Experimental investigation of contextual robustness and coherence in state discrimination

Jingyan Lin,[1,2] Kunkun Wang,[3] Lei Xiao,[2] Dengke Qu,[4] Gaoyan Zhu,[5] Yong Zhang,[1,*] and Peng Xue[2,†]
[1]*School of Science, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2]*School of Physics, Southeast University, Nanjing 211189, China*
[3]*School of Physics and Optoelectronic Engineering, Anhui University, Hefei 230601, China*
[4]*Beijing Computational Science Research Center, Beijing 100084, China*
[5]*College of Electronic and Information Engineering, Shangdong University of Science and Technology, Qingdao 266590, China*

(Received 19 September 2023; revised 15 April 2024; accepted 19 April 2024; published 8 May 2024)

Given that contextuality and coherence are significant resources in quantum physics, exploring the intricate interplay between these two factors presents a compelling avenue for research. Here, an experiment is presented to investigate the nuanced relationship between contextual robustness and coherence in a scenario of state discrimination. Specifically, two types of noises—depolarizing and dephasing—are introduced in the measurement procedure to assess the robustness of contextuality. By varying the overlap between the states to be discriminated, we explore the variations in contextual robustness across different levels of coherence. Importantly, our findings demonstrate that contextuality can persist under any degree of coherence. Notably, when coherence approaches zero (while remaining nonzero), contextuality can still endure under arbitrary amounts of partial dephasing. These results underscore the pivotal role of coherence in contextual phenomena, offering significant insights into the intricate interplay between coherence and contextuality.

DOI: 10.1103/PhysRevA.109.052208

I. INTRODUCTION

Distinguishing nonclassical from classical phenomena is instrumental in effectively harnessing quantum resources to achieve superior performance in practical tasks. The concept of noncontextuality, initially proposed by Bell [1], Kochen and Specker [2], and subsequently refined by Spekkens [3], represents a manifestation of classical phenomena. On the contrary, contextuality, manifested through the violation of noncontextuality inequalities, has emerged as a resource that provides quantum advantages in various domains such as cryptography [4,5], communication [6–9], and computation [10–13]. Quantum coherence [14,15] as a valuable resource in quantum information [16–18] is another significant nonclassical phenomenon that piques our interest.

Kochen-Specker noncontextuality [2] originates from the noncontextual hidden variable model. The historical definition of context has confined the exploration of contextuality to Hilbert spaces with dimensions three and above, and to measurements limited to sharp measurements [19–30]. However, Spekkens [3] introduced modifications to the concept of context and noncontextuality, building upon noncontextual ontological models. This innovative approach allows for the extension of contextuality to two-dimensional Hilbert spaces and unsharp measurements, significantly broadening the scope for exploring contextuality in a wider range of quantum systems [31–34].

Given that both contextuality and coherence are quantum resources with the potential to offer quantum advantages in various tasks, exploring the relationship between these two phenomena is inherently intriguing. Furthermore, when considering the preparation contextuality defined by Spekkens, it becomes necessary to account for at least two states, implying the existence of state-independent coherence. Consequently, basis-independent results can be established for both contextuality and coherence [35,36]. Thus, investigating basis-independent relationships between contextuality and coherence represents a compelling avenue of research.

Building upon Spekkens' extension of noncontextuality [3], we focus on the realm of minimum-error state discrimination within two-dimensional Hilbert spaces. In this scenario, it has been theoretically recognized that contextuality provides advantages in the successful discrimination of nonorthogonal states [37]. Subsequently, measurement errors are introduced by considering angular deviations from the ideal basis, leading to experimental proof of the contextual advantage for state discrimination in the presence of errors [38]. Note that the discrimination of nonorthogonal states inherently introduces nonzero coherence under the basis transformation. We further investigate both contextuality and coherence in the presence of experimentally applied depolarizing [37,39,40] and dephasing [41] noise.

Considering that quantum state discrimination serves as a cornerstone in quantum information processing tasks, it has garnered significant attention in both theoretical [42–45] and experimental [46–48] studies. Moreover, this critical aspect finds applications in various quantum cryptography protocols [49] and communication protocols [50]. In this case, the experimental exploration of contextual robustness and

*zhyong98@bupt.edu.cn
†gnep.eux@gmail.com

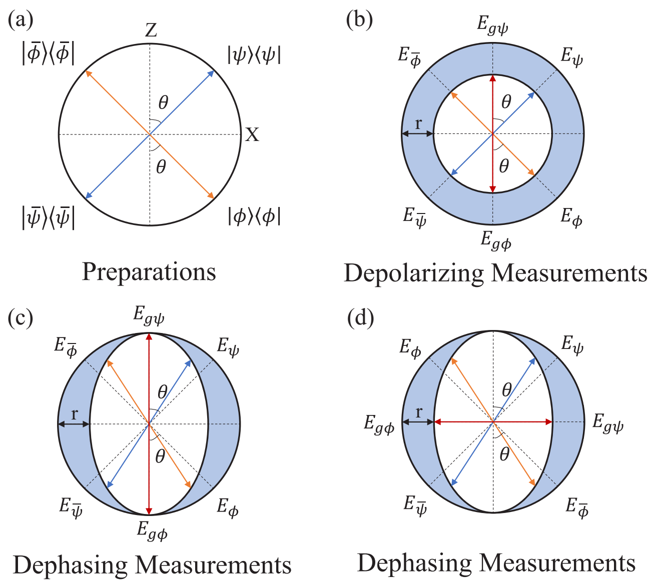

FIG. 1. Preparations and measurements in the zx plane of the Bloch sphere. (a) Preparations without noise. (b) Depolarizing measurements with the Helstrom measurement performed along the z axis. (c) Dephasing measurements with the Helstrom measurement performed along the z axis. (d) Dephasing measurements with the Helstrom measurement performed along the x axis.

coherence in quantum state discrimination assumes substantial significance.

In this work, by considering adjustable noise in measurements and coherence of prepared states within the noncontextuality framework, our experimental findings unveil a nuanced relationship. We derive two conclusions: Firstly, contextuality implies the presence of coherence in the system, while the reverse is not necessarily true under sufficient noise levels. Additionally, even when coherence approaches zero, a system can still exhibit contextuality in the presence of dephasing, even with arbitrarily large noise intensities. These findings contribute valuable insights into the complex interrelationship between contextual robustness and coherence, contributing to an enhanced understanding of quantum state discrimination.

II. CONTEXTUALITY ROBUSTNESS AND COHERENCE IN MINIMUM-ERROR STATE DISCRIMINATION

We focus on the scenario of minimum-error state discrimination (MESD) in two-dimensional Hilbert spaces [37], specifically under depolarizing and dephasing [41]. We consider a system prepared in one of two nonorthogonal states, $|\psi\rangle$ and $|\phi\rangle$, each having an equal prior probability. The measurement $M_g \equiv \{E_{g\psi}, E_{g\phi}\}$ is used to determine the state of the system. The probability of successfully discriminating the system's state s can be quantified by $s \equiv \frac{1}{2}\mathrm{Tr}[E_{g\psi}|\psi\rangle\langle\psi|] + \frac{1}{2}\mathrm{Tr}[E_{g\phi}|\phi\rangle\langle\phi|]$. When choosing M_g to be the Helstrom measurement [51], it implies that the direction of the measurement basis $\{E_{g\psi}, E_{g\phi}\}$ aligns with the direction of the vector $(|\phi\rangle - |\psi\rangle)$ on the Bloch sphere, as depicted in Fig. 1. This particular measurement ensures that the error rate of the two-state discrimination is minimized, fulfilling the requirement of MESD.

TABLE I. The probability of specific measurement outcomes occurring for particular states. Since $\mathcal{D}(\mathbb{I} - E) = \mathbb{I} - \mathcal{D}(E)$, the table is sufficient to obtain all measurement outcomes.

