University of California

Santa Barbara

Robust Ways of Witnessing Nonclassicality in the Simplest Scenario

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy

in

Mathematics

by

Massy Khoshbin

Committee in charge:

Professor David R. Morrison, Chair

Professor Stephen Bigelow

Professor Charles A. Akemann

Dr. Lorenzo Catani

March 2024

The dissertation of Massy Khoshbin is approved.

Professor Stephen Bigelow

Professor Charles A. Akemann

Dr. Lorenzo Catani

Professor David R. Morrison, Chair

March 2024

Robust Ways of Witnessing Nonclassicality in the Simplest Scenario

Dedicated to Professor Sadegh Angha

Acknowledgements

To say the path towards this dissertation has been a long and arduous trudge would be an understatement. The obstacles and challenges faced was exceeded by the support received from those around me. Without them, this would have been impossible.

First, I would like to acknowledge Medina Price. You always had conviction I would arrive at this point, despite my countless moments of doubt. Thank you for your consistent academic and moral support throughout my entire graduate journey.

I am incredibly grateful to my doctoral committee. Each of you has shown me more kindness and patience than I could have expected. Dr. Akemann and Dr. Bigelow, my sincerest thanks for all the time and support you have given in this process. Dr. Catani, I cannot thank you enough for your dedicated mentorship and providing me an opportunity to work and learn alongside you. Dr. Morrison, my heartfelt gratitude to you for taking me under your wing when I was most academically vulnerable. It has been a rare fortune to receive academic supervision from someone such as yourself.

Thank you to my peers and cohort at UCSB, especially Christian, Garo and Nadir. Your friendship made Santa Barbara, and particularly the grad tower in South Hall, home. I have enough wonderful memories with you all to last a lifetime.

A special thanks to Dr. Leifer, who prompted the work contained in this thesis.

To Matthew Brown, thank you so much for introducing me to the world of quantum foundations, and for the genuine interest, support and insight you imparted.

My warmest gratitude to all my friends in Texas for your deep love and support. You have made Dallas a home away from home and become a second family to me.

To my beloved parents and siblings, I am thankful beyond words for your undying love and support. Dad, my passion stems from your fervent spirit and wealth of wisdom.

Above all, I am eternally indebted and grateful to Professor Nader Angha. It is with your will that this has been accomplished. Your love and grace is immeasurable.

Massy Khoshbin

University of California
Santa Barbara, CA 93106

(562) 394-3922
massy@math.ucsb.edu

EDUCATION

Ph.D. Mathematics - University of California, Santa Barbara | March 2024
• Thesis: *Alternative robust ways of witnessing nonclassicality in the simplest scenario*
• Advisor: David R. Morrison
M.A. Mathematics - University of California, Santa Barbara | May 2017
B.A. Mathematics - University of California, San Diego| June 2008

TEACHING EXPERIENCE

Associate Instructor - UC Santa Barbara	Summer 2017, Summer 2020
Teaching Assistant - UC Santa Barbara	Fall 2014 - Summer 2021
Faculty/Math Specialist - Crane Country Day School	Fall 2019 - Spring 2020
Teaching Assistant and Calculus Lab Tutor - UC San Diego	Fall 2004 - Winter 2005

PUBLICATIONS

1. *Alternative robust ways of witnessing nonclassicality in the simplest scenario* (with L. Catani and M. Leifer). Phys. Rev. A, 2024, Volume 109, pp. 032212. eprint arXiv:2311.13474

RESEARCH EXPERIENCE

Ph.D. Research - UC Santa Barbara
Quantum information processing tasks in the presence of experimental noise
Foundations of quantum theory, measures of quantum contextuality
Generalized stabilizer subtheory within notions of nonclassicality
Quantum double hamiltonians for finite groups, gapped boundaries in phases of matter
Topological quantum computation and topological quantum field theory

Undergraduate Research - UC San Diego
Quantum computation, Grover's search algorithm, computational complexity
Presentation on "Quantum Computation and Grover's Search Algorithm" at the UCSD Faculty Mentor Program Research Symposium

Abstract

Robust Ways of Witnessing Nonclassicality in the Simplest Scenario

by

Massy Khoshbin

Motivated by treating the $2 \to 1$ parity-oblivious multiplexing under the influence of experimental noise, with its success probability serving as an indicator of nonclassicality, this thesis establishes connections between various notions of nonclassicality within the context of what is commonly referred to as the simplest nontrivial scenario (a prepare and measure scenario comprised of four preparations and two binary-outcome tomographically complete measurements). Specifically, we relate the established method developed by Pusey in [20] to witness a violation of preparation noncontextuality, which is not suitable in experiments where the operational equivalences to be tested are specified in advance, with a novel approach based on the notion of bounded ontological distinctness for preparations, defined by Chaturvedi and Saha in [9]. In our approach, we test bounded ontological distinctness for two particular preparations that are relevant in certain information processing tasks in that they are associated with the even- and odd-parity of the bits to be communicated. When there exists an ontological model where this distance is preserved we term this *parity preservation*. Our main result provides a noise threshold under which violating parity preservation (and so bounded ontological distinctness) agrees with the established method for witnessing preparation contextuality in the simplest nontrivial scenario. This is achieved by first relating the violation of parity preservation to the quantification of contextuality in terms of inaccessible information as developed by Marvian in [17], that we also show, given the way we quantify noise, to be more robust in witnessing contextuality than Pusey's noncontextuality inequality.

As an application of our findings, we treat the case of two-bit parity-oblivious multiplexing in the presence of noise. Specifically, leveraging the identified noise threshold for the existence of preparation contextuality, we establish a condition for which preparation contextuality is present in the case where the probability of success exceeds that achieved by any classical strategy. Overall, our results highlight that, below a certain threshold, all the different methods to witness nonclassicality agree. Consequently, an experimenter can choose the most suitable method based on their specific needs.

Contents

List of Figures

List of Tables

Chapter 1

Overview

1.1 Introduction

In the realm of quantum foundations, it is important to have an appropriate definition of what it means for a given quantum feature to resist an explanation within the classical worldview, *i.e.*, an appropriate notion of nonclassicality. One leading concept in this regard is preparation contextuality [24], which refers to the impossibility of a theory to admit of an ontological model that represents operationally equivalent preparations as identical probability distributions on the ontic state space.

A primary challenge in experimentally witnessing this notion is that the specified *a priori* target operational equivalences to test are not generally verified. To overcome this, Mazurek et al. in [18] introduced a solution which was later adopted by Pusey in [20]. This approach involves testing *a posteriori* operational equivalences that hold for the noisy preparations obtained in the experiment, successfully applied to witness preparation contextuality in the *simplest nontrivial scenario*, involving four preparations and two binary-outcome tomographically complete measurements (Fig. 2.1).

Despite the success of Pusey's approach, it proves unsuitable for certain information processing tasks in the simplest nontrivial scenario, powered by preparation contextuality, such as the two-bit parity-oblivious multiplexing (POM) [26], in the presence of noise. The reason lies in the necessity of specifying preparations and operational equivalences in advance for these tasks, preventing reanalysis of data using different sets of preparations. In this work, we present an alternative approach to detect nonclassicality in the simplest nontrivial scenario, offering a connection to Pusey's approach. Specifically, we introduce a novel approach based on a broader classicality concept named bounded ontological distinctness for preparations (BOD_P) [9], which still references the *a priori* ideal preparations.

Bounded ontological distinctness for preparations generalizes the notion of preparation noncontextuality, that demands that operational equivalences are mapped to ontological equivalences, insofar as it demands that also *differences* are preserved between the operational theory and the ontological model of the theory. In contrast to [9], we define differences in terms of distances: the operational distance as the maximal gap over all measurements between the outcome probabilities of the preparations, and the ontological distance as the total variational distance between the ontic distributions [17]. With this notion of classicality, we can now treat the simplest nontrivial scenario in the presence of noise and still refer to the a priori ideal preparations, by addressing the distance the noisy ones have with respect to them.

We test BOD_P for the distance between the even- and odd-parity mixtures of the preparations (*i.e.*, P_+ and P_- in Fig. 2.1) in the simplest nontrivial scenario, a relevant distance quantifying the communicated parity in POM in the presence of noise. We term the case with zero difference between the operational and ontological parity distances as satisfying *parity preservation.*

To connect our parity preservation approach with Pusey's approach based on preparation noncontextuality, we employ a third method developed by Marvian in [17]. Marvian's approach quantifies preparation contextuality in terms of the "inaccessible information" of an ontological model, defined as the largest distance between a pair of ontic distributions associated with operationally equivalent preparations.

We note that the work presented in this thesis contains material from the following manuscript:

M. Khoshbin, L. Catani, and M. Leifer, *Alternative robust ways of witnessing non-classicality in the simplest scenario*, Phys. Rev. A, 2024, Volume 109, pp. 032212.

1.2 Structure of thesis and outline of results

In Chapter 2, we provide the basic background and terminology regarding operational theories, their ontological models, and the property of noncontextuality. We also describe the simplest nontrivial scenario which is the setting our work takes place in, and define the distances we employ both at the operational and ontological level. We assume the reader to be familiar with the basic mathematical formulation of quantum mechanics and quantum computation ([27] and [19] are excellent references).

In Chapter 3, we describe Pusey's approach, Marvian's approach, bounded ontological distinctness for preparations (BOD_P), and our approach based on BOD_P. In Chapter 4, we introduce our characterization of experimental noise with respect to a noise bound δ, and present several results regarding the connection between these three approaches and their applications:

– Focusing on the simplest nontrivial scenario, we compare the noise thresholds for witnessing preparation contextuality in Marvian's and Pusey's approaches. We find that Marvian's is more robust (due to how to characterize noise, Marvian's inequality allows for better precision in bounding the relevant parameters in terms of δ), in that it detects contextuality given $\delta < 0.1$, while Pusey's requires $\delta < 0.06$ (Theorems 2 and 1). Here, the noise parameter δ represents the maximum allowed deviation of the noisy experimental preparations from the corresponding noiseless target preparations in terms of maximum distinguishability with a one-shot measurement. This means that, in a one-shot measurement, the noisy preparations cannot be distinguished from the ideal preparations with probability greater than $\frac{1+\delta}{2}$.

- We relate preparation contextuality as witnessed in Marvian's approach with violations of parity preservation. More precisely, we provide bounds defined in terms of the operational statistics for which a violation of Marvian's equality implies a violation of parity preservation and vice versa (Corollary 3).

- We rewrite the above result in terms of a noise bound by taking into account the noise parameter δ (Corollary 4). As a consequence, we also find that under the threshold $\delta = 0.007$, the three approaches – Pusey's, Marvian's and ours – agree on detecting nonclassicality; therefore an experimenter can choose the approach that they prefer in order to test nonclassicality (Corollary 6).

In Chapter 5, we apply the above results to the case of POM in the presence of noise, which is played in the simplest nontrivial scenario. It has been shown that winning the game with a probability of success greater than $\frac{3}{4} + \frac{\varepsilon}{4}$ (which can be achieved with a strategy involving classical bits) implies a violation of BOD_P [9], and more precisely, of parity preservation (Theorem 9). Here ε corresponds to the parity communicated as a consequence of the noise. Given that the threshold for violating the parity preservation is $\delta < 0.007$, and that below this threshold preparation noncontextuality is also violated (Theorem 6), we have a condition for which POM in the presence of noise is powered by preparation contextuality, thus extending the result of the noiseless case [26].

We note that a violation of BOD_P does not, in general, imply a violation of preparation noncontextuality, and a probability of success greater than $\frac{3}{4} + \frac{\varepsilon}{4}$ has been proven only to imply a violation of BOD_P, but not of preparation noncontextuality. Moreover, we note that in order to obtain such a result, it is insufficient to combine the no-go theorem of the noiseless case (*i.e.*, that a probability of success greater than $\frac{3}{4}$ implies preparation contextuality) with the fact that preparation contextuality holds whenever $\delta \leq 0.1$. This is because in the noisy POM one can in principle exploit the noise to communicate some parity also in classical strategies, thus improving the probability of success from $\frac{3}{4}$ to $\frac{3}{4} + \frac{\varepsilon}{4}$.

Section 5.3 presents a new argument providing an enhanced threshold ($\delta \leq 0.06$) for violating parity preservation, and subsequently for detecting nonclassicality in all three approaches. We note that this result extends beyond that which is in [15].

We conclude with Chapter 6, where we provide a discussion and outline future research.

Chapter 2

Preliminaries

In this chapter we provide the relevant background material to discuss preparation noncontextuality in the simplest scenario as well as the definitions of the distances that we use. In Sections 2.1, 2.2 and 2.3, we follow the treatment in [24].

2.1 Operational theories

The operational approach to a physical theory provides a mathematical framework for predicting the statistics of outcomes of experimental procedures. More formally, an *operational theory* is defined by a list of possible preparation procedures P, transformation procedures T, measurement procedures M, and experimental statistics $\mathcal{P}(k|P, T, M)$ indicating the conditional probability of observing the outcome k when the preparation P, transformation T, and measurement M are performed sequentially. A *prepare and measure* scenario indicates an operational theory where there are no transformation procedures, or when they are considered to be part of the preparations or the measurements. In this case, the operational statistics simplify to $\mathcal{P}(k|P, M)$.

For example, in an operational formulation of quantum theory, each preparation is associated with a density operator ρ on a Hilbert space, each transformation is associated with a completely positive and trace-preserving map Φ, each measurement is associated with a positive operator-valued measure $\{E_k\}$, and the probability of obtaining outcome k is given by the generalized Born rule, $\mathcal{P}(k|\rho, \Phi, E_k) = \mathrm{tr}[E_k \Phi(\rho)]$.

The mathematics of an operational theory is formalized by having an ontological model for it, which we examine next.

2.2 Ontological models

An *ontological model* of an operational theory is meant to provide a realist explanation of the operational predictions of the theory [13]. It is defined by a measurable space Λ of possible physical states, known as the *ontic state space*, with an associated σ-algebra Σ, and a set of measures or measurable functions over Λ that represent preparations, transformations and measurements. The points $\lambda \in \Lambda$ represents *ontic states* which encode all the physical properties of the system, and which are probabilistically assigned values. Ontic state spaces may or may not be finite or discrete.

For each preparation procedure P in the operational theory, there is an associated probability density over the state space in the ontological model, $\mu_P : \Lambda \to [0, 1]$. The μ_P are called *epistemic states* as they represent states of knowledge about the underlying ontic states. The probability that the physical state λ was prepared via the procedure P is given by $\mu_P(\lambda)$. Since a system is always in some physical state, the epistemic states satisfy $\int_\Lambda \mu_P(\lambda) d\lambda = 1$.

Similarly, the ontological model associates each transformation procedure T in the operational theory with a map $\Gamma_T : \Lambda \times \Lambda \to [0, 1]$. The conditional probability that some λ is sent to another λ' by T is given by $\Gamma_T(\lambda', \lambda)$. Since every physical state is mapped to some physical state by a transformation, we have $\int_\Lambda \Gamma_T(\lambda', \lambda) d\lambda' = 1$ for all $\lambda \in \Lambda$.

Next, each measurement procedure M is associated with a set of conditional probability functions $\{\xi_k^M : \Lambda \to [0, 1]\}_k$, where $\xi_k^M(\lambda)$ represents the probability of obtaining outcome k given that measurement M is implemented on a system in the ontic state λ. Since some measurement outcome always occurs, the sum over all possible outcomes satisfies $\sum_k \xi_k^M(\lambda) = 1$ for all $\lambda \in \Lambda$.

To correctly reproduce the experimental statistics of the operational theory, the distributions must satisfy the classical law of total probability,

$$\mathcal{P}(k|P,T,M) = \int_\Lambda \xi_k^M(\lambda')\Gamma_T(\lambda',\lambda)\mu_P(\lambda)d\lambda d\lambda^{'}.$$

We note that ontological models are assumed to be convex-linear, that is, a probabilistic implementation of a set of operations is represented by the probabilistic mixture of the corresponding probability densities.

