



## PAPER

## Contextual advantages across two-state discrimination strategies

## OPEN ACCESS

## RECEIVED

3 December 2025

## REVISED

5 March 2026

## ACCEPTED FOR PUBLICATION



1 April 2026

## PUBLISHED

17 April 2026

Original content from  
this work may be used  
under the terms of the  
[Creative Commons  
Attribution 4.0 licence](#).

Any further distribution  
of this work must  
maintain attribution to  
the author(s) and the title  
of the work, journal  
citation and DOI.

Kieran Flatt\*  and Joonwoo Bae 

School of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

\* Author to whom any correspondence should be addressed.

E-mail: [kflatt@kaist.ac.kr](mailto:kflatt@kaist.ac.kr) and [joonwoo.bae@kaist.ac.kr](mailto:joonwoo.bae@kaist.ac.kr)

**Keywords:** contextuality, state discrimination, quantum communications, quantum foundations

**Abstract**

Quantum state discrimination, alongside its other applications, has recently found use as a tool for witnessing generalised contextuality. In this article, we derive noncontextuality inequalities for both conclusive and inconclusive outcomes across various guessing strategies. For minimum-error discrimination, the advantage is in terms of the confidences of individual outcomes, while for unambiguous state discrimination, it is in terms of the average guessing probability. For maximum-confidence discrimination, we show that contextual advantages occur not only for the confidence but also their average, the guessing probability, as well as the inconclusive outcome rate. Our results show that contextual advantages are observed across all two-state discrimination schemes and figures of merit. We envisage that various quantum information applications based on state discrimination may offer advantages over noncontextual theories.

**1. Introduction**

If gaps exist between the behaviour of classical and quantum systems, it is possible to take advantage of the differences for enhanced information processing. Contextuality, which tells us that measurements do not merely reveal pre-existing features of quantum states, has long been considered one such distinguishing feature. Despite this, it proved difficult for a long to know how to use this property for practical applications. This was in part due to the fact that non-projective measurements, which are required in realistic, noisy scenarios, were difficult to describe in the original framework of Kochen-Specker contextuality [1]. For this reason, amongst others, a generalised form of contextuality was later put forward [2].

Taking contextuality as a defining feature of quantum theory implies a natural definition of classicality as the set of statistics which can be generated by theories lacking this property. While Kochen-Specker contextuality is reliant on the Hilbert space description of quantum theory, the generalised form is defined on the basis of observable probabilities. This allows one to define a set of noncontextual statistics, the violation of which signals a quantum advantage. With this in mind, recent work in the field of generalised contextuality has provided a series of noncontextual inequalities [3, 4]. These consist of proofs that the upper bound of a given operationally meaningful quantity is lower in theories lacking contextuality compared with the equivalent quantum figure of merit. Of particular importance here, as a fundamental communication primitive, is state discrimination [5–7]. It was first shown that quantum advantages may be witnessed in the setting of minimum-error state discrimination (MESD) [8] and subsequent works generalised this to aspects of unambiguous state discrimination (USD) and maximum confidence measurements (MCMs) [9–15]. Our aim in this paper is to complete this picture for two-state ensembles. We show that the ability to witness quantum advantages is universal across all forms of two-state discrimination, and in all measures of success. Note that, while ensembles with three-or-more states may be studied in this framework, it is not possible to construct *all* such ensembles in a noncontextual theory [2]. Two-state ensembles are thus the only cases which can be studied in a general manner.

Demonstrations of quantum advantage take one of two approaches. On the one hand, one can propose a measure of the average performance of a given task and then show that quantum implementations can attain a higher value than classical ones. The relevant notion of classicality in this case will be statistics that can be realised in non-contextual theories. In contrast, one can consider deterministic phenomena that can be witnessed in a single-shot scenario. The proof of the PBR theorem [16], for example, follows an analogous procedure by showing the existence of a task which can be performed perfectly in  $\psi$ -epistemic models but not in quantum theory [17].

State discrimination, the problem of determining by measurement which state, taken from some ensemble, was prepared, allows one to witness both forms of advantage. MESD is concerned with the success probability observed on average over a run of experiments. USD and MCM, on the other hand, are concerned with inferring preparations from a single detection event, as what one learns does not depend on averaged data. More generally, we may also ask about the single-shot properties of the former scheme as well as the averaged properties of the latter two.

In order to do this, we note that the three different schemes may each be analysed according to three distinct figures of merit: the average probability of a correct detection event, the probability of conclusive measurement events, and the retrodictive confidence that a state was prepared given a measurement outcome. Optimisation of these three quantities gives rise to MESD, USD and MCM respectively. Less often noted is that the remaining figures of merit also play a role in each scenario. In this work, we show that they are also witnesses of contextuality. We show that, for MESD, the confidences of each individual outcome give rise to an associated contextual advantage. The same is true for the average guessing probability of USD. We also investigate the average guessing probability and inconclusive outcome rate of an MCM and derive further noncontextual inequalities.

We begin by summarising both what is known so far and what the precise contributions are of this article. After reminding the reader of relevant background information, we then proceed with our main results: demonstrations that the confidences in MESD as well as the average success rates in USD and MCMs give rise to contextual inequalities. We close with remarks on the relevancy of these results to realistic experiments and applications.

## 2. Summary of results

State discrimination tasks, reviewed in more detail in the next section, are characterised by three basic figures of merit: the average guessing probability  $P_g$ , the inconclusive outcome rate  $P_0$  and the set confidences  $C(i)$  associated with measurement outcomes. These quantities can be associated with each of the three main versions of state discrimination: minimum error, unambiguous and maximum confidence. One also considers the resulting values of these quantities given optimisation of another.

Table 1 displays the various figures of merit which may be associated with each scheme. An advantage for the minimum error guessing probability was shown in [8], and advantages for the maximum confidences and the inconclusive outcome rate were shown in [11]. We show, in the current article, the results for the remaining cases.

