = (\|0\rangle + i\|2\rangle)/\sqrt{2}$	0.05871(292)	0.04536(260)	0.07411(314)	0.05590(295)	0.04792(276)	0.99970(3)	111	0.00161(39)
$\|\Psi_9\rangle = (\|1\rangle + \|3\rangle)/\sqrt{2}$	0.05254(281)	0.05410(287)	0.06923(310)	0.03832(253)	0.05248(284)	0.99990(1)	116	0.00217(47)
$\|\Psi_{10}\rangle = (\|2\rangle + \|3\rangle)/\sqrt{2}$	0.04617(271)	0.04837(276)	0.05828(293)	0.05866(294)	0.04302(263)	0.99984(2)	119	0.00346(59)
$\|\Psi_{11}\rangle = (\|1\rangle + i\|3\rangle)/\sqrt{2}$	0.04520(267)	0.05823(298)	0.07497(318)	0.05412(288)	0.06347(301)	0.99976(2)	107	0.00349(59)
$\|\Psi_{12}\rangle = (\|2\rangle + i\|3\rangle)/\sqrt{2}$	0.04606(271)	0.05062(286)	0.07166(316)	0.05692(296)	0.05170(282)	0.99985(2)	111	0.00451(66)
$\|\Psi_{13}\rangle =$ $(\|0\rangle + \|1\rangle + \|2\rangle + \|3\rangle)/2$	0.05151(282)	0.05975(298)	0.05688(290)	0.03355(242)	0.03762(251)	0.99984(2)	124	0.00023(14)
$\|\Psi_{14}\rangle =$ $(\|0\rangle + i\|1\rangle + i\|2\rangle - \|3\rangle)/2$	0.03465(241)	0.05502(283)	0.04541(271)	0.05303(281)	0.04487(269)	0.99969(3)	127	0.00111(29)
$\|\Psi_{15}\rangle =$ $(\|0\rangle + \|1\rangle + i\|2\rangle + i\|3\rangle)/2$	0.02958(229)	0.05632(290)	0.08011(330)	0.06789(309)	0.06646(309)	0.99974(3)	106	0.00059(23)
$\|\Psi_{16}\rangle =$ $(\|0\rangle + i\|1\rangle + \|2\rangle + i\|3\rangle)/2$	0.04131(256)	0.03732(249)	0.04559(270)	0.05769(297)	0.07529(322)	0.99973(3)	119	0.00159(35)
$\rho_{17} = (\|0\rangle\langle 0\| + \|1\rangle\langle 1\|)/2$	0.05109(280)	0.05873(295)	0.06309(302)	0.08482(331)	0.05938(294)	0.99983(2)	101	0.00205(45)
$\rho_{18} = (\|0\rangle\langle 0\| + \|2\rangle\langle 2\|)/2$	0.05583(289)	0.04845(274)	0.05832(294)	0.06057(298)	0.06926(311)	0.99981(2)	108	0.00138(37)
$\rho_{19} = (\|0\rangle\langle 0\| + \|3\rangle\langle 3\|)/2$	0.06085(298)	0.05591(289)	0.05506(290)	0.05101(281)	0.06898(312)	0.99982(2)	108	0.00035(19)
$\rho_{20} = (\|1\rangle\langle 1\| + \|2\rangle\langle 2\|)/2$	0.06748(311)	0.06960(312)	0.05140(281)	0.07073(315)	0.07581(322)	0.99987(2)	96	0.00041(20)
$\rho_{21} = (\|1\rangle\langle 1\| + \|3\rangle\langle 3\|)/2$	0.05623(291)	0.06646(308)	0.07303(317)	0.06852(310)	0.08076(329)	0.99978(2)	94	0.00145(38)
$\rho_{22} = (\|2\rangle\langle 2\| + \|3\rangle\langle 3\|)/2$	0.06200(302)	0.06095(297)	0.07231(315)	0.06083(299)	0.05911(295)	0.99985(2)	101	0.00203(45)
$\rho_{23} =$ $(\|0\rangle\langle 0\| + \|1\rangle\langle 1\| + \|2\rangle\langle 2\|)/3$	0.05917(295)	0.06265(301)	0.05262(283)	0.07067(313)	0.06372(301)	0.99984(2)	103	0.00108(33)
$\rho_{24} =$ $(\|0\rangle\langle 0\| + \|1\rangle\langle 1\| + \|3\rangle\langle 3\|)/3$	0.04312(263)	0.05509(288)	0.06152(298)	0.08014(329)	0.05773(292)	0.99979(2)	107	0.00112(33)
$\rho_{25} =$ $(\|0\rangle\langle 0\| + \|2\rangle\langle 2\| + \|3\rangle\langle 3\|)/3$	0.05330(285)	0.06296(300)	0.07202(316)	0.06612(307)	0.06202(301)	0.99976(3)	101	0.00082(29)
$\rho_{26} =$ $(\|1\rangle\langle 1\| + \|2\rangle\langle 2\| + \|3\rangle\langle 3\|)/3$	0.05467(287)	0.06193(301)	0.08139(329)	0.05878(296)	0.06957(314)	0.99981(2)	99	0.00069(26)
$\rho_{27} = (\|0\rangle\langle 0\| + \|1\rangle\langle 1\| +$ $\|2\rangle\langle 2\| + \|3\rangle\langle 3\|)/4$	0.07377(321)	0.06236(301)	0.07873(325)	0.07640(324)	0.07578(318)	0.99974(3)	89	0.00068(26)