2.3 Noncontextuality

The principle of noncontextuality states that operationally equivalent laboratory procedures must be represented as identical stochastic processes in one's ontological model. Given the rules for determining probabilities of outcomes, one can define a notion of equivalence among experimental procedures; two procedures are deemed *operationally equivalent* if they produce the same outcome statistics in every setting.

Specifically, two preparation procedures are equivalent if their conditional probabilities are equal for every possible transformation and measurement procedure:

$$P \simeq P' \text{ if } \mathcal{P}(k|P,T,M) = \mathcal{P}(k|P',T,M) \text{ for all } T, M.$$

The analogous condition holds for transformation and measurement procedures:

$$T \simeq T' \text{ if } \mathcal{P}(k|P,T,M) = \mathcal{P}(k|P,T',M) \text{ for all } P, M,$$

$$M \simeq M' \text{ if } \mathcal{P}(k|P,T,M) = \mathcal{P}(k|P,T,M') \text{ for all } P, T.$$

With this notion of operational equivalence established, we can now give a definition of noncontextuality at the ontological level. An ontological model is *preparation noncontextual* if any two such operationally equivalent preparations are represented by identical epistemic states: $P \simeq P' \implies \mu_P = \mu_{P'}$ (we note the converse trivially holds). If this fails to hold for all operationally equivalent pairs of procedures, we say the model is *preparation contextual*. Similarly, an ontological model is *transformation noncontextual* if operationally equivalent transformations are always represented by identical ontological maps: $T \simeq T' \implies \Gamma_T = \Gamma_{T'}$, and a failure results in transformation contextuality. Lastly, an ontological model is *measurement noncontextual* if operationally equivalent measurement procedures are always represented by identical functions on the ontological side, $M \simeq M' \implies \xi_k^M = \xi_k^{M'}$ for all k, and a failure constitutes measurement contextuality.

With this in mind, an operational theory is termed *preparation noncontextual* if there exists a preparation noncontextual ontological model for the theory. If no preparation noncontextual model exists for the theory, we say the operational theory is *preparation contextual*. The analogous notion applies for transformation and measurement noncontextuality at the operational level.

In this thesis, we work with operational theories of a specific type - those that can be associated with what is known as the simplest scenario - which we define in the following section. Accordingly, we only consider ontological models representing theories associated with the simplest scenario.

2.4 The simplest scenario

The simplest scenario is a prepare and measure scenario defined by four chosen preparations $\{P_{ij}\} = \{P_{00}, P_{01}, P_{10}, P_{11}\}$ (we note, given that all procedures in operational theories are naturally closed under probabilistic mixtures, this is to say that preparations in the theory belong to the convex hull of the four preparations $\{P_{ij}\}$), and two binary-outcome *tomographically complete* measurements $\{X, Y\}$. Tomographically complete measurements are able to uniquely identify the operational statistics of a preparation. In the quantum setting, this is the stipulation that operators must form an operator basis on the Hilbert space of the system, providing all the information about the state.

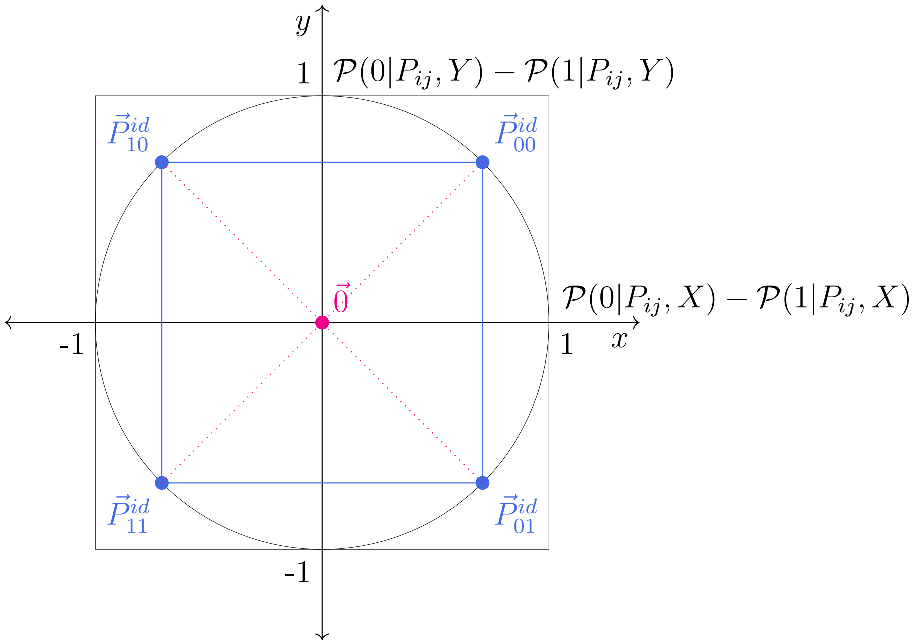
Theories associated with experiments consisting of four preparations and two binary-outcome tomographically complete measurements are termed the *simplest* scenario because they constitute the simplest nontrivial example of operational theories that can witness a violation of preparation contextuality. It is shown in Appendix B of [20] that any other scenario with fewer preparations or measurements always admits the existence of a preparation noncontextual model.

A geometrical representation of the scenario is depicted in Fig. 2.1, where the preparations are represented as vectors in the Cartesian plane with the x-axis specifying the difference between the probabilities of obtaining outcomes 0 and 1 for that preparation given the measurement X, and the y-axis specifying the same expression but given the measurement Y, *i.e.*, for $i, j \in \{0, 1\}$,

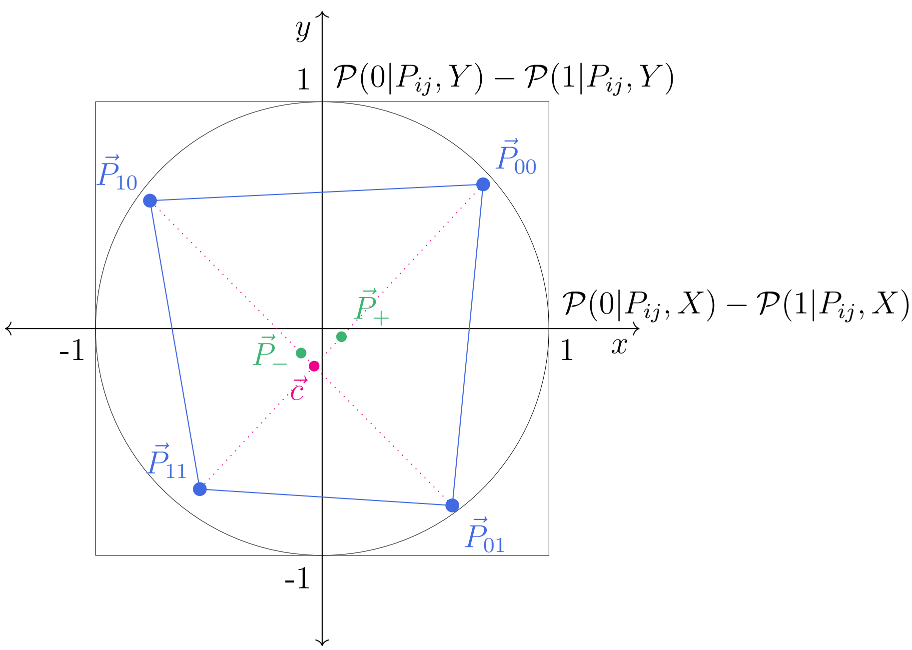
$$x_{ij} = \mathcal{P}(0|P_{ij}, X) - \mathcal{P}(1|P_{ij}, X), \tag{2.1}$$
$$y_{ij} = \mathcal{P}(0|P_{ij}, Y) - \mathcal{P}(1|P_{ij}, Y), \tag{2.2}$$
$$\vec{P}_{ij} = (x_{ij}, y_{ij}). \tag{2.3}$$

Notice that these coordinates take values in $[-1, 1]$, and so the preparations can at maximum span the square $\vec{P}_{00} = (1, 1), \vec{P}_{01} = (-1, 1), \vec{P}_{10} = (1, -1), \vec{P}_{11} = (-1, -1)$. This square corresponds to the *gbit* state space [3, 23], describing a non-physical theory beyond qubit quantum theory.

The simplest scenario is an extension of the Bloch disk (corresponding to the inscribed unit disk in the figure) representation of qubit quantum theory (this disc equates to X-Y planar cross section of Bloch sphere) to the full solid square, which entails states whose operational statistics encompass all possible theories.



(a)



(b)

Figure 2.1: **The simplest nontrivial scenario.** Four preparations (vertices of the blue square) and two tomographically complete measurements (corresponding to the x and y axes) are represented within the Bloch circle and the gbit square (in black). Fig. (a) represents the noiseless case. The a priori operational equivalence of Eq. (2.4) is represented by the vector $\vec{0}$ (in red). Fig. (b) represents the noisy case. The a posteriori operational equivalence of Eq. (2.5) is represented by the vector $\vec{c}$ (in red); the midpoints $\vec{P}_{+}, \vec{P}_{-}$ (in green) represent even- and odd-parity mixtures, respectively.

We note that X and Y are general measurements and do not represent the Pauli X and Y measurements unless we are explicitly in the case of qubit quantum theory. We adopt this notation for compatibility with the operational statistics being expressed via Cartesian coordinates. Nevertheless, the X and Y measurements do indeed coincide with the Pauli X and Y in the case where preparations belong to qubit quantum theory.

With the case of quantum theory, the four preparations giving the maximum violation of noncontextuality inequalities [20, 26, 7, 5] are (modulo the application of a unitary transformation on both preparations and measurements) denoted by $\{P_{00}^{id}, P_{01}^{id}, P_{10}^{id}, P_{11}^{id}\}$, and correspond to the vectors

$$\vec{P}_{00}^{id} \equiv \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \vec{P}_{01}^{id} \equiv \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \vec{P}_{10}^{id} \equiv \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \vec{P}_{11}^{id} \equiv \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

These are associated to the pure quantum states in the equator of the Bloch sphere located at angles $\frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$ with respect to the $+1$ eigenstate of the Pauli X measurement. The superscript *id* stands for “ideal”, stressing that these are the preparations one aims to prepare in a test of preparation contextuality in the simplest scenario.

Notably, this choice of preparations and measurements also provides the optimal quantum strategy in protocols like the two-bit parity-oblivious multiplexing [26], two-bit quantum random access codes [1], the CHSH* game [14] and several others [6]. Given that contextuality plays the role of resource for performing better than classical strategies in those protocols, we consider this choice in order to test the preparation contextuality of quantum theory for one qubit, and characterize experimental noise based on deviation from these points.

The coordinates (operational statistics) of the vector uniquely determine the preparation it represents. It follows that two preparations P_a and P_b are operationally equivalent – denoted by $P_a \simeq P_b$ – if and only if their coordinate vectors are equal, $P_a \simeq P_b \iff \vec{P}_a = \vec{P}_b$. Notice how the four ideal points $\vec{P}_{00}^{id}, \vec{P}_{01}^{id}, \vec{P}_{10}^{id}$, and $\vec{P}_{11}^{id}$ form two operationally equivalent decompositions of the preparation represented by the vector $\vec{0} = (0, 0)$ at the intersection of the x and y axes (the completely mixed state $\frac{I}{2}$), *i.e.*,

$$\frac{1}{2}\vec{P}_{00}^{id} + \frac{1}{2}\vec{P}_{11}^{id} = \vec{0} = \frac{1}{2}\vec{P}_{01}^{id} + \frac{1}{2}\vec{P}_{10}^{id}. \tag{2.4}$$

This operational equivalence is termed, the *a priori operational equivalence*, as it is the target operational equivalence one aims to obtain in an ideal experimental test of preparation contextuality in the simplest scenario prior to performing the actual realistic experiment.

In realistic experimental scenarios, one cannot exactly prepare the ideal preparations, but necessarily obtains some noisy version of them, which we denote by $\{\vec{P}_{00}, \vec{P}_{01}, \vec{P}_{10}, \vec{P}_{11}\}$. [1]

These obey an operational equivalence different from that of Eq. (2.4), and which we denote by the vector $\vec{c}$,

$$\underbrace{p\vec{P}_{00} + (1-p)\vec{P}_{11}}_{\vec{P}_p} = \vec{c} = \underbrace{q\vec{P}_{01} + (1-q)\vec{P}_{10}}_{\vec{P}_q}, \tag{2.5}$$

for two probability weights $p, q \in [0, 1]$, where we denote the two operationally equivalent preparations as P_p and P_q. This latter operational equivalence is termed the *a posteriori operational equivalence.*

If the preparations P_{ij} lie within quantum theory, we can view them as the image of P_{ij}^{id} under some completely positive trace-preserving map. In general theories, we can view the noise geometrically as having displaced the coordinates (operational statistics) of the preparation, and consider the degree of noise as treated by the radius (with respect to an operational distance) around the ideal points.

2.5 Distances

In this section we review the ontological and operational distances employed in our work. Specifically, we define the distance between preparation procedures as well as the distance

[1] We emphasize that in this work we assume only the preparations, and not the measurements, are subject to noise.

between their associated epistemic states. We start with the latter. In order to quantify the distance between two probability distributions μ_a and μ_b in the ontological model, we use the total variational distance, which is the standard choice (see also [17]):

$$d(\mu_a, \mu_b) \equiv \frac{1}{2} \int_\Lambda |\mu_a(\lambda) - \mu_b(\lambda)|.$$

On the other hand, we define the distances between two preparations P_a and P_b in the operational theory by the largest difference in the probabilities of measurement outcomes between two preparations across all available measurements and respective outcomes in the theory:

$$d(P_a, P_b) \equiv \max_{k,M} \{|\mathcal{P}(k|P_a, M) - \mathcal{P}(k|P_b, M)|\}.$$

For binary-outcome measurements, this quantity is the same for both the 0 and 1 outcome:

$$\begin{aligned} d(P_a, P_b) &= \max_{M} \{|\mathcal{P}(0|P_a, M) - \mathcal{P}(0|P_b, M)|\} \\ &= \max_{M} \{|\mathcal{P}(1|P_a, M) - \mathcal{P}(1|P_b, M)|\}. \end{aligned}$$

Given that there are only two measurements, X and Y, the operational distance can be equivalently expressed in the simplest scenario as follows,

$$d(P_a, P_b) \equiv \frac{1}{2} \max\{|x_a - x_b|, |y_a - y_b|\}. \tag{2.6}$$

The motivations for employing this specific operational distance are that it provides clear geometrical intuitions and that it makes the calculations tractable, as we will discuss in Chapter 6.

We recall that the definition of preparation noncontextuality stipulates

$$\forall P_a, P_b : \;\; P_a \simeq P_b \implies \mu_a(\lambda) = \mu_b(\lambda).$$

Thus, given the notions of distances just defined, we may recast the above definition with the following equivalent condition in terms of distances,

$$\forall P_a, P_b : \; d(P_a, P_b) = 0 \implies d(\mu_a, \mu_b) = 0.$$

Chapter 3

Approaches to Nonclassicality in the Simplest Scenario

In this chapter we describe the three approaches considered in this work: the first is due to Pusey [20] and provides robust noncontextuality inequalities for the simplest scenario; the second is due to Marvian [17] and witnesses preparation contextuality as the degree of what he defines as inaccessible information; the last is our novel approach which rephrases bounded ontological distinctness [9] in terms of the difference between the operational and ontological distances, and which specifically considers the preparations associated with the even- and odd-parity mixtures (P_+ and P_- in Fig. 2.1), thereby introducing the notion of parity preservation [15].

