Experimental Test of Contextuality Based on State Discrimination with a Single Qubit

Qiuxin Zhang[1], Chenhao Zhu[1], Yuxin Wang[1], Liangyu Ding[1], Tingting Shi[1], Xiang Zhang[1,2], Shuaining Zhang[1,2*], and Wei Zhang[1,2,3*]

[1]*Department of Physics, Renmin University of China, Beijing 100872, China*
[2]*Beijing Academy of Quantum Information Sciences, Beijing 100193, China*
[3]*Beijing Key Laboratory of Opto-electronic Functional Materials and Micro-nano Devices, Renmin University of China, Beijing 100872, China*

(Received 4 May 2022; accepted manuscript online 22 June 2022)

Exploring quantum phenomena beyond predictions of any classical model has fundamental importance to understand the boundary of classical and quantum descriptions of nature. As a typical property that a quantum system behaves distinctively from a classical counterpart, contextuality has been studied extensively and verified experimentally in systems composed of at least three levels (qutrit). Here we extend the scope of experimental test of contextuality to a minimal quantum system of only two states (qubit) by implementing the minimum error state discrimination on a single $^{171}Yb^{+}$ ion. We observe a substantial violation of a no-go inequality derived by assuming non-contextuality, and firmly conclude that the measured results of state discrimination cannot be reconciled with any non-contextual description. We also quantify the contextual advantage of state discrimination and the tolerance against quantum noises.

DOI: 10.1088/0256-307X/39/8/080301

In 1935, Einstein, Podolsky, and Rosen [1] firstly posed the question of whether quantum observables have physical reality. It was then realized that such a question has significant physical indication rather than a conceptual or metaphysical notion. A variety of statements called *no-go theorems* have been formulated in mathematical terms to exclude from our natural world some different assumptions of reality. The pioneering works by Bell [2,3] and Kochen and Specker (KS) [4,5] provided two successful quantitative descriptions, where the violation of Bell inequality effectively precludes local causation, and the no-go theorem of KS theory proves that quantum mechanics will give conflicting predictions from any classical non-contextual model. [4,5] These statements are of great importance since they are both rigorous in mathematics and realistic in experiments, [6] where the test of Bell inequality usually requires entangled states, and KS inequality can be violated for a more general state of a quantum system.

The test of KS inequality has been demonstrated in various physical platforms composed of photons, [7,8] superconducting systems, [9] nuclear spins, [10] ions, [11–13] or neutrons. [14] In fact, in the conventional formalism of KS theory, a *qutrit* system including at least three levels is the simplest case to show quantum contextuality. Previous works are then demonstrated for either two-qubit systems or a qutrit with multiple observables. [7–20] Recently, a research about contextual advantage for quantum state discrimination (SD) suggests that a test of contextuality may be extended to even simpler systems and with fewer observables. [21]

In this work, we show the first experimental test of quantum contextuality in the simplest possible system of *a single qubit* with only *two observable measurements* based on the non-classicality of state discrimination. State discrimination focuses on how to effectively determine the state of a quantum system according to a set of measured probabilities with a known prior distribution. For a Hermitian system, one can prove that two non-orthogonal pure states cannot be entirely discriminated within either a classical or a quantum theory. However, a quantum description can significantly improves the probability of success, such that the research along this line is of great interest in both the study of fundamental quantum physics [22] and the search of potential applications in quantum information processing. [23,24] This advantage occurs exclusively in non-classical theory and has been studied in various areas such as quantum clone theorem, [25–27] dimension witness, [28] and mutual information. [29] Using a trapped $^{171}Yb^{+}$ ion, we realize for the first time an experimental test of contextuality under the minimal error state discrimination (MESD) scenario with equal prior probability, where a no-go inequality that tolerates experimental error in non-contextual notation can be derived. [21] With high-fidelity single-qubit logic gates and highly distinguishable state-dependent fluorescence detection, a significant violation of the inequality is observed and the experimental results well agree with theoretical predictions of a quantum description.

Preliminary. We consider the problem of distinguishing two pure states $|\psi\rangle$ and $|\phi\rangle$ with equal priori probability, as shown in Fig. 1(a). The states $|\bar{\psi}\rangle$ and $|\bar{\phi}\rangle$ are orthogonal to $|\psi\rangle$ and $|\phi\rangle$, respectively, and are introduced as complement to construct operational equivalence [6] (see *Data analysis and results*) and derive the no-go inequal-

*Corresponding authors. Email: wzhangl@ruc.edu.cn; zhangshuaining@ruc.edu.cn
© 2022 Chinese Physical Society and IOP Publishing Ltd

ity with the hypothesis of non-contextuality. By defining the Helstrom measurement [30] basis set (the dashed line) as $\{M_{d|\psi\rangle} = 1/2(I + \sigma_z), M_{d|\phi\rangle} = 1/2(I - \sigma_z)\}$ with $M_{d|\psi\rangle} + M_{d|\phi\rangle} = I$, we can write the probability of a successful differentiation between $|\psi\rangle$ and $|\phi\rangle$ as $s = \frac{1}{2}(\mathrm{Tr}[M_{d|\psi\rangle}|\psi\rangle\langle\psi|] + \mathrm{Tr}[M_{d|\phi\rangle}|\phi\rangle\langle\phi|])$. It is reasonable and inevitable for maximizing the successful probability to assume that the measurements have the symmetry property $\mathrm{Tr}[M_{d|\psi\rangle}|\psi\rangle\langle\psi|] = \mathrm{Tr}[M_{d|\phi\rangle}|\phi\rangle\langle\phi|]$, so that $s = \mathrm{Tr}[M_{d|\psi\rangle}|\psi\rangle\langle\psi|] = \mathrm{Tr}[M_{d|\phi\rangle}|\phi\rangle\langle\phi|]$. With that, the quantity s is related to the overlap of $|\phi\rangle$ and $|\psi\rangle$ as

$$s = \frac{1}{2}(1 + \sqrt{1 - |\langle\phi|\psi\rangle|^2}) = \frac{1}{2}(1 + \sqrt{1 - c}), \tag{1}$$

where the overlap $c \equiv |\langle\phi|\psi\rangle|^2$ is usually referred as confusability. It is obvious that when the direction of the measurement basis $M_{d|\psi(\phi)\rangle}$ and the direction of the vector $(|\psi\rangle - |\phi\rangle)$ are parallel as shown in Fig. 1(a), the error rate of the two-state discrimination is the lowest, and we call the minimum-error state discrimination (MESD).