## 3. Background

### 3.1. State discrimination

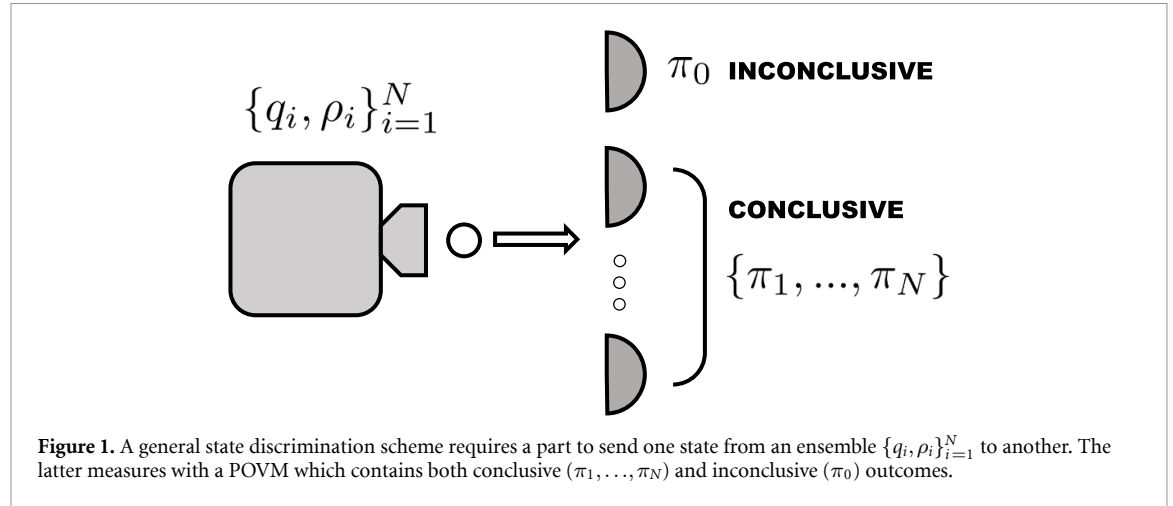
Any state discrimination task [5–7] is characterised by a finite, discrete ensemble  $\{q_i, \rho_i\}_{i=1}^N$  of  $N$  quantum states  $\rho_i$  each prepared with a probability  $q_i$ . The average state produced is denoted  $\rho = \sum_{i=1}^N q_i \rho_i$ . One party, Alice, prepares a system in one of the states and then passes it on to a receiver, Bob. Bob's task is to determine the state using measurements, represented by a positive-operator-valued-measure (POVM)  $\{\pi_j\}$ , such that  $0 \leq \pi_i \leq 1$  and  $\sum_j \pi_j = \mathbf{1}$ , in which the outcomes are associated with either a given state or an inconclusive detection event. This set-up is displayed in figure 1. Bob will design his measurement to be optimal according to a chosen figure of merit, giving rise to MESD, USD and MCM depending on the choice. We summarise the key points of each scheme in table 1.

Three basic quantities can be used to measure the success of a protocol. One is the *average guessing probability*, given by

$$P_g = \sum_{i=1}^N q_i P_{M|P}(i|i) = \sum_{i=1}^N q_i \text{Tr}[\rho_i \pi_i], \quad (1)$$

**Table 1.** Three schemes may be used to implement quantum state discrimination. In this table, we display firstly the basic properties defining them. In the last three columns, we summarise the contextual advantages associated with each scheme. Cells containing citations were previously shown; cells containing numbers are definitional; cells stating ‘This work’ are shown in the current work.

Scheme	Overview			Contextual advantages		
	Optimisation	Inc. Outcome?	All ensembles?	$P_g$	$P_0$	$C(i)$
MESD	$\max P_g$	×	✓	[8]	0	This work
USD	$\min P_0$	✓	×	This work	[11]	1
MCM	$\max C(i) \forall i$	✓	✓	This work	This work	[11]



which is the weighted average that measurement outcome  $i$  occurs given that state  $\rho_i$  was prepared. We employ (as throughout this article) the notation  $P_{M|P}(i|k)$  for the probability that measurement  $M$  resulted in outcome  $i$  given that the preparation  $P$  resulted in state  $k$ .

The second figure of merit which is involved is the *inconclusive outcome rate*. In certain protocols, it is necessary to include an 0th outcome signifying that the measurement did not reveal anything about the prepare state. This may be by design or due to experimental flaws. The rate at which this happens is given by

$$P_0 = \text{Tr}[\rho\pi_0]. \quad (2)$$

Finally, the set of *confidences*,

$$C(i) = P_{P|M}(i|i) = \frac{q_i P_{P|M}(i|i)}{P_M(i)}. \quad (3)$$

are the retrodictive probabilities that a particular state was prepared given that the relevant measurement outcome was observed.

Optimisation over each of these three quantities, with the potential of additional constraints, can be associated with a different form of state discrimination. If the average guessing probability, equation (1) is maximised over an  $N$  outcome measurement, one is said to perform minimum error state discrimination. In the simplest case of two states the maximum value is known to be the Helstrom bound [18].

USD [19–21] refers to scenarios in which all conclusive outcomes, those labelled 1 to  $N$ , determine the prepared state with certainty. This result is achieved by the  $i$ th POVM element projecting onto a subspace outside of the support of the span of the other states. A necessary condition, referred to as Chefles’ criterion [22], for such a measurement to be possible is that the states in the ensemble are all linearly independent. POVMs implementing this will typically include an inconclusive outcome, and minimising its occurrence rate, equation (2), will be optimal.

A MCM [23–25] is defined as one in which every confidence, equation (3), reaches its maximum value for all  $i$ . This is known to happen if each POVM element  $\pi_i$  is a rank-one measurement proportional to the projector onto the eigenvector of  $\rho^{-1}\rho_i\rho^{-1}$  with the highest associated eigenvalue. In general, one will maximise the set of confidences, and then minimise the inconclusive outcome rate associated with the optimal measurement. For scenarios involving two-dimensional systems, a semi-definite programming approach has been developed for this optimisation task [25]. Note that, under

this approach, we can redefine USD as the subset of MCMs in which it is possible to attain  $C(i) = 1$  for all  $i$  in the ensemble.

### 3.2. Generalised contextuality

Different definitions of contextuality have been proposed and here we focus on the ‘generalised contextuality’ approach developed by Spekkens [2]. This framework shows that operational theories subject to a constraint of non-contextuality cannot reproduce all of statistics given by quantum theory.

An operational theory represents experimental procedures as mathematical functions on an ontic state space  $\Lambda$ , in which each state is associated with a point  $\lambda$  encoding all physical properties of the system. An experiment is then decomposed into preparations, transformations and measurements. Preparations  $P$  are defined as epistemic states,  $\mu_j(\lambda)$ , probability distributions over  $\Lambda$ . Measurements  $M$  are treated as sets of positive response functions  $\{\xi_{i|M}(\lambda)\}$  normalised so that  $\sum_i \xi_{i|M}(\lambda) = 1 \forall \lambda$ . The probability of a measurement outcome  $i$  given measurement  $M$  and preparation  $j$  in such models is given by

$$P_{M|P}(i|j) = \int_{\Lambda} d\lambda \mu_j(\lambda) \xi_{i|M}(\lambda). \quad (4)$$

Two epistemic states are said to be *preparation equivalent* if every choice of measurement gives the same statistics for both states:

$$\int_{\Lambda} d\lambda \mu_1(\lambda) \xi_{i|M}(\lambda) = \int_{\Lambda} d\lambda \mu_2(\lambda) \xi_{i|M}(\lambda) \quad \forall \{i|M\}. \quad (5)$$

We can now define a noncontextual theory: a theory is said to be preparation noncontextual if it assigns the same epistemic state to all equivalent preparations in a model. A similar definition holds for measurement noncontextuality.