	$\lvert\phi\rangle\langle\phi\rvert$	$\lvert\psi\rangle\langle\psi\rvert$	$\lvert\bar{\phi}\rangle\langle\bar{\phi}\rvert$	$\lvert\bar{\psi}\rangle\langle\bar{\psi}\rvert$
$\mathcal{D}(E_{\phi})$	$1-\epsilon$	c	ϵ	$1-c$
$\mathcal{D}(E_{\psi})$	c	$1-\epsilon$	$1-c$	ϵ
$\mathcal{D}(E_{g\phi})$	s	$1-s$	$1-s$	s

Without loss of generality, we can confine the preparations and measurements of the system states to the zx plane of the Bloch sphere, as illustrated in Fig. 1. The Helstrom measurement satisfies the symmetric relation $\mathrm{Tr}[E_{g\psi}|\psi\rangle\langle\psi|] = \mathrm{Tr}[E_{g\phi}|\phi\rangle\langle\phi|]$, allowing us to calculate s as

$$s = \mathrm{Tr}[E_{g\psi}|\psi\rangle\langle\psi|] = \mathrm{Tr}[E_{g\phi}|\phi\rangle\langle\phi|]. \tag{1}$$

To study contextuality, the orthogonal states of $|\psi\rangle$ and $|\phi\rangle$ are taken into account in Fig. 1 to establish the preparation equivalence

$$\tfrac{1}{2}|\psi\rangle\langle\psi| + \tfrac{1}{2}|\bar{\psi}\rangle\langle\bar{\psi}| = \tfrac{1}{2}|\phi\rangle\langle\phi| + \tfrac{1}{2}|\bar{\phi}\rangle\langle\bar{\phi}| = \frac{\mathbb{I}}{2}. \tag{2}$$

The measurement $\{E_{\psi}, E_{\bar{\psi}}, E_{\phi}, E_{\bar{\phi}}\}$ represents the projective measurements on the corresponding subscript state. The concept of context is defined as the set of features that are not specified by specifying the equivalence class. Therefore, preparation noncontextuality implies that two states belonging to the same preparation equivalence class cannot be distinguished by any quantity. This enables us to formulate the noncontextuality inequalities, which serve as tools for assessing whether the system exhibits contextual or noncontextual behavior. The noncontextuality inequality in MESD is given by [37]

$$s \leqslant 1 - \frac{c}{2}, \tag{3}$$

where $c = \mathrm{Tr}[|\psi\rangle\langle\psi|\,|\phi\rangle\langle\phi|] = |\langle\psi|\phi\rangle|^2$ represents the overlap between $|\psi\rangle$ and $|\phi\rangle$ and indicates the level of confusability between these two states. When the process of state discrimination is influenced by noise introduced through imperfect preparations or measurements, the inequality is relaxed to [37]

$$s \leqslant 1 - \frac{c-\epsilon}{2}, \tag{4}$$

where ϵ denotes the noise intensity. The violation of the inequality is denoted as $\Delta s' = s - 1 + \frac{c-\epsilon}{2}$, which can be utilized to assess the contextuality of the system.

Here, we only consider noise introduced by imperfect measurements with errors, represented by $\mathcal{D}(E)$. The probability of a specific measurement outcome occurring for a particular state is documented in Table I. We first focus on the measurement with depolarizing. As illustrated in Fig. 1(b), the imperfect measurements can be described as

$$\mathcal{D}_r^{\mathrm{depol}}(E) \equiv (1-r)E + \frac{r}{2}\mathbb{I}, \tag{5}$$

where r signifies the intensity of depolarizing, and it is linked to the noise intensity ϵ through the equation $\epsilon = \frac{r}{2}$. The minimum level of noise r preventing the system from violating the noncontextuality inequality can be readily quantified as [41]

$$r_{\min}^{\text{depol}} = 1 - \frac{1}{\sin^2\theta + \cos\theta}, \tag{6}$$

serving as a metric for contextual robustness. It serves as the threshold for r that demarcates the boundary between contextuality and noncontextuality under depolarizing noise. When $r < r_{\min}^{\text{depol}}$, the noncontextuality inequality (4) is violated, revealing the advantage of quantum measurements over noncontextual measurements.

For dephasing, the imperfect measurements can be described as

$$\mathcal{D}_r^{\text{deph}}(E) \equiv (1-r)E + r\sum_{i\in\{0,1\}} \langle i|E|i\rangle |i\rangle\langle i|, \tag{7}$$

where $\{|i\rangle\}_{i\in\{0,1\}}$ represents the basis states along the z axis. Similarly, r signifies the intensity of dephasing, and it is linked to the noise intensity ϵ through the equation $\epsilon = \frac{r}{2}\sin^2\theta$. We consider two measurement schemes shown in Figs. 1(c) and 1(d), respectively. The first scheme is similar to the depolarization case, with the Helstrom measurement performed along the z axis. The contextual robustness can be expressed as [41]

$$r_{\min}^{\text{deph}} = 1 - \frac{1-\cos\theta}{\sin^2\theta}, \tag{8}$$

where $r_{\min}^{\text{depol}}$ serves as the threshold for r that demarcates the boundary between contextuality and noncontextuality under dephasing measurements. In the second scheme, the direction of the Helstrom measurement is rotated from the z axis to the x axis. The contextual robustness can be expressed as [41]

$$r_{\min}^{\text{deph}} = 1 - \sin\theta. \tag{9}$$

When $r < r_{\min}^{\text{deph}}$, the noncontextuality inequality (4) is violated, revealing the advantage of quantum measurements over noncontextual measurements.

As we exclude noise in the preparation procedure illustrated in Fig. 1(a), the l_1 norm of coherence can be quantified as [15]

$$C(\rho) = \min_{\sigma\in I} \|\rho - \sigma\|_1 = \sum_{i\neq j} |\rho_{ij}|. \tag{10}$$

Here, due to the symmetry among the states, coherence of $|\psi\rangle$ and $|\phi\rangle$ satisfies

$$C(|\psi\rangle\langle\psi|) = C(|\phi\rangle\langle\phi|) = \sin\theta, \tag{11}$$

which is governed by $\theta \in [0, \frac{\pi}{2}]$. Hence, manipulating the system state via changes in θ can influence the coherence of states. Noting that it is also valuable to consider the effect of noise on coherence, we present the relevant results in Appendix C by examining effective coherence. Given that coherence serves as a defining characteristic distinguishing quantum behavior from the classical realm, it becomes intriguing to explore contextual robustness under different amounts of coherence.

III. EXPERIMENTAL EXPLORATION OF CONTEXTUALITY ROBUSTNESS AND COHERENCE

We employ a single-photon source [52–54] to prepare the qubit states and measure them under arbitrary intensity of depolarizing and dephasing within an optical system shown in Fig. 2. The qubit is encoded by the polarization state of the photons as $\{|0\rangle \equiv |H\rangle, |1\rangle \equiv |V\rangle\}$, where $|H\rangle$ ($|V\rangle$) denotes the horizontal (vertical) polarization. The initial state $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$ is obtained by passing the photon through HWP1. The key point of our experiment is to achieve a measurement process with adjustable noise intensity, as depicted within the black dashed box in Fig. 2. By manipulating the angle of HWP2, we can achieve the correct or incorrect discrimination results in the middle path of the output. Furthermore, by manipulating the angles of HWP3 and HWP4 between the two beam displacers (BD), measurements can be conducted under varying noise intensities. Details of the experiment are presented in Appendix A.

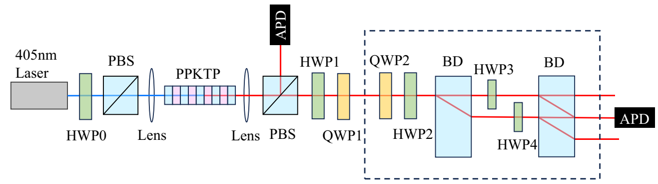

FIG. 2. Experimental setup. Polarized photon pairs, characterized by orthogonal polarization directions, are generated through a periodically poled potassium titanyl phosphate (PPKTP) crystal pumped by a 405-nm laser. Following the second polarization beam splitter's (PBS) division of these photon pairs, the vertically polarized photons are utilized as the trigger and detected by a single photon avalanche diode (APD). Simultaneously, the horizontally polarized photons serve as the signal, encoding the qubit in our experimental setup. Subsequently, a half-wave plate (HWP) denoted as HWP1 is employed to prepare the state. The measurement procedure, enclosed within a delineated black dashed box, facilitates measurements amidst the presence of noise.