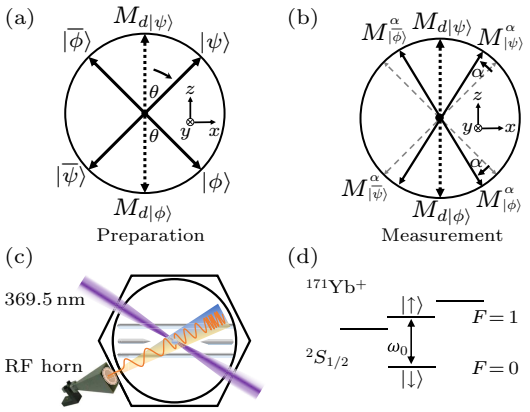

Fig. 1. (a) The scheme of state preparations on the x–z Bloch plane. The state $|\psi(\phi)\rangle$ (dark arrow) has an angle θ with measurement basis $M_{d|\psi(\phi)\rangle}$ (dashed line). We change the value of θ from 0 to $\pi/2$ to prepare different state ensembles, such that the two states change from being orthogonal to parallel. The state $|\bar{\psi}(\bar{\phi})\rangle$ is orthogonal to the state $|\psi(\phi)\rangle$. (b) The measurement scheme for states on the x–z Bloch plane. The ideal measurement bases are shown with gray dashed lines. A measurement error ϵ is considered in the basis $M^{\alpha}_{|\psi(\phi)\rangle}$ ($\alpha \in [0, \pi/2]$), which has an angle deviation α from the ideal basis $M_{|\psi(\phi)\rangle}$. The measurement basis $M^{\alpha}_{|\phi(\psi)\rangle}$ and $M^{\alpha}_{|\bar{\phi}(\bar{\psi})\rangle}$ are orthogonal. (c) The trapped single ^{171}Yb$^+$ ion system with a four-rod Paul trap. The purple beam is the 369.5 nm laser for cooling, state initialization and state readout. The single-qubit gate is realized by a microwave horn antenna. (d) The energy level diagram of ^{171}Yb$^+$.

As proved in Ref. [21], any ontological model of an operational theory with non-contextuality must lead to the following no-go theorem for MESD:

$$s \leq 1 - \frac{c}{2}. \tag{2}$$

This inequality would be relaxed owing to the error of quantum operations or measurement basis under realistic experimental conditions. To this end, one obtains

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

where ϵ is a quantitative indicator of measurement incorrectness, which is introduced just like the one appeared in the Clauser–Horne–Shimony–Holt (CHSH) inequality [31] to describe the error of measurement in the Bell inequality. In our experiment, ϵ is related to the deviation of the actual measurement basis $\{M^{\alpha}_{|\psi\rangle}, M^{\alpha}_{|\phi\rangle}\}$ [solid lines in Fig. 1(b)] from the ideal measurement axes (dashed lines) with an angle $\alpha \in [0, \pi/2]$. We further choose that the directions of deviation for $M^{\alpha}_{|\psi\rangle}$ and $M^{\alpha}_{|\phi\rangle}$ are opposite to each other, i.e., they are tilted either away or close to the x-axis simultaneously.

Taking the measurement error into account, the quantum bound of success probability of SD can be derived as [21]

$$s = \frac{1}{2}\left[1 + \sqrt{1 - \epsilon + 2\sqrt{\epsilon(1-\epsilon)c(1-c)} + c(2\epsilon - 1)}\right]. \tag{4}$$

Apparently, expression (4) reduces to Eq. (1) when $\epsilon = 0$, and the no-go result of Eq. (2) is the corresponding limiting case of Eq. (3). The success probability s of discrimination in quantum description is higher than the upper bound of non-contextual theories. We use the same terminology as in Ref. [21] to label the measurements and outcomes in a natural way. For instance, the outcome of $M^{\alpha}_{|\phi\rangle}$ is more likely to occur than the outcome of $M^{\alpha}_{|\psi\rangle}$ if one begins with $P_{|\phi\rangle}$ rather than $P_{|\bar{\psi}\rangle}$ preparation, and so on. Thus, Eqs. (3) and (4) must satisfy the constraint

$$\epsilon \leq c \leq 1 - \epsilon. \tag{5}$$

We also calculate the theoretical predications of $c = \langle\psi|M^{\alpha}_{|\phi\rangle}|\psi\rangle$, $s = \langle\phi|M_{d|\phi\rangle}|\phi\rangle$, and $1 - \epsilon = \langle\phi|M^{\alpha}_{|\phi\rangle}|\phi\rangle$ under a quantum description for different state preparations and measurements. The precise predictions are listed in Table 1. Note that two additional preparations along the y-axis are introduced as an extra dimension to correct the experimental preparation inaccuracy and supplement the measurement observable σ_y to ensure the completeness of tomographic measurement.

Table 1. Theoretical predictions for different state preparations and measurements. Each pair of preparation P^{θ}_{i} (with $i \in \{|\phi\rangle, |\bar{\phi}\rangle, |\psi\rangle, |\bar{\psi}\rangle\}$) and measurements M_j (with $M_j \in \{M^{\alpha}_{|\psi\rangle}, M^{\alpha}_{|\phi\rangle}, M_{d|\phi\rangle}\}$) yields a set of parameters $c = \sin^2[(\alpha - 2\theta)/2]$, $\epsilon = \sin^2(\alpha/2)$ and $s = \cos^2(\theta/2)$. To ensure that the preparations and measurements are tomographically complete, we add the preparations $\{P_y, P_{\bar{y}}\}$ and measurement M_y in the y-axis.