3.1 Pusey's approach

This approach allows one to test preparation contextuality in the simplest scenario considering the a posteriori operational equivalence of Eq.(2.5) and provides eight noncontextuality inequalities obtained by exploiting a connection to Bell inequalities in the famous Clauser-Horne-Shimony-Holt (CHSH) scenario [10]. The requirement of preparation noncontextuality for this a posteriori operational equivalence implies the existence of an ontological model for which the epistemic states μ_p, μ_q associated with the operationally equivalent preparations P_p and P_q are such that $\mu_p = \mu_q$, where $\mu_p = p\mu_{00} + (1-p)\mu_{11}$ and $mu_q = q\mu_{01} + (1-q)\mu_{10}$. It is shown in [20] that this would be sufficient to make the ontological model preparation noncontextual, *i.e*, that any other operational equivalence within the convex hull of $\{P_{ij}\}$ corresponds to an ontological equivalence in the model. This lucrative property is not true in general, as one typically one cannot conclude preparation noncontextuality from a single pair of equivalent procedures, however in the particular case of our simplest scenario it holds.

Pusey's inequalities are denoted by $S(x_{ij}, y_{ij}) \leq 0$. They correspond to the analogue of the eight CHSH inequalities in the prepare and measure scenario we are considering. In terms of the coordinates (x_{ij}, y_{ij}) of the noisy preparations, one can write a representative of them as follows, where $S(x_{ij}, y_{ij})$ denotes the expression on the left:

$$p(x_{00} + y_{00} + x_{11} + y_{11}) + q(x_{01} - y_{01} + x_{10} - y_{10}) + (y_{10} - x_{10} - x_{11} - y_{11}) - 2 \leq 0. \quad (3.1)$$

This inequality is maximally violated, in quantum theory, by the choice of states and measurements described in the Section 2.4 (see Fig. 2.1a), and the expression takes the value $S(x_{ij}^{id}, y_{ij}^{id}) = 4\left(\frac{1}{\sqrt{2}}\right) - 2 \approx 0.82$. When the (x_{ij}, y_{ij}) form a square centered at the origin with radius a, one can rewrite the expression $S(x_{ij}, y_{ij})$ as $4(a) - 2$. Therefore the algebraic maximum value of $S(x_{ij}, y_{ij})$ is 2, which is obtained when considering the vertices of the *gbit* (generalized bit) square [3], which represents the quantification of contextuality by the gbit theory with this notion of nonclassicality.

We emphasize again that Pusey's approach, which considers the a posteriori operational equivalence, is not suitable when considering cases that require the operational equivalence to be specified in advance. An example is the case of the parity-oblivious multiplexing protocol, where the a posteriori operational equivalence does not embody the requirement of parity-obliviousness, unlike the a priori operational equivalence.

3.2 Marvian's approach

Marvian's approach quantifies preparation contextuality through the notion of *inaccessible information*, which is defined to be the largest distance between pairs of epistemic states associated to equivalent preparation procedures, minimized over all possible ontological models

$$C_{\text{prep}}^{\min} \equiv \inf_{\text{Models}} \sup_{P_a \simeq P_b} d(\mu_a, \mu_b). \tag{3.2}$$

We note that the inaccessible information $C_{\text{prep}}^{\min}$ of an operational theory is zero if and only if the theory admits of a preparation noncontextual model. That is,

$$C_{\text{prep}}^{\min} = 0 \iff \text{preparation noncontextuality.}$$

This equality is referred to as *Marvian's preparation noncontextuality equality.*

In [17], Marvian provides a lower bound on $C_{\text{prep}}^{\min}$ in terms of operational quantities to witness preparation contextuality. In the simplest scenario, these quantities reduce to a function of the preparations – here denoted with $\gamma(x_{ij}, y_{ij})$ – the details of which are provided in Section 4.2: $C_{\text{prep}}^{\min} \geq \gamma(x_{ij}, y_{ij})$. In contrast to the notion of bounded ontological distinctness and parity preservation that are defined in the next sections, the approaches of Marvian and Pusey both refer to the same notion: preparation noncontextuality. Therefore, they always agree in detecting contextuality in the noiseless case. However, in the noisy case and given the way we quantify noise, they provide different thresholds below which they are guaranteed to be violated, as we will see at the end of Section 4.2. In short, we say that they provide a different *robustification.* This is why they are treated separately and it is important for our purposes to consider both.

3.3 Bounded ontological distinctness for preparations (BOD_P)

The generalized notion of noncontextuality reflects the natural principle that operational equivalences should be mapped to ontological identities. The approach of bounded ontological distinctness for preparations (BOD_P) extends this idea from operational and ontological equivalences to operational distinguishability and ontological distinctness.

BOD_P was introduced in [9] and it is a criterion of classicality that requires the equivalence of the operational distinguishability between any two preparations in the operational theory and the ontological distinctness between their ontic representations, generalizing the notion of preparation noncontextuality and being based on the same credentials (a methodological principle inspired by Leibniz [25], or, equivalently, a principle of no operational fine tuning [4]). The *operational distinguishability* $s_{\mathcal{O}}^{P_a,P_b}$ of a pair of preparations P_a and P_b is defined as

$$s_{\mathcal{O}}^{P_a,P_b} \equiv \frac{1}{2}\max_{M}\left\{\mathcal{P}(0|P_a,M)+\mathcal{P}(1|P_b,M)\right\}, \tag{3.3}$$

where the maximum is over all measurements in the operational theory.

The *ontological distinctness* $s_{\Lambda}^{\mu_a,\mu_b}$ for the corresponding pair of epistemic states μ_a and μ_b is defined as

$$s_{\Lambda}^{\mu_a,\mu_b} \equiv \frac{1}{2}\int_{\Lambda}\max\{\mu_a(\lambda),\mu_b(\lambda)\}. \tag{3.4}$$

We say that an operational theory admits of a model satisfying BOD_P if the value of operational distinguishability for any pair of preparations in the theory equals the value of ontological distinctness for the pair of epistemic states. That is, *bounded ontological distinctness for preparations* demands the following equality to hold for *all pairs* of preparations and corresponding epistemic states:

$$BOD_P \iff s_{\mathcal{O}}^{P_a,P_b} = s_{\Lambda}^{\mu_a,\mu_b}\ \forall a,b. \tag{3.5}$$

To see that a model satisfying BOD_P is preparation noncontextual, we consider any operationally equivalent pair P_a and P_b and observe that the expression in Eq. (3.3) reduces to $\frac{1}{2}$ when $P_a \simeq P_b$. Given that $s_{\mathcal{O}}^{P_a,P_b}$ represents the maximum probability of distinguishing the two preparations, the value $\frac{1}{2}$ asserts the fact that operationally equivalent procedures are completely indistinguishable. The BOD_P criterion in Eq. (3.5) would then imply that the summation in Eq. (3.4) equals one, which entails that $\mu_a(\lambda) = \mu_b(\lambda)\ \forall \lambda \in \Lambda$. We therefore have $BOD_P \implies$ preparation noncontextuality. While this implication is always true, we emphasize that the converse does not necessarily hold.

In general, an operational theory may not admit of an ontological model satisfying BOD_P, thus implying that there exists a pair of preparations for which $s_{\Lambda}^{\mu_a,\mu_b} - s_{\mathcal{O}}^{P_a,P_b} > 0$. The difference $s_{\Lambda}^{\mu_a,\mu_b} - s_{\mathcal{O}}^{P_a,P_b}$ is not explicitly treated by the authors of [9], however it is of crucial relevance in the present work given that we want to quantify this notion of nonclassicality and consider how it robustifies in the case of realistic noisy scenarios, like the simplest scenario. We quantify the violation of BOD_P as the difference of operational and ontological *distances*. Such quantification can be simply related to the original definition of BOD_P in terms of distinguishability and distinctness, as we now show. Recalling equations (2.1), (2.2), (2.6), and considering the available measurements in the simplest scenario, the relationship between operational distinguishability and distance is given by

$$
\begin{aligned}
s_{\mathcal{O}}^{P_a,P_b} &= \frac{1}{2}\max_{M}\{\mathcal{P}(0|P_a,M) + \mathcal{P}(1|P_b,M)\} \\
&= \max\left\{\frac{1+\frac{1}{2}|x_a - x_b|}{2}, \frac{1+\frac{1}{2}|y_a - y_b|}{2}\right\} \\
&= \frac{1+d(P_a,P_b)}{2}.
\end{aligned}
\tag{3.6}
$$

A similar relationship holds between the ontological distinctness and distance:

$$
\begin{aligned}
s_{\Lambda}^{\mu_a,\mu_b} &= \frac{1}{2}\int_{\Lambda} \max\left\{\mu_a(\lambda), \mu_b(\lambda)\right\} \\
&= \frac{1}{2}\left(1 + \frac{1}{2}\int_{\Lambda} |\mu_a(\lambda) - \mu_b(\lambda)|\right) \\
&= \frac{1 + d(\mu_a, \mu_b)}{2}.
\end{aligned}
\tag{3.7}
$$

We denote with $\mathcal{D}_{P_a,P_b}$ the difference between the ontological and operational distances for the pair of preparations P_a and P_b and their associated epistemic states. The expression $\mathcal{D}^{\min}_{P_a,P_b}$ denotes the difference $\mathcal{D}_{P_a,P_b}$ minimized over all possible ontological models. That is,

$$
\begin{aligned}
\mathcal{D}_{P_a,P_b} &\equiv d(\mu_a, \mu_b) - d(P_a, P_b), \\
\mathcal{D}^{\min}_{P_a,P_b} &\equiv \inf_{\text{Models}} \mathcal{D}_{P_a,P_b}.
\end{aligned}
\tag{3.8}
$$

Combining equations (3.6), (3.7) and (3.8), we view the difference of operational distinguishability and ontological distinctness as half the difference in operational and ontological distances:

$$
s_{\Lambda}^{\mu_a,\mu_b} - s_{\mathcal{O}}^{P_a,P_b} = \frac{1}{2}\mathcal{D}_{P_a,P_b}.
$$

Therefore $s_{\Lambda}^{\mu_a,\mu_b} - s_{\mathcal{O}}^{P_a,P_b} = 0$ if and only if $\mathcal{D}_{P_a,P_b} = 0$ and bounded ontological distinctness for preparations (BOD_P) can be equivalent expressed as the difference between operational and ontological distances being zero for all pairs:

$$
BOD_P \iff \mathcal{D}^{\min}_{P_a,P_b} = 0 \ \forall P_a, P_b. \tag{3.9}
$$

3.4 Parity preservation

In this thesis we test BOD_P for the difference between the operational and ontological distance, $\mathcal{D}_{P_+,P_-}$, of the even- and odd-parity mixtures (see Fig. 2.1),

$$
P_+ = \frac{P_{00} + P_{11}}{2}, P_- = \frac{P_{01} + P_{10}}{2}.
$$

The operational distance $d(P_+, P_-)$ indeed codifies the information about the parity between the bits i and j labeling the four preparations $\{P_{ij}\}$ if one measures the preparation with the measurements X and Y. For example, if $d(P_+, P_-) = 0$, then by measuring X and Y on P_+ and P_- one would always get the same outcome, thus obtaining a probability $\frac{1}{2}$ for distinguishing between them and no information about the parity between the bits i and j.

If $d(P_+, P_-) \neq 0$ then some information about the parity between the bits i and j can be obtained by measuring X and Y. If $d(P_+, P_-)$ is preserved in the ontological model, meaning $\mathcal{D}_{P_+,P_-} = 0$, we say that there is *parity preservation*. Clearly, a violation of parity preservation implies a violation of BOD_P, but not vice versa. If an operational theory does not admit of a parity preserving ontological model, then $\mathcal{D}^{\min}_{P_+,P_-} > 0$. The focus on parity preservation is relevant in the context of parity-oblivious multiplexing and will allow us to connect a violation of BOD_P with a violation of preparation noncontextuality given certain bounds (Corollaries 3 and 4).

Chapter 4

Results

In this chapter we state the main results of our work (see Table 4.1 for a concise summary). We begin with Theorem 1, which reformulates Pusey's preparation noncontextuality inequality with the specification of a noise parameter δ (see Fig. 4.1). This enables us to determine, in Corollary 1, a noise threshold below which the noncontextuality inequality is violated. The same is done for Marvian's preparation noncontextuality equality (Theorem 2) and a noise threshold is found below which it is still violated (Corollary 2). We continue by establishing a connection between Marvian's noncontextuality equality and parity preservation in Theorem 3. As a consequence of this theorem, in Corollary 3, we provide conditions for which a violation of one notion implies a violation of the other. That is, we show that with enough preparation contextuality, one is certain to violate parity preservation, and vice versa. The parameters for the conditions in Corollary 3 are defined from the experimental data and can be rewritten, with the aid of Theorem 4, in terms of the noise parameter δ. In this way, Corollary 3 can be rephrased into Corollary 4. The latter is used to quantify the amount of noise needed to guarantee a violation of parity preservation – Corollary 5. Finally, having obtained the noise thresholds for violating each notion of nonclassicality considered in this work, we provide what is arguably the most relevant result – Corollary 6 – that establishes a noise threshold below which all three approaches agree in witnessing nonclassicality.

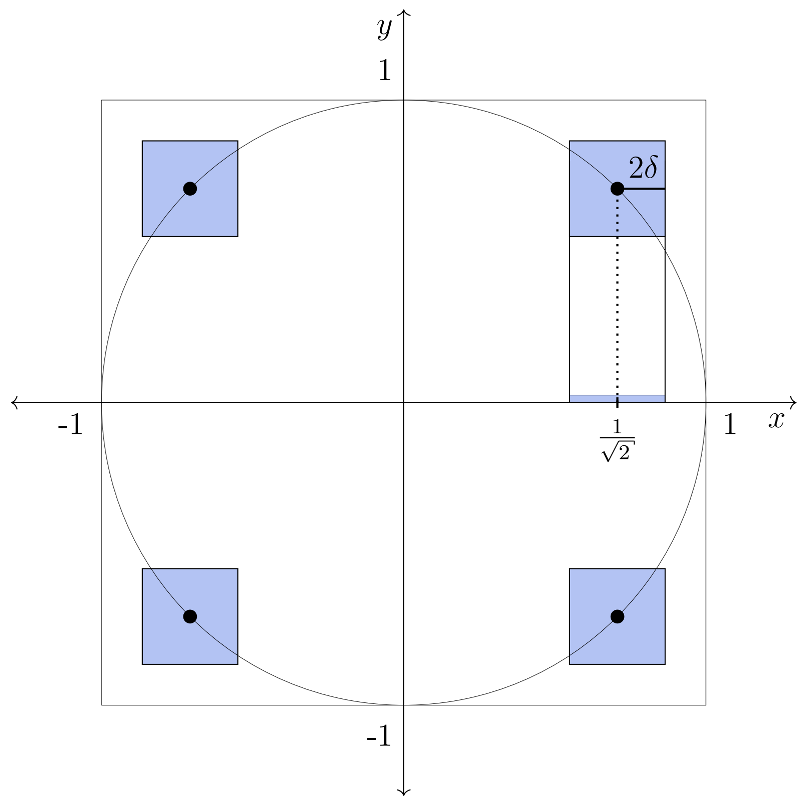


Figure 4.1: **Noise bound of δ.** Ideal points are indicated at the center of the blue squares. These squares of radius 2δ represent regions where noisy points reside, and, in accordance with Eq. (2.6), contain all preparations with an operational distance of *at most* δ from the ideal points, *i.e.*, $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \delta$. In other words, noisy points in the blue squares cannot be distinguished from their ideal counterparts with a probability greater than δ in any one-shot measurement.

4.1 Threshold for violating Pusey's inequality

Theorem 1. *Suppose the preparations $\{P_{ij}\}$ of the simplest scenario satisfy a noise bound $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \delta$, where δ is the noise parameter and $\{P_{ij}^{id}\}$ are the ideal a priori preparations. Pusey's expression $S(x_{ij}, y_{ij})$ of Eq. (3.1) satisfies the following lower bound in terms of the noise parameter δ:*

$$S(x_{ij}, y_{ij}) \geq 2\sqrt{2} - 2 - 16\delta + 32\sqrt{2}\,\delta^2, \tag{4.1}$$

where $\{x_{ij}, y_{ij}\}$ are the coordinates of the preparations $\{P_{ij}\}$.