The final piece of terminology required is a measure for the distance between two pure states. In an operational model this is called the confusability,  $c_{1,2}$ ,

$$\begin{aligned} c_{1,2} &= \int_{\Lambda} d\lambda \mu_2(\lambda) \xi_1(\lambda) \\ &= \int_{\text{supp}[\mu_1]} d\lambda \mu_2(\lambda) \end{aligned} \quad (6)$$

where the second line follows from imposing noncontextuality and  $\xi_1(\lambda)$  is a response function giving  $P_{M|P}(1|1) = 1$ . This function is symmetric such that  $c_{1,2} = c_{2,1}$ . When comparing noncontextual theories with quantum theory, this value is to be interpreted as the squared overlap of the state. That is, if the quantum state  $|\psi_1\rangle$  and  $|\psi_2\rangle$  are associated with the epistemic states  $\mu_1(\lambda)$  and  $\mu_2(\lambda)$  respectively, then the two theories are compared by fixing  $c_{1,2} = |\langle \psi_1 | \psi_2 \rangle|^2$ .

Two additional assumptions are typically made of the epistemic model. One is that, for each pure state there exists a deterministic response function. The second is that probabilistic mixing over quantum states maps onto probabilistic mixing over the associated epistemic states:

$$p\rho_1 + (1-p)\rho_2 \mapsto p\mu_1(\lambda) + (1-p)\mu_2(\lambda), \quad (7)$$

in which  $0 \leq p \leq 1$  is a probability. These assumptions are used to construct a mapping between quantum theory and any operational model, whether noncontextual or otherwise.

## 4. Contextual advantages in state discrimination

We begin in earnest by reviewing the previously known contextual advantages in state discrimination.

### 4.1. MESD: success probability

The first contextual advantage to be shown in a state discrimination task was the guessing probability of a minimum error measurement [8]. The specific scenario investigated was an equiprobable two-state discrimination task, represented in quantum theory by two pure states  $|\psi_1\rangle$  and  $|\psi_2\rangle$ . The quantum upper bound of this problem is given by the well-known Helstrom bound [18]

$$P_g^{(Q)} = \frac{1}{2} \left( 1 + \sqrt{1 - |\langle \psi_1 | \psi_2 \rangle|^2} \right). \quad (8)$$

This is compared with the maximum probability of distinguishing two preparations  $\mu_1(\lambda)$  and  $\mu_2(\lambda)$  with a confusability  $c_{1,2}$  in a noncontextual theory.

In order to impose the constraint of noncontextuality, preparation equivalence is required. There, however, do not exist any equivalences if we invoke only two pure states. In order to get around this, the existence of a ‘mirror ensemble’  $\{|\bar{\psi}_1\rangle, |\bar{\psi}_2\rangle\}$  consisting of the pair of states orthogonal to the initial ensemble  $\langle\psi_i|\bar{\psi}_i\rangle = 0$  is assumed. With this addition we now have an equivalence between alternative preparations of the maximally mixed state:

$$\frac{1}{2}|\psi_1\rangle\langle\psi_1| + \frac{1}{2}|\bar{\psi}_1\rangle\langle\bar{\psi}_1| = \frac{1}{2}|\psi_2\rangle\langle\psi_2| + \frac{1}{2}|\bar{\psi}_2\rangle\langle\bar{\psi}_2| \quad (9)$$

which then implies that in a noncontextual model the following constraint will apply:

$$\frac{1}{2}\mu_1(\lambda) + \frac{1}{2}\bar{\mu}_1(\lambda) = \frac{1}{2}\mu_2(\lambda) + \frac{1}{2}\bar{\mu}_2(\lambda) \quad (10)$$

in which  $\bar{\mu}_i(\lambda) \cong |\bar{\psi}_i\rangle\langle\bar{\psi}_i|$ . This relation constrains the relationship between the epistemic states in the considered model. In particular, one finds that

$$\mu_1(\lambda) = \mu_2(\lambda) \quad \forall \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\mu_2]. \quad (11)$$

Application of this constraint along with the assumption that all epistemic states are associated with a deterministic response function then allows one to bound the guessing probability. The result is that

$$P_g^{(NC)} = 1 - \frac{c_{1,2}}{2} \quad (12)$$

which is strictly smaller than the Helstrom bound.

#### 4.2. USD: success probability and inconclusive outcome rate

We now turn our attention to the quantum advantages associated with USD, which were first demonstrated in [11]. Full technical details can be found in that article. These quantum advantages are associated with the inconclusive outcome rate, which is strictly higher in a non-contextual theory, and the guessing probability. We begin by noting that, as the probability of incorrectly guessing will be equal to zero, these two quantities are directly related to one another in USD,  $P_g = 1 - P_0$ , and it suffices to consider just one property.

It is trivial to show that in the binary discrimination scenario, USD is possible in a noncontextual theory. The response functions which satisfy these will be of the form

$$\xi_{i|M_{\text{USD}}}(\lambda) = \begin{cases} \gamma, & \lambda \in \text{supp}[\bar{\mu}_{i+1}] \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

where addition in the subscripts is modular and  $0 < \gamma \leq 1$  and  $M_{\text{USD}}$  labels the measurement. Such response functions are equivalent to POVM elements of the form  $\pi_i = \gamma|\psi_{i+1}\rangle\langle\psi_{i+1}|$ . This observation is furthermore supported by a recent theorem [26] stating that a sufficient condition for a set of states to be preparable in a preparation-noncontextual model is that they are linearly independent. This is in fact the same criterion as that for states to be unambiguously discriminable.

One can then optimise the choice of  $\gamma$  in order to minimise the inconclusive outcome rate, given that the response functions are constrained by preparation noncontextuality. The ‘mirror ensemble’ approach, as taken in the minimum error calculation, is again used and the result is found to be

$$P_0^{(NC)} = \frac{1}{2}(1 + c_{1,2}). \quad (14)$$

This value is strictly larger than or equal to the inconclusive outcome rate for quantum USD. In this manner, the contextuality of state discrimination was extended to USD.

### 4.3. MCM: confidences

The next result to review is the advantage associated with the confidence in MCMs. As noted in the background, MCMs and USD coincide if the latter form is possible. The simplest case in which it is not is a dephasing of the binary pure state ensemble. That is, two states with the form

$$\rho_i = (1-p)|\psi_i\rangle\langle\psi_i| + p\frac{\mathbb{1}}{2} \quad i = 1, 2 \quad (15)$$

where  $0 \leq p \leq 1$  quantifies the noise existing in each state. The quantum problem can be solved using typical methods and the maximum confidence for each state is found to be

$$C^{(Q)}(i) = \frac{1}{2} \left( 1 + \frac{(1-p)\sqrt{1-|\langle\psi_1|\psi_2\rangle|^2}}{\sqrt{1-(1-p)^2|\langle\psi_1|\psi_2\rangle|^2}} \right). \quad (16)$$

Using the same methods as discussed above, the noncontextual bound to the confidence can also be calculated. This is found to be

$$C^{(NC)}(i) = \frac{1}{2} \left( 1 + \frac{(1-p)(1-c_{1,2})}{1-(1-p)c_{1,2}} \right) \quad (17)$$

which again is smaller than the equivalent quantum confidence seen above.