	$P^{\theta}_{\|\psi\rangle}$	$P^{\theta}_{\|\bar{\psi}\rangle}$	$P^{\theta}_{\|\phi\rangle}$	$P^{\theta}_{\|\bar{\phi}\rangle}$	P_y	$P_{\bar{y}}$
$M^{\alpha}_{\|\psi\rangle}$	$1-\epsilon$	ϵ	c	$1-c$	0.5	0.5
$M^{\alpha}_{\|\phi\rangle}$	c	$1-c$	$1-\epsilon$	ϵ	0.5	0.5
$M_{d\|\phi\rangle}$	$1-s$	s	s	$1-s$	0.5	0.5
M_y	0.5	0.5	0.5	0.5	1	0

To test the no-go inequality (3), we need to prepare all states and measure all observables as given in Table 1, under the condition of operational equivalence. Then we can obtain the values of $\{c, \epsilon, s\}$ from the relations given above and substitute them into Eq. (3). If a violation of such an inequality is observed beyond experimental error, one can conclude that no non-contextual model can account for the experimental observation. Obviously this situation will be

more difficult to occur with increasing noise. In this sense, the no-go inequality (3) provides a criterion for the boundary between quantum and classical regimes.

Experimental Setup. The experiment is performed on a single $^{171}Yb^{+}$ ion trapped in a four-rod Paul trap, as shown in Fig. 1(c). A pair of clock states $|F=0, m_F=0\rangle$ and $|F=1, m_F=0\rangle$ of the $^2S_{1/2}$ ground-state manifold are used to encode a qubit as $\{|\downarrow\rangle, |\uparrow\rangle\}$ with transition frequency $\omega_0 = 2\pi \times 12.64281\,\text{GHz}$ [Fig. 1(d)]. The qubit can be initialized by Doppler cooling and optical pumping, and detected by shining a 369.5 nm laser to excite the state $|\uparrow\rangle$ to the $^2P_{1/2}$ manifold and collecting the fluorescent photons with a photo-multiplier tube (PMT).[32] The error of detecting $|\uparrow\rangle$ for $|\downarrow\rangle$ is 1.71% and the other one is 2.08%. We calibrate the results to remove the detection errors as proposed in Ref. [33].

The states $|\psi\rangle$ and $|\phi\rangle$ we want to prepare lie in the x–z plane, and can be denoted as

$$\begin{aligned} |\psi\rangle &= \cos(\theta/2)|\downarrow\rangle + \sin(\theta/2)|\uparrow\rangle, \\ |\phi\rangle &= \cos[(\pi-\theta)/2]|\downarrow\rangle + \sin[(\pi-\theta)/2]|\uparrow\rangle \end{aligned} \tag{6}$$

on the bases of $\{|\downarrow\rangle, |\uparrow\rangle\}$. When θ changes from 0 to $\pi/2$, the two states become closer to the x-axis on the Bloch sphere with increasing confusability c. To prepare such states, single-qubit quantum gates are implemented via microwave with resonant frequency. By controlling the duration and phase of the microwave, we can achieve a native single-qubit rotation

$$R(\beta,\gamma) = \begin{pmatrix} \cos(\beta/2) & -ie^{-i\gamma}\sin(\beta/2) \\ -ie^{i\gamma}\sin(\beta/2) & \cos(\beta/2) \end{pmatrix}, \tag{7}$$

where β and γ are the rotation angle and phase, respectively. The π time of the Rabi oscillation is 7.51 μs, corresponding to a Rabi frequency of $2\pi \times 66.6\,\text{kHz}$.

The experimental process is as follows: After 500 μs Doppler cooling, the qubit is initialized to the $|\downarrow\rangle$ state by a 6 μs optical pumping. Then, the set of states $\{P^{\theta}_{|\psi\rangle}, P^{\theta}_{|\bar{\psi}\rangle}, P^{\theta}_{|\phi\rangle}, P^{\theta}_{|\bar{\phi}\rangle}\}$ are prepared through the single-qubit rotations $\{R_y(\theta), R_y(\pi+\theta), R_y(\pi-\theta), R^{\dagger}_y(\theta)\}$, where $R_y(\theta) = R(\theta, \pi/2)$ and $R^{\dagger}_y(\theta) = R(\theta, 3\pi/2)$. The observable set is chosen as $\{M^{\alpha}_{|\psi\rangle}, M^{\alpha}_{|\phi\rangle}, M_{d|\phi\rangle}\}$, which are realized by adding single-qubit operations $\{R^{\dagger}_y(\theta-\alpha), R^{\dagger}_y(\pi-\theta+\alpha), R^{\dagger}_y(\pi)\}$ to rotate the measurement basis from the original z-basis to the desired one. By scanning the parameter region $\{\theta, \alpha\}$, we can obtain the dataset $\{c, \epsilon, s\}$.

The parameter ϵ is a function of the variable α as $\epsilon = \sin^2(\alpha/2)$. Similarly, we have $c = \sin^2[(\alpha-2\theta)/2]$. When ϵ has a certain value, c must meet the inequality (5), which can be reduced to

$$\sin^2(\alpha/2) \le \sin^2[(\alpha-2\theta)/2] \le \cos^2(\alpha/2). \tag{8}$$

Note that $\sin^2[(\alpha-2\theta)/2] \le \cos^2(\alpha/2)$ is always true with $\theta, \alpha \in [0, \pi/2]$. Therefore, when α is fixed, the range of θ can be determined by solving $\sin^2(\alpha/2) \le \sin^2[(\alpha-2\theta)/2]$ with $\theta, \alpha \in [0, \pi/2]$. In our experiment, the parameters θ and α are scanned with an interval of 0.1, resulting in 136 sets of data $\{c, \epsilon, s\}$. We repeat the experiment for 1000 trials for each choice of the data set.

To correct the state preparation inaccuracy on the Bloch sphere and to ensure that the measurements are tomographically complete, we also supplement extra state preparations and observable measurements through the similar quantum operations. Details of this procedure will be given in the subsequent discussion and the Supplemental Materials. As a result, the total state preparations are $\{P^{\theta}_{|\psi\rangle}, P^{\theta}_{|\bar{\psi}\rangle}, P^{\theta}_{|\phi\rangle}, P^{\theta}_{|\bar{\phi}\rangle}, P_y, P_{\bar{y}}\}$, and observable measurements are $\{M^{\alpha}_{|\psi\rangle}, M^{\alpha}_{|\phi\rangle}, M_{d|\phi\rangle}, M_y\}$, as shown in Table 1. We emphasize that although more than two state preparations and observable measurements are carried in the experiment, only two observables with binary outcome are obtained from the set $\{M^{\alpha}_{|\psi\rangle}, M^{\alpha}_{|\phi\rangle}, M_{d|\phi\rangle}\}$, and are subsequently used to test the no-go inequality and show the boundary between quantum and classical regimes.

