The toy theory is the unique largest noncontextual theory satisfying A_1^3-symmetry

Daniel J. Brod
Instituto de Fisica, Universidade Federal Fluminense, 24210-346 Niteroi, RJ, Brazil

Lorenzo Catani
International Iberian Nanotechnology Laboratory (INL),
Av. Mestre Jose Veiga s/n, 4715-330 Braga, Portugal

Robert W. Spekkens
Perimeter Institute for Theoretical Physics, 31 Caroline St. N, Waterloo, Ontario, N2L 2Y5, Canada

The extended abstract below is based on the attached manuscript, which will be expanded in an updated version with additional applications of our main result on the uniqueness of the toy theory.

In order to understand in which respects quantum theory truly resists classical explanation, it is crucial to define a well-motivated, rigorous, experimentally testable, and practically useful notion of classical explainability. Generalized noncontextuality has been shown to provide such a notion [1]. It refers to the requirement that a theory admits of a generalized-noncontextual ontological model. This means that experimental procedures that are operationally equivalent are represented identically in the ontological model of the theory. The ontological model constitutes a formal framework for providing a realist explanation of the theory's predictions [2]. The assumption of noncontextuality is motivated by a methodological principle inspired by Leibniz's principle of the identity of indiscernibles [3], or, equivalently, by the principle of no operational fine-tuning [4]. The violation of noncontextuality – which we refer to as *contextuality* – has been subjected to experimental tests [5–9] and has been shown to constitute a resource in information-processing tasks [10–14].

Despite its importance, determining which aspects of a given quantum phenomenon witness contextuality is often nontrivial. Results have been established in state discrimination [15–21], interference [22, 23], uncertainty relations [24], cloning [25], measurement compatibility [26, 27], coherence [22, 28], metrology [29], thermodynamics [29, 30], weak values [31], and quantum Darwinism [32]. However, although this list is extensive, it is still outpaced by the number of quantum phenomena for which such analysis has not been implemented. Moreover, there is currently no systematic method to address the question in general. Each new case typically requires the development of a dedicated proof.

In this work, we describe a method for deriving contextuality proofs in theories with a tomographically complete set of three two-outcome measurements satisfying A_1^3-symmetry. Specifically, we show that, under these assumptions and for a single system, the toy theory of [33] is the unique largest theory admitting of a noncontextual ontological model—every such theory being a subtheory of it— modulo a gauge freedom in the ontological model that we refer to as *junk ontology*, and which can be ruled out by appealing to the same methodological principle (Leibnizianity) that motivates noncontextuality. The precise statement is as follows:

Theorem 1. *For a single system, the toy theory of [33] is the unique largest theory that admits of a noncontextual ontological model and satisfies A_1^3-symmetry with respect to three two-outcome measurements forming a tomographically complete set; that is, every such theory is a subtheory of the toy theory, up to the gauge freedom (junk ontology). The underlying ontological model is, in turn, unique up to this gauge freedom.*

Thus, for experiments in which A_1^3-symmetry is tested and tomographic completeness assumed, if the corresponding quantum phenomenology can be reproduced within the toy theory, then it admits of a noncontextual explanation; otherwise, it constitutes a proof of contextuality. There is no need to examine all possible noncontextual ontological models to determine whether they can reproduce the phenomenology under consideration.

An operational theory satisfies A_1^3-symmetry with respect to three two-outcome measurements on a single system, which we denote by X, Y, Z, if, for every state in the theory, there exist seven other states with the same X-, Y-, and Z-predictabilities, such that the resulting set of eight states spans the eight possible combinations of expectation values $\langle X\rangle$, $\langle Y\rangle$, and $\langle Z\rangle$. Many operational theories satisfy this condition, including qubit quantum theory, its stabilizer subtheory, and gbit theory. The assumption, which implies a nontrivial set of operational equivalences among single-system states, was originally employed in [24] and [23].

A theory that satisfies A_1^3-symmetry is the toy theory[1], an example of a classical statistical theory with an epistemic restriction—namely, a constraint on the allowed probability distributions over the underlying phase-space points. A more formal treatment than [33], appealing to symplectic structure and applicable to both discrete and continuous-variable systems, can be found in [34]. The toy theory demonstrates that many phenomena previously regarded as genuinely nonclassical features of quantum theory can in fact be reproduced within such a noncontextual theory, and it has inspired a substantial body of subsequent research.