For the measurements under depolarizing, we select the Helstrom measurement along the z axis, as shown in Fig. 1(b). Varying ten angles of θ in states and eight depolarizing noisy intensities r^{depol} in measurements, the experimental results are illustrated in Fig. 3(a). To fulfill the rigorous requirements mentioned in Ref. [37], we postprocessed the experimental data through data analysis, as detailed in Appendix B. Note that despite considering four states $\{|\psi\rangle\langle\psi|, |\bar{\psi}\rangle\langle\bar{\psi}|, |\phi\rangle\langle\phi|, |\bar{\phi}\rangle\langle\bar{\phi}|\}$ during the construction of preparation equivalence in Eq. (2), the resulting noncontextuality inequality (4) can be related to only two states $\{|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|\}$. This simplification is attributed to the high symmetry present in Table I. However, since achieving this symmetry experimentally is challenging, we relax it to Table II using the method outlined in Appendix B. In this case,

TABLE II. The probability of specific measurement outcomes occurring for particular states with no symmetries assumed.

	$P_{\|\phi\rangle}$	$P_{\|\psi\rangle}$	$P_{\|\bar{\phi}\rangle}$	$P_{\|\bar{\psi}\rangle}$	P_y	$P_{\bar{y}}$
M_ϕ	$1-\epsilon_\phi$	c_ψ	$\epsilon_{\bar{\phi}}$	$1-\epsilon_\phi+\epsilon_{\bar{\phi}}-c_\psi$	0.5	0.5
M_ψ	c_ϕ	$1-\epsilon_\psi$	$1-c_{\bar{\phi}}$	$c_\phi-c_{\bar{\phi}}+\epsilon_\psi$	0.5	0.5
$M_{g\phi}$	s_ϕ	$1-s_\psi$	$1-s_{\bar{\phi}}$	$s_\phi-s_{\bar{\phi}}-s_\psi$	0.5	0.5
M_y	0.5	0.5	0.5	0.5	$1-\epsilon_y$	ϵ_y

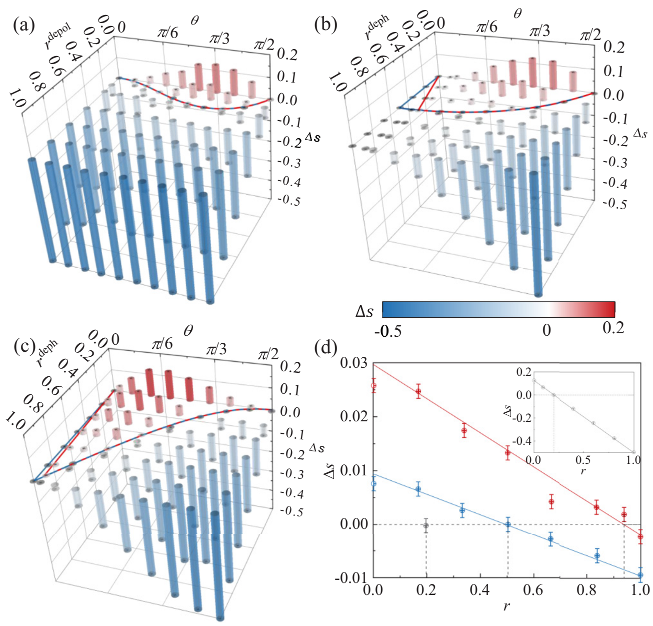

FIG. 3. Experimental tests of noncontextuality and contextuality. Each cylinder in (a)–(c) represents the chosen states θ and noisy intensity r^{depol} (r^{deph}) in our experiment. The color filling inside the cylinder represents the violation of the noncontextuality inequality Δs. The experimental (theoretical) boundaries, represented by the red (blue) curve, mark the transition from noncontextuality to contextuality. (a) Helstrom measurement performed along the z axis under depolarizing. The robustness of contextual reaches a maximum of $r_{\text{min}}^{\text{deol}} = 0.1976 \pm 0.0008$ at $\theta = \frac{\pi}{3}$. (b) Helstrom measurement performed along the z axis under dephasing noise. The robustness of contextual reaches a maximum of $r_{\text{min}}^{\text{deph}} = 0.504 \pm 0.001$ at $\theta = \pi/25$, higher than under depolarizing measurements. (c) Helstrom measurement performed along the x axis under dephasing noise. The robustness of contextual reaches a maximum of $r_{\text{min}}^{\text{deph}} = 0.938 \pm 0.001$ at $\theta = \pi/50$. (d) The maximum contextual robustness in (a)–(c). The circles represent the experimental results, and the lines depict the corresponding theoretical values. The black circles and line in the inset depict the situation of (a). Only one black circle falls within the scope of the main figure. Blue and red circles and line correspond to the situation of (b) and (c), respectively. Error bars indicate the statistical uncertainty

the noncontextual inequalities are no longer reducible to the single inequality (4). Consequently, we actually calculate the violation of all inequalities related to the probability of successful discrimination as shown in inequalities (B11). If the violation of these eight noncontextual inequalities is denoted as $\{\Delta s_i\}_{i=1}^{8}$, we just need to study the one with the biggest violation, which is $\Delta s = \frac{1}{2}\max\{\Delta s_i\}_{i=1}^{8}$. This is because the violation of any one of these inequalities signifies the contextual nature of the system.

When $\Delta s < 0$, the preparation and measurement procedures can be described by noncontextual models. Consequently, the probability of successfully discriminating states in this scenario does not exceed that of the classical case. Conversely, when $\Delta s > 0$, contextuality is present within the system. This contextual nature enhances the success probability of state discrimination beyond the classical case. In particular, when $\Delta s = 0$, a curve aligned with Eq. (6) emerges, illustrating the boundary between noncontextuality and contextuality.

The boundary curve is closely connected to the coherence of prepared states in Eq. (11), determined by θ, and the contextuality robustness, denoted by $r_{\text{min}}^{\text{depol}}$. Thus, it also serves to depict the robustness of contextuality with varying degrees of coherence. As $\theta = 0$ and $\theta = \frac{\pi}{2}$ correspond to orthogonal (fully distinguishable) and parallel (completely indistinguishable) states, respectively, the classical and quantum boundary coincide, resulting in no violation of the noncontextuality inequality. Consequently, the boundary curve demonstrates that contextuality can exist under any degree of coherence, excluding 0 and 1.

However, the robustness under depolarizing is not particularly high. Even under the scenario where $r_{\text{min}}^{\text{depol}}$ is maximized ($\theta = \frac{\pi}{3}$), the robustness only reaches 0.1976 ± 0.0008. Furthermore, as coherence tends to zero, contextual robustness also approaches zero. It is important to note that the presence of contextuality guarantees the existence of coherence, but the reverse is not necessarily true. This is because if $r^{\text{depol}} > r_{\text{min}}^{\text{depol}}$, the system remains noncontextual regardless of the degree of coherence.

To compare the relationship between contextual robustness and coherence under different types of noise, we employ a controlled variable approach. Without altering the direction of Helstrom measurement M_g, we change depolarizing to dephasing, as illustrated in Fig. 1(c). In Fig. 3(b), the theoretical contextual robustness (blue curve) would exhibit a discontinuity near $\theta = 0$, where the robustness drops directly from 0.5 to 0. However, based on the experimental results (red curve), the violation of inequality Δs near this point is too small to be reliably detected within the allowed error range (the standard deviation is 0.001). Consequently, the discontinuity point occurs near $\theta = \pi/25$.

It can be observed that for $\theta > \pi/25$, the contextual robustness monotonically decreases as coherence increases. This can be understood as smaller coherence of prepared states approaching identity measurement in dephasing measurements, resulting in better robustness against noise. The highest robustness is achieved at $\theta = \pi/25$, where $r_{\text{min}}^{\text{deph}} = 0.504 \pm 0.001$. Comparing Figs. 3(a) and 3(b), contextuality exhibits different levels of robustness to depolarizing and dephasing. Specifically, the optimal robustness under dephasing ($r_{\text{min}}^{\text{deph}} = 0.504 \pm 0.001$) surpasses that under depolarizing ($r_{\text{min}}^{\text{depol}} = 0.1976 \pm 0.0008$).