Data Analysis and Results. The test of no-go inequality derived from the non-contextual model requires the data meet operational equivalence in a general framework regardless of quantum or classic. For this purpose, we need to process the raw experimental data in two stages[34,35] (see details in the Supplemental Materials). The first phase fits the data of raw preparations using generalized probability theory (GPT) to ensure that the states are not exclusively dependent on quantum theory under the condition of complete tomographic measurements. The results of this stage, refereed as the primary preparations, are then supplied to the second phase which treats the data to perfectly comply to operational equivalence, i.e.,

$$\mathcal{P}_{\frac{1}{2}\phi+\frac{1}{2}\bar{\phi},M_i} = \mathcal{P}_{\frac{1}{2}|\psi\rangle+\frac{1}{2}|\bar{\psi}\rangle,M_i} = \frac{1}{2}. \tag{9}$$

Here, $\mathcal{P}_{\frac{1}{2}|\phi\rangle+\frac{1}{2}|\bar{\phi}\rangle,M_i}$ denotes a process associated with a particular $M_i \in \{M_{|\phi\rangle}, M_{|\psi\rangle}, M_{d|\phi\rangle}\}$, which is randomly selected from the average random mid-sample of the $P_{|\phi\rangle}$ and $P_{|\bar{\phi}\rangle}$ preparation process in theory. The case is analogous for $\mathcal{P}_{\frac{1}{2}\psi+\frac{1}{2}\bar{\psi},M_i}$. In experiment, $\mathcal{P}_{\frac{1}{2}|\phi\rangle+\frac{1}{2}|\bar{\phi}\rangle,M_i}$ is realized by a linear combination of the statistics of $|\phi\rangle$ and $|\bar{\phi}\rangle$ under measurement M_i, i.e.,

$$\mathcal{P}_{\frac{1}{2}|\phi\rangle+\frac{1}{2}|\bar{\phi}\rangle,M_i} = \frac{1}{2}[P(1 \mid P_{|\phi\rangle}, M_i) + P(1 \mid P_{|\bar{\phi}\rangle}, M_i)]. \tag{10}$$

Here, $P(1|P_{\phi(\bar{\phi})}, M_i)$ denotes the frequency when the measurement axis is M_i and the outcome is 1. Using the maximum likelihood estimation method, we obtain the secondary state preparations by minimizing the target function as the summation of distances from their corresponding primaries.

The two stages of data processing alter the position of states in the x–z plane as illustrated in Fig. 2(a). The data processing outcomes are plotted in Figs. 2(b)–2(d). Notice that the raw and primary data of state preparations are slightly off 1/2 owing to experimental errors in operation and measurement. The primary preparations fitted from the GPT theory have a slightly larger variance, which can be hardly observed in Figs. 2(b)–2(d). The results of secondary preparations show that all the 136 points fulfill the operational equivalence criterion with an exactly equal distribution 1/2. The similarity between states before and after data processing is evaluated as the average of mixture weights, which is determined to be $97.13\% \pm 2.3\%$ in our experiment. Further details and results can be found in the Supplemental Materials.