Proof. Let us consider Pusey's noncontextuality inequality of Eq. (3.1) for the case where we insert the coordinates of the preparations P_{ij} satisfying the operational equivalence (2.5). Suppose $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq \delta$. Given the coordinates of the ideal preparations $\vec{P}^{id}_{ij}$, it follows that $l_\delta \leq |x_{ij}|, |y_{ij}| \leq u_\delta$, where $l_\delta = \frac{1}{\sqrt{2}} - 2\delta$ and $u_\delta = \frac{1}{\sqrt{2}} + 2\delta$. This can be seen in Fig. 4.1, where any noisy point $\vec{P}_{ij}$ within an operational distance of δ from the ideal points $\vec{P}^{id}_{ij}$ has coordinates whose absolute values lie within the range $\left[\frac{1}{\sqrt{2}} - 2\delta, \frac{1}{\sqrt{2}} + 2\delta\right]$. Further, $p, q \geq \frac{1-4\sqrt{2}\delta}{2}$ (see Fig. 4.2).

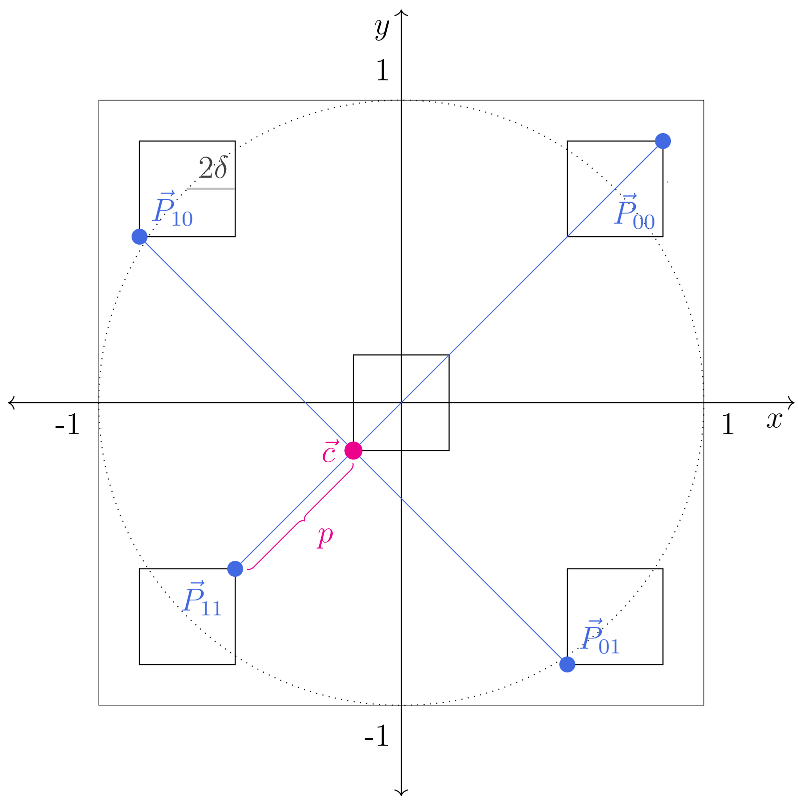


Figure 4.2: **Finding a minimum weight p.** An example where the weight p in Eq. (2.5) takes the minimum possible value, corresponding to $p = \frac{1-4\sqrt{2}\delta}{2}$. This can be seen by noting that p is the weight associated to the distance between $\vec{P}_{11}$ and $\vec{c}$ and observing that the Euclidean distance between $\vec{P}_{11}$ and $\vec{c}$ equals $1 - 4\sqrt{2}\delta$, while the Euclidean distance between $\vec{P}_{11}$ and $\vec{P}_{00}$ equals 2.

Applying these bounds to Eq. (3.1), we obtain

$$\begin{aligned}
&S(x_{ij}, y_{ij}) \\
&= p(x_{00} + y_{00} + x_{11} + y_{11}) + q(x_{01} - y_{01} + x_{10} - y_{10}) + (y_{10} - x_{10} - x_{11} - y_{11}) - 2 \\
&\geq \frac{1 - 4\sqrt{2}\,\delta}{2}(l_\delta + l_\delta - u_\delta - u_\delta) + \frac{1 - 4\sqrt{2}\,\delta}{2}(l_\delta + l_\delta - u_\delta - u_\delta) + (l_\delta + l_\delta + l_\delta + l_\delta) - 2 \\
&= 2(1 - 4\sqrt{2}\,\delta)(l_\delta - u_\delta) + 4l_\delta - 2 \\
&= 2(1 - 4\sqrt{2}\,\delta)\left(\frac{1}{\sqrt{2}} - 2\delta - \frac{1}{\sqrt{2}} - 2\delta\right) + 4\left(\frac{1}{\sqrt{2}} - 2\delta\right) - 2 \\
&= 2\sqrt{2} - 2 - 16\delta + 32\sqrt{2}\,\delta^2.
\end{aligned}$$

□

Given that the corresponding preparation noncontextuality inequality is $S(x_{ij}, y_{ij}) \leq 0$, solving for the right hand side in Eq. (3.1) results in a violation threshold of $\delta \approx 0.06$, which leads to the following corollary.

Corollary 1. *If $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq 0.06$, then $S(x_{ij}, y_{ij}) > 0$ and Pusey's preparation noncontextuality inequality is violated.*

4.2 Threshold for violating Marvian's equality

Theorem 2. *Suppose the preparations $\{P_{ij}\}$ of the simplest scenario satisfy a noise bound $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \delta$, where δ is the noise parameter and $\{P_{ij}^{id}\}$ are the ideal a priori preparations. Marvian's inaccessible information of Eq. (3.2) of the scenario satisfies the following lower bound in terms of the noise parameter δ:*

$$C_{\text{prep}}^{\min} \geq \frac{\sqrt{2} - 4\delta - 1}{4(\sqrt{2} - 4\delta)}. \tag{4.2}$$

Proof. We begin by using an inequality proven in [17], that provides a lower bound for the inaccessible information of an operational theory:

$$C_{\text{prep}}^{\min} \geq \frac{P_{\text{guess}} - (1 - \frac{d-1}{d}\beta_{\min}^{-1})}{(d-1)d^{n-1}}. \tag{4.3}$$

The values n and d represent the number of measurements in the theory and number of outcomes for each measurement, respectively. In the simplest scenario, we have $d = n = 2$. The expression P_{guess} can be interpreted as a probability of success for a particular type of information processing task game, and is defined in the general setting in [17]. In the simplest scenario, P_{guess} simplifies to

$$\frac{1}{4}[P(0|X, R_{00}) + P(1|X, R_{01}) + P(0|Y, R_{10}) + P(1|Y, R_{11})], \tag{4.4}$$

for a choice of preparations $R_{00}, R_{01}, R_{10}, R_{11}$ from the operational theory. The value of P_{guess} therefore varies based upon the choice, and one aims to maximize this value; to find the maximum of P_{guess} in our scenario, we choose the following preparations:

$$\vec{R}_{00} = (1, 0), \vec{R}_{01} = (-1, 0), \vec{R}_{10} = (0, 1), \vec{R}_{11} = (0, -1). \tag{4.5}$$

Applying (4.5) to (4.4) gives $P_{\text{guess}} = 1$. Although these four points are not in our a priori simplest scenario, we can augment the convex hull of $\{\vec{P}_{ij}\}$ with the set of these additional four points, $U = \{(\pm 1, 0), (0, \pm 1)\}$, as they do not affect or contribute to the contextuality of the scenario. They indeed correspond to the stabilizer states in qubit theory, that are known not to violate preparation noncontextuality inequalities [7]. This latter fact can be also seen through Pusey's preparation noncontextuality expression, whereby a) the calculated value of $S(x_{ij}, y_{ij})$ for U is zero, and b) the maximal value of contextuality via $S(x_{ij}, y_{ij})$ is always attained with the initial four noisy preparations P_{ij}. We augment the set of preparations in our simplest scenario with U for the purposes of utilizing Marvian's guessing probability; the inclusion of U allows us to employ the inequality in Eq. (4.3) in a manner that detects nonclassicality appropriately. That is, with $P_{\text{guess}} = 1$, the right-hand side of Eq. (4.3) is positive precisely when there is preparation contextuality.

With these values applied to Eq. (4.3), we have the following reduction:

$$C_{\text{prep}}^{\min} \geq \frac{1}{4}\beta_{\min}^{-1}. \tag{4.6}$$

The remaining term, $\beta_{\min}$, is obtained from the operational statistics and so the right hand side in Eq. (4.6) is solely a function of the preparations - it is what we referred to as $\gamma(x_{ij}, y_{ij})$ in Eq. (4.3). We now define and bound $\beta_{\min}$. Following [17], we have

$$\beta_{\min} \equiv \inf_{P} \max_{i,j} 2^{d_{\max}(P, Q_{ij})}, \tag{4.7}$$

where the infimum is taken over all preparations P in the theory, $Q_{ij} \equiv \frac{1}{2}R_{0i} + \frac{1}{2}R_{1j}$ for $i, j \in \{0, 1\}$, and the distance $d_{\max}$ is the operational max relative entropy for preparations:

$$d_{\max}(P_a, P_b) \equiv -\log_2 \sup\{u : u \leq 1, \exists P_{a'} : P_b \simeq uP_a + (1-u)P_{a'}\}.$$

Considering our choices in Eq. (4.5), we have $\vec{Q}_{00} = (\frac{1}{2}, \frac{1}{2}), \vec{Q}_{01} = (\frac{1}{2}, -\frac{1}{2}), \vec{Q}_{10} = (-\frac{1}{2}, \frac{1}{2})$, and $\vec{Q}_{11} = (-\frac{1}{2}, -\frac{1}{2})$. We can now bound $\beta_{\min}$ from above:

$$\begin{aligned}
\beta_{\min} &= \inf_{P} \max_{i,j} 2^{d_{\max}(P, Q_{ij})} \\
&= \inf_{P} \max_{i,j} 2^{-\log_2 \sup\{u : u \leq 1, \exists P_a : Q_{ij} \simeq uP + (1-u)P_a\}} \\
&= \inf_{P} \max_{i,j} (\sup\{u : u \leq 1, \exists P_a : Q_{ij} \simeq uP + (1-u)P_a\})^{-1} \\
&= \max_{i,j} (\sup\{u : u \leq 1, \exists P_a : Q_{ij} \simeq u\frac{I}{2} + (1-u)P_a\})^{-1} \\
&\leq \max_{i,j} (\{u : Q_{ij} \simeq u\frac{I}{2} + (1-u)S_{ij}\})^{-1} \\
&= (\{u : Q_{00} \simeq u\frac{I}{2} + (1-u)S_{00})^{-1} \\
&= (\{u : \frac{1}{\sqrt{2}} = u \cdot 0 + (1-u)(1 - 2\sqrt{2}\delta)\})^{-1} \\
&= \left(\frac{\sqrt{2} - 4\delta - 1}{\sqrt{2} - 4\delta}\right)^{-1} \\
&= \frac{\sqrt{2} - 4\delta}{\sqrt{2} - 4\delta - 1}.
\end{aligned}$$

In the fourth line we use the fact that, as also showed in [17], the infimum is achieved for the completely mixed state $\frac{I}{2}$. The upper bound in the fifth line arises from finding the smallest value that $\sup\{u\}$ can be guaranteed to achieve from mixing $\frac{I}{2}$ with a preparation P_a in our theory to output a fixed Q_{ij}. Any $\vec{P}_a$ that lies within the convex hull of the noisy points is a candidate to mix with $\frac{I}{2}$, and the optimal value is achieved with the preparations indicated by $\vec{S}_{ij}$ (see Fig. 4.3). Due to symmetry, the value of u is the same for any pair of $\vec{Q}_{ij}$ and $\vec{S}_{ij}$. In line 6 we choose, without loss of generality, $ij = 00$ to calculate the u value. We conclude with solving $\frac{1}{\sqrt{2}} = (1-u)(1-2\sqrt{2}\delta)$ for u. Therefore $\beta_{\min}^{-1} \geq \frac{\sqrt{2}-4\delta-1}{\sqrt{2}-4\delta}$. This inequality applied to Eq. (4.6) establishes the result.

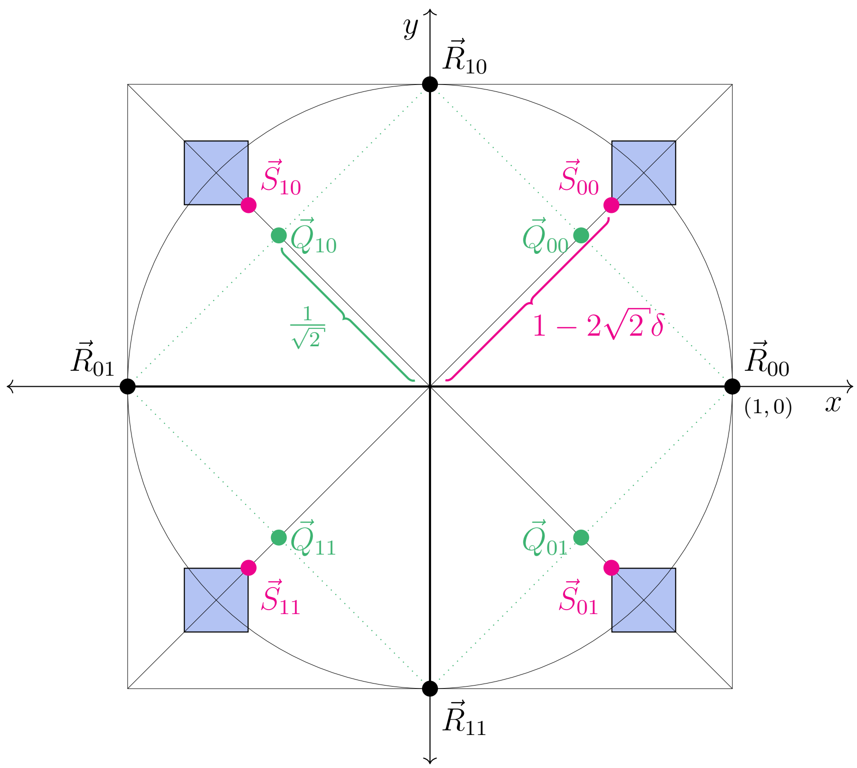


Figure 4.3: **Bounding the quantity $\beta_{\min}$.** The points $\vec{R}_{ij}$ corresponding to the states R_{ij} used to evaluate P_{guess} are shown in black. Their equal mixtures, denoted $\vec{Q}_{ij}$, which are used in calculating $\beta_{\min}$, are indicated in green. The red $\vec{S}_{ij}$ denote the points that are radially furthest from the $\vec{Q}_{ij}$ which are guaranteed to lie within the convex hull of the noisy preparations. That the $\vec{S}_{ij}$ are radially furthest away from the origin (the completely mixed state) is what ensures $d_{\max}$ to be the largest distance. Note that with less noise (smaller δ), the value of $\sup\{u\}$ increases, since the stationary $\vec{Q}_{ij}$ would then be (relatively) closer to $\vec{0}$ than the $\vec{S}_{ij}$; this in turn decreases $\frac{1}{u}$, which bounds $\beta_{\min}$ from above.

□

Given that preparation noncontextuality coincides with $C_{\text{prep}}^{\min} = 0$, and we obtain a threshold of violation $\delta \approx 0.1$ when solving for the right hand side in Eq. (4.2), we have the following corollary (see Fig. 4.4).