To maximize the robustness of contextuality, we rotate the Helstrom measurement M_g from the z axis to x under dephasing conditions, as illustrated in Fig. 1(d). In Fig. 3(c), the theoretical contextual robustness (blue curve) indicates that as θ approaches 0, $r_{\text{min}}^{\text{deph}}$ converges toward 1. However, constrained by the limited precision of our experiment, we are only able to set $\theta = \pi/50$ to achieve the maximum contextual robustness of $r_{\text{min}}^{\text{deph}} = 0.938 \pm 0.001$. This implies that when the system's coherence approaches zero and dephasing can be of any magnitude (up to 0.938 ± 0.001), the violation of the noncontextuality inequality can still occur, indicating the persistence of contextuality.

Finally, to depict the transition of the system from contextuality to noncontextuality under the influence of noise more clearly, we have chosen representative points from Figs. 3(a)–3(c), where the contextual robustness achieves its maximum value, as illustrated in Fig. 3(d). For depolarizing measurements, represented by the black circles and line, we have $\theta = \frac{\pi}{3}$. The system intersects the boundary (dashed black line) at $r = 0.1976 \pm 0.0008$ with a violation quantity $\Delta s = 0 \pm 0.001$. For dephasing measurements with the Helstrom measurement performed along the z axis, depicted by the blue circles and line, we set $\theta = \pi/25$. The system intersects the boundary at $r = 0.504 \pm 0.001$ with $\Delta s = 0 \pm 0.001$. For dephasing measurements with the Helstrom measurement conducted along the x axis, shown by the red circles and line, we have $\theta = \pi/50$. The system intersects the boundary (dashed black line) at $r = 0.938 \pm 0.001$ with $\Delta s = 0.002 \pm 0.001$.

IV. CONCLUSION

Our investigation centers on the scenario of minimum-error state discrimination within two-dimensional Hilbert spaces. We conduct an experiment involving depolarizing and dephasing measurements with adjustable noise intensity to explore the relationship between contextual robustness and coherence. The experimental results demonstrate that, for both depolarizing and dephasing, contextuality can persist under any degree of coherence, excluding the extreme cases of coherence 0 and 1. Particularly, in the scenario of dephasing, the robustness of contextuality is superior to that in the case of depolarizing. Remarkably, as the coherence approaches zero (yet remains nonzero), contextuality can exist under arbitrary dephasing. However, it is important to note that while the presence of contextuality implies the existence of coherence, the opposite relationship does not always apply. Furthermore, these conclusions also hold under the basis-independent coherence presented in Appendix C. Consequently, contextuality can serve as a reliable basis-independent coherence witness. Finally, the conclusion that the failure of noncontextuality cannot be achieved without set coherence of all the states (or all the measurements) bears similarity to the relationship between entanglement and nonlocality. It appears that coherence plays a pivotal role in inducing contextuality, akin to the role of entanglement in nonlocality. Hence, we would like to provide evidence or quantitative discussion to support this claim in future investigation.

ACKNOWLEDGMENTS

This work is supported by the National Key R&D Program of China (Grant No. 2023YFA1406701) and the National Natural Science Foundation of China (Grants No. 12025401, No. 92265209, No. 12104009, No. 12104036, and No. 12305008). D.K.Q. acknowledges the support from the China Postdoctoral Science Foundation (Grant No. 2023M730198) and the fellowship of China National Postdoctoral Program for Innovative Talents (Grant No. BX20230036).

APPENDIX A: EXPERIMENTAL DETAILS

Here, we describe the procedure for adjusting the angle of the wave plate to prepare the desired states and conduct measurements with adjustable noise intensity. By choosing $\theta = 4\varphi_1$, where φ_1 represents the angle of HWP1, we can prepare the state $|\psi\rangle = \cos\frac{\theta}{2}|H\rangle + \sin\frac{\theta}{2}|V\rangle$. This state is employed in the three scenarios outlined in the main text.

Utilizing noise-free measurements, we conduct quantum state tomography processes [55] on the prepared states to obtain their density matrix. This involves the projection of the prepared states onto four basis $\{|H\rangle, |V\rangle, \frac{|H\rangle-|V\rangle}{\sqrt{2}}, \frac{|H\rangle-i|V\rangle}{\sqrt{2}}\}$ and recording the number of photons counted by detectors. Subsequently, we compare the experimentally derived density matrix with theoretical values to assess the quality of the prepared states based on fidelity. Across all measured states, our experimental results consistently demonstrate a fidelity exceeding $99.1\% \pm 0.3\%$. Additionally, coherence C and set coherence R_1 can be deduced by plugging the experimentally derived density matrix into Eqs. (10) and (C1).

However, unlike the preparation phase, measurements with adjustable noise intensity exhibit variation across these three cases. Fortunately, we have identified a universal approach for determining the wave plate angle that is applicable to any trace preserved type of noise. Consider the noise-free projection basis as $|\xi\rangle\langle\xi|$; in the presence of noise, it can be denoted as $\mathcal{D}(|\xi\rangle\langle\xi|)$. The trace preserved property ensures that $\mathcal{D}(|\xi\rangle\langle\xi|)$ can always be diagonalized and expressed as $\mathcal{D}(|\xi\rangle\langle\xi|) = (1-\lambda)|\zeta\rangle\langle\zeta| + \lambda|\bar{\zeta}\rangle\langle\bar{\zeta}|$, where λ represents the noise intensity, and $|\zeta\rangle\langle\zeta|$ and $|\bar{\zeta}\rangle\langle\bar{\zeta}|$ form an orthogonal basis.

Subsequently, we elaborate on the implementation of this form of noise measurement by utilizing the dashed box segment of the experimental setup as depicted in the main text. Through the adjustment of the angles of QWP2 and HWP2, a simultaneous transformation of $\{|\zeta\rangle\langle\zeta|, |\bar{\zeta}\rangle\langle\bar{\zeta}|\}$ to $\{|H\rangle\langle H|, |V\rangle\langle V|\}$ is achievable. As the photon traverses the first BD, we attain the projection onto $|\zeta\rangle\langle\zeta|$ in the upper path and onto $|\bar{\zeta}\rangle\langle\bar{\zeta}|$ in the middle path. Subsequently, by employing HWP3 (HWP4), the photon is directed through the second BD into the middle path with a probability of $1-\lambda$ (λ), facilitating the acquisition of the desired imperfect measurement in this path.

This approach allows us to obtain imperfect measurements for all elements in Table II, enabling the assessment of the contextuality of the system. Consequently, we experimentally derive the corresponding noise intensity ϵ, overlap c, and the success probability of state discrimination s. Additionally, the noise intensity r mentioned in the main text can be derived from ϵ.

Furthermore, to calculate effective coherence, we conduct quantum state tomography processes [55] on the prepared states with imperfect measurements. This involves the projection of the prepared states onto four basis $\{\mathcal{D}(|H\rangle), \mathcal{D}(|V\rangle), \mathcal{D}(\frac{|H\rangle-|V\rangle}{\sqrt{2}}), \mathcal{D}(\frac{|H\rangle-i|V\rangle}{\sqrt{2}})\}$. The effective coherence C_{eff} and set coherence R_{eff} can then be deduced by plugging the effective density matrix into Eqs. (10) and (C1). The resulting outcome is theoretically equivalent to the scenario of a perfect measurement preceded by a noisy channel.

This approach enables us to investigate decoherence in the Appendix C.