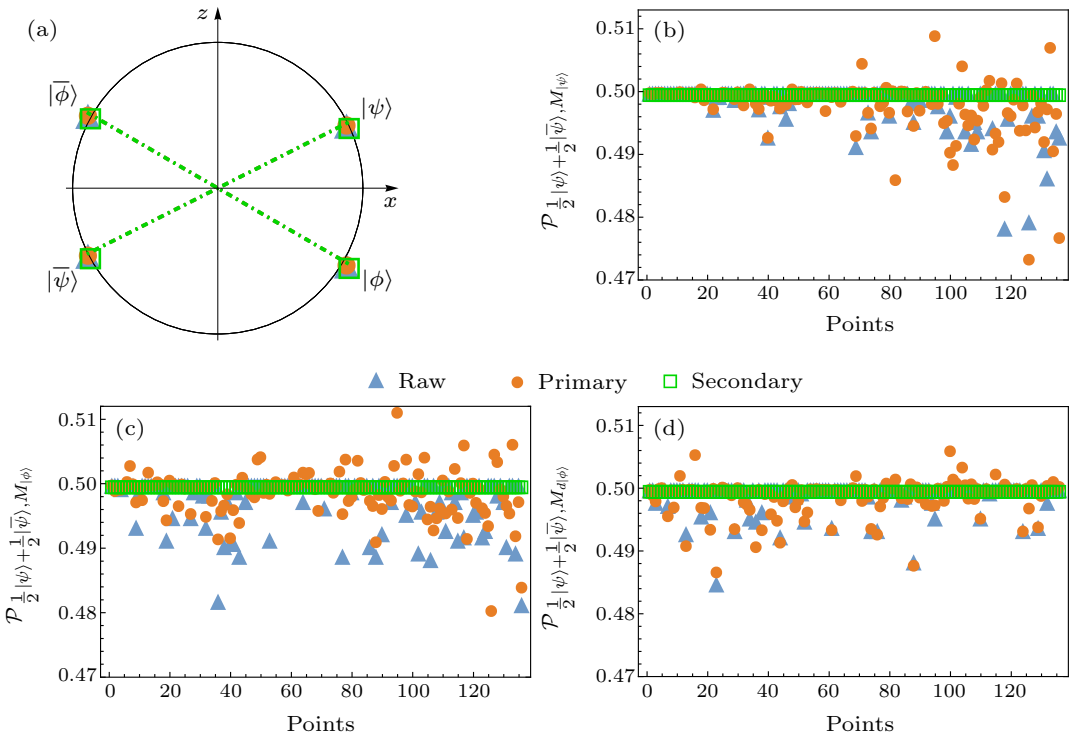

Fig. 2. The results of data analysis for operational equivalence. (a) The states $\{|\phi\rangle, |\bar{\phi}\rangle, |\psi\rangle, |\bar{\psi}\rangle\}$ prepared in experiments are plotted on the x–z plane of the Bloch sphere with $\theta = 1.1$ rad and $\alpha = 0$ rad, corresponding to the point with index 12 in panels (b), (c) and (d). Raw preparations (blue triangles) are given from the experimental data. Primary preparations (orange dots) are the results of fitting with the GPT theory. Secondary preparations (green rectangles) are chosen to satisfy the assumption of operational equivalence. Here, (b), (c), and (d) show that the probabilities $\mathcal{P}_{\frac{1}{2}|\psi\rangle+\frac{1}{2}|\bar{\psi}\rangle,M_i}$ reach statistical operational equivalence, i.e. $\mathcal{P}_{\frac{1}{2}|\psi\rangle+\frac{1}{2}|\bar{\psi}\rangle,M_i} = \frac{1}{2}$ after data processing. The results of $\mathcal{P}_{\frac{1}{2}|\phi\rangle+\frac{1}{2}|\bar{\phi}\rangle,M_i}$ are given in the Supplemental Materials.

We use the secondary preparation data which satisfies operational equivalence to calculate the results of $\{c, \epsilon, s\}$ and show the violation of the no-go inequality (3). In Fig. 3(a), we plot the probability of successful state discrimination s by varying the confusability c with fixed measurement error $\epsilon = \{\sin(0/2)^2, \sin(0.3/2)^2, \sin(0.785/2)^2, \sin(1.2/2)^2\} = \{0, 0.022, 0.146, 0.315\}$. The measured $s_{\rm exp}$ is clearly above the upper bound $s_{\rm non-con} = 1 - (c - \epsilon)/2$ for any non-contextual model (dashed lines), and agree well with theoretical predictions $s_{\rm theory}$ of Eq. (4) given by quantum theory with contextuality (solid lines). The shaded region between the dashed and solid lines thus denotes the advantage of a contextual theory in SD. We also notice that in the limiting case of $c = \epsilon$, the success probability of SD can reach $s = 1$ even for a pair non-orthogonal state preparations with $c > 0$. This is because in the presence of a finite ϵ, a small confusability between two preparations with $c \leq \epsilon$ can always be attributed to the measurement error, such that the two states are perfectly distinguished within an ontological description of the system, regardless of contextuality or non-contextuality.

To show the contextual advantage of SD more clearly, we plot in Fig. 3(b) the exceeding of s over the non-contextual bound $\Delta s_{\rm theory/exp} = s_{\rm theory/exp} - s_{\rm non-con}$ for each choice of ϵ. We emphasize that the highest value of $\Delta s_{\rm theory} = s_{\rm theory} - s_{\rm non-con}$ is 0.207 for the parameters (c, ϵ)=(0.853, 0.146), which clearly exceeds the maximal value of 0.125 for the case with no measurement error achieved at (c,ϵ)=(0.75,0). These two points are marked in 3(b) with a red pentagram and a hexagonal star, respectively. It is obvious that the addition of a finite ϵ significantly improves the contextual advantage of SD over any non-contextual model. In experiment, the location of the pentagram and hexagonal star are determined at $\{c, \epsilon, s\} = \{0.750 \pm 0.010, 0.0000 \pm 0.0006, 0.125 \pm 0.007\}$ and $\{c, \epsilon, s\} = \{0.853 \pm 0.009, 0.146 \pm 0.011, 0.206 \pm 0.006\}$, respectively. The quantity Δs also manifests the tolerance against noise in SD. For instance, one can prove that in a contextual theory s decreases with increasing depolarization error of state preparation and measurement,[21] which eventually compromises the violation of inequality Eq. 3. Thus, a higher value of Δs achieved at a finite ϵ indicates that a quantum SD is more robust against depolarization noises.

The results of Δs for all the 136 points obtained by theory and experiment are plotted in false color in Figs. 3(c) and 3(d), respectively. To quantify the resemblance between theory and experiment, we define the average fidelity as

$$f = 1 - \frac{1}{N} \sum_{i=1}^{N=136} \frac{|\Delta s_{i,\rm exp} - \Delta s_{i,\rm theory}|}{\Delta s_{i,\rm theory}}, \tag{11}$$

and obtain $f = 98.82\% \pm 0.12\%$. Since the gate infidelity induced by the fluctuation of microwave parameters is very low, which is $\sim (2.0 \pm 0.3) \times 10^{-4}$ as charac-

terized by the randomized benchmarking method,[36] the main error source is the sampling error with an average of 1.29%±0.45% (see details in the Supplemental Materials).

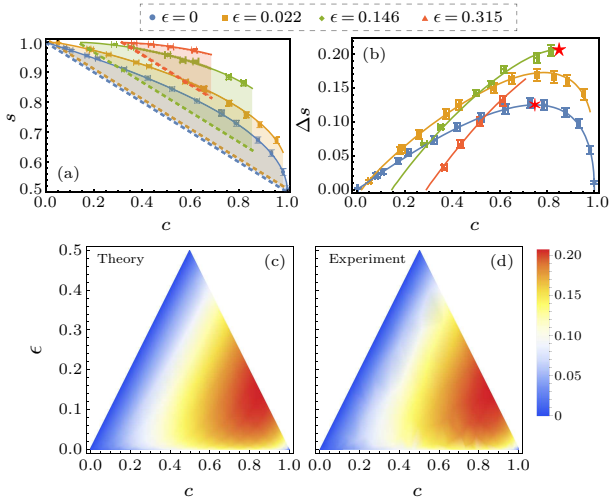

Fig. 3. The results of contextual advantage for state discrimination. (a) The results of success probability s versus confusability c for $\epsilon = \{0, 0.022, 0.146, 0.315\}$. The points of dots, squares, diamonds and triangles are the experimental data. The dashed lines represent the non-contextual upper bound $s_{\mathrm{non-con}}$ given by Eq. (3), and the solid lines are the predictions by a quantum theory s_{theory}. The shaded areas mark the contextual advantage $\Delta s_{\mathrm{theory}} = s_{\mathrm{theory}} - s_{\mathrm{non-con}}$ for state discrimination. (b) The results of Δs versus c with the same ϵ's and line shapes as in (a). The red pentagram and hexagonal star are the highest points with and without ϵ, locating at $(c, \epsilon) = (0.853, 0.146)$ and $(c, \epsilon) = (0.75, 0)$, respectively. (c) and (d) The false-color plots of Δs by changing c and ϵ.