Corollary 2. *If $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq 0.1$, then $C_{\text{prep}}^{\min} > 0$ and Marvian's preparation noncontextuality equality is violated.*

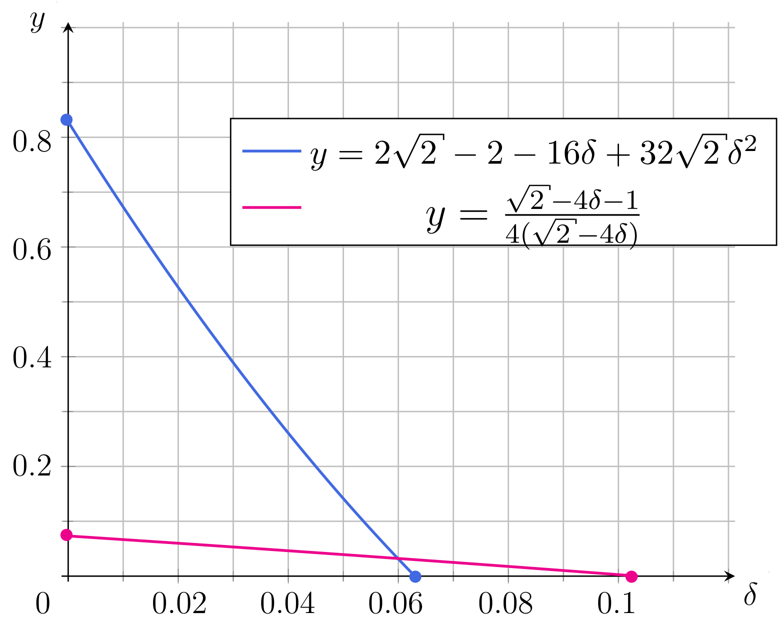


Figure 4.4: **Preparation contextuality witnessed by Pusey's and Marvian's approaches.** The blue and red curves plot the functions on the right hand sides of equations (3.1) and (4.2), respectively. They provide upper bounds to Pusey's and Marvian's expressions, respectively, in terms of the noise parameter δ. What is relevant in the figure is the range of δ in which each function takes a positive value, thus detecting contextuality. The specific values of the two functions are not to be compared as they are meaningful only within the scope of each approach.

4.3 Relating Marvian's equality to parity preservation

We begin by establishing the following inequality, proven below, connecting $\mathcal{D}_{P_+,P_-}^{\min}$ to $C_{\text{prep}}^{\min}$, thus relating our approach to Marvian's.

Theorem 3. *Given the simplest scenario with even- and odd-parity preparations P_+, P_- and inaccessible information $C_{\text{prep}}^{\min}$, there exist functions α_1, α_2, and α_3 of the preparations $\{P_{ij}\}$ satisfying*

$$\mathcal{D}_{P_+,P_-}^{\min} \geq \alpha_1 C_{\text{prep}}^{\min} - \alpha_2,$$
$$\mathcal{D}_{P_+,P_-}^{\min} \leq \alpha_1 C_{\text{prep}}^{\min} + \alpha_3.$$

Proof. We first recall that

$$\underbrace{p\vec{P}_{00} + (1-p)\vec{P}_{11}}_{\vec{P}_p} = \vec{c} = \underbrace{q\vec{P}_{01} + (1-q)\vec{P}_{10}}_{\vec{P}_q},$$

and that $\vec{P}_+ = \frac{1}{2}\vec{P}_{00} + \frac{1}{2}\vec{P}_{11}$, $\vec{P}_- = \frac{1}{2}\vec{P}_{01} + \frac{1}{2}\vec{P}_{10}$.

We will proceed to define a weight r and preparations $P_{+'}$ and $P_{-'}$ that pair with P_+ and P_- such that the following two criteria hold:

- $P_{+'}$ and $P_{-'}$ are *also* convex combinations of $\{P_{00}, P_{11}\}$ and $\{P_{01}, P_{10}\}$, respectively.
- $\vec{P}_p$ and $\vec{P}_q$ can be recast using one common weight r as

$$\underbrace{(1-r)\vec{P}_+ + r\vec{P}_{+'}}_{\vec{P}_p} = \vec{c} = \underbrace{(1-r)\vec{P}_- + r\vec{P}_{-'}}_{\vec{P}_q}. \tag{4.8}$$

With Eq. (4.8), we are able to write $\mu_p = (1-r)\mu_+ + r\mu_{+'}$ and $\mu_q = (1-r)\mu_- + r\mu_{-'}$. We can constructively define the weight r as follows. To start, observe that $\vec{c}$ is either a convex combination of $\{\vec{P}_+, \vec{P}_{00}\}$ or $\{\vec{P}_+, \vec{P}_{11}\}$, and that the weight p determines which combination of the two pairs gives $\vec{c}$. The same is true for $\vec{c}$ being expressed as either a convex combination of $\{\vec{P}_-, \vec{P}_{01}\}$ or $\{\vec{P}_-, \vec{P}_{10}\}$, with the weight q which pair. With this in mind, we define the even- and odd-parity weights, r_+ and r_-, to be (here all magnitudes $||\cdot||$ indicate Euclidean length),

$$r_+ \equiv \begin{cases} \frac{||\vec{P}_+ - \vec{c}||}{||\vec{P}_+ - \vec{P}_{00}||} & p \geq \frac{1}{2} \\ \frac{||\vec{P}_+ - \vec{c}||}{||\vec{P}_+ - \vec{P}_{11}||} & p \leq \frac{1}{2} \end{cases}, \quad r_- \equiv \begin{cases} \frac{||\vec{P}_- - \vec{c}||}{||\vec{P}_- - \vec{P}_{01}||} & q \geq \frac{1}{2} \\ \frac{||\vec{P}_- - \vec{c}||}{||\vec{P}_- - \vec{P}_{10}||} & q \leq \frac{1}{2} \end{cases}. \tag{4.9}$$

We now take r to be the maximum of these two values, as this ensures that our first criterion outlined earlier is satisfied, namely that the preparations $P_{+'}$ and $P_{-'}$ can be expressed as convex combinations of $\{P_{00}, P_{11}\}$ and $\{P_{01}, P_{10}\}$, respectively. In practice, one of either $P_{+'}$ or $P_{-'}$ will equate to the original noisy P_{ij}, with the other corresponding to a strict convex combination (see Fig. 4.5). The specifics will depend on the weights in Eq. (2.5).

Next, we show the quantities just defined allow us lower and upper bound the distance $d(\mu_p, \mu_q)$, and subsequently $C_{\text{prep}}^{\text{min}}$, in terms of $d(\mu_+, \mu_-) - d(P_+, P_-)$, thus obtaining the wanted result. We have

$$\begin{aligned} d(\mu_p, \mu_q) &= d((1-r)\mu_+ + r\mu_{+'}, (1-r)\mu_- + r\mu_{-'}) \\ &\leq (1-r)d(\mu_+, \mu_-) + rd(\mu_{+'}, \mu_{-'}) \\ &\leq (1-r)d(\mu_+, \mu_-) + r, \end{aligned}$$

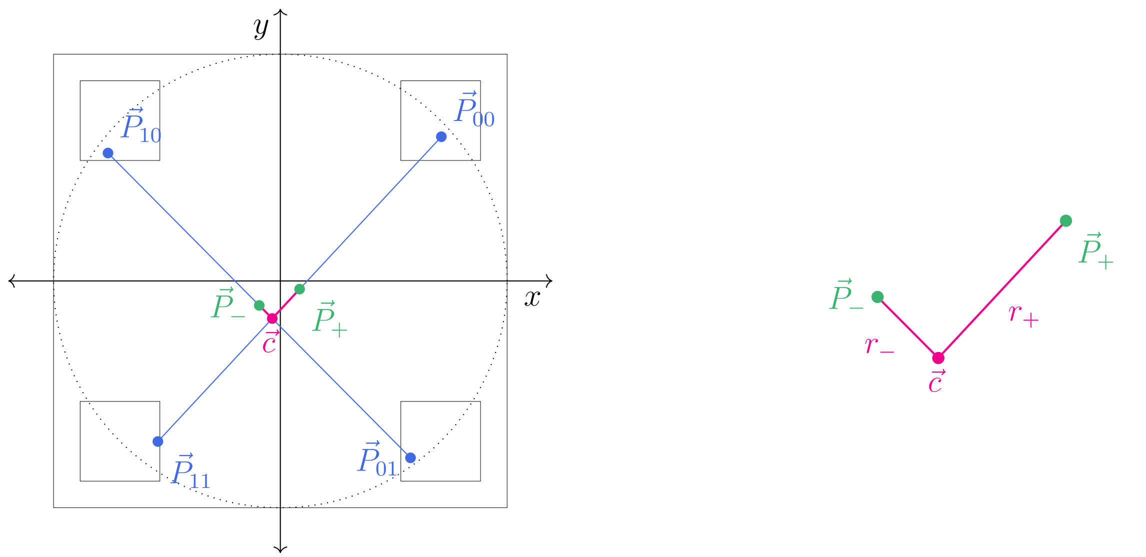
where the first inequality follows from the triangle inequality and the second inequality follows from $d(\mu_{+'}, \mu_{-'}) \leq 1$. It therefore follows that

$$\begin{aligned} \frac{1}{1-r} d(\mu_p, \mu_q) &\leq d(\mu_+, \mu_-) + \frac{r}{1-r} \\ \underbrace{\frac{1}{1-r}}_{\alpha_1} d(\mu_p, \mu_q) - \underbrace{\left[\frac{r}{(1-r)} + d(P_+, P_-)\right]}_{\alpha_2} &\leq d(\mu_+, \mu_-) - d(P_+, P_-). \end{aligned} \tag{4.10}$$

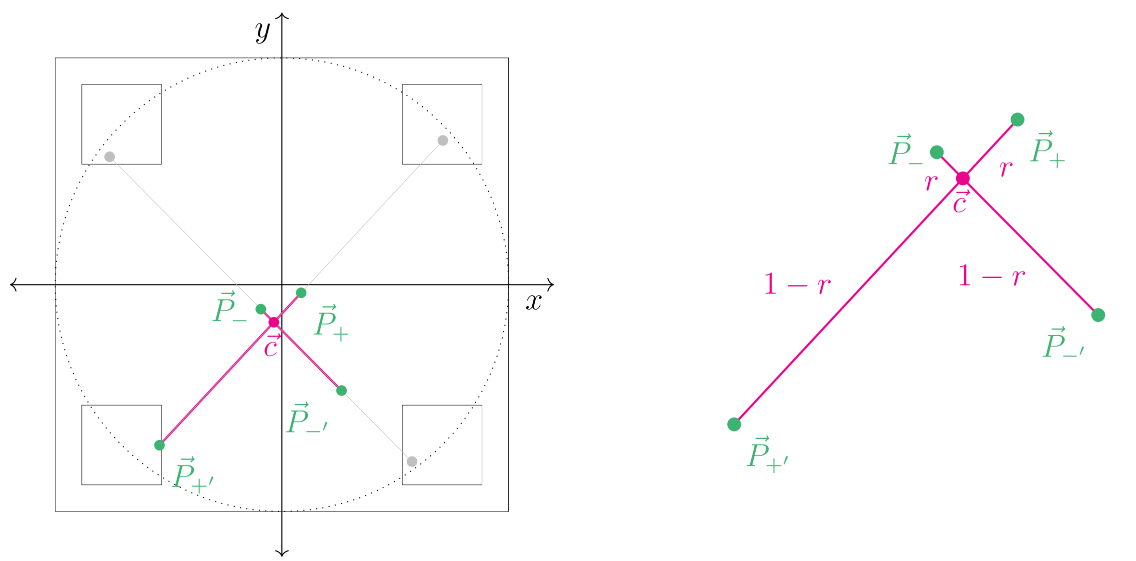
Similarly, we have

$$\begin{aligned} d(\mu_p, \mu_q) &= d((1-r)\mu_+ + r\mu_{+'}, (1-r)\mu_- + r\mu_{-'}) \\ &\geq (1-r)d(\mu_+, \mu_-) - rd(\mu_{+'}, \mu_{-'}) \\ &\geq (1-r)d(\mu_+, \mu_-) - r, \end{aligned}$$

which gives us



(a)



(b)

Figure 4.5: **An optimal common ratio r.** (a) In this example $p < \frac{1}{2}$ and $q > \frac{1}{2}$, implying that $\vec{c}$ is a combination of $\{\vec{P}_+, \vec{P}_{11}\}$ and $\{\vec{P}_-, \vec{P}_{01}\}$, respectively, as shown on the left side. Note that $r_+ > r_-$, as shown on the right side, so that r_+ will be used as the ratio r. (b) The intersection $\vec{c}$ has been recast (left side) in terms of two new convex combinations $\{\vec{P}_-, \vec{P}_{-'}\}$ and $\{\vec{P}_+, \vec{P}_{+'}\}$ that use the same weight r (right side). Since these combinations are themselves mixtures of the a priori preparations in Eq. (2.5), they can be used to express $\vec{P}_p$ and $\vec{P}_q$. In this example, $\vec{P}_{+'} = \vec{P}_{11}$, while $\vec{P}_{-'}$ is now a strictly convex combination of $\vec{P}_-$ and $\vec{P}_{01}$. (Choosing r to have been the minimum rather than maximum would have retained $\vec{P}_{-'} = \vec{P}_{01}$; however, $\vec{P}_{+'}$ would have then been outside of the convex hull of preparations and thus outside the operational theory.)

$$\frac{1}{1-r}d(\mu_p,\mu_q) \geq d(\mu_+,\mu_-) - \frac{r}{1-r}$$

$$\underbrace{\frac{1}{1-r}}_{\alpha_1} d(\mu_p,\mu_q) + \underbrace{\left[\frac{r}{(1-r)} - d(P_+,P_-)\right]}_{\alpha_3} \geq d(\mu_+,\mu_-) - d(P_+,P_-). \tag{4.11}$$

In the expressions above – equations (4.10) and (4.11) – we have identified the functions $\alpha_1, \alpha_2, \alpha_3$. Taking the infimum of equations (4.10) and (4.11) over all ontological models, and noting that $C^{\min}_{\text{prep}}$ is the infimum of $d(\mu_p, \mu_q)$ across all models, the desired result is proven. $\square$

By rearranging the terms, these relationships lead to the following corollary.

Corollary 3. *Given the simplest scenario with even- and odd-parity preparations P_+, P_- and inaccessible information $C^{\min}_{\text{prep}}$, there exist functions α_1, α_2, and α_3 of the preparations $\{P_{ij}\}$ such that*

$$C^{\min}_{\text{prep}} > \frac{\alpha_2}{\alpha_1} \implies \mathcal{D}^{\min}_{P_+,P_-} > 0, \tag{4.12a}$$

$$\mathcal{D}^{\min}_{P_+,P_-} > \alpha_3 \implies C^{\min}_{\text{prep}} > 0. \tag{4.12b}$$

In other words, we have established that a sufficient criterion for violating parity preservation is $C^{\min}_{\text{prep}} > \frac{\alpha_2}{\alpha_1}$ and, in turn, a sufficient criterion for violating Marvian's preparation noncontextuality equality is $\mathcal{D}^{\min}_{P_+,P_-} > \alpha_3$. The parameters α_1, α_2, and α_3 are evaluated solely from the experimental data and are obtained without making reference to δ. Given any four preparations $\{P_{ij}\}$, one can directly calculate the values of α_1, α_2, and α_3 and refer to Corollary 3 to see if there is a violation of one notion of classicality given enough violation of the other. Corollary 3 can be rephrased in terms of the noise parameter δ via the following theorem, as proven below.