Finally, we provide a concrete example to illustrate the imperfect measurement process. Let us consider measurements under dephasing noise with the Helstrom measurement along the x axis. The Helstrom measurement is expressed as $\mathcal{D}_r^{\mathrm{deph}}(|+\rangle\langle+|) = (1-\frac{r^{\mathrm{deph}}}{2})|+\rangle\langle+| + \frac{r^{\mathrm{deph}}}{2}|-\rangle\langle-|$, where $|+\rangle = \frac{1}{\sqrt{2}}(|H\rangle + |V\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|H\rangle - |V\rangle)$. By adjusting HWP2 to an angle of $\varphi_2 = 22.5^\circ$, the transformation from $\{|+\rangle\langle+|, |-\rangle\langle-|\}$ to $\{|H\rangle\langle H|, |V\rangle\langle V|\}$ is achieved. Subsequently, the angles of HWP3 and HWP4 are chosen such that $\varphi_4 = \frac{\pi}{4} - \varphi_3$, enabling measurements under the noise intensity $r^{\mathrm{deph}} = 2\sin^2(2\varphi_3)$. Two counts are considered: the count of correctly identified photons n_0 and the count of incorrectly identified photons n_1. These counts are measured with two different angles of HWP2: $\varphi_2 = 22.5^\circ$ and $\varphi_2 = -22.5^\circ$. The success probability of state discrimination is determined as $s = n_0/(n_0+n_1)$. Applying the same method, we can determine the noise intensity ϵ and the overlap c for imperfect measurements.

APPENDIX B: DATA ANALYSIS

In adherence to the rigorous standards of Ref. [37], we adopt the data-processing method initially proposed in Ref. [31], comprising two sequential steps. The first step, known as the primary procedure, assumes the experimental data to be noncontextual and applies fitting through the generalized probability theory (GPT) [56]. This approach facilitates the verification of contextuality without reliance on the validity of quantum theory. The second step, termed the secondary procedure, ensures that the data conform to the requirement of preparation equivalence.

By relaxing the assumed symmetries among the state preparations, we redefine Table I in the main text within the framework of GPT, as illustrated in Table II. It is noteworthy that two extra preparations along the y axis are introduced, creating an additional dimension to bring the data in the primary procedure closer to our raw experimental data. Moreover, the observable σ_y is included to ensure the completeness of tomographic measurement.

All entries in the table are experimentally obtained and utilized to construct a 6×4 matrix D^{r}, defined as $D^{\mathrm{r}} =$

$$\begin{pmatrix} P^{\mathrm{r}}_{|\phi\rangle,M_\phi} & P^{\mathrm{r}}_{|\psi\rangle,M_\phi} & P^{\mathrm{r}}_{|\bar\phi\rangle,M_\phi} & P^{\mathrm{r}}_{|\bar\psi\rangle,M_\phi} & P^{\mathrm{r}}_{y,M_\phi} & P^{\mathrm{r}}_{\bar y,M_\phi} \\ P^{\mathrm{r}}_{|\phi\rangle,M_\psi} & P^{\mathrm{r}}_{|\psi\rangle,M_\psi} & P^{\mathrm{r}}_{|\bar\phi\rangle,M_\psi} & P^{\mathrm{r}}_{|\bar\psi\rangle,M_\psi} & P^{\mathrm{r}}_{y,M_\psi} & P^{\mathrm{r}}_{\bar y,M_\psi} \\ P^{\mathrm{r}}_{|\phi\rangle,M_{g\phi}} & P^{\mathrm{r}}_{|\psi\rangle,M_{g\phi}} & P^{\mathrm{r}}_{|\bar\phi\rangle,M_{g\phi}} & P^{\mathrm{r}}_{|\bar\psi\rangle,M_{g\phi}} & P^{\mathrm{r}}_{y,M_{g\phi}} & P^{\mathrm{r}}_{\bar y,M_{g\phi}} \\ P^{\mathrm{r}}_{|\phi\rangle,M_y} & P^{\mathrm{r}}_{|\psi\rangle,M_y} & P^{\mathrm{r}}_{|\bar\phi\rangle,M_y} & P^{\mathrm{r}}_{|\bar\psi\rangle,M_y} & P^{\mathrm{r}}_{y,M_y} & P^{\mathrm{r}}_{\bar y,M_y} \end{pmatrix}. \tag{B1}$$

In the primary procedure, according to the proposition in Supplementary Note 4 of Ref. [31], a matrix D^{p} can emerge from a GPT with three two-outcome measurements that are tomographically complete if and only if every column of D^{p} satisfies the subsequent equation for $j \in \{\phi, \psi, \bar\phi, \bar\psi, y, \bar y\}$ and $i \in \{\phi, \psi, g\phi, y\}$,

$$aP^{\mathrm{p}}_{p_j,M_1} + bP^{\mathrm{p}}_{p_j,M_2} + cP^{\mathrm{p}}_{p_j,M_3} + dP^{\mathrm{p}}_{p_j,M_4} - 1 = 0, \tag{B2}$$

where $P^{\mathrm{p}}_{p_j,M_i}$ represents an element of matrix D^{p}, and $\{a, b, c, d\}$ are real numbers. Geometrically, the proposition stipulates that the eight columns of D^{p} are situated on a three-dimensional hyperplane defined by the constants $\{a, b, c, d\}$. To determine the GPT of best fit, we fit a three-dimensional hyperplane to the eight four-dimensional points constituting the columns of D^{r}. Subsequently, each column of D^{r} is mapped to its nearest point on the hyperplane, with these eight points forming the columns of D^{r}.

The GPT of best fit is established by minimizing the weighted distance χ_i, which gauges the deviation between raw data and primary data, defined as

$$\chi_i = \sqrt{\sum_{j=1}^{4} \frac{P^{\mathrm{p}}_{p_j,M_i} - P^{\mathrm{r}}_{p_j,M_i}}{\left(\Delta P^{\mathrm{r}}_{p_j,M_i}\right)^2}}. \tag{B3}$$

Here, $\Delta P^{\mathrm{r}}_{p_j,M_i}$ denotes the statistical uncertainty in the experiment. Thus, the primary procedure is formulated as the following minimization problem:

$$\begin{gathered} \min_{\{P^{\mathrm{p}}_{p_j,M_i},a,b,c,d\}} \chi^2 = \min_{\{a,b,c,d\}} \sum_{i=1}^{6} \chi_i^2 \\ s.t. \quad aP^{\mathrm{p}}_{p_j,M_1} + bP^{\mathrm{p}}_{p_j,M_2} + cP^{\mathrm{p}}_{p_j,M_3} + dP^{\mathrm{p}}_{p_j,M_4} - 1 = 0, \\ \forall j = 1, \dots, 6. \end{gathered} \tag{B4}$$

Using the method of Lagrange multipliers [57], this problem can be reformulated as a four-variable optimization problem:

$$\begin{gathered} \min_{\{a,b,c,d\}} \chi^2 = \min_{\{a,b,c,d\}} \sum_{i=1}^{6} \chi_i^2 \\ s.t. \quad aP^{\mathrm{p}}_{p_j,M_1} + bP^{\mathrm{p}}_{p_j,M_2} + cP^{\mathrm{p}}_{p_j,M_3} + dP^{\mathrm{p}}_{p_j,M_4} - 1 = 0, \\ \forall j = 1, \dots, 6, \end{gathered} \tag{B5}$$

where $\chi_i^2 =$

$$\frac{\left(aP^{\mathrm{p}}_{p_j,M_1} + bP^{\mathrm{p}}_{p_j,M_2} + cP^{\mathrm{p}}_{p_j,M_3} + dP^{\mathrm{p}}_{p_j,M_4} - 1\right)^2}{\left(a\Delta P^{\mathrm{p}}_{p_j,M_1}\right)^2 + \left(b\Delta P^{\mathrm{p}}_{p_j,M_2}\right)^2 + \left(c\Delta P^{\mathrm{p}}_{p_j,M_3}\right)^2 + \left(d\Delta P^{\mathrm{p}}_{p_j,M_4} - 1\right)^2}. \tag{B6}$$

The χ^2 parameter resulting from the fitting procedure serves as a metric for the goodness of fit of the hyperplane to the data. Given that we are fitting eight data points to a hyperplane defined by four fitting parameters a, b, c, d, we anticipate the χ^2 parameter to follow a χ^2 distribution with four degrees of freedom [58], with a mean of 4. If the fitting procedure had yielded significantly higher χ^2 values, it would suggest that the theoretical depiction of the preparation and measurement procedures necessitated more than three degrees of freedom. Conversely, if the fitting had produced an average χ^2 substantially lower than 4, it would indicate an overestimation of the uncertainty in data.