It was shown in these three results that quantum advantages are common to different forms of state discrimination. What we will see in the remaining sections is that this advantage can also be observed for other figures of merit.

## 5. Characterising contextual advantages

In this section, we derive further noncontextual inequalities for each state discrimination strategy. We do this by performing the following set of optimisation tasks in both quantum and noncontextual theories. First, we maximise the guessing probability and then calculate the confidences of the outcomes. Second, we optimise the set of confidences and find the minimum inconclusive outcome rate and guessing probability for the optimal measurement.

### 5.1. MESD: confidences

Let us consider two pure states that are prepared with arbitrary *a priori* probabilities  $q_1, q_2 > 0$ , that is, the ensemble  $\{q_i, |\psi_i\rangle\langle\psi_i|\}_{i=1}^2$ . We construct the measurements that maximise the guessing probability and then calculate confidences of these measurements.

Let us begin with the quantum scenario. Two pure states may be parameterised without loss of generality as

$$\begin{aligned} |\psi_1\rangle &= \cos\left(\frac{\theta}{2}\right)|0\rangle - \sin\left(\frac{\theta}{2}\right)|1\rangle \\ |\psi_2\rangle &= \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)|1\rangle \end{aligned} \quad (18)$$

for which the squared overlap is  $|\langle\psi_1|\psi_2\rangle|^2 = \cos^2\theta = c_{1,2}$ . An optimal measurement that maximises the guessing probability can be obtained by finding the eigenvectors of the operator,  $q_1|\psi_1\rangle\langle\psi_1| - q_2|\psi_2\rangle\langle\psi_2|$ . The POVM which results in maximum success is then given by  $\pi_i = |\phi_i\rangle\langle\phi_i|$ , with

$$\begin{aligned} |\phi_1\rangle &= \cos\left(\frac{\phi}{2}\right)|0\rangle - \sin\left(\frac{\phi}{2}\right)|1\rangle \\ |\phi_2\rangle &= \cos\left(\frac{\phi}{2}\right)|0\rangle + \sin\left(\frac{\phi}{2}\right)|1\rangle \end{aligned} \quad (19)$$

and

$$\tan\left(\frac{\phi}{2}\right) = \frac{(q_2 - q_1)\cos(\theta) - \sqrt{1 - 4q_1q_2\cos^2(\theta)}}{\sin(\theta)}. \quad (20)$$

The resulting maximum success probability is

$$P_g^{(Q)} = \frac{1}{2} \left( 1 + \sqrt{1 - 4q_1q_2 \cos^2(\theta)} \right) \tag{21}$$

The maximum confidence given by MESD in quantum theory can therefore be evaluated straightforwardly for

$$C^{(Q)}(1) = \frac{q_1 \cos^2\left(\frac{\theta+\phi}{2}\right)}{q_1 \cos^2\left(\frac{\theta+\phi}{2}\right) + q_2 \cos^2\left(\frac{\theta-\phi}{2}\right)}$$

$$C^{(Q)}(2) = \frac{q_2 \cos^2\left(\frac{\theta+\phi}{2}\right)}{q_1 \cos^2\left(\frac{\theta-\phi}{2}\right) + q_2 \cos^2\left(\frac{\theta+\phi}{2}\right)}, \tag{22}$$

in which  $\phi$  is evaluated using equation (20). Here, unlike for the equiprobable case, the confidences associated with each outcome differ.

We compare this with the equivalent quantity in a noncontextual. We first calculate the upper-bound for the MESD success probability for the relevant ensemble. This was first done in [10], where

$$P_g^{(NC)} = 1 - (1 - \max\{q_1, q_2\}) c_{1,2}. \tag{23}$$

This figure of merit is seen to display a contextual advantage. We require, alongside the upper-bound, the particular set of response functions that reach this bound in order to evaluate the confidence. This can be done constructively. The set of two-outcome measurements in a two-state discrimination task can be represented by two sets of response functions. First,  $M_1$  is given by

$$\xi_{1|M_1}(\lambda) = \begin{cases} 1, & \lambda \in \text{supp}[\mu_1] \\ 0, & \text{otherwise} \end{cases}$$

$$\xi_{2|M_1}(\lambda) = \begin{cases} 0, & \lambda \in \text{supp}[\mu_1] \\ 1, & \text{otherwise} \end{cases} \tag{24}$$

and  $M_2$  is given by

$$\xi_{1|M_2}(\lambda) = \begin{cases} 0, & \lambda \in \text{supp}[\mu_2] \\ 1, & \text{otherwise} \end{cases}$$

$$\xi_{2|M_2}(\lambda) = \begin{cases} 1, & \lambda \in \text{supp}[\mu_2] \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

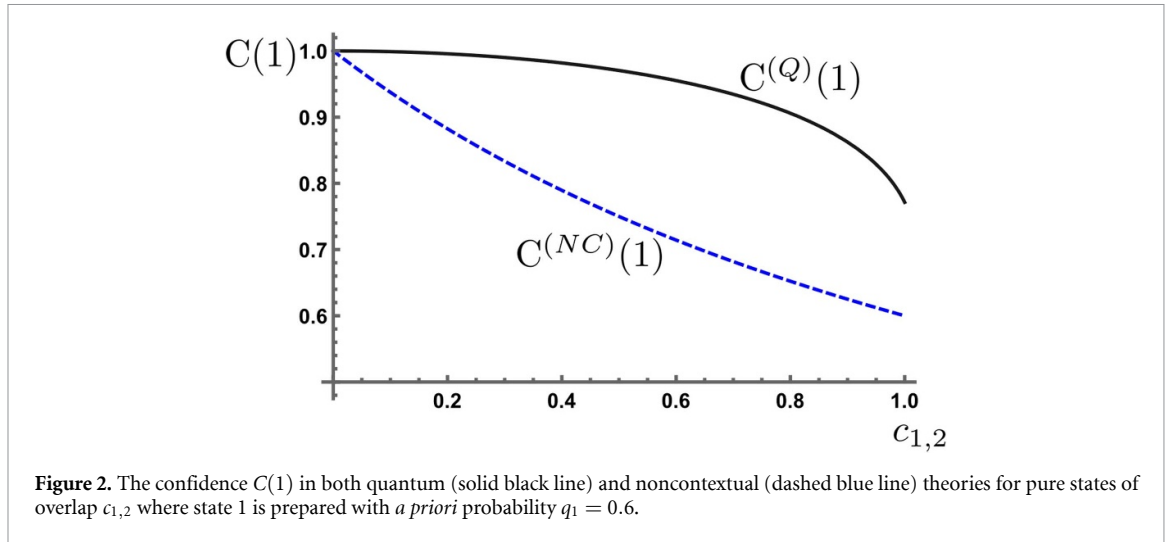
Any set of response functions is then given by a convex combination of these two measurements. Operationally speaking, this means that we can implement any two-outcome measurement in a non-contextual theory by implementing  $M_1$  with probability  $\omega$  and  $M_2$  with probability  $1 - \omega$ . The response function resulting from this procedure will be