Theorem 4. *Given the functions α_1, α_2, and α_3, if each noisy preparation P_{ij} satisfies $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq \delta$, the following upper bounds hold:*

$$\frac{\alpha_2}{\alpha_1} \leq \frac{2(1+2\sqrt{3})\delta - 4\sqrt{2}\delta^2}{1-2\sqrt{2}\delta}, \tag{4.13a}$$

$$\alpha_3 \leq \frac{4\sqrt{3}\delta}{1-2\sqrt{2}\delta - 4\sqrt{3}\delta}. \tag{4.13b}$$

Proof. From Eq. (4.10), we have $\alpha_1 = \frac{1}{1-r}$ and $\alpha_2 = \frac{r}{1-r} + d(P_+, P_-)$, so that

$$\frac{\alpha_2}{\alpha_1} = d(P_+, P_-)(1-r) + r. \tag{4.14}$$

Note that $\vec{P}_+$ and $\vec{P}_-$ are averages of points contained in δ-neighborhoods of the ideal points. Therefore, $\vec{P}_+$ and $\vec{P}_-$ are each contained in a δ-neighborhood of the origin, which implies that $d(P_+, P_-) \leq 2\delta$. Next, we observe that $\vec{c}$ is the intersection point of two line segments contained in diagonal strips of radius $2\sqrt{2}\delta$ centered in the origin. Thus $\vec{c}$ is contained within a tilted square of side length $4\sqrt{2}\delta$ centered in the origin. It follows that the maximal Euclidean distance between $\vec{P}_+$ or $\vec{P}_-$ and $\vec{c}$ is $4\sqrt{3}\delta$ (see Fig. 4.6a). This yields

$$||\vec{P}_+ - \vec{c}||, ||\vec{P}_- - \vec{c}|| \leq 4\sqrt{3}\delta. \tag{4.15}$$

Moreover, since $||\vec{P}_{00} - \vec{P}_+|| = \frac{1}{2}||\vec{P}_{00} - \vec{P}_{11}||$ (and similarly for the case with $\vec{P}_-$), the minimum value of the denominators in Eq. (4.9) occurs when $||\vec{P}_{00} - \vec{P}_{11}||$ is minimized (see Fig. 4.6b), that corresponds to the value $2 - 4\sqrt{2}\delta$. This leads to

$$||\vec{P}_+ - \vec{P}_{00}||, ||\vec{P}_+ - \vec{P}_{11}||, ||\vec{P}_- - \vec{P}_{01}||, ||P_0 - P_{10}|| \geq 1 - 2\sqrt{2}\delta. \tag{4.16}$$

Combining Eqs. (4.15) and (4.16) with Eq. (4.9), it follows that

$$r = \max\{r_+, r_-\} \leq \frac{4\sqrt{3}\delta}{1-2\sqrt{2}\delta}.$$

Referring back to Eq. (4.14), this yields

$$
\begin{aligned}
\frac{\alpha_2}{\alpha_1} &= d(P_+, P_-)(1-r) + r \\
&\leq d(P_+, P_-) + r \\
&\leq 2\delta + \frac{4\sqrt{3}\,\delta}{1 - 2\sqrt{2}\,\delta} \\
&= \frac{2(1+2\sqrt{3})\delta - 4\sqrt{2}\,\delta^2}{1 - 2\sqrt{2}\,\delta}.
\end{aligned}
$$

This establishes the first bound. For the second bound, we refer to Eq. (4.11):

$$
\begin{aligned}
\alpha_3 &= \frac{r}{1-r} - d(P_+, P_-) \\
&\leq \frac{r}{1-r} \\
&\leq \frac{\frac{4\sqrt{3}\,\delta}{1-2\sqrt{2}\,\delta}}{1 - \frac{4\sqrt{3}\,\delta}{1-2\sqrt{2}\,\delta}} \\
&= \frac{4\sqrt{3}\,\delta}{1 - 2\sqrt{2}\,\delta - 4\sqrt{3}\,\delta}.
\end{aligned}
$$

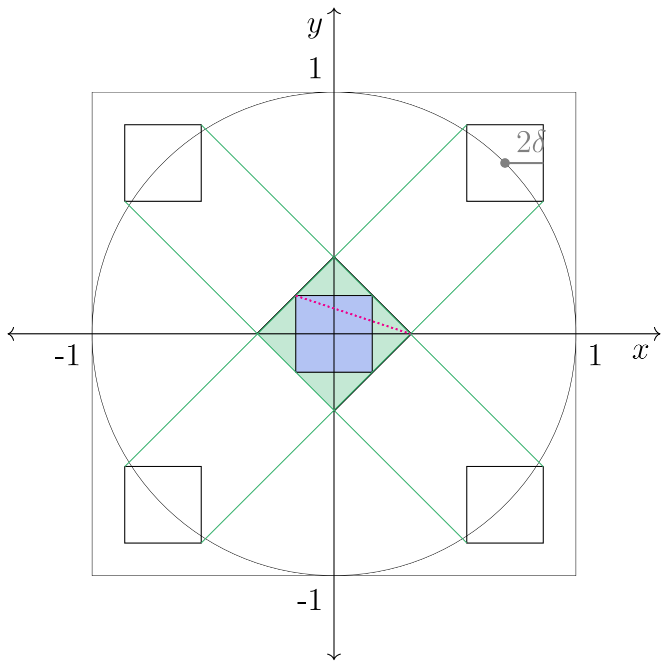
□

Combining the previous two results, we arrive at the following statement, which recasts Corollary 3 in terms of the noise bound δ.

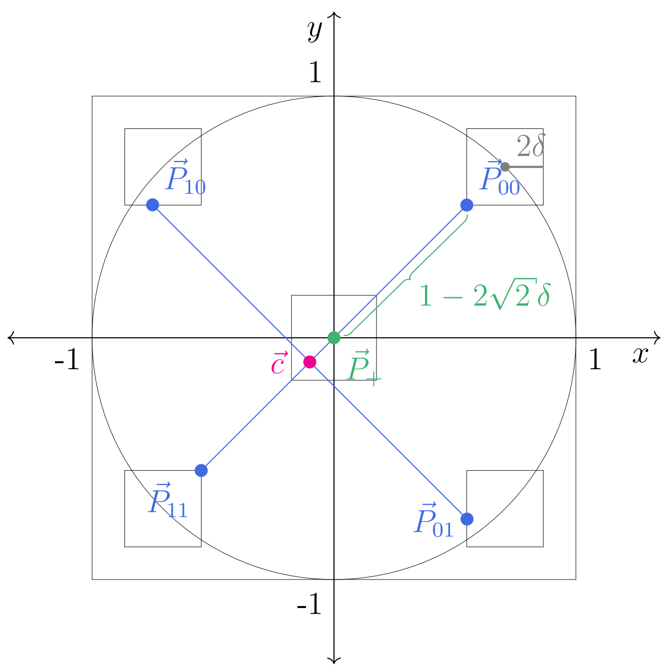
Corollary 4. *Given the simplest scenario with even- and odd-parity preparations P_+, P_- and inaccessible information $C^{\min}_{\text{prep}}$, if each P_{ij} satisfies $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq \delta$, then the following implications hold:*

$$
C^{\min}_{\text{prep}} > \frac{2(1+2\sqrt{3})\delta - 4\sqrt{2}\,\delta^2}{1 - 2\sqrt{2}\,\delta} \implies \mathcal{D}^{\min}_{P_+,P_-} > 0, \tag{4.17a}
$$

$$
\mathcal{D}^{\min}_{P_+,P_-} > \frac{4\sqrt{3}\,\delta}{1 - 2\sqrt{2}\,\delta - 4\sqrt{3}\,\delta} \implies C^{\min}_{\text{prep}} > 0. \tag{4.17b}
$$



(a)



(b)

Figure 4.6: **Bounding distances in terms of noise.** (a) The point $\vec{c}$ lies within the region shown in green. The points $\vec{P}_{+}, \vec{P}_{-}$ lie within the inscribed blue square. The red line illustrates the maximal distance between points lying in each region, which is of length $4\sqrt{3}\,\delta$. (b) An example of the minimum possible Euclidean distance between a preparation vector $\vec{P}_{ij}$ and midpoint vector $\vec{P}_{+}$ or $\vec{P}_{-}$, here shown with the case of $\vec{P}_{00}$ and $\vec{P}_{+}$, where $||\vec{P}_{+} - \vec{P}_{00}|| = 1 - 2\sqrt{2}\,\delta$.

Approach	Notion of nonclassicality	ideal preparations	Noise threshold of violation
Pusey's	Preparation contextuality	No	$\delta < 0.06$
Marvian's	Preparation contextuality	No	$\delta < 0.1$
Novel (this work)	Violation of BOD_P	Yes	$\delta < 0.007$

Table 4.1: **Three robust ways of witnessing nonclassicality in the simplest scenario.** Violating Pusey's inequality and Marvian's equality are ways of witnessing preparation contextuality. Our novel approach is based on the notion of parity preservation. A violation of parity preservation is an instance of a violation of BOD_P. Violating parity preservation under the indicated threshold means also a violation of preparation noncontextuality.

4.4 Threshold for nonclassicality

The results in the previous subsection ensure that for any given value of $C^{\min}_{\text{prep}}$, one can find a noise bound δ such that if noisy preparations P_{ij} lie within δ distance of the ideal preparations P^{id}_{ij}, then $\frac{\alpha_2}{\alpha_1}$ is reduced sufficiently to guarantee that indeed $C^{\min}_{\text{prep}} > \frac{\alpha_2}{\alpha_1}$ and therefore parity preservation is violated.

Equating the right hand sides of inequalities (4.2) and (4.13a), we find the threshold for which Marvian's inequality gives a sufficient lower bound to violate parity preservation as in Eq. (4.12a), which results to be $\delta \approx 0.007$ (see Fig. 4.7).

That is, if $\delta \leq 0.007$, then Eq. (4.13a) gives us $\frac{\alpha_2}{\alpha_1} < 0.063$ whereas Eq. (4.2) gives us $C^{\min}_{\text{prep}} > 0.069$, so that indeed $C^{\min}_{\text{prep}} > \frac{\alpha_2}{\alpha_1}$ and $\mathcal{D}^{\min}_{P_+,P_-}$ takes on positive values. We have now established the following corollary.

Corollary 5. *If $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq 0.007$, then $\mathcal{D}^{\min}_{P_+,P_-} > 0$ and parity preservation is violated.*

Given that Marvian's and Pusey's approaches both exhibit a violation if $\delta \leq 0.06$, we can therefore conclude sufficient conditions for equivalency in witnessing nonclassicality between all three approaches as follows.

Corollary 6. *If the noise parameter δ satisfies $\delta \leq 0.007$, then $S(x_{ij}, y_{ij}) > 0, C^{\min}_{\text{prep}} > 0$, and $\mathcal{D}^{\min}_{P_+,P_-} > 0$. Therefore, all three criteria of classicality are violated.*

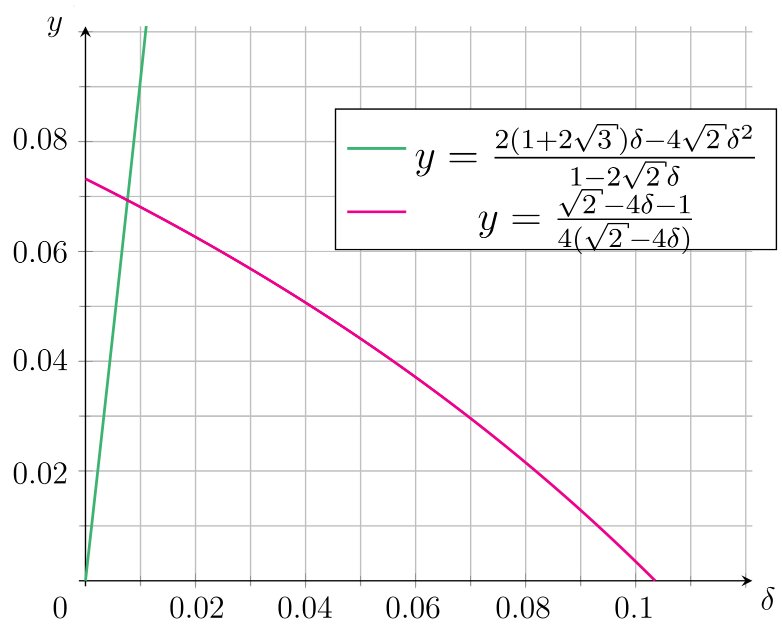


Figure 4.7: **Violation of parity preservation.** The green curve indicates the function of δ in Eq. (4.17a). For values of δ smaller than about 0.007 (see Theorem 5), the red curve upper bounding $C_{\text{prep}}^{\min}$ from Eq. (4.2) exceeds the green curve and implies, via Eq. (4.17a) in Corollary 4, that parity preservation is violated, *i.e.*, $\mathcal{D}_{P_+,P_-}^{\min} > 0$.

4.5 The case of quantum depolarizing noise

In this section we focus on the simplest nontrivial scenario in the case where preparations are contained within the unit Bloch disc and the experimental noise is assumed to be modeled by a quantum depolarizing channel. In this case, noisy preparations P_{ij} are mixtures of the ideal preparations P_{ij}^{id} with the completely mixed state $\frac{I}{2}$ (see Fig. 4.8). We begin with an updated bound on Pusey's expression and subsequent threshold for violating preparation noncontextuality based on this bound.

Theorem 5. *Suppose the preparations $\{P_{ij}\}$ of the simplest scenario satisfy a noise bound $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \delta$, where δ is the noise parameter and $\{P_{ij}^{id}\}$ are the ideal a priori preparations. Pusey's expression $S(x_{ij}, y_{ij})$ of Eq. (3.1) satisfies the following lower bound in terms of the noise parameter δ:*

$$S(x_{ij}, y_{ij}) \geq 2\sqrt{2} - 2 - 8\delta + \frac{8\sqrt{2}\delta^2 - 4\delta}{1 - \sqrt{2}\delta}, \tag{4.18}$$

where $\{x_{ij}, y_{ij}\}$ *are the coordinates of the preparations* $\{P_{ij}\}$*.*

The proof follows that of Theorem 1, with the following changes in values: $u_\delta = \frac{1}{\sqrt{2}}$ and $p, q \geq \frac{1-2\sqrt{2}\delta}{2-2\sqrt{2}\delta}$.

Given that the corresponding preparation noncontextuality inequality is $S(x_{ij}, y_{ij}) \leq 0$, solving for the right hand side in Eq. (4.18) results in a violation threshold of $\delta \approx 0.07$, which leads to the following theorem.

Corollary 7. *If* $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq 0.07$*, then* $S(x_{ij}, y_{ij}) > 0$ *and Pusey's preparation noncontextuality inequality is violated.*

In the case of quantum depolarizing noise, Theorem 2 and Corollary 2 are left unchanged. This is not the case for Theorem 4, that provides a different bound than the one in Eq. (4.13a). We provide the modified statement below.

Lemma 1. *Given the functions* α_1 *and* α_2*, if each noisy preparation* P_{ij} *satisfies* $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \delta$*, the following upper bound holds:*

$$\frac{\alpha_2}{\alpha_1} \leq \delta + \frac{\sqrt{2}\delta}{1 - 2\sqrt{2}\delta}. \tag{4.19}$$

The proof follows that of Theorem 4 above, with the following changes in values:

- $d(P_+, P_-) \leq \delta$
- $\vec{c} = \vec{0}$
- $||\vec{P}_+ - \vec{c}||, ||\vec{P}_- - \vec{c}|| \leq \sqrt{2}\delta$
- $||\vec{P}_+ - \vec{P}_{00}||, ||\vec{P}_+ - \vec{P}_{11}||, ||\vec{P}_- - \vec{P}_{01}||, ||P_0 - P_{10}|| \geq 1 - 2\sqrt{2}\delta$
- $r \leq \frac{\sqrt{2}\delta}{1-2\sqrt{2}\delta}$

Combining Eq. (4.19) with Eq. (4.12a) leads to the following theorem.

Theorem 6. *Given the simplest scenario with even- and odd-parity preparations* P_+, P_- *and inaccessible information* $C^{\min}_{\text{prep}}$*, if each* P_{ij} *satisfies* $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq \delta$*, then the following implication holds.*

$$C^{\min}_{\text{prep}} > \delta + \frac{\sqrt{2}\,\delta}{1 - 2\sqrt{2}\,\delta} \implies \mathcal{D}^{\min}_{P_+,P_-} > 0. \tag{4.20}$$

Equating the right hand sides of inequalities (4.2) and (4.19), we find the threshold for which Marvian's inequality provides a sufficient lower bound to violate parity preservation as in Eq. (4.20), which results in $\delta \approx 0.02$. Consequently, the following theorem holds.