and thus it can be neglected. The coincidence counts are used to calculate the measured probabilities of eight outcomes of the each set of measurements.

In principle, entropic inequalities only provide a necessary but not sufficient criterion for noncontextuality [10]. However, entropic inequalities turn also to be sufficient, since any contextual probabilistic model will display entropic violations if properly mixed with a classical model. To find such a noncontextual and realistic distribution, we first define

$$\alpha = Aa, \quad \beta = Bb, \quad C = AB, \quad c = ab, \quad \gamma = ABab. \tag{6}$$

Under this definition, we have the classical distribution satisfying $q_1' = \cdots = q_6' = 1$ which is analogous to (2), but takes into account the definition (6). With a similar experimental setup, we can implement four sets of measurements $\{A, a\}$, $\{B, b\}$, $\{A, B\}$, and $\{a, b\}$ and obtain the classical distribution. By equally mixing the quantum and classical distributions, one obtains a distribution $\tilde{q}_i = \frac{1}{2}(q_i + q_i')$. Note that quantum theory predicts

$$H(\tilde{q}_1) = \cdots = H(\tilde{q}_5) = 0, \quad H(\tilde{q}_6) = 1, \tag{7}$$

therefore for this mixed distribution the entropic inequality is maximally violated ($1 \leqslant 0$).

The experimental results are shown in Fig. 4. We repeat the experiment on 27 qudit states including 16 pure states as a tomographically complete set of qudit states and 11 mixed states. The reason why we choose mixed states to test the inequality is that experimentally it is not easy to generate pure states with perfect purity. We test mixed states to make a better case for the state-independent property. The results in Table I show that a state-independent violation of the entropic noncontextuality inequality (4) occurs by 89 standard deviations (at least). Due to the imperfections in the

experiment there is a bit of difference between experimental results and theoretical predictions. The differences between the experimental data and theoretical predictions are caused by the fluctuations of photon numbers, the accuracy of wave plates, and the dephasing introduced via the misalignment of BDs, two of which form a Mach-Zehnder interferometer. We show the error analysis in the Supplemental Material [23]. On the other hand, the derivation of entropy near probability 0 or 1 is infinite, which makes the entropy extremely sensitive to the noise of probability when it should be 0 theoretically. There is also a bit of difference between the score for different states (pure states and mixed states). That is because compared to pure states, the generation of desired mixed states is a little complicated and the unideal fidelities of the states influence the experimental results as well.

In our experiment photon loss opens up a detection efficiency loophole. Thus, a fair-sampling assumption is taken here which assumes that the event selected out by the photonic coincidence is an unbiased representation of the whole sample [14,24–26].

The compatibility of measurements in contexts is an important issue in the experimental test of contextuality. We calculate the distance $\sum_{i=1}^{9}(p_i - p_i')^2$, where p_i is the measured probability of the observable measured in one context and p_i' is the measured probability of the same observable measured in the other context. As shown in Table I (last column), the distances for all the states being tested are small enough (<0.005), which indicates (almost) perfectly compatible measurements are realized in our experiment.