We apply the result above to identify contextual aspects of the SWAP-test phenomenology. The SWAP test is a central primitive in quantum information processing: it provides a way to compare two quantum states by estimating their overlap (equivalently, by testing their symmetry/antisymmetry properties), and it underlies a variety of state-comparison and overlap-estimation tasks.

Although Theorem 1 is stated for a single system, it is all that is needed for the two-particle analysis below. Assuming that the composite ontic state space is the Cartesian product of the single-particle ones, each subsystem is a toy bit and the relevant tests can be taken to be arbitrary response functions on the resulting 16-cell joint ontic space; the argument then uses only the ontic representations of product states. A full characterization of the two-particle state space—only partially resolved in the attached manuscript, where the partially entangled case remains open—is not required here.

Assuming A_1^3-symmetry with respect to three two-outcome measurements forming a tomographically complete set, we first establish an ideal-case result: in the toy theory, and therefore in any noncontextual theory, any two-particle test that is passed by all states of the form $\phi \otimes \phi$ with probability 1 must be trivial, i.e., it is passed with probability 1 by all products of two pure states.

Theorem 2. *Consider any noncontextual model satisfying A_1^3-symmetry with respect to a tomographically complete set of three two-outcome observables. Suppose there is a two-particle test that is passed with probability 1 by all states of the form $\phi \otimes \phi$. Then it is the trivial test, and thus is also passed with probability 1 by all states $\psi \otimes \phi$, for $\psi \neq \phi$.*

We then prove a noise-robust extension: if $\psi \otimes \psi$ passes a test with sufficiently high probability, then (under the same assumptions) any pair $\psi \otimes \phi$ must also pass with a correspondingly high probability.

Theorem 3. *Consider any noncontextual model satisfying A_1^3-symmetry with respect to a tomographically complete set of three two-outcome observables. If there exists a test such that all six pairs of pure states $\psi \otimes \psi$ pass with probability $p \geq 3/4 + \delta$, then any pair of states $\psi \otimes \phi$, with $\psi \neq \phi$, must pass the same test with probability $p \geq 1/2 + 2\delta$.*

These noncontextual predictions differ from those of the SWAP test in quantum theory, where a gap is present between the output probabilities for the cases of identical and orthogonal states. This gap allows us to identify the SWAP-test phenomenology as a witness of contextuality. Since this phenomenology already appears in the stabilizer subtheory, our result constitutes a new proof of contextuality for the two-qubit stabilizer theory. In the article we also show that the same noncontextual SWAP-test phenomenology arises in the 8-state model of Wallman and Bartlett [35], which is preparation- and measurement-noncontextual for the single-system case but transformation-contextual [36].

Our uniqueness theorem implies that, in any setting where A_1^3-symmetry can be certified (and the measurements are tomographically complete), noncontextual predictions can be read off directly from the toy theory, thereby providing an easier and more systematic route to demonstrating contextual/noncontextual separations. In an updated version of this work, we plan to show how a range of existing contextuality proofs can be reproduced via our result, including those found in the contexts of state discrimination [15], conclusive exclusion [21], uncertainty relations [24], state-dependent cloning [25], and two-state overlap [37]. Beyond re-deriving known separations, it would be particularly interesting to use the toy-theory reduction as a search tool for identifying new contextuality witnesses in phenomena that have not yet been analysed.

A second direction for future research concerns our contextuality witness based on the SWAP-test phenomenology. Since the SWAP test has been shown to be operationally equivalent to the Hong–Ou–Mandel effect [38], our results suggest a possible route to studying contextual aspects of bosonic indistinguishability. More broadly, it would be valuable to understand to what extent the present approach can be extended beyond the assumption of A_1^3-symmetry.

---

[1] More precisely, it is the convexified version of the toy theory of [33], i.e., the version in which any state that is a convex mixture of valid pure states is itself a valid state. We will refer to this as “the toy theory”.

I. BIBLIOGRAPHY

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

[2] N. Harrigan and R. W. Spekkens, Einstein, Incompleteness, and the Epistemic View of Quantum States, Foundations of Physics **40**, 125 (2010).

[3] R. W. Spekkens, The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein, arXiv:1909.04628 (2019).

[4] L. Catani and M. Leifer, A mathematical framework for operational fine tunings, Quantum **7**, 948 (2023).

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

[6] X. Zhan, E. G. Cavalcanti, J. Li, Z. Bian, Y. Zhang, H. M. Wiseman, and P. Xue, