Contextuality witness inspired by optimal state discrimination

Carles Roch i Carceller[*] and Jonatan Bohr Brask
Center for Macroscopic Quantum States bigQ, Department of Physics, Technical University of Denmark, Fysikvej 307, 2800 Kgs. Lyngby, Denmark

(Received 1 November 2023; revised 15 February 2024; accepted 16 February 2024; published 4 March 2024)

Many protocols and tasks in quantum information science rely inherently on the fundamental notion of contextuality to provide advantages over their classical counterparts, and contextuality represents one of the main differences between quantum and classical physics. In this work we present a witness for preparation contextuality inspired by optimal two-state discrimination. The main idea is based on finding the accessible averaged success and error probabilities in both classical and quantum models. We can then construct a noncontextuality inequality and associated witness which we find to be robust against depolarising noise and loss in the form of inconclusive events.

DOI: 10.1103/PhysRevA.109.032202

I. INTRODUCTION

In classical physics, the properties of physical objects can be assumed to exist independently of any observation. However, quantum mechanics shows that attributes of physical systems do not exist predeterminedly, in the sense that it is not generally possible to consistently assign values to measurable quantities that are independent of which other quantities are jointly measured. This impossibility of reproducing the predictions of quantum mechanics with models that assign values independent of the measurement context is known as quantum contextuality [1]. This concept originates with the Bell-Kochen-Specker theorem [2,3], which demonstrates that quantum theory is incompatible with noncontextual hidden-variable models. It has been demonstrated that contextuality constitutes a resource for various applications in quantum information including magic states [4], quantum key distribution [5], device-independent security [6], and quantum randomness certification [7,8]. The traditional definition of contextuality requires a composite system, and its standard proof applies to Hilbert spaces of dimension three or higher [9,10]. The notion of (non)contextualilty has been further generalized in Ref. [11], based on operational equivalences and ontological models. Similarly to Kochen-Specker, generalized contextuality has also been proven to provide a resource for certain quantum information tasks. For instance, parity-oblivious multiplexing [12,13], random-access codes [14], quantum randomness certification [15], communication [16–18], and state discrimination [19,20]. Quantum theory has also been shown to be less preparation contextual than the general operational theory known as box world [21].

In this work we aim to find a simple witness for generalised contextuality in the sense introduced in Ref. [11]. While a number of contextuality witnesses exist in the literature [22–27], here we benefit from the simplicity of prepare-and-measure scenarios with two preparations and a single measurement. Such scenarios are of wide importance for both fundamental studies of quantum mechanics and applications in quantum technology including sensing, communication, and randomness generation [28–30]. We then find a contextuality witness in this framework, which is inspired by optimal two-state discrimination. Our results show that contextuality can be witnessed in the presence of both significant depolarizing noise and loss.

The rest of the paper is organized as follows. In Sec. II we give a brief introduction to the basic notions in state discrimination in a theory-independent manner. We continue in Sec. III presenting the prepare-and-measure scenario and the goal that defines the main state discrimination task to properly define the witness independently in both quantum and noncontextual models. Finally, we discuss the main results of the paper in Sec. IV and conclude the work in Sec. V.

II. BASIC NOTIONS IN STATE DISCRIMINATION

Any state discrimination scenario is formed by state preparations and effects [31,32]. The first are labeled by preparation procedures $x \in X$ and the second by the possible answers $b \in B$ to the questions in X. The gathered data are usually expressed as conditional probabilities (correlations) $p(b|x)$. The goal in state discrimination is to determine x from the transmitted states, i.e., to achieve $b = x$. For any particular model (e.g., quantum or noncontextual), an optimization problem can be built obeying the constraints of that model. As is customarily done in state discrimination settings, we denote the probability $p(b = x|x)$ of a correct answer the *success* probability, whereas $p(b \neq x|x)$ for $b \in X$ is called the *error* probability. One must also consider events where the answer b is not in the set of questions X (i.e., $X \subset B$). We group answers not in X and label them by $b = \emptyset$. We denote $p(b = \emptyset|x)$ the *inconclusive* probability.

Success, error, and inconclusive probabilities each play a different role in the discrimination scenario [28,33,34].

[*]crica@dtu.dk

Different state discrimination tasks can be defined by different figures of merits, which are functions of the observed conditional probabilities, and different constraints on these probabilities. For example, the goal in minimum-error state discrimination (MESD) is to maximize the success probability whilst inconclusive events do not occur [35–37] (hence converting the goal into a minimization of the error probability due to normalization). On the other hand, in unambiguous state discrimination (USD), the goal is also to maximize the success probability, with the main constraint that error probabilities must vanish [38–40] (thus converting the goal into a minimization of inconclusive probabilities). Lastly, in maximum-confidence state discrimination (MCSD), the goal is to maximize the confidence C, i.e., the probability of receiving input x given the outcome $b = x$, which can be expressed as the success probability divided by the rate of events of interest [41–46]. Concretely, $C := p_x p(x|x)/\eta_x$, for $\eta_b = \sum_x p_x p(b|x)$, where p_x are the prior probabilities for each preparation x. No further constraints are applied to MCSD, making it rather a more general approach. Also, it can be reduced to MESD and USD as particular cases. If $C = 1$ the input x must be unambiguously identified, resulting in USD, while MESD is recovered by adopting $\sum_x \eta_x C_x$ as the figure of merit.

III. SCENARIO

In the following, we focus on two-state discrimination, characterized by the sets of preparations $X = \{0, 1\}$, considered equiprobable, and outcomes $B = \{0, 1, \o\}$. We also introduce the averaged success p_{suc}, error p_{err}, and inconclusive p_{inc} probabilities as

$$p_{\mathrm{suc}} := \tfrac{1}{2}[p(0|0) + p(1|1)], \tag{1}$$

$$p_{\mathrm{err}} := \tfrac{1}{2}[p(1|0) + p(0|1)], \tag{2}$$

$$p_{\mathrm{inc}} := \tfrac{1}{2}[p(\o|0) + p(\o|1)] = 1 - p_{\mathrm{suc}} - p_{\mathrm{err}}. \tag{3}$$

We will fix p_{inc} and ask the following question: which regions in correlation space, parameterized by p_{suc} and p_{err}, are feasible in quantum mechanics or in a noncontextual model? The answer to this question is not trivial if state preparations are not perfectly distinguishable. For a fixed inconclusive rate, the sum $p_{\mathrm{suc}} + p_{\mathrm{err}} = 1 - p_{\mathrm{inc}}$ is fixed and we can focus on the difference. We therefore define the following witness on the level of probabilities

$$\mathcal{W} := \tfrac{1}{2}(p_{\mathrm{suc}} - p_{\mathrm{err}}). \tag{4}$$

For each model, we will separately formulate an optimization problem and find a bound on $\mathcal{W}$. The feasible region is necessarily convex since, for two different measurement strategies producing different behaviors, probabilistically choosing between them (using local randomness) defines another valid measurement strategy. The corresponding behavior will then be the convex combination of the first two behaviors. We can thus use techniques in convex optimization to efficiently solve the maximization problem for each model.

A. Quantum model

Consider an ensemble of two noisy states $\rho_x = r_s |\psi_x\rangle\langle\psi_x| + (1 - r_s)\mathbb{1}/2$ for $x = 0, 1$, with distinguishability characterized by the overlap $\delta = |\langle\psi_0|\psi_1\rangle|$. Let $\hat{\pi}_b$ represent a valid positive operator-valued measure (POVM) for $b = 0, 1, \o$, such that $\mathrm{Tr}[\rho_x\hat{\pi}_b] = p(b|x)$. Our goal is to find the maximum difference between success and error probabilities for a fixed inconclusive rate. To do so, let us introduce the following operator

$$\hat{\Delta}_x := \frac{(-1)^x}{2}(\hat{\pi}_0 - \hat{\pi}_1). \tag{5}$$

We must find the maximum difference

$$\mathcal{W}^{\mathrm{Q}} := \max \frac{1}{2}\left(p^{\mathrm{Q}}_{\mathrm{suc}} - p^{\mathrm{Q}}_{\mathrm{err}}\right) = \max \sum_x \mathrm{Tr}[\hat{\Delta}_x\rho_x], \tag{6}$$

where the optimization is over all measurements forming valid POVMs, $\hat{\pi}_b \geqslant 0$ and $\sum_b \hat{\pi}_b = \mathbb{1}$ and subject to $p_{\mathrm{inc}} = \frac{1}{2}\mathrm{Tr}[(\rho_0 + \rho_1)\hat{\pi}_\o]$. This maximization can be rendered as a semi-definite program (SDP) [47].

In Ref. [48] we found an analytical form of the optimal measurement. The solution to Eq. (6) is given by

$$\mathcal{W}^{\mathrm{Q}} = \frac{r_s}{2}\sqrt{(1 - \delta^2)\left(1 - \frac{2p_{\mathrm{inc}}}{1 + r_s\delta}\right)} \quad \text{for } p_{\mathrm{inc}} \leqslant r_s\delta,$$
$$\mathcal{W}^{\mathrm{Q}} = \frac{r_s}{2}\sqrt{1 - \delta^2}\sqrt{1 - r_s^2\delta^2}\,\frac{1 - p_{\mathrm{inc}}}{1 - r_s^2\delta^2} \quad \text{for } p_{\mathrm{inc}} \geqslant r_s\delta. \tag{7}$$

One can also write down the optimal success and error probabilities. For $p_{\mathrm{inc}} \leqslant r_s\delta$,

$$p^{\mathrm{Q}}_{\mathrm{suc}} = \frac{1}{2}\left(1 + 2\mathcal{W}^{\mathrm{Q}} - p_{\mathrm{inc}}\frac{1 + \delta}{1 + r_s\delta}\right), \tag{8}$$

and for $p_{\mathrm{inc}} \geqslant r_s\delta$

$$p^{\mathrm{Q}}_{\mathrm{suc}} = \tfrac{1}{2}(1 + 2\mathcal{W}^{\mathrm{Q}} - p_{\mathrm{inc}}), \tag{9}$$

and $p^{\mathrm{Q}}_{\mathrm{err}} = 1 - p^{\mathrm{Q}}_{\mathrm{suc}} - p_{\mathrm{inc}}$.

The success and error probabilities we found are the maximal and minimal probabilities according to quantum theory in a qubit state-discrimination problem. Interestingly, one can recover the bounds from other protocols as specific cases. For instance, if the experiment only produces conclusive outcomes ($p_{\mathrm{inc}} = 0$), the problem is reduced to the usual MESD. Then, one recovers the Helstrom bound as a minimum error rate p_{err} [36,49,50]. On the other hand, if the experiment is designed with a null error rate ($p_{\mathrm{err}} = 0$) and zero noise ($r_s = 1$), one recovers USD. In that case, the maximal success probability is $p_{\mathrm{suc}} = 1 - \delta$, leaving a minimal rate of inconclusive events $p_{\mathrm{inc}} = \delta$, the minimal value for USD [51,52]. Finally, one can directly compute the maximum confidence of the whole ensemble by writing $C = p_{\mathrm{suc}}/(p_{\mathrm{suc}} + p_{\mathrm{err}})$. One recovers the maximum confidence obtained in Ref. [20] if $p_{\mathrm{inc}} \leqslant r_s\delta$. For larger values of the inconclusive rate, one can still compute the maximum confidence with the same formula since MCSD and the present scheme share the exact same goal (maximize the success and minimize the error probabilities).

B. Noncontextual model

We now outline a noncontextual ontological model for the prepare-and-measure scenario [11,19,53]. The system is associated with an ontic state space Ω in which each point λ completely defines all physical properties, i.e., the outcomes of all possible measurements. Each state preparation x samples the ontic state space according to a probability distribution $\mu_x(\lambda)$, referred to as the *epistemic state*. Each measurement is defined by a set of *response functions*, that is, nonnegative functions $\xi_b(\lambda)$ over the ontic space, such that $\sum_b \xi_b(\lambda) = 1$ for all $\lambda \in \Omega$. The probability of obtaining the outcome b when state μ_x was prepared is then

$$p(b|x) = \int_\Omega d\lambda\ \mu_x(\lambda)\xi_b(\lambda). \tag{10}$$

While distinct ontic states can be perfectly discriminated, epistemic states with overlapping distributions cannot. Distinguishability can be quantified in terms of the *confusability* between two epistemic states μ_x and μ_y:

$$c_{xy} := \int_{\mathrm{supp}[\mu_x(\lambda)]} d\lambda\ \mu_y(\lambda). \tag{11}$$

It is the discrimination of epistemic states which we compare against quantum state discrimination.

Furthermore, we require the ontological model to be preparation-noncontextual. Two preparations are said to be operationally equivalent if they cannot be distinguished by any measurement, and an ontological model is said to be preparation-noncontextual if all operationally equivalent preparations are assigned to the same epistemic state. To impose noncontextuality on the ontological model, we assume the existence of a particular pair of pure states $S := \{\mu_0, \mu_1\}$ and complementary states $S^\perp := \{\mu_0^\perp, \mu_1^\perp\}$, i.e., μ_x and $\mu_x^\perp$ have nonoverlapping supports $\mu_x(\lambda)\mu_x^\perp(\lambda) = 0\ \forall\,\lambda$. The pairwise confusability c_{01} of μ_0 and μ_1 is the same as for $\mu_0^\perp$ and $\mu_1^\perp$. Preparation noncontextuality implies that preparing μ_x and $\mu_x^\perp$ with equal probability for $x = 0$ or $x = 1$ must be equivalent [11,19], that is, $\frac{1}{2}\mu_0 + \frac{1}{2}\mu_0^\perp = \frac{1}{2}\mu_1 + \frac{1}{2}\mu_1^\perp$. This statement implies that any pair from the set of states $\{\mu_0, \mu_1, \mu_0^\perp, \mu_1^\perp\}$ are equal on their overlap, i.e.,

$$\mu_0(\lambda) = \mu_1(\lambda)\ \forall\lambda \in \mathrm{supp}[\mu_0(\lambda)] \cup \mathrm{supp}[\mu_1(\lambda)]\ \forall\,\lambda, \tag{12}$$

and similarly for the other pairs. This, in turn, results in symmetric confusabilities, $c_{01} = c_{10} := c$. Quantum and noncontextual models can be then compared through $\delta^2 = c$. The noncontextual model we use in this work can be understood as an attempt to describe quantum theory, and in general, it will reproduce some quantum correlations but not all, as we explore below.

We now present the main problem in a noncontextual model. The two preparations are represented by the following epistemic states affected by depolarizing noise

$$\tilde{\mu}_0(\lambda) = r_s\mu_0(\lambda) + (1 - r_s)\mu_{\mathbb{1}/2}(\lambda),$$
$$\tilde{\mu}_1(\lambda) = r_s\mu_1(\lambda) + (1 - r_s)\mu_{\mathbb{1}/2}(\lambda). \tag{13}$$

These can be characterized by the confusability of the noiseless epistemic states $c := c_{10}$ from Eq. (11). We also consider a single measurement with two conclusive outcomes $b = 0, 1$ and an inconclusive result $b = \emptyset$, represented by the response functions $\xi_b(\lambda)$. Let us define the analogous observable to Eq. (5)

$$\Delta_x^{\mathrm{NC}}(\lambda) := \frac{(-1)^x}{2}[\xi_0(\lambda) - \xi_1(\lambda)]. \tag{14}$$

Then, we can rewrite the problem as a maximization

$$\mathcal{W}^{\mathrm{NC}} := \max \frac{1}{2}\left(p_{\mathrm{suc}}^{\mathrm{NC}} - p_{\mathrm{err}}^{\mathrm{NC}}\right)$$
$$= \max \sum_x \int_\Omega d\lambda\ \tilde{\mu}_x(\lambda)\Delta_x^{\mathrm{NC}}(\lambda), \tag{15}$$

subject to $\xi_b(\lambda)$ being valid response functions, $\xi_b(\lambda) \geqslant 0$ and $\sum_b \xi_b(\lambda) = 1$, $\forall\lambda$, and a given rate of inconclusive events $p_{\mathrm{inc}} = \frac{1}{2}\int d\lambda[\tilde{\mu}_0(\lambda) + \tilde{\mu}_1(\lambda)]\xi_\emptyset(\lambda)$. In Ref. [48] we showed how this maximization can be rendered as a simple linear problem, for which we are able to find an analytical solution

$$\mathcal{W}^{\mathrm{NC}} = \frac{r_s}{2}(1 - c)\left(1 - \frac{p_{\mathrm{inc}}}{1 + r_s c}\right) \quad \text{for } p_{\mathrm{inc}} \leqslant (1 + r_s c)/2,$$
$$\mathcal{W}^{\mathrm{NC}} = \frac{1}{2}(1 - p_{\mathrm{inc}})\frac{r_s(1 - c)}{1 - r_s c} \quad \text{for } p_{\mathrm{inc}} \geqslant (1 + r_s c)/2. \tag{16}$$

This results in the following success and error probabilities. For $p_{\mathrm{inc}} \leqslant (1 + r_s c)/2$,

$$p_{\mathrm{suc}}^{\mathrm{NC}} = \frac{1 + r_s}{2}\left(1 - \frac{p_{\mathrm{inc}}}{1 + r_s c}\right) - \frac{r_s c}{2}, \tag{17}$$

and for $p_{\mathrm{inc}} \geqslant (1 + r_s c)/2$,

$$p_{\mathrm{suc}}^{\mathrm{NC}} = \tfrac{1}{2}(1 + 2\mathcal{W}^{\mathrm{NC}} - p_{\mathrm{inc}}), \tag{18}$$

and $p_{\mathrm{err}}^{\mathrm{NC}} = 1 - p_{\mathrm{suc}}^{\mathrm{NC}} - p_{\mathrm{inc}}$.

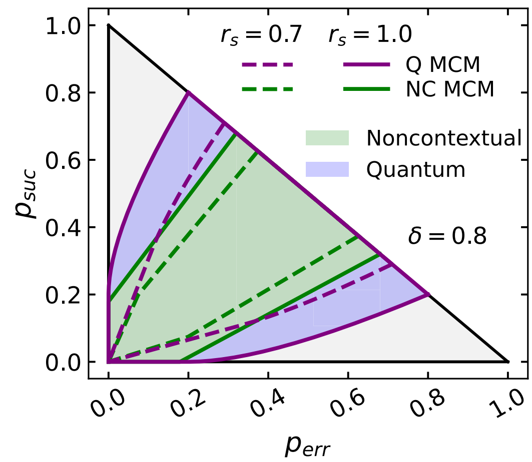

FIG. 1. Space of probabilities corresponding to a two-state discrimination setting. Continuous lines denote maximum confidence measurements in both quantum (purple) and noncontextual (green) models. Even with a bounded value of noise ($r_s = 0.7$), the MCM line according to the quantum model falls outside the noncontextual region.