Corollary 8. *If* $d(\vec{P}_{ij}, \vec{P}^{id}_{ij}) \leq 0.02$*, then* $\mathcal{D}^{\min}_{P_+,P_-} > 0$ *and parity preservation is violated.*

Given that both Marvian's and Pusey's approaches exhibit a violation if $\delta \leq 0.07$, we obtain the following noise threshold below which all methods agree in witnessing nonclassicality in the simplest scenario in the case of quantum depolarizing noise.

Corollary 9. *If the noise parameter* δ *satisfies* $\delta \leq 0.02$*, then* $S(x_{ij}, y_{ij}) > 0, C^{\min}_{\text{prep}} > 0$*, and* $\mathcal{D}^{\min}_{P_+,P_-} > 0$*. Therefore, all three criteria of classicality are violated.*

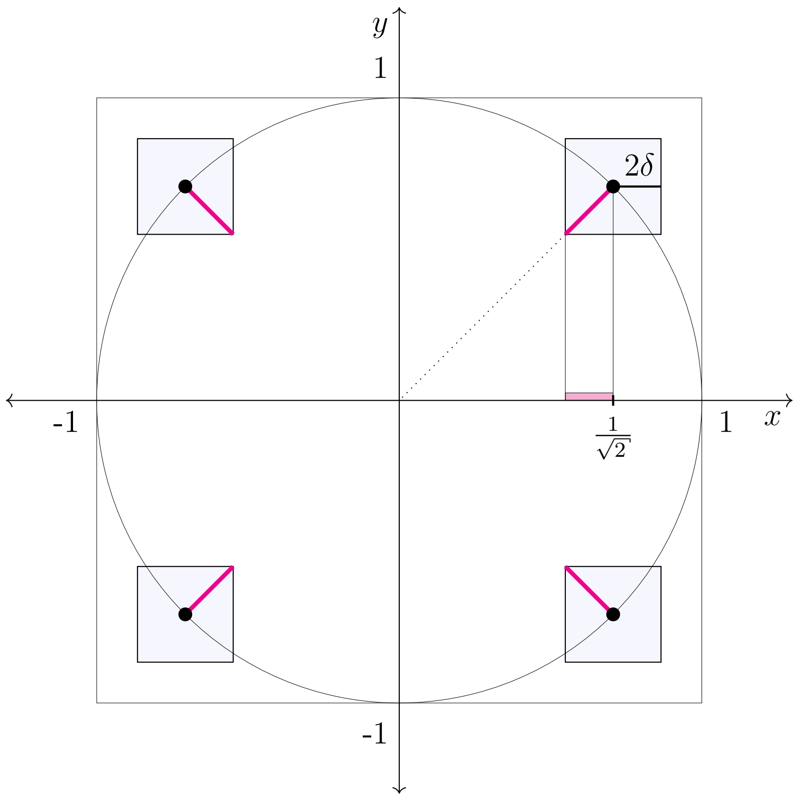


Figure 4.8: **Depolarizing noise constrained to a bound of δ.** Ideal preparations are indicated at the center of the shaded squares. Noisy preparations are assumed to lie on the red segments. In accordance with Eq. (2.6), they have an operational distance of *at most* δ from the ideal ones.

Chapter 5

Parity-oblivious multiplexing

In this chapter we treat the $m-$bit parity-oblivious multiplexing protocol in both the noiseless and noisy scenarios. In particular, we focus on the $m = 2$ case. Indeed, in this instance, the protocol's setting involving four preparations and two measurements corresponds to the simplest scenario of Fig. 2.1.

5.1 Noiseless case

The $m-$bit parity-oblivious multiplexing protocol was first introduced in 2009 by Spekkens *et al* [26]. Since then, it has motivated further investigation and substantial research on relating it to other scenarios [8, 11, 2], and on developing protocols with preparation contextuality as a resource for the computational advantage [12, 22, 21, 16, 28].

We here consider the $m = 2$ case. Let us imagine that Alice prepares a two-bit string, that we denote with x. Bob wishes to learn the value of a single bit among the two (without Alice knowing which one) with a probability at least p. Alice and Bob can try to achieve this by agreeing on a strategy that consists of Alice sending some information carriers and Bob performing certain measurements. However, the task contains an additional constraint, called *parity-obliviousness*: Alice cannot communicate the parity of the two-bit string x to Bob. Let us denote the bit that Bob outputs as b.

The integer y denotes which of the two bits b should correspond to, and x_y denotes the actual bit in Alice's string.

The probability of success of the game takes, in general, the following form in terms of the probabilities $P(b|P_x, M_y)$, where P_x denotes the preparations of Alice, given the bit string x she has to communicate, and M_y denotes the measurements of Bob, given the bit y to be guessed,

$$\begin{aligned} p(b = x_y) = \frac{1}{8}[&\mathcal{P}(0|P_{00}, M_0) + \mathcal{P}(0|P_{01}, M_0) + \mathcal{P}(1|P_{10}, M_0) + \mathcal{P}(1|P_{11}, M_0) \\ &+ \mathcal{P}(0|P_{00}, M_1) + \mathcal{P}(0|P_{10}, M_1) + \mathcal{P}(1|P_{01}, M_1) + \mathcal{P}(1|P_{11}, M_1)]. \end{aligned}$$

The optimal classical probability of success satisfies $p(b = x_y) \leq \frac{3}{4}$, as the only classical encoding that transfers some information to Bob without violating parity-obliviousness consists of encoding only a single bit x_i. Given that y is chosen at random, any bit x_i would perform the same. Therefore, Alice and Bob can agree on Alice always sending x_1 and Bob outputting $b = x_1$. The probability of success is given by the probability that $y = 1$, which is $\frac{1}{2}$, and the probability that Bob outputs correctly (at random, with probability $\frac{1}{2}$) in the other case where $y \neq 1$, that occurs with probability $\frac{1}{2}$. For this optimal classical strategy we obtain $p(b = x_y) = \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$, as stated. This value is the same as the Bell bound of the CHSH game and two-bit quantum random access codes [10, 1] . In [26], Spekkens *et al* proved the following theorem (here stated only for the case $m = 2$).

Theorem 7. *The optimal success probability in two-bit parity-oblivious multiplexing of any operational theory that admits of a preparation non-contextual ontological model satisfies* $p(b = x_y) \leq \frac{3}{4}$.

This theorem indicates that preparation contextuality is a necessary resource for performing the two-bit parity-oblivious multiplexing protocol with higher success probability than the one achievable by optimal classical strategies. Moreover, it turns out that, by using the same optimal quantum strategy of two-bit quantum random access codes [1] (associated with the preparations P^{id}_{ij} and measurements X, Y of Fig. 2.1a), the probability of success is $\omega_Q(\text{POM}) = \cos^2(\frac{\pi}{8}) \approx 0.85$. It can be shown that this value is the maximum achievable with quantum strategies [26]. Notice that, when Bob measures on the X or Y basis, he cannot gain any information about the parity, as the parity 0 and parity 1 mixtures – $\frac{1}{2}P^{id}_{00} + \frac{1}{2}P^{id}_{11}$ and $\frac{1}{2}P^{id}_{01} + \frac{1}{2}P^{id}_{10}$, respectively – correspond to the same quantum state (the completely mixed state).

5.2 Noisy case

The two-bit parity-oblivious multiplexing protocol in realistic scenarios necessarily involves noisy preparations, that allow some parity to possibly be communicated to Bob. If one wants to generalize Theorem 7 of the noiseless case and show that preparation contextuality still powers the protocol in the noisy scenario, one faces a couple of challenges. First, one must assume that the noise could potentially be used to communicate parity, thus allowing classical strategies to achieve a probability of success greater than the ones of the noiseless case. Then, once one finds the value of the optimal classical probability of success in the noisy case, one must show that performing better than this implies a proof of preparation contextuality.

We denote the two-bit parity-oblivious multiplexing protocol in the presence of noise with $\varepsilon-$POM, where ε denotes the noise in terms of the maximum probability of parity communicated, *i.e.*, the operational distance $d(P_+, P_-)$ between the even- and odd-parity mixtures. We first notice that, in order to witness the possible nonclassicality associated with a certain probability of success, Pusey's approach is not ideal. It considers a posteriori operational equivalences, thus not explicitly referring to the parity that can be communicated as a consequence of the noise and which may be part of the reason for the given probability of success. We now show how our approach based on the notion of parity preservation, that explicitly refers to the violation of parity-obliviousness, is more suitable for the task.

In order to find an optimal classical strategy in $\varepsilon-$POM, we can mirror the optimal strategy described in the noiseless case of the previous subsection, this time taking into account the ε parity that can be communicated. Namely, we take into account that Bob knows the parity of x, that is, $x_1 + x_2$ (the addition is modulo 2), with probability ε. Without loss of generality, we can assume that Alice and Bob always agree on Alice sending the first bit, x_1. If $y = 1$, then Bob outputs $b = x_1$ and wins with probability 1. If $y = 2$, then with probability ε Bob knows the parity and therefore the value of x_2 (*i.e.*, he outputs $b = x_1 + (x_1 + x_2) = x_2$ and wins with probability 1), and with probability $(1-\varepsilon)$ he does not know the parity and can at best make a random guess about the value of x_2 and so win with probability $\frac{1}{2}$. In summary, this strategy results in the probability of success $p(b = x_y) = \frac{1}{2}\cdot 1+\frac{1}{2}\left[\varepsilon \cdot 1 + (1-\varepsilon)\cdot \frac{1}{2}\right] = \frac{3}{4}+\frac{\varepsilon}{4}$. Notice that there is no classical strategy that can perform better than the one just described; we indeed assumed that Bob uses the knowledge of parity to his maximum possible advantage and that Alice communicates one bit of information, which is the best she can do. This leads to the following lemma.

Theorem 8. *The optimal probability of success in 2-bit ε-parity-oblivious multiplexing using a classical strategy satisfies* $p(b = x_y) \leq \frac{3}{4} + \frac{\varepsilon}{4}$.

Here we prove the following theorem, which is an application of Proposition 2 and the results of Section 4.2 of [9].

Theorem 9. *The optimal success probability in two-bit ε-parity-oblivious multiplexing of any operational theory that admits of a parity preserving ontological model satisfies* $p(b = x_y) \leq \frac{3}{4} + \frac{\varepsilon}{4}$.

Proof. Suppose $\{P_+, P_-\}$ can be distinguished in a single-shot measurement with probability ε. That is,

$$\max_{M=X,Y}\{|\mathcal{P}(k|P_+,M)-\mathcal{P}(k|P_-,M)|\}=\varepsilon. \tag{5.1}$$

This expression can be recast as $\max\{\frac{1}{2}|x_+ - x_-|, \frac{1}{2}|y_+ - y_-|\} = \varepsilon$. Therefore, by definition we have $d(P_+,P_-)=\varepsilon$. Considering Eq. (3.6), this means that $s_{\mathcal{O}}^{P_+,P_-} = \frac{1+\varepsilon}{2}$. Our assumption of parity preservation entails that $d(\mu_+,\mu_-)=\varepsilon$. Thus, Eq. (3.7) now reads as $s_{\Lambda}^{\mu_+,\mu_-} = \frac{1+\varepsilon}{2}$, which also means that $s_{\mathcal{O}}^{P_+,P_-} = s_{\Lambda}^{\mu_+,\mu_-}$. In other words, we have

$$s_{\Lambda}^{\mu_{00}+\mu_{11},\mu_{01}+\mu_{10}} = s_{\mathcal{O}}^{P_{00}+P_{11},P_{01}+P_{10}} = \frac{1+\varepsilon}{2}.$$

It follows from the results of Proposition 2 in [9] that

$$\frac{1}{2}\left(s_{\mathcal{O}}^{P_{00}+P_{01},P_{10}+P_{11}}\right)+\frac{1}{2}\left(s_{\mathcal{O}}^{P_{00}+P_{10},P_{01}+P_{11}}\right) \leq \frac{1+\frac{1+\varepsilon}{2}}{2} = \frac{3}{4}+\frac{\varepsilon}{4}. \tag{5.2}$$

The probability of success is the averaged sum of all possible ways Bob can win, given a randomly chosen measurement X, Y and state $P_{00}, P_{01}, P_{10}, P_{11}$ that Alice has, and an output $0, 1$ that Bob guesses based on the outcome of his measurement. This amounts to

$$\begin{aligned} p(b=x_y) &= \frac{1}{8}[\mathcal{P}(0|P_{00},X)+\mathcal{P}(0|P_{01},X)+\mathcal{P}(1|P_{10},X)+\mathcal{P}(1|P_{11},X) \\ &\quad +\mathcal{P}(0|P_{00},Y)+\mathcal{P}(0|P_{10},Y)+\mathcal{P}(1|P_{01},Y)+\mathcal{P}(1|P_{11},Y)] \\ &= \frac{1}{2}\left(\frac{\mathcal{P}(0|\frac{1}{2}P_{00}+\frac{1}{2}P_{01},X)}{2}+\frac{\mathcal{P}(1|\frac{1}{2}P_{10}+\frac{1}{2}P_{11},X)}{2}\right) \\ &\quad +\frac{1}{2}\left(\frac{\mathcal{P}(0|\frac{1}{2}P_{00}+\frac{1}{2}P_{10},Y)}{2}+\frac{\mathcal{P}(1|\frac{1}{2}P_{01}+\frac{1}{2}P_{11},Y)}{2}\right) \\ &= \frac{1}{2}\left(s_{\mathcal{O}}^{P_{00}+P_{01},P_{10}+P_{11}}\right)+\frac{1}{2}\left(s_{\mathcal{O}}^{P_{00}+P_{10},P_{01}+P_{11}}\right). \end{aligned} \tag{5.3}$$

By combining equations (5.2) and (5.3) we obtain $p(b=x_y) \leq \frac{3}{4}+\frac{\varepsilon}{4}$. □

In other words, if one obtains a probability of success greater than the classical probability of success, then parity preservation is violated. Moreover, we have shown in Theorem 6 that under a certain threshold, this implies that Marvian's equality (as well as Pusey's inequality) is also violated. Therefore, we have shown that as long as the noise is below a certain threshold, preparation contextuality is still present whenever parity-oblivious multiplexing manifests computational advantage over classical strategies.