$$\xi_{1|M_\omega}(\lambda) = \begin{cases} \omega, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\mu_2] \\ 1, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\bar{\mu}_2] \\ 0, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\mu_2] \\ 1 - \omega, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\bar{\mu}_2] \end{cases}$$

$$\xi_{2|M_\omega}(\lambda) = \begin{cases} 1 - \omega, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\mu_2] \\ 0, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\bar{\mu}_2] \\ 1, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\mu_2] \\ \omega, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\bar{\mu}_2] \end{cases} \tag{26}$$

and we now evaluate the success probability using this measurement, in order to optimise over  $\omega$ :

$$P_g^{(NC)} = q_1 \int_{\Lambda} d\lambda \mu_1(\lambda) \xi_{1|\omega}(\lambda) + q_2 \int_{\Lambda} d\lambda \mu_2(\lambda) \xi_{2|\omega}(\lambda). \tag{27}$$



**Figure 2.** The confidence  $C(1)$  in both quantum (solid black line) and noncontextual (dashed blue line) theories for pure states of overlap  $c_{1,2}$  where state 1 is prepared with a *a priori* probability  $q_1 = 0.6$ .

We can evaluate this expression using equation (26) and begin from the first term:

$$\int_{\Lambda} d\lambda \xi_{1|M_\omega}(\lambda) \mu_1(\lambda) = \omega \int_{\text{supp}[\mu_1] \cap \text{supp}[\mu_2]} d\lambda \mu_1(\lambda) + \int_{\text{supp}[\mu_1] \cap \text{supp}[\bar{\mu}_2]} d\lambda \mu_1(\lambda), \tag{28}$$

which can be simplified upon noticing the following relation:

$$\begin{aligned} \int_{\text{supp}[\mu_1] \cap \text{supp}[\mu_2]} d\lambda \mu_1(\lambda) &= \int_{\text{supp}[\mu_1] \cap \text{supp}[\mu_2]} d\lambda \mu_1(\lambda) + \int_{\text{supp}[\bar{\mu}_1] \cap \text{supp}[\mu_2]} d\lambda \mu_1(\lambda) \\ &= \int_{\text{supp}[\mu_2]} d\lambda \mu_1(\lambda) \\ &= c_{1,2}. \end{aligned} \tag{29}$$

and an analogous procedure gives the other values. The resulting success probability is

$$P_g^{(NC)} = 1 - (1 - \omega q_1 - (1 - \omega) q_2) c_{1,2} \leq 1 - (1 - \max\{q_1, q_2\}) c_{1,2} \tag{30}$$

which reproduces the upper bound in equation (23). The measurement which maximises this is to choose  $\omega = 1$ , that is, implement  $M_1$ , if  $q_1 > q_2$ . If, on the other hand,  $q_2 > q_1$ , then the optimal choice is  $\omega = 0$ , that is, to implement  $M_2$ . The quantum analogy of these two measurements would be to project onto the basis defined by  $|\psi_1\rangle$  ( $|\psi_2\rangle$ ) for  $M_1$  ( $M_2$ ). Note that, if  $q_1 = q_2$  then all choices of  $\omega$  will give the same, previously known, upper bound for the non-contextual success probability. We are now in a position to evaluate the confidence for each measurement outcome. We may assume without loss of generality that we are in a scenario in which  $q_1 > q_2$ , and so use  $M_1$  to maximise the success probability.

The evaluation of the confidences using this measurement can be performed

$$\begin{aligned} C^{(NC)}(1) &= \frac{q_1}{c_{1,2} + (1 - c_{1,2}) q_1} \\ C^{(NC)}(2) &= 1, \end{aligned} \tag{31}$$

which are distinct from the quantum counterparts in equation (22). We plot the quantum and noncontextual values of  $C(1)$  in figure 2.

We observe that, out of two arms for outcomes 1 and 2, only the arm detecting a state with larger *a priori* probability may contain contextual advantage whereas the other not. This contrasts with the result in [11] where the *a priori* probabilities are equal; the values of confidence for two outcomes are identical and thus both have a contextual advantage. One may point out that the state space structure may give rise to the distinction. The set of quantum states has extremal points continuously connected by state transformations, in contrast to which, that of a noncontextual theory has a finite number of extremal ones. Consequently, when *a priori* probabilities are not equal, optimal quantum measurements are also continuously transformed whereas their noncontextual counterparts are biased to some specific extreme points.

## 5.2. MCM: inconclusive outcome rate

We now turn our attention to the inconclusive rate in MCMs, and will study the noisy discrimination scenario, when the aim in the quantum case is to discriminate the states in equation (15) and the analogy of this in the ontological model is to discriminate the epistemic states

$$\tilde{\mu}_i(\lambda) = (1-p)\mu_i(\lambda) + p\mu_{\mathbb{1}/2}(\lambda), \quad i = 1, 2. \quad (32)$$

Here,  $\mu_i(\lambda)$  are two extremal epistemic states associated with the two pure states in the quantum framework and  $\mu_{\mathbb{1}/2}(\lambda)$  is the epistemic-state representation of the maximally mixed quantum state  $\mathbb{1}/2$ . Note that in noncontextual theories the latter takes the same representation for all methods of preparation (i.e. all decompositions in the quantum framework).

The quantum task has been studied in detail [25]. If the states are parameterised as

$$\begin{aligned} |\psi_1\rangle &= \cos\left(\frac{\theta}{2}\right)|0\rangle - \sin\left(\frac{\theta}{2}\right)|1\rangle \\ |\psi_2\rangle &= \cos\left(\frac{\theta}{2}\right)|0\rangle + \sin\left(\frac{\theta}{2}\right)|1\rangle \end{aligned} \quad (33)$$

then optimal POVM elements will be rank-one and follow the structure  $\pi_i = \alpha|\phi_i\rangle\langle\phi_i|$   $i = 1, 2$ , where  $0 \leq \alpha \leq 1$  is a free parameter and

$$\begin{aligned} |\phi_i\rangle &= \cos\left(\frac{\phi}{2}\right)|0\rangle + (-1)^i \sin\left(\frac{\phi}{2}\right)|1\rangle \\ \tan(\phi) &= p \cos(\theta) \sqrt{\frac{1-p^2 \cos^2(\theta)}{1-\cos^2(\theta)}} \end{aligned} \quad (34)$$

The important point here is to note that the optimal quantum measurement is a function of the amount of dephasing noise. From this construction we can write the inconclusive POVM element as  $\mathbb{1} - \pi_1 - \pi_2$  and minimise over the parameters  $\alpha$  the inconclusive outcome rate, which is found to be

$$P_0^{(Q)} = (1-p)|\langle\psi_1|\psi_2\rangle|. \quad (35)$$

This is to be compared with the equivalent quantity in a noncontextual model of MCMs. The feature which we must use is that the noncontextually optimal measurement is represented by the same measurement as is used for USD:

$$\begin{aligned} \xi_{1|M_{\text{USD}}}(\lambda) &= \begin{cases} \gamma_1, & \lambda \in \text{supp}[\bar{\mu}_2] \\ 0, & \text{otherwise} \end{cases} \\ \xi_{2|M_{\text{USD}}}(\lambda) &= \begin{cases} \gamma_2, & \lambda \in \text{supp}[\bar{\mu}_1] \\ 0, & \text{otherwise} \end{cases} \end{aligned} \quad (36)$$

where the response function representing an inconclusive outcome will be given by  $\xi_{0|M_{\text{USD}}}(\lambda) = 1 - \xi_{1|M_{\text{USD}}}(\lambda) - \xi_{2|M_{\text{USD}}}(\lambda)$  and  $0 < \gamma_i \leq 1$  for  $i = 1, 2$ . This response function in full will have the structure

$$\xi_0(\lambda) = \begin{cases} 1, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\mu_2] \\ 1 - \gamma_1, & \lambda \in \text{supp}[\mu_1] \cap \text{supp}[\bar{\mu}_2] \\ 1 - \gamma_2, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\mu_2] \\ 1 - \gamma_1 - \gamma_2, & \lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\bar{\mu}_2]. \end{cases} \quad (37)$$

From this it is possible to calculate the inconclusive outcome rate. First note that the average density matrix for the noisy epistemic states will be

$$\tilde{\mu}(\lambda) = \frac{1-p}{2}\mu_1(\lambda) + \frac{1-p}{2}\mu_2(\lambda) + p\mu_{\mathbb{1}/2}(\lambda). \quad (38)$$

Then the quantity we wish to evaluate is

$$\begin{aligned} P_0^{(\text{NC})} &= \int_{\Lambda} d\lambda \tilde{\mu}(\lambda) \xi_{0|M_{\text{USD}}}(\lambda) \\ &= \frac{1-p}{2} \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_1(\lambda) + \frac{1-p}{2} \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_2(\lambda) + p \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_{\mathbb{1}/2}(\lambda). \end{aligned} \quad (39)$$

We can evaluate each of these integrals in turn using equation (37). First, we find

$$\begin{aligned} \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_1(\lambda) &= \int_{\text{supp}[\mu_1] \cap \text{supp}[\mu_2]} d\lambda \mu_1(\lambda) + (1 - \gamma_1) \int_{\text{supp}[\mu_1] \cap \text{supp}[\bar{\mu}_2]} d\lambda \mu_2(\lambda) \\ &= c_{1,2} + (1 - \gamma_1) c_{1,\bar{2}} \\ &= 1 - \gamma_1 + \gamma_1 c_{1,2}. \end{aligned} \quad (40)$$

A similar process gives

$$\int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_2(\lambda) = 1 - \gamma_2 + \gamma_2 c_{1,2}. \quad (41)$$

The final piece to evaluate is the integral over the maximally mixed epistemic state. In order to do this, we can decompose the epistemic state into either preparation, noting that the two will be the same due to the model's noncontextuality. We can thus evaluate:

$$\begin{aligned} \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \mu_{\mathbb{1}/2}(\lambda) &= \int_{\Lambda} d\lambda \xi_{0|M_{\text{USD}}}(\lambda) \left( \frac{1}{2} \mu_1(\lambda) + \frac{1}{2} \bar{\mu}_1(\lambda) \right) \\ &= 1 - \frac{1}{2} (\gamma_1 + \gamma_2) \end{aligned} \quad (42)$$

where the second line follows using similar techniques as above.

Bringing the previous few results together we now have

$$P_0^{(NC)} = 1 - \frac{1}{2} (1 - (1 - p) c_{1,2}) (\gamma_1 + \gamma_2). \quad (43)$$

We need to find the minimum value of this subject to the constraints imposed on the non-negativity of the response function. Clearly, this function reaches a minimum when  $\gamma_1 + \gamma_2$  as is allowed. Examining the inconclusive response function, equation (37), we can see that the piece  $\xi_0(\lambda)$  on the domain  $\lambda \in \text{supp}[\bar{\mu}_1] \cap \text{supp}[\bar{\mu}_2]$  requires that  $\gamma_1 + \gamma_2 \leq 1$  to be non-negative. Therefore we can take the upper bound  $\gamma_1 + \gamma_2 = 1$  and the minimum inconclusive rate in a non-contextual model will be

$$P_0^{(NC)} = \frac{1}{2} (1 + (1 - p) c_{1,2}). \quad (44)$$

Which is to be compared with equation (35) using  $c_{1,2} = |\langle \psi_1 | \psi_2 \rangle|^2$ . We display the two functions in figure 3 with both fixed  $c_{1,2}$  and fixed noise  $p$  where it can be seen that the inconclusive rate for a non-contextual model is greater than or equal to that of quantum theory.

### 5.3. MCM: guessing probability

The final piece which we fill in is the guessing probability, equation (1), of a MCM. We can recast the guessing probability into a form more amenable to the current task by writing it in terms of the confidences and outcome rates of the ensemble:

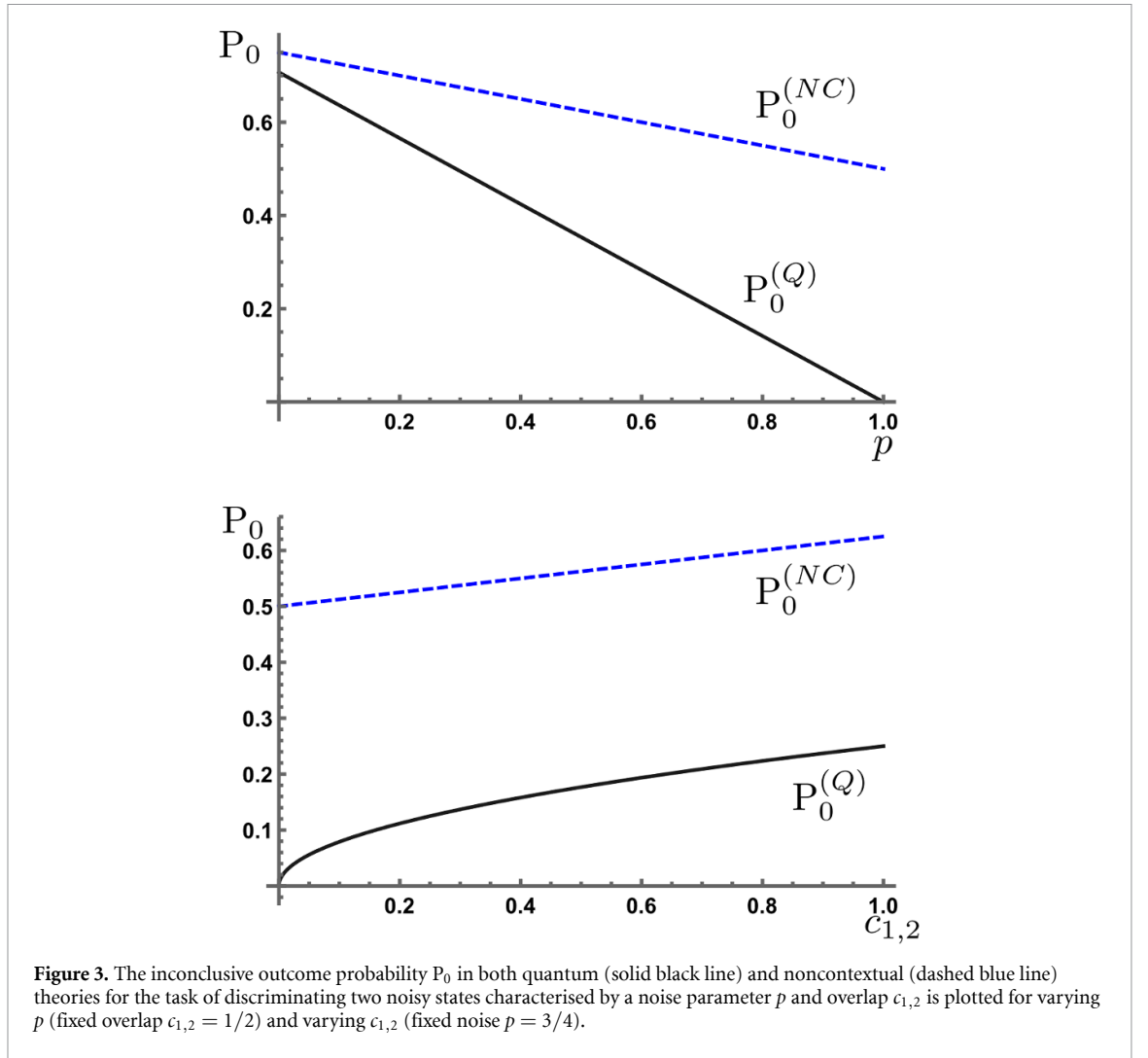
$$P_g = \sum_i P_M(i) C(i), \quad (45)$$

which follows simply from the definition and Bayes rule. The index  $i$  here runs over the states in the ensemble.

We can again note that for pure quantum states and extremal epistemic states, the MCM coincides with USD. For these cases, we need simply to repeat that the guessing probability in USD is  $P_g = 1 - P_0$  and we have the same advantage shown above. For this reason, we need only to focus on the noisy ensemble discussed in the previous subsection.

The simplest way to demonstrate the contextual advantage here is to note that both confidences in a binary discrimination task are equal, as can be seen in equations (16) and (17). We can therefore write  $C(1) = C(2) = C$  and equation (45) becomes

$$\begin{aligned} P_g &= \sum_i P_M(i) C \\ &= (1 - P_0) C. \end{aligned} \quad (46)$$



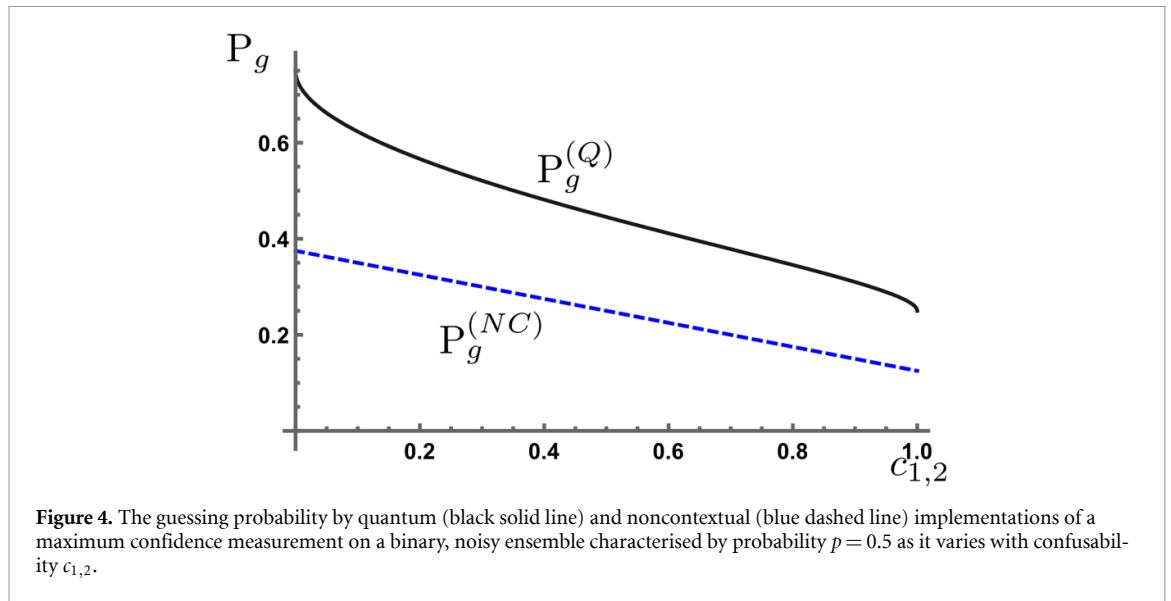
Both  $P_0$  and  $C$  have been previously calculated for both quantum and contextual theories, and both were seen to display a quantum advantage. As we have  $P_0^Q \leq P_0^{(NC)}$  and  $C^{(Q)}(i) \geq C^{(NC)}(i) \forall i$ , it trivially follows that  $P_g^{(Q)} \geq P_g^{(NC)}$  and a quantum advantage therefore exists also for this parameter. For completeness, the full expressions for these two quantities will be:

$$\begin{aligned}
 P_g^{(NC)} &= \frac{1}{2} \left( 1 - \frac{p}{2} - (1-p)c_{1,2} \right) \\
 P_g^{(Q)} &= \frac{1}{2} \left( 1 - (1-p)|\langle \psi_1 | \psi_2 \rangle| (1-p) \sqrt{\frac{1 - (1-p)|\langle \psi_1 | \psi_2 \rangle|}{1 + (1-p)|\langle \psi_1 | \psi_2 \rangle|}} \sqrt{1 - |\langle \psi_1 | \psi_2 \rangle|^2} \right)
 \end{aligned} \tag{47}$$

which are displayed in figure 4.