In summary, we have demonstrated an experimental no-go test of contextuality derived within the frame work of minimum error state discrimination in a single qubit system with only two observable measurements. With high fidelities of state initialization, manipulation and detection, we observe a probability of successful differentiation between two non-orthogonal state preparations, which significantly exceeds the upper bound for any non-contextual theories. Our results represent the first experimental ruling out of non-contextual descriptions in a simple qubit system with only two observables. We also introduce the measurement error to quantify the contextual advantage and the tolerance over noise of quantum state discrimination. Our work not only deepen the understanding of the quantum nature of contextuality, but also motivate future research on noise effect from quantitative aspects.

Acknowledgments. We are thankful for support from the Beijing Natural Science Foundation (Grant No. Z180013), the National Natural Science Foundation of China (Grant Nos. 12074427 and 12074428), and the National Key R&D Program of China (Grant No. 2018YFA0306501). S. Zhang also thanks the China Postdoctoral Science Foundation (Grant Nos. BX20200379 and 2021M693478) for support.

References

[1] Einstein A, Podolsky B and Rosen N 1935 *Phys. Rev.* **47** 777
[2] Bell J S 1964 *Physics* **1** 195
[3] Brunner N, Cavalcanti D, Pironio S, Scarani V and Wehner S 2014 *Rev. Mod. Phys.* **86** 419
[4] Kochen S and Specker E P 1967 *J. Math. Mech.* **17** 59
[5] Budroni C, Cabello A, Güuhne O, Kleinmann M and Larsson J 2021 arXiv:2102.13036 [quant-ph]
[6] Spekkens R W 2005 *Phys. Rev. A* **71** 052108
[7] Zhang A, Xu H, Xie J, Zhang H, Smith B J, Kim M S and Zhang L 2019 *Phys. Rev. Lett.* **122** 080401
[8] Qu D, Wang K, Xiao L, Zhan X and Xue P 2021 *npj Quantum Inf.* **7** 154
[9] Jerger M, Reshitnyk Y, Oppliger M, Potočnik A, Mondal M, Wallraff A, Goodenough K, Wehner S, Juliusson K, Langford N K and Fedorov A 2016 *Nat. Commun.* **7** 12930
[10] van Dam S B, Cramer J, Taminiau T H and Hanson R 2019 *Phys. Rev. Lett.* **123** 050401
[11] Leupold F M, Malinowski M, Zhang C, Negnevitsky V, Cabello A and Alonso J 2018 *Phys. Rev. Lett.* **120** 180401
[12] Wang P F, Zhang J H, Luan C Y, Um M, Wang Y, Qiao M, Xie T, Zhang J N, Cabello A and Kim K 2022 *Sci. Adv.* **8** eabk1660
[13] Zhang X, Um M, Zhang J, An S, Wang Y, Deng D L, Shen C, Duan L M and Kim K 2013 *Phys. Rev. Lett.* **110** 070401
[14] Hasegawa Y, Loidl R, Badurek G, Baron M and Rauch H 2003 *Nature* **425** 45
[15] Mermin N D 1990 *Phys. Rev. Lett.* **65** 1838
[16] Peres A 1991 *J. Phys. A* **24** L175
[17] Cabello A 2008 *Phys. Rev. Lett.* **101** 210401
[18] Kirchmair G, Zähringer F, Gerritsma R, Kleinmann M, Gühne O, Cabello A, Blatt R and Roos C F 2009 *Nature* **460** 494
[19] Klyachko A A, Can M A, Binicioğlu S and Shumovsky A S 2008 *Phys. Rev. Lett.* **101** 020403
[20] Kleinmann M, Budroni C, Larsson J Gühne O and Cabello A 2012 *Phys. Rev. Lett.* **109** 250402
[21] Schmid D and Spekkens R W 2018 *Phys. Rev. X* **8** 011015
[22] Pusey M F, Jonathan B and Terry R 2012 *Nat. Phys.* **8** 475
[23] Bennett C H 1992 *Phys. Rev. Lett.* **68** 3121
[24] Bae J and Kwek L C 2015 *J. Phys. A* **48** 083001
[25] Bennett C H and Brassard G 1984 *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India, p 175
[26] Wootters W K and Zurek W H 1982 *Nature* **299** 802
[27] Gisin N and signaling Q C W 1998 *Phys. Lett. A* **242** 1
[28] Brunner N, Navascués M and Vértesi T 2013 *Phys. Rev. Lett.* **110** 150501
[29] Konig R, Renner R and Schaffner C 2009 *IEEE Trans. Inf. Theory* **55** 4337
[30] Helstrom C W 1969 *J. Stat. Phys.* **1** 231
[31] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880
[32] Olmschenk S, Younge K C, Moehring D L, Matsukevich D N, Maunz P and Monroe C 2007 *Phys. Rev. A* **76** 052314
[33] Shen C and Duan L M 2012 *New J. Phys.* **14** 053053
[34] Mazurek M D, Pusey M F, Kunjwal R, Resch K J and Spekkens R W 2016 *Nat. Commun.* **7** ncomms11780
[35] Mazurek M D, Pusey M F, Resch K J and Spekkens R W 2021 *PRX Quantum* **2** 020302
[36] Knill E, Leibfried D, Reichle R, Britton J, Blakestad R B, Jost J D, Langer C, Ozeri R, Seidelin S and Wineland D J 2008 *Phys. Rev. A* **77** 012307

Supplemental Material: Experimental Test of Contextuality based on State Discrimination with a Single Qubit

Qiuxin Zhang,[1] Chenhao Zhu,[1] Yuxin Wang,[1] Liangyu Ding,[1] Tingting Shi,[1] Xiang Zhang,[1,2] Shuaining Zhang,[1,2,*] and Wei Zhang[1,2,3,†]

[1]*Department of Physics, Renmin University of China, Beijing 100872, China*
[2]*Beijing Academy of Quantum Information Sciences, Beijing 100193, China*
[3]*Beijing Key Laboratory of Opto-electronic Functional Materials and Micro-nano Devices, Renmin University of China, Beijing 100872, China*

I. DATA ANALYSIS

Our data analysis is composed of two steps. The first step is called primary procedure, where the experimental data are assumed to be non-contexual and fitted by the generalized probability theory (GPT) theory. The second step, called secondary procedure, is then employed to ensure operational equivalence. The results of the two stages are substituted to test the non-contextual inequality Eq. (3) in the main text. If the inequality is violated, we conclude that the experimental data cannot be explained by any non-contexual theory. Since the derivation of the inequality and the data analysis procedure do not assume *a priori* the validity of quantum theory, an experimental violation would be considered as a general ruling out of non-contexuality, independent of any specific form of contexual theory.