In the secondary procedure, since the columns of the D^{p} matrix define the GPT states, we denote the vector defined by the ith column as $\mathbf{P}^{\mathrm{p}}_i$. The secondary preparation is characterized by a probabilistic mixture of the primary preparations. Consequently, the GPT state of the secondary preparation is represented by a vector $\mathbf{P}^{\mathrm{s}}_i$, which is a probabilistic mixture of the $\mathbf{P}_i^{\mathrm{p}}$, defined as

$$\mathbf{P}_i^{\mathrm{s}} = \sum_{i'=1}^{6} u_{i'}^{i} \mathbf{P}_{i'}^{\mathrm{p}}, \tag{B7}$$

where the $u_{i'}^{i}$ are the weights in the mixture.

Therefore, the 4×4 matrix D^{s} is formed by the linear transformation of D^{p}, denoted as

$$D^{\mathrm{s}} = UD^{\mathrm{p}}, \tag{B8}$$

where U is a 4×6 matrix representing the linear transformation. The preparation equivalence is enforced by the condition

$$P^{\mathrm{s}}_{|\phi\rangle,M_i} + P^{\mathrm{s}}_{|\bar{\phi}\rangle,M_i} = P^{\mathrm{s}}_{|\psi\rangle,M_i} + P^{\mathrm{s}}_{|\bar{\psi}\rangle,M_i}, \quad \forall j = 1, 2, 3, \tag{B9}$$

which is essential for establishing noncontextual inequalities. Here, $P^{\mathrm{s}}_{|\phi\rangle,M_i}, P^{\mathrm{s}}_{|\bar{\phi}\rangle,M_i}, P^{\mathrm{s}}_{|\psi\rangle,M_i}, P^{\mathrm{s}}_{|\bar{\psi}\rangle,M_i}$ are elements of D^{s}. Additionally, we aim to ensure that the new states formed by linear combination are as close to the initial states as possible. Therefore, the secondary procedure can be summarized as the following maximization problem:

$$\begin{aligned} &\max D_{\mathrm{ps}} = \max \frac{1}{4} \sum_{v=1}^{4} U_{vv} \\ &s.t. \quad P^{\mathrm{s}}_{|\phi\rangle,M_i} + P^{\mathrm{s}}_{|\bar{\phi}\rangle,M_i} = P^{\mathrm{s}}_{|\psi\rangle,M_i} + P^{\mathrm{s}}_{|\bar{\psi}\rangle,M_i}, \\ &\qquad \forall j = 1, 2, 3. \end{aligned} \tag{B10}$$

Solving this problem yields all elements of D^{s}. The first three rows of the matrix correspond to rows 2–4 and columns 2–5 in Table II. Consequently, it becomes feasible to compute all the elements $\{\epsilon_\phi, \epsilon_{\bar{\phi}}, \epsilon_\psi, c_\phi, c_{\bar{\phi}}, c_\psi, s_\phi, s_{\bar{\phi}}, s_\psi\}$ required for noncontextual inequalities given afterward based on this relation.

In Fig. 4(a), for clarity, we depict the construction of secondary preparations within the zx plane of the Bloch sphere. Indeed, to refine the primary data toward the raw data, we incorporate processes within the bulk of the Bloch sphere using a similar methodology. Equation (B9) dictates that all preparations need not converge to the center of the Bloch sphere, but rather to the same state. Consequently, the two pairs need not be coplanar in the Bloch sphere. Based on this principle, supplementing the original set with just the two eigenstates of σ_y offers a suitable compromise between maintaining a low number of preparations and ensuring proximity of the secondary preparations to the ideal. As the σ_y eigenstates are maximally distant from the zx plane, they facilitate moving any point close to that plane in the $\pm y$ direction, while inducing only modest motion within the zx plane.

Given the absence of assumed symmetry, the noncontextual inequalities in this scenario are not reducible to a single inequality, as detailed in the full set of 15 noncontextuality inequalities provided in Appendix D of Ref. [37]. Consequently, we calculate the violation of all inequalities related to the probability of successful discrimination listed below and

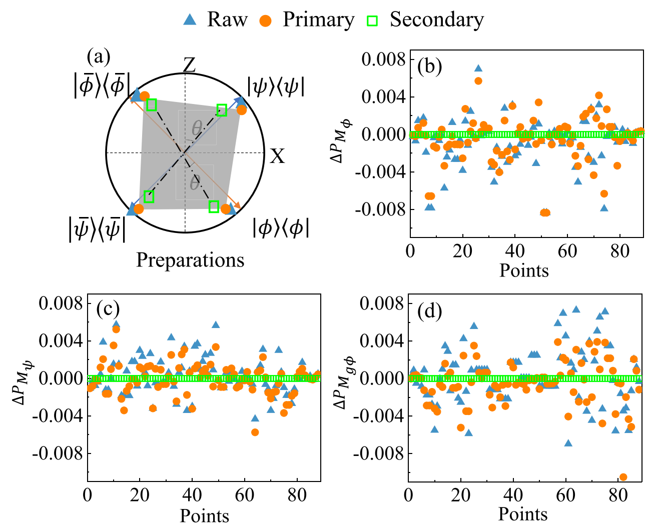

FIG. 4. The results of data analysis for preparation equivalence. Raw preparations (blue triangles) are derived from experimental data. Primary preparations (orange dots) represent the outcomes of fitting with the GPT theory. Secondary preparations (green rectangles) are specifically chosen to satisfy the assumption of preparation equivalence. (a) The formation of secondary preparations occurs within the zx plane of the Bloch sphere. If the four primary preparations are achievable, then any preparation within their convex hull can be simulated. The scenario where preparation equivalence is met corresponds to the alignment of the center point of the black dotted line between two pairs of states. In (b)–(d), the secondary data satisfies $\Delta P_{M_i} = 0$, where $\Delta P_{M_i} = P^{\mathrm{s}}_{|\phi\rangle,M_i} + P^{\mathrm{s}}_{|\bar{\phi}\rangle,M_i} - P^{\mathrm{s}}_{|\psi\rangle,M_i} - P^{\mathrm{s}}_{|\bar{\psi}\rangle,M_i}$. Each point in (b)–(d) corresponds to one preparation and measurement setup within the contextual region of Fig. 3 in the main text.

identify the largest violation.

$$\begin{aligned} 0 &\leqslant c_\psi + s_{\bar{\phi}} - s_\psi + \epsilon_\phi \\ 0 &\leqslant c_\psi - s_{\bar{\phi}} + s_\psi + \epsilon_\phi \\ 0 &\leqslant -c_\psi + s_\phi + s_\psi + \epsilon_{\bar{\phi}} \\ 0 &\leqslant c_\phi + s_{\bar{\phi}} - s_\psi + \epsilon_\psi \\ 0 &\leqslant -c_{\bar{\phi}} + s_\phi + s_\psi + \epsilon_\phi \\ 0 &\leqslant c_\phi - s_{\bar{\phi}} + s_\psi + \epsilon_\phi \\ 0 &\leqslant 2 - c_\psi - s_\phi - s_\psi + \epsilon_{\bar{\phi}} \\ 0 &\leqslant 2 - c_{\bar{\phi}} - s_\phi - s_\psi + \epsilon_\psi. \end{aligned} \tag{B11}$$

For consistency with the noncontextual inequality violations mentioned in the main text, we multiply the violations of these equations by one-half. The violation of any one of these inequalities signifies the contextual nature of the system.

In situations where the noise intensity approaches 1, the secondary process shows substantial deviation from the raw data at certain points, primarily because the preparations and measurements are very close together. Consequently, while all data undergo primary processing, secondary processing is selectively applied to points that reveal contextuality, preserving the rigor of our conclusions. Additionally, for the scenario where both $|\phi\rangle\langle\phi|$ and $|\psi\rangle\langle\psi|$ are close to the z axis in the zx plane of the Bloch sphere in the main text, two extra preparations along the x axis are introduced to bring the data in the primary procedure closer to the raw data. Data analysis for the 88 points revealing contextuality is presented in Figs. 4(b)–4(d), where all secondary points satisfy the condition of preparation equivalence.