## 6. Discussion

Demonstrating advantages of quantum systems over classical ones plays a central role in the developing of useful applications for quantum technologies. As state discrimination is the communication primitive underlying many applications, it is important to firmly state which property of quantum theory allow for it to perform above the classical limit. Here, we show that it is contextuality. While the contextual advantages associated with state discrimination were previously known, it was not clear which aspects were able to show a contextual advantage. The main lesson of this paper is that all figures of merit are



reliant on this property. This was shown by demonstrating advantages for the guessing probability and inconclusive outcome rate in a MCM, as well as showing that under certain circumstances the confidence associated with a minimum error measurement also witness contextuality.

MCMs, and the measurement confidences more generally, play an important role in this work. One reason is that, in experimental settings, one is forced to deal with preparation and measurement noise, including non-detection events. These can be naturally taken into account using the confidence as a figure of merit, unlike the average success probability. Another reason is that they allow us to witness two notions of quantum advantage. In a single-shot setting, we can make no claims about the average statistics. The confidence, however, is a well-defined quantity in this setting, and is able to see beyond-classical statistics.

Our results should find use in experimental tests of contextuality, in which data are unavoidably noisy. It would be interesting to connect the inequalities found in state discrimination, a form of prepare-and-measure scenario, to nonlocal inequalities through the recent works bridging the two topics [27, 28], as well as understand the noise conditions under which contextuality can be witnessed [29] in state discrimination settings. As state discrimination plays an important role in many uses of quantum theory, for example randomness generation [12, 30] and sequential communications [31], we, furthermore, believe that this work will find applications in practical use cases.

## Acknowledgments

This work was supported by the National Research Foundation of Korea (RS-2024-00408613) and the Institute for Information & Communication Technology Promotion (IITP) (RS-2023-00229524, RS-2025-02304540, RS-2025-25464876, RS-2025-25464616).

## Data availability statement

No new data were created or analysed in this study.

## ORCID iDs

Kieran Flatt  0009-0001-5816-1320

Joonwoo Bae  0000-0002-2345-1619

## References

- [1] Budroni C, Cabello A, Gühne O, Kleinmann M and Larsson J 2022 Kochen-specker contextuality *Rev. Mod. Phys.* **94** 045007
- [2] Spekkens R W 2005 Contextuality for preparations, transformations and unsharp measurements *Phys. Rev. A* **71** 052108
- [3] Schmid D, Spekkens R W and Wolfe E 2018 All the noncontextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences *Phys. Rev. A* **97** 062103
- [4] Lostaglio M and Senno G 2020 Contextual advantage for state-dependent cloning *Quantum* **4** 258
- [5] Bae J and Kwek L-C 2015 Quantum state discrimination and its applications *J. Phys. A: Math. Theor.* **48** 083001
- [6] Bergou J A 2010 Discrimination of quantum states *J. Mod. Opt.* **57** 160–80
- [7] Barnett S M and Croke S 2009 Quantum state discrimination *Adv. Opt. Photon.* **1** 238–78
- [8] Schmid D and Spekkens R W 2018 Contextual advantage for state discrimination *Phys. Rev. X* **8** 011015
- [9] Leifer M and Duarte C 2020 Noncontextuality inequalities from antidistinguishability *Phys. Rev. A* **101** 062113
- [10] Shin J, Donghoon H and Kwon Y 2021 Quantum contextual advantage depending on nonzero prior probabilities in state discrimination of mixed qubit states *Entropy* **23** 1583
- [11] Flatt K, Lee H, Carceller C R I, Brask J B and Bae J 2022 Contextual advantages and certification for maximum-confidence discrimination *PRX Quantum* **3** 030337
- [12] Carceller C R I, Flatt K, Lee H, Bae J and Brask J B 2022 Quantum vs noncontextual semi-device-independent randomness certification *Phys. Rev. Lett.* **129** 050501
- [13] Giordani T *et al* 2023 Experimental certification of contextuality, coherence and dimension in a programmable universal photonic processor *Sci. Adv.* **9** eadj4249
- [14] Wagner R, Baldijão R D, Tezzin A and Amaral B 2023 Using a resource theoretic perspective to witness and engineer quantum generalized contextuality for prepare-and-measure scenarios *J. Phys. A: Math. Theor.* **56** 505303
- [15] Carceller C R I and Brask J B 2024 Contextuality witness inspired by optimal state discrimination *Phys. Rev. A* **109** 032202
- [16] Pusey M F, Barrett J and Rudolph T 2012 On the reality of the quantum state *Nat. Phys.* **8** 475–8
- [17] Leifer M S 2014 Is the quantum state real? an extended review of  $\psi$ -ontology theorems *Quanta* **3** 67–155
- [18] Helstrom C W 1976 *Quantum Detection and Estimation Theory Mathematics in Science and Engineering : A Series of Monographs and Textbooks* (Academic)
- [19] Ivanovic I D 1987 How to differentiate between non-orthogonal states *Phys. Lett. A* **123** 257–9
- [20] Dieks D 1988 Overlap and distinguishability of quantum states *Phys. Lett. A* **126** 303–6
- [21] Peres A 1988 How to differentiate between non-orthogonal states *Phys. Lett. A* **128** 19
- [22] Chefles A 1998 Unambiguous discrimination between linearly independent quantum states *Phys. Lett. A* **239** 339–47
- [23] Croke S, Andersson E, Barnett S M, Gilson C R and Jeffers J 2006 Maximum confidence quantum measurements *Phys. Rev. Lett.* **96** 070401
- [24] Mosley P J, Croke S, Walmsley I A and Barnett S M 2006 Experimental realization of maximum confidence quantum state discrimination for the extraction of quantum information *Phys. Rev. Lett.* **97** 193601
- [25] Lee H, Flatt K, Carceller C R I, Brask J B and Bae J 2022 Maximum-confidence measurement for qubit states *Phys. Rev. A* **106** 032422
- [26] Zhang Y, Schmid D, Ying Y and Spekkens R W 2025 Reassessing the boundary between classical and nonclassical for individual quantum processes
- [27] Plávala M and Gühne O 2024 Contextuality as a precondition for quantum entanglement *Phys. Rev. Lett.* **132** 100201
- [28] Wright V J and Farkas M'e 2023 Invertible map between bell nonlocal and contextuality scenarios *Phys. Rev. Lett.* **131** 220202
- [29] Khoshbin M, Catani L and Leifer M 2024 Alternative robust ways of witnessing nonclassicality in the simplest scenario *Phys. Rev. A* **109** 032212
- [30] Brask J B, Martin A, Esposito W, Houlmann R, Bowles J, Zbinden H and Brunner N 2017 Megahertz-rate semi-device-independent quantum random number generators based on unambiguous state discrimination *Phys. Rev. Appl.* **7** 054018
- [31] Lee H, Flatt K and Bae J 2025 Sequential quantum maximum-confidence discrimination *Phys. Rev. A* **112** 052206