Primary Procedure.– We use $D^r(\theta,\alpha)$ to denote a 6×4 matrix of the experimental raw data with parameters $\{\theta,\alpha\}$ as shown in Table 1 in the main text. For brevity, $D^r(\theta,\alpha)$ is abbreviated as D^r. The elements of the matrix are $D^r_{i,j}=P^r_{p_j,M_i}$, where p_j with $j\in\{|\phi\rangle,|\bar{\phi}\rangle,|\psi\rangle,|\bar{\psi}\rangle,y,\bar{y}\}$ is a preparation, and M_i with $i\in\{|\psi\rangle,|\phi\rangle,d|\phi\rangle,y\}$ is a measurement. According to operational measurement theory, each column $\{P^r_{p_j,M_{|\phi\rangle}},P^r_{p_j,M_{|\phi\rangle}},P^r_{p_j,M_{d|\phi\rangle}},P^r_{p_j,M_y}\}$ needs to satisfy the tomographical completeness of the measurement if and only if

$$aP^r_{p_j,M_{|\phi\rangle}}+bP^r_{p_j,M_{|\phi\rangle}}+cP^r_{p_j,M_{d|\phi\rangle}}+dP^r_{p_j,M_y}-1=0, \tag{1}$$

where $\{a,b,c,d\}$ are real numbers. However, in an actual experiment with inevitable errors such as imperfect state readout, the above limitations can not be strictly met.

To solve this issue, we adopt the methods given in Ref. [1] to express this process as a minimization problem

$$\min_{\left\{P^p_{p_j,M_i},a,b,c,d\right\}}\chi^2=\sum_{i=1}^{6}\chi_i^2 \tag{2}$$

subject to the condition

$$aP^p_{p_j,M_{|\phi\rangle}}+bP^p_{p_j,M_{|\phi\rangle}}+cP^p_{p_j,M_{d|\phi\rangle}}+dP^p_{p_j,M_y}-1=0 \tag{3}$$

for $j\in\{|\phi\rangle,|\bar{\phi}\rangle,|\psi\rangle,|\bar{\psi}\rangle\}$. The parameter χ_i is defined as

$$\chi_i=\sqrt{\sum_{j=1}^{4}\frac{P^p_{p_j,M_i}-P^r_{p_j,M_i}}{\left(\Delta P^p_{p_j,M_i}\right)^2}}, \tag{4}$$

where $\Delta P^p_{p_j,M_i}$ is the uncertainty determined by $\sqrt{P^p_{p_j,M_i}(1-P^p_{p_j,M_i})/N}$ with N the repeat times of the same sequence. In our experiment, we take $N=1000$ for all data sets. The coefficients $\{a,b,c,d\}$ can be obtained according to the Langrange multipliers method, and $P^p_{p_j,M_i}$'s are the elements of new matrix D^p which satisfies the condition of tomographical completeness in statistics.

* zhangshuaining@ruc.edu.cn

† wzhangl@ruc.edu.cn

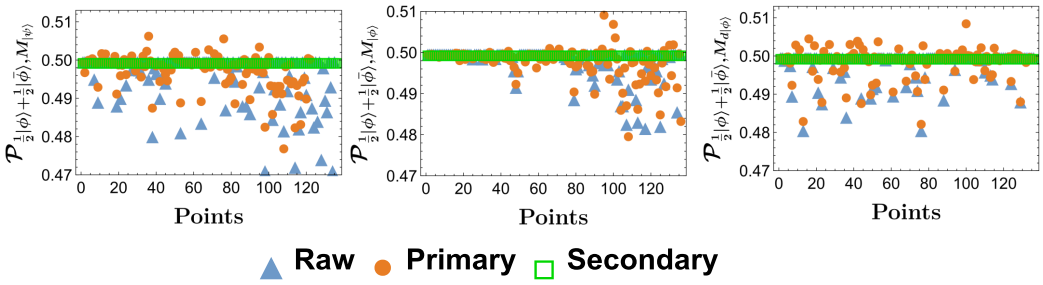

FIG. 1. The results of data analysis for operational equivalence. $\mathcal{P}_{\frac{1}{2}|\phi\rangle+\frac{1}{2}|\bar{\phi}\rangle,M_i}$ satisfies operational equivalence for all the points and all the three measurements, that is, they are all equal to 1/2. Each point represents one preparation and measurement setup, i.e., a specific choice of the angle of $\pi - 2\theta$ between the two states and the deviation angle α of the measurement axes, as illustrated in Figs. 1(a) and 1(b) of the main text. The data points are labeled in the following sequence: first scan θ within the range of inequality Eq. 5 of the main text with a fixed α, then vary the angle α afterwards.

Secondary Procedure.– This process is to satisfy operational equivalence for data with

$$\mathcal{P}_{\frac{1}{2}\phi+\frac{1}{2}\bar{\phi},M_i} = \mathcal{P}_{\frac{1}{2}\psi+\frac{1}{2}\bar{\psi},M_i} = \frac{1}{2}, \tag{5}$$

where

$$\mathcal{P}_{\frac{1}{2}|\psi(\phi)\rangle+\frac{1}{2}|\psi(\bar{\phi})\rangle,M_i} = \frac{1}{2}\left[P\left(1 \mid P_{|\psi(\phi)\rangle}, M_i\right) + P\left(1 \mid P_{|\psi(\bar{\phi})\rangle}, M_i\right)\right]$$

and $M_i \in \left\{M_{|\psi\rangle}, M_{|\phi\rangle}, M_{d|\phi\rangle}\right\}$. Because of state readout error, the experimental raw data in general cannot satisfy Eq. (5). Besides, the primary procedure will make slight changes of the data under constraint in Eq. (2), which may also compromise the requirement of operational equivalence. This can be seen in Fig. 2 of the main text and Fig. (1).