IV. DISCUSSION

In Fig. 1 we show the achievable probabilities in quantum and noncontextual models. The white region delimited by

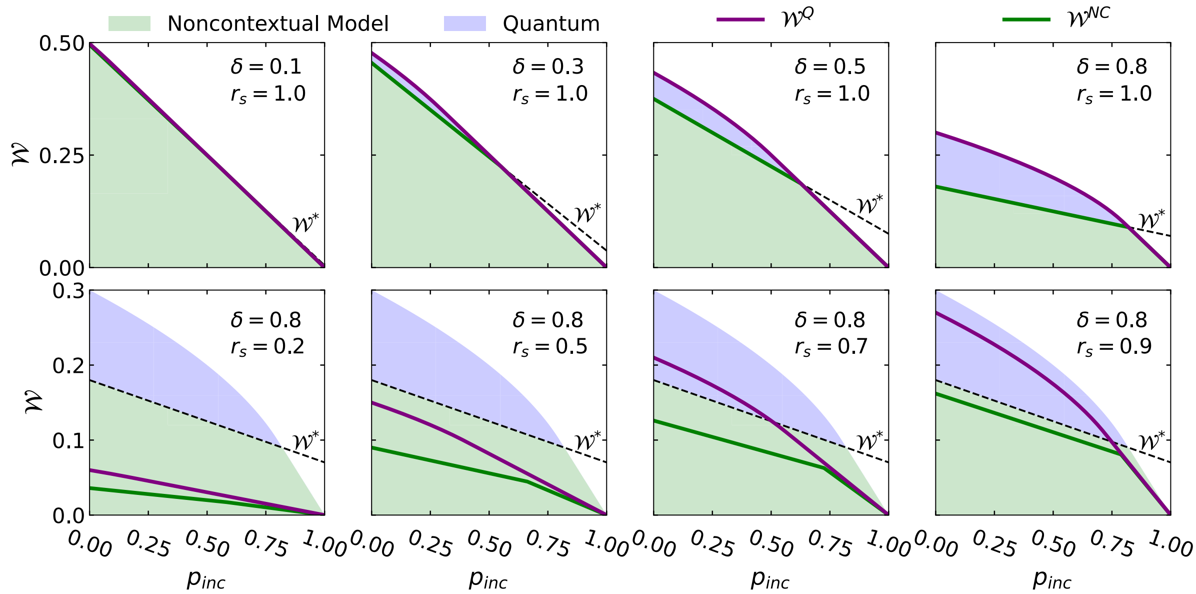

FIG. 2. Bounds on the witness $\mathcal{W}$ according to quantum and noncontextual models. On the first row we show noiseless cases with different overlaps. Below, in the second row, we fix a particular overlap and show the effects of depolarising noise on the preparation. The green area denotes the feasible values according to quantum and noncontextual models and the blue region solely for the quantum model. The black-dashed line shows the contextuality witness $\mathcal{W}^*$ in Eq. (19). Any behavior above $\mathcal{W}^*$ is an evidence of contextuality.

the black contour shows the feasible space in the case of fully distinguishable preparations. That is, when states can be directly identified with ontic states λ. The area shaded in blue shows the feasible space according to quantum theory. In its contour we find $p_{\rm suc}^{\rm Q}$ from Eq. (8), for $r_s = 1$. The region reproducible by the noncontextual model (green area) is contained in the quantum set. Similarly, we find $p_{\rm suc}^{\rm NC}$, from Eq. (17), in its contour, also for $r_s = 1$. We see that the quantum predictions depart from classical (noncontextual) interpretations. Increasing the overlap $\delta = \sqrt{c}$, this distinction becomes more pronounced, and at the same time both the quantum and noncontextual feasible spaces shrink.

Moreover, we can identify some extremes of the quantum region with the bounds found in each state discrimination protocol. The diagonal line that delimits the upper-right part of the feasible regions covers the state discrimination scenarios with zero inconclusive rates. The vertices of the quantum region on that line reproduce the Helstrom bound [49,54,55] obtained in MESD. The same applies for the vertices corresponding to the noncontextual line, which reproduce the maximal success probabilities in MESD obtained in Ref. [19]. Also, the maximal $p_{\rm err}$ and $p_{\rm suc}$ on the flat part of the bottom and left-most boundary, respectively, reproduce the maximal unambiguous error and success rates for quantum [38] and noncontextual [20] models, obtained in USD. Finally, the entire quantum boundary (purple line) corresponds to a maximum confidence measurement (MCM) [42,56] for the ensemble of qubit states (i.e., maximizing the average confidence). MCM thus provides optimal success probability in any (qubit-)state discrimination scenario. Similarly, the noncontextual boundary is obtained by a noncontextual MCM. We can see, by writing $C = p_{\rm suc}/(p_{\rm suc} + p_{\rm err})$, that the confidence coincides with the bounds found in the literature [20,41–43,56].

When depolarizing noise is included, the bounds on all protocols depart from the borders of the quantum region. The Helstrom bound from both quantum and noncontextual MESD comes closer to the center of the probability space as noise increases. Also, when noise is taken into account, USD is not possible as here we can see that the bottom and left-most borders are not reachable. The space enclosed by the MCM lines also narrows. Indeed, noise makes the prepared physical states less distinguishable in both models. For a given noisy ensemble, the points on the quantum region outside the MCM lines are not accessible.