APPENDIX C: COHERENCE, SET COHERENCE, AND EFFECTIVE COHERENCE

The definition of contextuality used in the main text is independent of the basis, whereas coherence is basis dependent. This distinction might raise concerns about the potential influence of coordinate choices on our conclusions. To address this, we illustrate below that our two main conclusions in the main text can still be drawn when examining the relationship between contextuality and set coherence, both of which are independent of bases.

The concept of set coherence, as introduced in Ref. [59], offers a basis-independent quantification of coherence by minimizing the coherence of states in the set across all possible basis choices. For a pair of qubit states ρ_1 and ρ_2, the mean robustness of set coherence R_1 characterizes the intrinsic coherence of this set and is given by the expression

$$R_1(\vec{p}) = \min_{\vec{p}\in S^2} \tfrac{1}{2}\|\vec{q}_1\||\sin(\vec{p},\vec{q}_1)| + \tfrac{1}{2}\|\vec{q}_2\||\sin(\vec{p},\vec{q}_2)|, \quad \text{(C1)}$$

where $\vec{q}_1$ ($\vec{q}_2$) is the Bloch vector of ρ_1 (ρ_2), and $(\vec{p},\vec{q}_1)$ represents the angle between the Bloch vector $\vec{p}$ and $\vec{q}_1$. For pairs of pure qubit states, the minimum in Eq. (C1) is achieved when $\vec{p}$ coincides with either $\vec{q}_1$ or $\vec{q}_2$. In the case of mixed states, the optimal $\vec{p}$ aligns with the Bloch vector of the purest state, corresponding to the longest Bloch vector.

Firstly, Fig. 5(a) explores the relationship between contextuality and set coherence without the influence of noise. The horizontal axis signifies the system state, while the vertical axis represents the values of each physical quantity. The expressions for these four physical quantities in Fig. 5(a) can be formulated as

$$\begin{aligned} R_1 &= \tfrac{1}{2}\sin(2\theta),\\ C &= \sin\theta,\\ \Delta s_z &= \tfrac{1}{2}(\cos\theta + \sin^2\theta - 1),\\ \Delta s_x &= \tfrac{1}{2}(\sin\theta + \cos^2\theta - 1). \end{aligned} \quad \text{(C2)}$$

When comparing the connection between contextuality and set coherence in both scenarios, despite the difference in coordinate selection, the relationship remains consistent. To emphasize this point, Fig. 5(b) offers a clearer depiction of the association between contextuality and set coherence, where the experimental data in two cases align closely with the same theoretical values. It is noteworthy that a R_1 value corresponds to two Δs, while the values of C and Δs have a one-to-one correspondence. Therefore, investigating the relationship between coherence and contextuality is more advantageous for drawing conclusions, which is a central focus of our research in the main text.

The conclusion, similar to the main text, can still be drawn here: contextuality implies the presence of set coherence in

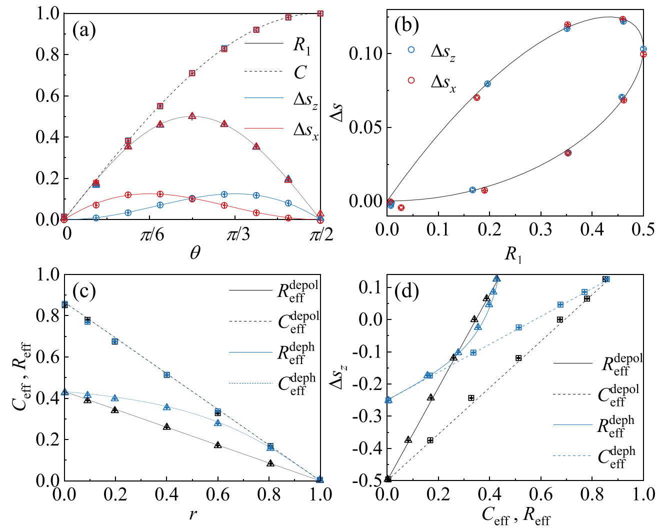

FIG. 5. Contextuality and different measures of coherence-coherence C, set coherence R_1, and efficient coherence $C_{\rm eff}$ (set coherence $R_{\rm eff}$). All data points in the figure correspond to experimental results. For all states measured, the fidelity consistently remains above $99.1\% \pm 0.3\%$. (a) The system with two coordinate choices in the absence of noise. Blue points represent the case of Helstrom measurement performed along the z axis, while red points represent the case of the Helstrom measurement performed along the x axis. Different shapes represent different physical quantities: circles represent the violation of noncontextual inequalities Δs, triangles represent R_1, and squares represent C. The curve represents the corresponding theoretical values of each physical quantity. (b) The relationship between contextuality and set coherence under the change of parameter θ. (c) The system with one coordinate choice in the present of two types of noise. The system state θ is set to 1.0472, near the maximum of Δs in the absence of noise. Black represents the case of depolarizing noise, and blue represents the case of dephasing noise. The solid line represents $R_{\rm eff}$, and the dashed line represents $C_{\rm eff}$. (d) The relationship between contextuality and efficient coherence (set coherence) under the change of noise intensity r.

the system, as depicted in Figs. 5(a) and 5(b). This relationship is also theoretically established in Appendix B of Ref. [41]. However, the reverse is not necessarily true under sufficient levels of noise. Additionally, it is noteworthy that the conclusion holds when set coherence approaches zero (while remaining nonzero); contextuality can endure under arbitrary amounts of partially dephasing, as observed by comparing the set coherence curve in Fig. 5(a) here and the contextuality boundary curve in Fig. 3(c) in the main text.

In the subsequent analysis shown in Fig. 5(c), we investigate decoherence through effective coherence (set coherence) under two types of noise, specifically focusing on the case of the Helstrom measurement performed along the z axis. The effective coherence $C_{\rm eff}$ (set coherence $R_{\rm eff}$) is deduced from the density matrix obtained through state tomography with imperfect measurements in our scenario. This is equivalent to quantifying the coherence (set coherence) of the system after it has undergone the same type of noisy channel. The equations for $R_{\rm eff}$, $C_{\rm eff}$, and Δs under depolarizing noise are as follows:

$$\begin{aligned} R_{\text{eff}}^{\text{depol}} &= \tfrac{1}{2}(1-r)\sin(2\theta), \\ C_{\text{eff}}^{\text{depol}} &= (1-r)\sin\theta, \\ \Delta s_z^{\text{depol}} &= \tfrac{1}{2}[(1-r)(\cos\theta+\sin^2\theta)-1]. \end{aligned} \tag{C3}$$

The expressions for R_{eff}, C_{eff}, and Δs under dephasing noise are formulated as

$$\begin{aligned} R_{\text{eff}}^{\text{deph}} &= \frac{(1-r)\sin(2\theta)}{2\sqrt{\cos^2\theta+(1-r)^2\sin^2\theta}}, \\ C_{\text{eff}}^{\text{deph}} &= (1-r)\sin\theta, \\ \Delta s_z^{\text{deph}} &= \tfrac{1}{2}[\cos\theta+(1-r)\sin^2\theta-1]. \end{aligned} \tag{C4}$$

As the intensity of noise increases, the efficient coherence C_{eff} (set coherence R_{eff}) decreases, as illustrated in Fig. 5(c). Building upon this observation of decoherence, we further study the performance of contextuality in Fig. 5(d). The behaviors of R_{eff}, C_{eff}, and Δs exhibit both linear and nonlinear relationships, depending on the type of noise. In the case of decoherence, this implies that the smaller the efficient coherence (set coherence), the less pronounced the violation of the noncontextuality inequality, making the existence of contextuality less likely.

APPENDIX D: ADDITIONAL DISCUSSION REGARDING CONTEXTUALITY AND COHERENCE

Here, we explore the feasible method to establish a basis-independent relationship between contextuality and coherence in the context of Spekkens' categorization [3]. Spekkens distinguishes between preparation noncontextuality and measurement noncontextuality, with Kochen-Specker noncontextuality falling into the latter category.

For preparation noncontextuality, the construction of preparation equivalence, akin to the first equality in Eq. (2) in the main text, is essential. This implies that the derived noncontextual inequalities involve at least two states and their orthogonal counterparts. In such cases, the basis-independent definition of coherence, known as set coherence, becomes relevant and can be employed to explore the basis-independent relationship between contextuality and set coherence of all the states, as demonstrated in our work.