5.3 A new threshold for nonclassicality

We can calculate what the probability of success in the game is, in the presence of noise, when still using the same quantum protocol that is used in the noiseless case. We stress that this strategy does not use noise to one's advantage, and therefore does not give rise to the optimal probability of success. The protocol used in the noiseless quantum case is the averaged sum of all possible ways Bob can win, given a randomly chosen measurement X, Y, a (noisy) quantum state $P_{00}, P_{01}, P_{10}, P_{11}$ that Alice prepares and sends Bob, and an output $0, 1$ that Bob guesses based on the outcome of his measurement [26]. This would amount to

$$\begin{aligned} p(b = x_y) = \frac{1}{8}[&\mathcal{P}(0|P_{00}, X) + \mathcal{P}(0|P_{01}, X) + \mathcal{P}(1|P_{10}, X) + \mathcal{P}(1|P_{11}, X) \\ &+ \mathcal{P}(0|P_{00}, Y) + \mathcal{P}(0|P_{10}, Y) + \mathcal{P}(1|P_{01}, Y) + \mathcal{P}(1|P_{11}, Y)]. \end{aligned}$$

Now, we suppose that Alice and Bob are playing ε-parity-oblivious multiplexing. That is, Alice's noisy preparations satisfy a parity constraint of ε. We therefore have $d(P_+, P_-) \leq \varepsilon$. This is equivalent to the condition $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \frac{\varepsilon}{2}$. Given the coordinates of $\vec{P}_{ij}^{id}$, it follows that $\frac{1}{\sqrt{2}} - \varepsilon \leq |x_{ij}|, |y_{ij}| \leq \frac{1}{\sqrt{2}} + \varepsilon$. This can be seen Fig. 4.1 where any noisy point $\vec{P}_{ij}$ within an operational distance of δ from the ideal points $\vec{P}_{ij}^{id}$ has coordinates whose absolute values can lie within a range spanning twice that parameter: $\left[\frac{1}{\sqrt{2}} - 2\delta, \frac{1}{\sqrt{2}} + 2\delta\right]$. Applying this to the summands in Eq. (5.3), we have

$p(b = x_y)$
$= \frac{1}{8}\left(\frac{x_{00}+1}{2}+\frac{x_{01}+1}{2}+\frac{1-x_{10}}{2}+\frac{1-x_{11}}{2}+\frac{y_{00}+1}{2}+\frac{y_{10}+1}{2}+\frac{1-y_{01}}{2}+\frac{1-y_{11}}{2}\right)$
$= \frac{1}{16}\left(x_{00}+x_{01}-x_{10}-x_{11}+y_{00}+y_{10}-y_{01}-y_{11}\right)+\frac{1}{2}$
$\geq \frac{1}{16}\left[8\cdot\left(\frac{1}{\sqrt{2}}-\varepsilon\right)\right]+\frac{1}{2}$
$= \frac{2+\sqrt{2}}{4}-\frac{\varepsilon}{2}.$

We have just proven the following

Theorem 10. *If $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq \frac{\varepsilon}{2}$, then $p(b = x_y) > \frac{2+\sqrt{2}}{4} - \frac{\varepsilon}{2}$.*

We see that if the noise is reduced to zero, then this theorem provides the standard optimal quantum probability of success $\frac{2+\sqrt{2}}{4} \approx 0.85$ as has been established in the noiseless case [26]. We recast Theorem 9 via contraposition and recall we have the following fact:

Theorem 11. *If $p(b = x_y) > \frac{3}{4} + \frac{\varepsilon}{4}$, then an operational theory does not admit of a parity preservation ontological model.*

Giving that solving $\frac{3}{4} + \frac{\varepsilon}{4} = \frac{2+\sqrt{2}}{4} - \frac{\varepsilon}{2}$ provides a solution of $\varepsilon = \frac{\sqrt{2}-1}{3} \approx .138$, we have established the below result as a consequence of the previous two theorems.

Theorem 12. *If $d(\vec{P}_{ij}, \vec{P}_{ij}^{id}) \leq 0.06$, then parity preservation is violated.*

Given that Marvian's and Pusey's approaches both exhibit a violation if $\delta \leq 0.06$, we can therefore conclude sufficient conditions for equivalency in witnessing nonclassicality between all three approaches as follows.

Corollary 10. *If the noise parameter δ satisfies $\delta \leq 0.06$, then all three criteria of classicality are violated.*

This result is striking in that it gives a possibility of investigating whether there might be an equivalency between these different notions of nonclassicality, and, in particular, between Pusey's approach and parity preservation (given that they exhibit very similar thresholds of violation). That is, one may seek to establish whether or not it is the case that an operational theory will witness a violation via Pusey's approach if and only if it will violate parity preservation. If this is the case, one would need to use Pusey's expression directly (Eq. (3.1)) rather than the inequality presented in terms of noise (Eq. (4.1)), as well as derive an optimal quantum probability of success in the noisy case (which has yet to be found), to check whether parity preservation is violated by way of the noisy coordinates giving rise to a success probability greater than the classical bound.

Chapter 6

Conclusion

To establish the credibility of a notion of nonclassicality, it is imperative that it be experimental testable. With this in mind, we explored three approaches to assess nonclassicality within the simplest nontrivial scenario featuring four noisy preparations and two tomographically complete measurements. Specifically, we investigated Pusey's and Marvian's approaches for detecting preparation contextuality, along with a novel approach for witnessing a violation of BOD_P.

We showed that these three approaches align in detecting nonclassicality as long as the level of experimental noise remains below a certain threshold, $\delta < 0.007$ (in the case of quantum depolarizing noise, this improves to $\delta < 0.02$, and with the results presented in Section 5.3 we have a significant improvement to $\delta < 0.06$ that applies to any type of experimental noise). Therefore, experimenters have the flexibility to choose the approach that best suits their needs when testing for nonclassicality in their experiments, provided the noise remains within this range. This flexibility becomes particularly relevant in scenarios where certain approaches are not suitable, such as in the noisy parity-oblivious multiplexing protocol. In the latter case, we argued that the appropriate notion to test is parity preservation, which refers to the a priori ideal preparations and explicitly allows one to quantify the violation of the parity constraint. Nevertheless, by virtue of our results, below the noise threshold, Marvian's and Pusey's approaches can also be employed to detect nonclassicality in the experiment. Indeed, below this threshold, we also established that preparation contextuality is still present when one performs the protocol with a success probability greater than what can be achieved with classical strategies.

Crucial to obtaining our results is in the way we characterized noise through the noise parameter δ. The latter quantifies – via the operational distance – the deviation in the measurement statistics between the experimentally realized and the ideal target preparations. Our choice of operational distance corresponds to the maximum difference over the x and y coordinates of the statistics between the preparations. There are two reasons for employing it. First, it is geometrically intuitive, as witnessed by the fact that preparations with δ noise distance from the ideal ones belong to a square of radius 2δ around them (see Fig. 4.1). Second, it makes calculations tractable. We give a couple of examples.

1) Distances over the coordinates are easily identified within Pusey' expression (Eq. 3.1), making it possible to bound it.

2) The operational distance between P_+ and P_- (Eq. (5.1)) coincides with the parity that can be communicated in $\varepsilon-$POM, thus allowing for a straightforward proof of Theorem 9.

With alternative definitions of operational distance like the maximum relative entropy or the operational total variational distance [17] we would have not exploited the above lucrative features. We leave for future research the question of how the results change by using these other ways of characterizing noise.

We emphasize that determining a mathematical threshold below which both preparation noncontextuality and parity preservation are violated (Theorems 5 and 6) was neither straightforward nor necessarily anticipated. While we expected the existence of a noise threshold below which both parity preservation (and consequently BOD_P)

and preparation noncontextuality are violated, it was not clear how and if this could be found. In particular, it was not immediately obvious how to *quantitatively relate* preparation contextuality, that deals with operational equivalences and corresponding ontological inequivalences, with violations of BOD_P (and parity preservation), which deal with operational distances and corresponding greater ontological distances. The key to our achievement in obtaining such a threshold hinges on Theorem 3. The latter ultimately leads to a function of the noise parameter δ in Corollary 4 (Eq. (4.17a)) which intersects Marvian's witness of contextuality $C_{\text{prep}}^{\text{min}}$ (as shown in Fig. 4.7), revealing a region where not only preparation noncontextuality but also parity preservation is violated. Our success in finding this function lies in the way we recast the operational equivalence in terms of even- and odd-parity mixtures. We managed not only to allow for a connection between preparation noncontextuality and parity preservation but also to retain the amount of parity preservation violation, as shown in the proof of Theorem 3.

Our method based on parity preservation can be seen as an application of the results contained in [9] in the context of the simplest scenario. In connection with this previous work, we provided a reformulation using the concept of distances instead of distinguishabilities, which offered a clear alternative interpretation of both the operational and ontological differences. In addition, we re-obtained the proof presented in [9] about the $\varepsilon-$POM being powered by a violation of BOD_P, stressing that the violation is in terms of parity preservation. Consequently, we noticed, by virtue of Theorem 6, that $\varepsilon-$POM is also powered by preparation contextuality as long as the noise parameter δ remains below a threshold.

The results of this work are relevant for applications in information processing tasks that aim to witness nonclassicality and that are set in the simplest nontrivial scenario. We recall how these tasks – examples of which are the two-bit parity-oblivious multiplexing treated here and other versions of the two-bit quantum random access codes

– are of central importance because they are the primitive communication tasks where nonclassicality can be certified in a device or semi-device independent way. The question remains open as to whether the methods presented in this manuscript can be extended to scenarios beyond the simplest nontrivial case. We leave this interesting avenue for future research.

Bibliography

[1] Andris Ambainis et al. “Dense Quantum Coding and a Lower Bound for 1-Way Quantum Automata”. In: *Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing*. STOC ’99. Atlanta, Georgia, USA: Association for Computing Machinery, 1999, pp. 376–383. DOI: 10.1145/301250.301347. URL: https://doi.org/10.1145/301250.301347.

[2] Andris Ambainis et al. “Parity oblivious d-level random access codes and class of noncontextuality inequalities”. In: *Quantum Information Processing* 18.4 (2019), p. 111. DOI: 10.1007/s11128-019-2228-3. URL: https://doi.org/10.1007/s11128-019-2228-3.

[3] Jonathan Barrett. “Information processing in generalized probabilistic theories”. In: *Phys. Rev. A* 75 (3 Mar. 2007), p. 032304. DOI: 10.1103/PhysRevA.75.032304. URL: https://link.aps.org/doi/10.1103/PhysRevA.75.032304.

[4] Lorenzo Catani and Matthew Leifer. “A mathematical framework for operational fine tunings”. In: *Quantum* 7 (Mar. 2023), p. 948. ISSN: 2521-327X. DOI: 10.22331/q-2023-03-16-948. URL: https://doi.org/10.22331/q-2023-03-16-948.

[5] Lorenzo Catani et al. “Aspects of the phenomenology of interference that are genuinely nonclassical”. In: *Phys. Rev. A* 108 (2 Aug. 2023), p. 022207. DOI: 10.1103/PhysRevA.108.022207. URL: https://link.aps.org/doi/10.1103/PhysRevA.108.022207.

[6] Lorenzo Catani et al. “Connecting XOR and XOR* games”. In: *Phys. Rev. A* 109 (1 Jan. 2024), p. 012427. DOI: 10.1103/PhysRevA.109.012427. URL: https://link.aps.org/doi/10.1103/PhysRevA.109.012427.

[7] Lorenzo Catani et al. “What is Nonclassical about Uncertainty Relations?” In: *Phys. Rev. Lett.* 129 (24 Dec. 2022), p. 240401. DOI: 10.1103/PhysRevLett.129.240401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.129.240401.

[8] André Chailloux et al. “Optimal bounds for parity-oblivious random access codes”. In: *New Journal of Physics* 18.4 (Apr. 2016), p. 045003. DOI: 10.1088/1367-2630/18/4/045003. URL: https://doi.org/10.1088/1367-2630/18/4/045003.

[9] Anubhav Chaturvedi and Debashis Saha. “Quantum prescriptions are more ontologically distinct than they are operationally distinguishable”. In: *Quantum* 4 (Oct. 2020), p. 345. ISSN: 2521-327X. DOI: 10.22331/q-2020-10-21-345. URL: https://doi.org/10.22331/q-2020-10-21-345.

[10] John F. Clauser et al. “Proposed Experiment to Test Local Hidden-Variable Theories”. In: *Phys. Rev. Lett.* 23 (15 Oct. 1969), pp. 880–884.

[11] Shouvik Ghorai and A. K. Pan. “Optimal quantum preparation contextuality in an n-bit parity-oblivious multiplexing task”. In: *Phys. Rev. A* 98 (3 Sept. 2018), p. 032110. DOI: 10.1103/PhysRevA.98.032110. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.032110.

[12] Alley Hameedi et al. “Communication Games Reveal Preparation Contextuality”. In: *Phys. Rev. Lett.* 119 (22 Nov. 2017), p. 220402. DOI: 10.1103/PhysRevLett.119.220402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.119.220402.

[13] Nicholas Harrigan and Robert W. Spekkens. “Einstein, Incompleteness, and the Epistemic View of Quantum States”. In: *Foundations of Physics* 40.2 (Feb. 2010),

pp. 125–157. DOI: 10.1007/s10701-009-9347-0. URL: https://doi.org/10.1007/s10701-009-9347-0.

[14] Luciana Henaut et al. "Tsirelson's bound and Landauer's principle in a single-system game". In: *Phys. Rev. A* 98 (6 Dec. 2018), p. 060302. DOI: 10.1103/PhysRevA.98.060302. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.060302.

[15] Massy Khoshbin, Lorenzo Catani, and Matthew Leifer. "Alternative robust ways of witnessing nonclassicality in the simplest scenario". In: *Phys. Rev. A* 109 (Mar. 2024), p. 032212. DOI: 10.1103/PhysRevA.109.032212. URL: https://link.aps.org/doi/10.1103/PhysRevA.109.032212.

[16] Matteo Lostaglio and Gabriel Senno. "Contextual advantage for state-dependent cloning". In: *Quantum* 4 (Apr. 2020), p. 258. ISSN: 2521-327X. DOI: 10.22331/q-2020-04-27-258. URL: https://doi.org/10.22331/q-2020-04-27-258.

[17] Iman Marvian. "Inaccessible information in probabilistic models of quantum systems, non-contextuality inequalities and noise thresholds for contextuality". In: *arXiv:2003.05984v1 [quant-ph]* (2020). URL: https://arxiv.org/abs/2003.05984.

[18] Michael D. Mazurek et al. "An experimental test of noncontextuality without unphysical idealizations". In: *Nature Communications* 7.1 (2016), ncomms11780. DOI: 10.1038/ncomms11780. URL: https://doi.org/10.1038/ncomms11780.

[19] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. 10th Anniversary Edition. Cambridge University Press, 2000.

[20] Matthew F. Pusey. "Robust preparation noncontextuality inequalities in the simplest scenario". In: *Phys. Rev. A* 98 (2 Aug. 2018), p. 022112. DOI: 10.1103/PhysRevA.98.022112. URL: https://link.aps.org/doi/10.1103/PhysRevA.98.022112.

[21] Debashis Saha and Anubhav Chaturvedi. “Preparation contextuality as an essential feature underlying quantum communication advantage”. In: *Phys. Rev. A* 100 (2 Aug. 2019), p. 022108. DOI: 10.1103/PhysRevA.100.022108. URL: https://link.aps.org/doi/10.1103/PhysRevA.100.022108.

[22] David Schmid and Robert W. Spekkens. “Contextual Advantage for State Discrimination”. In: *Phys. Rev. X* 8 (1 Feb. 2018), p. 011015. DOI: 10.1103/PhysRevX.8.011015. URL: https://link.aps.org/doi/10.1103/PhysRevX.8.011015.

[23] Anthony J Short and Jonathan Barrett. “Strong nonlocality: a trade-off between states and measurements”. In: *New Journal of Physics* 12.3 (Mar. 2010), p. 033034. DOI: 10.1088/1367-2630/12/3/033034. URL: https://doi.org/10.1088/1367-2630/12/3/033034.

[24] R. W. Spekkens. “Contextuality for preparations, transformations, and unsharp measurements”. In: *Phys. Rev. A* 71 (5 May 2005), p. 052108. DOI: 10.1103/PhysRevA.71.052108. URL: https://link.aps.org/doi/10.1103/PhysRevA.71.052108.

[25] Robert W. Spekkens. “The ontological identity of empirical indiscernibles: Leibniz’s methodological principle and its significance in the work of Einstein”. In: *arXiv:1909.04628* (2019). URL: https://arxiv.org/abs/1909.04628.

[26] Robert W. Spekkens et al. “Preparation Contextuality Powers Parity-Oblivious Multiplexing”. In: *Phys. Rev. Lett.* 102 (1 Jan. 2009), p. 010401. DOI: 10.1103/PhysRevLett.102.010401. URL: https://link.aps.org/doi/10.1103/PhysRevLett.102.010401.

[27] Nik Weaver. *Mathematical Quantization*. Chapman and Hall, 2001.

[28] Shiv Akshar Yadavalli and Ravi Kunjwal. “Contextuality in entanglement-assisted one-shot classical communication”. In: *Quantum* 6 (Oct. 2022), p. 839. ISSN: 2521-

327X. DOI: 10.22331/q-2022-10-13-839. URL: https://doi.org/10.22331/q-2022-10-13-839.