A different perspective is plotted in Fig. 2. Contextual behavior is manifested in the blue -shaded region above the dashed black line, which corresponds to the inequality

$$\mathcal{W} \leqslant \mathcal{W}^* = \mathcal{W}^{\rm NC}|_{r_s=1} = \frac{1-\delta^2}{2}\left(1 - \frac{p_{\rm inc}}{1+\delta^2}\right), \qquad (19)$$

for noiseless preparations with distinguishability bounded by overlap δ (quantum) or confusability δ^2 (noncontextual). Here, $\mathcal{W}^*$ is the noncontextual bound without noise. Note that it is sufficient to lower-bound the confusability (or equivalently the overlap) because our bounds on $\mathcal{W}$ decrease as preparations become less distinguishable. Also, note that the witness places no assumptions on the measurement, which is completely uncharacterised. We see, from the two lower-right plots in Fig. 2 that noncontextuality can be witnessed in the presence of fairly high values of noise (for example, $r_s = 0.7$ equivalent to 30% depolarizing noise).

We finally look at the amount of depolarizing noise that our witness can tolerate while still being able to detect contextuality. Depolarizing noise in the preparation is parameterized through r_s [see above Eq. (5)]. In Fig. 3 we show the minimum tolerable r_s for which contextuality can be witnessed, i.e., for which $\mathcal{W}^{\rm Q} \geqslant \mathcal{W}^*$. Two noiseless preparations with

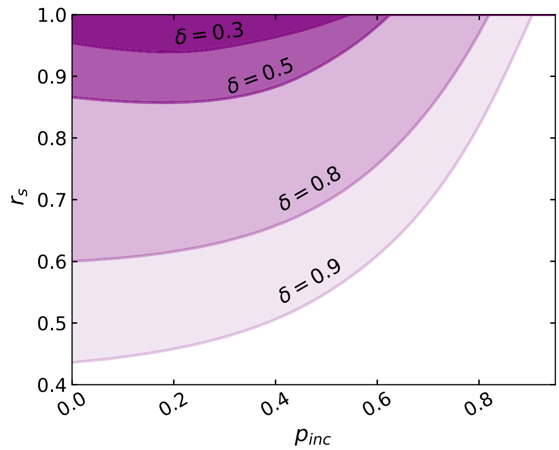

FIG. 3. Tolerable amount of depolarizing noise for which our witness can detect quantum contextuality as function of inconclusive rate p_{inc} for different overlaps δ. Contextuality is witnessed in the shaded regions above the solid lines (larger r_s means less noise).

confusability $c = \delta^2$ in a noncontextual model can reproduce all quantum correlations from a two-state discrimination scenario with fixed p_{inc}, as long as r_s is below the plotted lines. Noise tolerance is higher for larger confusabilities. The cases with null inconclusive outcomes are covered in Refs. [19,57]. Remarkably, observe that for low noise, the appearance of inconclusive events strengthens robustness, as is, e.g., the case for $\delta = 0.3$ or $\delta = 0.5$ in Fig. 3. An intuitive explanation for this might be that, in the absence of noise, the optimal rate of inconclusive events is $p_{\text{inc}} = \delta$ (USD is impossible for lower p_{inc}). With noise, however, the error probability cannot be zero, which allows for lower optimal inconclusive probabilities. This is reflected in Fig. 3, where the minima in each curve is given by an inconclusive probability lower than δ.

V. CONCLUSION

In this work, we presented a witness of contextuality in two-state discrimination scenarios. We started by formulating the problem of finding the optimal measurement in two-state discrimination settings. We consider a measurement to be optimal if it maximizes the difference between success and error probabilities. This leads to correlations reaching the boundary of the feasible space parameterized by success and error probabilities. Parametrizing the correlation space in this manner allows us to clearly distinguish the feasible sets for the quantum and noncontextual models. Additionally, allowing for inconclusive events we found that maximum-confidence measurements are optimal in both quantum and noncontextual models. That led to the definition of the witness $\mathcal{W}$ in Eq. (4) and the inequality Eq. (19). This inequality allows for a flexible rate of inconclusive events (e.g., due to losses) and is robust against depolarizing noise, as we show in Fig. 3. Even more importantly, our results show that, in some cases, incorporating inconclusive results can help strengthening the noise tolerance in terms of witnessing quantum contextuality. Thus, our results open new avenues for exploring contextual advantages in realistic scenarios, using inconclusive results as a benefit towards noise robustness. The results we present in this work are explicitly derived for two state discrimination scenarios. However, we strongly believe they can be generalized to multiple state discrimination scenarios. Although we left this as an open question out of the scope of this work, intuitively, it can be done in a twofold manner, depending on the goal in the discrimination task. If one is interested in discriminating all states equally, the problem can be reduced to pairwise state discrimination subtasks. Hence reducing an N state to an individual two-state discrimination scenario as presented in this work. Otherwise, if one aims to distinguish one state from the ensemble, the problem can again be reduced to the discrimination between the state of interest and the mixture of the rest of the ensemble. This is equivalent to our scenario, but only one of the states is affected by depolarizing noise. Thus, at the end of the day, one can think of more general cases and find our results still applicable.