In summary, we experimentally demonstrate an entropic test of state-independent contextuality on a single photonic four-level system. We show that 27 different single photonic states violate an entropic inequality which involves correlations between results of sequential compatible measurements by at least 89 standard deviations. Our results show that, even for a single system, and independent of its state, there is a universal set of tests whose results do not admit a noncontextual interpretation.

This work has been supported by the Natural Science Foundation of China (Grants No. 11674056 and No. U1930402) and the startup fund from Beijing Computational Science Research Centre. D.K. is supported by the National Research Foundation, Prime Minister's Office, Singapore and the Ministry of Education, Singapore under the Research Centres of Excellence program. P.K. is supported by the National Science Centre in Poland (NCN Project No. 2016/23/G/ST2/04273). S.R. is supported by the research grant system of Sharif University of Technology (G960219).

[1] E. Specker, Dialectica **14**, 239 (1960).
[2] J. S. Bell, Rev. Mod. Phys. **38**, 447 (1966).
[3] S. Kochen and E. P. Specker, J. Math. Mech. **17**, 59 (1967).
[4] N. D. Mermin, Rev. Mod. Phys. **65**, 803 (1993).
[5] A. Cabello, Phys. Rev. Lett. **101**, 210401 (2008).
[6] P. Badziag, I. Bengtsson, A. Cabello, and I. Pitowsky, Phys. Rev. Lett. **103**, 050401 (2009).
[7] C. H. Bennett and G. Brassard, Theor. Comput. Sci. **560**, 7 (2014).
[8] P. W. Shor, SIAM J. Comput. **26**, 1484 (1997).
[9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2010).
[10] R. Chaves, Phys. Rev. A **87**, 022102 (2013).
[11] P. Kurzyński, R. Ramanathan, and D. Kaszlikowski, Phys. Rev. Lett. **109**, 020404 (2012).
[12] R. Chaves and T. Fritz, Phys. Rev. A **85**, 032113 (2012).
[13] T. Fritz and R. Chaves, IEEE Trans. Inf. Theory **59**, 803 (2012).
[14] X. Zhan, P. Kurzyński, D. Kaszlikowski, K. Wang, Z. Bian, Y. Zhang, and P. Xue, Phys. Rev. Lett. **119**, 220403 (2017).
[15] S. Raeisi, P. Kurzyński, and D. Kaszlikowski, Phys. Rev. Lett. **114**, 200401 (2015).
[16] E. Amselem, M. Rådmark, M. Bourennane, and A. Cabello, Phys. Rev. Lett. **103**, 160405 (2009).
[17] G. Kirchmair, F. Zähringer, R. Gerritsma, M. Kleinmann, O. Gühne, A. Cabello, R. Blatt, and C. F. Roos, Nature (London) **460**, 494 (2009).
[18] C. Zu, Y.-X. Wang, D.-L. Deng, X.-Y. Chang, K. Liu, P.-Y. Hou, H.-X. Yang, and L.-M. Duan, Phys. Rev. Lett. **109**, 150401 (2012).
[19] A. Peres, Phys. Lett. A **151**, 107 (1990).
[20] N. D. Mermin, Phys. Rev. Lett. **65**, 3373 (1990).
[21] K. Wang, G. C. Knee, X. Zhan, Z. Bian, J. Li, and P. Xue, Phys. Rev. A **95**, 032122 (2017).
[22] K. Wang, X. Wang, X. Zhan, Z. Bian, J. Li, B. C. Sanders, and P. Xue, Phys. Rev. A **97**, 042112 (2018).
[23] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.101.060101 for the details of the configurations of the optical circuits for the measurements and the error analysis.
[24] R. Lapkiewicz, P. Li, C. Schaeff, N. K. Langford, S. Ramelow, M. Wieśniak, and A. Zeilinger, Nature (London) **474**, 490 (2011).
[25] X. Zhan, X. Zhang, J. Li, Y. Zhang, B. C. Sanders, and P. Xue, Phys. Rev. Lett. **116**, 090401 (2016)
[26] X. Zhan, E. G. Cavalcanti, J. Li, Z. Bian, Y. Zhang, H. M. Wiseman, and P. Xue, Optica **4**, 966 (2017).