To overcome this issue, we perform the secondary procedure by noticing that the state $P^s_{p_j,M_i}$ can be formed by linear combinations of $P^p_{p_j,M_i}$. This linear transformation can be expressed as a matrix $D^s = UD^p$, where U is a 4×6 matrix representing the linear transformation and the resulting D^s is a 4×4 matrix. We also require that the new states formed by linear combination are as close to the initial states as possible. This translates to maximizing

$$\max D_{\text{ps}} \equiv \frac{1}{4}\sum_{v=1}^{4} U_{vv} \tag{6}$$

with constrains of operational equivalence Eq. (5)) and $\sum_{v=1}^{6} U_{wv} = 1$. The quantity D_{ps} measures the *similarity* between primary matrix D^p and secondary matrix D^s. The results of data analysis are shown in Fig. 2 of the main text and Fig. 1, where all secondary points satisfy the operational equivalence. The average similarity of the total 136 points is 97.13% ± 2.3%.

II. ERROR ANALYSIS

Data correction and Sampling Error.– The detection error of the experiment is corrected by the error matrix [2], and the uncertainty will be corrected accordingly. The error matrix (M_E) is defined as

$$M_E = \begin{pmatrix} 1-p_0 & p_1 \\ p_0 & 1-p_1 \end{pmatrix}, \tag{7}$$

where p_0 represents the error of detecting $|\uparrow\rangle$, p_1 represents the error of detecting $|\downarrow\rangle$. In our experiment, $p_0 = 2.08\%$ and $p_1 = 1.71\%$. An ideal detection of $\{|\uparrow\rangle,|\downarrow\rangle\}$ can be denoted by $V_{\text{ideal}} = \{f^i_{|\uparrow\rangle}, f^i_{|\downarrow\rangle}\}^{-1}$. And an experimental detection of $\{|\uparrow\rangle,|\downarrow\rangle\}$ can be denoted by $V_{\text{exp}} = \{f^e_{|\uparrow\rangle}, f^e_{|\downarrow\rangle}\}^{-1}$. The principle of error correction is based on the formula $M_E V_{\text{ideal}} = V_{\text{exp}}$. The value of

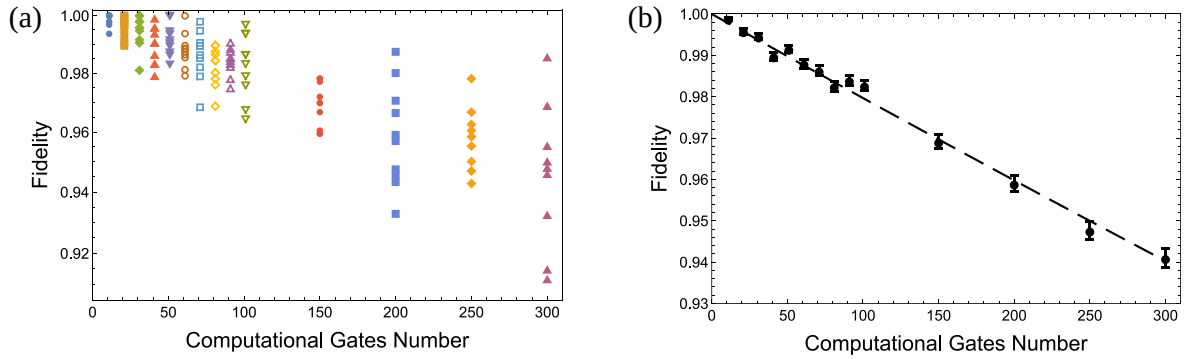

FIG. 2. The results of randomized benchmarking. We test random sequences of computational logic gates of length 11, 21, 31, 41, 51, 61, 71, 81, 91, 101, 150, 200, 250, 300, respectively. The horizontal coordinate is the computational logic gate number and the vertical coordinate is the fidelity. (a) For each computational logic gate number, we generate 10 different random sequences and obtain the results. (b) The result obtained by averaging the results of the 10 random sequences of the same computational logic gate in (a). We fit the data to an exponential function L^x whose result is drawn in a dashed line. According to the fitted parameter x, the error rate is $(2.0 \pm 0.3)\times10^{-4}$, implying that the average fidelity of a single logic gate is 99.980(3)%.

V_{ideal} can be found by maximum likelihood method with constraint $f^i_{|\uparrow\rangle} + f^i_{|\downarrow\rangle} = 1$. And the uncertainty is

$$\sigma = \sqrt{\frac{(M_E^{-1})^2 V_{\text{exp}} - V_{\text{ideal}}^2}{N}}, \tag{8}$$

where N is repetition times of the same experimental sequence. In our experiment, N is taken to be 1000, and the average σ is $1.29\% \pm 0.45\%$.

Single Qubit Gate Error.– We describe the performance of gate operation and quantify the gate error by the industry-standard randomized benchmarking technique [3]. Randomized benchmarking is a widely used and scalable technique which uses random sequences of varying length on quantum state to estimate per gate error.

A gate sequence of length L contains L computational gates uniformly drawn from $e^{i(\sigma_i)\pi/4}$ with $i = x, y$, and $L+1$ interleaving identity or Pauli operations chosen from $e^{i(\sigma_i)\pi/2}$. The purpose of introducing $L + 1$ Pauli randomization gates is to depolarize noise that relates to specific computation gate. For each sequence length L, we choose ten different sequences to act on the trapped ion system and measure the state fidelity between the ideal and prepared final states. We average the state fidelities over sequences of the same length and numerically fit the average fidelity with an exponential function with length up to 300. The fitted error rate per single qubit gate is $(2.0 \pm 0.3)\times10^{-4}$, as shown in Fig. 2.

[1] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, Nat. Commun. **7**, 11780 (2016).
[2] C. Shen and L.-M. Duan, New J. Phys. **14**, 053053 (2012).
[3] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).