ACKNOWLEDGMENTS

We gratefully acknowledge support from the Danish National Research Foundation, Center for Macroscopic Quantum States (bigQ, DNRF142), VILLUM FONDEN (research grant 40864), the Carlsberg Foundation CF19-0313, and a DTU-KAIST Alliance Stipend.

[1] C. Budroni, A. Cabello, O. Gühne, M. Kleinmann, and J.-A. Larsson, Kochen-Specker contextuality, Rev. Mod. Phys. **94**, 045007 (2022).

[2] S. Kochen and E. Specker, The problem of hidden variables in quantum mechanics, Indiana Univ. Math. J. **17**, 59 (1968).

[3] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966).

[4] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the 'magic' for quantum computation, Nature (London) **510**, 351 (2014).

[5] H. Bechmann-Pasquinucci and A. Peres, Quantum cryptography with 3-state systems, Phys. Rev. Lett. **85**, 3313 (2000).

[6] K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, M. Pawlowski, and M. Bourennane, Contextuality offers device-independent security, arXiv:1006.0468.

[7] A. A. Abbott, C. S. Calude, J. Conder, and K. Svozil, Strong Kochen-Specker theorem and incomputability of quantum randomness, Phys. Rev. A **86**, 062109 (2012).

[8] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, New J. Phys. **18**, 045003 (2016).

[9] A. A. Klyachko, M. A. Can, S. Binicioğlu, and A. S. Shumovsky, Simple test for hidden variables in spin-1 systems, Phys. Rev. Lett. **101**, 020403 (2008).

[10] A. Cabello, P. Badziąg, M. Terra Cunha, and M. Bourennane, Simple hardy-like proof of quantum contextuality, Phys. Rev. Lett. **111**, 180404 (2013).

[11] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[12] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation contextuality powers parity-oblivious multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).

[13] S. Ghorai and A. K. Pan, Optimal quantum preparation contextuality in an n-bit parity-oblivious multiplexing task, Phys. Rev. A **98**, 032110 (2018).

[14] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, Parity oblivious d-level random access codes and class of noncontextuality inequalities, Quant. Info. Proc. **18**, 111 (2019).

[15] C. Roch i Carceller, K. Flatt, H. Lee, J. Bae, and J. B. Brask, Quantum vs noncontextual semi-device-independent randomness certification, Phys. Rev. Lett. **129**, 050501 (2022).

[16] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Communication games reveal preparation contextuality, Phys. Rev. Lett. **119**, 220402 (2017).

[17] D. Saha, P. Horodecki, and M. Pawłowski, State independent contextuality advances one-way communication, New J. Phys. **21**, 093057 (2019).

[18] D. Saha and A. Chaturvedi, Preparation contextuality as an essential feature underlying quantum communication advantage, Phys. Rev. A **100**, 022108 (2019).

[19] D. Schmid and R. W. Spekkens, Contextual advantage for state discrimination, Phys. Rev. X **8**, 011015 (2018).

[20] K. Flatt, H. Lee, C. R. Roch i Carceller, J. B. Brask, and J. Bae, Contextual advantages and certification for maximum-confidence discrimination, PRX Quantum **3**, 030337 (2022).

[21] M. Banik, S. S. Bhattacharya, A. Mukherjee, A. Roy, A. Ambainis, and A. Rai, Limited preparation contextuality in quantum theory and its relation to the Cirel'son bound, Phys. Rev. A **92**, 030103(R) (2015).

[22] A. Krishna, R. W. Spekkens, and E. Wolfe, Deriving robust noncontextuality inequalities from algebraic proofs of the Kochen–Specker theorem: The Peres–Mermin square, New J. Phys. **19**, 123031 (2017).

[23] R. Kunjwal and R. W. Spekkens, From the Kochen-Specker theorem to noncontextuality inequalities without assuming determinism, Phys. Rev. Lett. **115**, 110403 (2015).

[24] R. Kunjwal and R. W. Spekkens, From statistical proofs of the Kochen-Specker theorem to Noise-Robust noncontextuality inequalities, Phys. Rev. A **97**, 052110 (2018).

[25] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, Nat. Commun. **7**, ncomms11780 (2016).

[26] M. F. Pusey, Robust preparation noncontextuality inequalities in the simplest scenario, Phys. Rev. A **98**, 022112 (2018).

[27] D. Schmid, R. W. Spekkens, and E. Wolfe, All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, Phys. Rev. A **97**, 062103 (2018).