For measurement noncontextuality, the construction of measurement equivalence, similar to preparation equivalence, is essential. This implies that the resulting noncontextual inequalities involve at least two measurements. In such cases, the set coherence of the measurement basis becomes relevant and can be employed to explore the basis-independent relationship between contextuality and set coherence for all measurements. For example, when verifying Kochen-Specker contextuality, at least one noncommutation pair of operators is required. In this context, it is intriguing to investigate the set coherence of the measurements. Moreover, the set coherence of states can also be investigated by constructing a set of states to test Kochen-Specker contextuality.

Finally, the conclusion that the failure of noncontextuality cannot be achieved without set coherence of all the states (or all the measurements), strictly proven in Ref. [41], bears similarity to the relationship between entanglement and nonlocality. We speculate that coherence plays a pivotal role in inducing contextuality, akin to the role of entanglement in nonlocality. Therefore, we aim to establish a link between coherence 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] R. W. Spekkens, Phys. Rev. A **71**, 052108 (2005).
[4] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Phys. Rev. Lett. **102**, 010401 (2009).
[5] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, New J. Phys. **18**, 045003 (2016).
[6] D. Saha, P. Horodecki, and M. Pawłowski, New J. Phys. **21**, 093057 (2019).
[7] S. Gupta, D. Saha, Z.-P. Xu, A. Cabello, and A. S. Majumdar, Phys. Rev. Lett. **130**, 080802 (2023).
[8] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Phys. Rev. Lett. **119**, 220402 (2017).
[9] S. A. Yadavalli and R. Kunjwal, Quantum **6**, 839 (2022).
[10] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).
[11] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey, Phys. Rev. Lett. **129**, 120403 (2022).
[12] R. Raussendorf, J. Bermejo-Vega, E. Tyhurst, C. Okay, and M. Zurel, Phys. Rev. A **101**, 012350 (2020).
[13] A. K. Daniel and A. Miyake, Phys. Rev. Lett. **126**, 090505 (2021).
[14] T. Baumgratz, M. Cramer, and M. B. Plenio, Phys. Rev. Lett. **113**, 140401 (2014).
[15] A. Streltsov, G. Adesso, and M. B. Plenio, Rev. Mod. Phys. **89**, 041003 (2017).
[16] A. Winter and D. Yang, Phys. Rev. Lett. **116**, 120404 (2016).
[17] J. H. Selby and C. M. Lee, Quantum **4**, 319 (2020).
[18] F. Ahnefeld, T. Theurer, D. Egloff, J. M. Matera, and M. B. Plenio, Phys. Rev. Lett. **129**, 120501 (2022).
[19] D. Qu, K. Wang, L. Xiao, X. Zhan, and P. Xue, npj Quantum Inf. **7**, 154 (2021).
[20] P. Wang, J. Zhang, C.-Y. Luan, M. Um, Y. Wang, M. Qiao, T. Xie, J.-N. Zhang, A. Cabello, and K. Kim, Sci. Adv. **8**, eabk1660 (2022).
[21] S. B. van Dam, J. Cramer, T. H. Taminiau, and R. Hanson, Phys. Rev. Lett. **123**, 050401 (2019).
[22] Z.-H. Liu, H.-X. Meng, Z.-P. Xu, J. Zhou, J.-L. Chen, J.-S. Xu, C.-F. Li, G.-C. Guo, and A. Cabello, Phys. Rev. Lett. **130**, 240202 (2023).
[23] P. Xue, L. Xiao, G. Ruffolo, A. Mazzari, T. Temistocles, M. T. Cunha, and R. Rabelo, Phys. Rev. Lett. **130**, 040201 (2023).
[24] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[25] A. Cabello, Phys. Rev. Lett. **101**, 210401 (2008).

[26] E. Amselem, M. Rådmark, M. Bourennane, and A. Cabello, Phys. Rev. Lett. **103**, 160405 (2009).
[27] X. Zhan, X. Zhang, J. Li, Y. Zhang, B. C. Sanders, and P. Xue, Phys. Rev. Lett. **116**, 090401 (2016).
[28] X. Zhan, E. G. Cavalcanti, J. Li, Z. Bian, Y. Zhang, H. M. Wiseman, and P. Xue, Optica **4**, 966 (2017).
[29] X. Zhan, P. Kurzyński, D. Kaszlikowski, K. Wang, Z. Bian, Y. Zhang, and P. Xue, Phys. Rev. Lett. **119**, 220403 (2017).
[30] D. Qu, P. Kurzyński, D. Kaszlikowski, S. Raeisi, L. Xiao, K. Wang, X. Zhan, and P. Xue, Phys. Rev. A **101**, 060101(R) (2020).
[31] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, ncomms11780 (2016).
[32] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Phys. Rev. Lett. **130**, 230201 (2023).
[33] G. Sharma, S. Sazim, and S. Mal, Phys. Rev. A **104**, 032424 (2021).
[34] A. Chaturvedi, M. Farkas, and V. J. Wright, Quantum **5**, 484 (2021).
[35] R. Wagner, R. D. Baldijão, A. Tezzin, and B. Amaral, J. Phys. A: Math. Theor. **56**, 505303 (2023).
[36] R. Wagner, A. Camillini, and E. F. Galvão, Quantum **8**, 1240 (2024).
[37] D. Schmid and R. W. Spekkens, Phys. Rev. X **8**, 011015 (2018).
[38] Q. Zhang, C. Zhu, Y. Wang, L. Ding, T. Shi, X. Zhang, S. Zhang, and W. Zhang, Chin. Phys. Lett. **39**, 080301 (2022).
[39] R. Kunjwal and R. W. Spekkens, Phys. Rev. Lett. **115**, 110403 (2015).
[40] I. Marvian, arXiv:2003.05984.
[41] V. P. Rossi, D. Schmid, J. H. Selby, and A. B. Sainz, Phys. Rev. A **108**, 032213 (2023).
[42] M. Lugli, P. Perinotti, and A. Tosini, Phys. Rev. Lett. **125**, 110403 (2020).
[43] H. Lee, K. Flatt, C. R. i Carceller, J. B. Brask, and J. Bae, Phys. Rev. A **106**, 032422 (2022).
[44] D. Ha and J. S. Kim, Phys. Rev. A **105**, 032421 (2022).
[45] S. M. Cohen, Phys. Rev. A **107**, 012401 (2023).
[46] I. A. Burenkov, M. V. Jabir, A. Battou, and S. V. Polyakov, PRX Quantum **1**, 010308 (2020).
[47] G. Zhu, O. Kálmán, K. Wang, L. Xiao, D. Qu, X. Zhan, Z. Bian, T. Kiss, and P. Xue, Phys. Rev. A **100**, 052307 (2019).
[48] X. W. Wang, G. Y. Zhu, L. Xiao, X. Zhan, and P. Xue, Quantum Inf. Process **23**, 8 (2024).
[49] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[50] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, Phys. Rev. Lett. **104**, 230503 (2010).
[51] C. W. Helstrom, J. Stat. Phys. **1**, 231 (1969).
[52] P. Xue, R. Zhang, H. Qin, X. Zhan, Z. H. Bian, J. Li, and B. C. Sanders, Phys. Rev. Lett. **114**, 140502 (2015).
[53] Z. Bian, J. Li, H. Qin, X. Zhan, R. Zhang, B. C. Sanders, and P. Xue, Phys. Rev. Lett. **114**, 203602 (2015).
[54] K. Wang, X. Qiu, L. Xiao, X. Zhan, Z. Bian, B. C. Sanders, W. Yi, and P. Xue, Nat. Commun. **10**, 2293 (2019).
[55] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
[56] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[57] M. Krystek and M. Anton, Meas. Sci. Technol. **18**, 3438 (2007).
[58] W. H. Press, *Numerical Recipes 3rd Edition: The Art of Scientific Computing* (Cambridge University, Cambridge, 2007).
[59] S. Designolle, R. Uola, K. Luoma, and N. Brunner, Phys. Rev. Lett. **126**, 220404 (2021).