[28] J. Bae and L.-C. Kwek, Quantum state discrimination and its applications, J. Phys. A: Math. Theor. **48**, 083001 (2015).

[29] S. M. Barnett, R. B. M. Clarke, V. M. Kendon, E. Riis, A. Chefles, and M. Sasaki, Experimental quantum state discrimination, in *Quantum Communication, Computing, and Measurement 3*, edited by P. Tombesi and O. Hirota (Springer US, Boston, 2002), pp. 59–67.

[30] S. Gómez, E. S. Gómez, O. Jiménez, A. Delgado, S. P. Walborn, and G. Lima, Experimental quantum state discrimination using the optimal fixed rate of inconclusive outcomes strategy, Sci. Rep. **12**, 17312 (2022).

[31] J. Bae, D.-G. Kim, and L.-C. Kwek, Structure of optimal state discrimination in generalized probabilistic theories, Entropy **18**, 39 (2016).

[32] G. Kimura, T. Miyadera, and H. Imai, Optimal state discrimination in general probabilistic theories, Phys. Rev. A **79**, 062306 (2009).

[33] S. M. Barnett and S. Croke, Quantum state discrimination, Adv. Opt. Photon. **1**, 238 (2009).

[34] J. A. Bergou, Quantum state discrimination and selected applications, J. Phys.: Conf. Ser. **84**, 012001 (2007).

[35] E. R. Loubenets, General lower and upper bounds under minimum-error quantum state discrimination, Phys. Rev. A **105**, 032410 (2022).

[36] J. Bae, Structure of minimum-error quantum state discrimination, New J. Phys. **15**, 073037 (2013).

[37] U. Herzog, Minimum-error discrimination between a pure and a mixed two-qubit state, J. Opt. B: Quantum Semiclassical Opt. **6**, S24 (2004).

[38] I. D. Ivanovic, How to differentiate between non-orthogonal states, Phys. Lett. A **123**, 257 (1987).

[39] D. Dieks, Overlap and distinguishability of quantum states, Phys. Lett. A **126**, 303 (1988).

[40] A. Peres, How to differentiate between non-orthogonal states, Phys. Lett. A **128**, 19 (1988).

[41] O. Jiménez, M. A. Solís-Prosser, A. Delgado, and L. Neves, Maximum-confidence discrimination among symmetric qudit states, Phys. Rev. A **84**, 062315 (2011).

[42] S. Croke, E. Andersson, S. M. Barnett, C. R. Gilson, and J. Jeffers, Maximum confidence quantum measurements, Phys. Rev. Lett. **96**, 070401 (2006).

[43] U. Herzog, Discrimination of two mixed quantum states with maximum confidence and minimum probability of inconclusive results, Phys. Rev. A **79**, 032323 (2009).

[44] U. Herzog, Optimized maximum-confidence discrimination of n mixed quantum states and application to symmetric states, Phys. Rev. A **85**, 032312 (2012).

[45] U. Herzog, Optimal measurements for the discrimination of quantum states with a fixed rate of inconclusive results, Phys. Rev. A **91**, 042338 (2015).

[46] E. Bagan, R. Muñoz-Tapia, G. A. Olivares-Rentería, and J. A. Bergou, Optimal discrimination of quantum states with a fixed rate of inconclusive outcomes, Phys. Rev. A **86**, 040303(R) (2012).

[47] L. Vandenberghe and S. Boyd, Semidefinite programming, SIAM Rev. **38**, 49 (1996).

[48] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevA.109.032202 for a contextuality witness inspired by optimal state discrimination.

[49] C. W. Helstrom, Detection theory and quantum mechanics (ii), Inf. Control **13**, 156 (1968).

[50] D. Qiu, Minimum-error discrimination between mixed quantum states, Phys. Rev. A **77**, 012328 (2008).

[51] M. Kleinmann, H. Kampermann, and D. Bruß, Unambiguous discrimination of mixed quantum states: Optimal solution and case study, Phys. Rev. A **81**, 020304(R) (2010).

[52] R. Salazar and A. Delgado, Quantum tomography via unambiguous state discrimination, Phys. Rev. A **86**, 012118 (2012).

[53] R. W. Spekkens, Negativity and contextuality are equivalent notions of nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[54] C. W. Helstrom, Detection theory and quantum mechanics, Inf. Control **10**, 254 (1967).

[55] C. W. Helstrom, Quantum detection and estimation theory, J. Stat. Phys. **1**, 231 (1969).

[56] H. Lee, K. Flatt, C. Roch i Carceller, J. B. Brask, and J. Bae, Maximum-confidence measurement for qubit states, Phys. Rev. A **106**, 032422 (2022).

[57] V. P. Rossi, D. Schmid, J. H. Selby, and A. B. Sainz, Contextuality with vanishing coherence and maximal robustness to dephasing, Phys. Rev. A **108**, 032213 (2023).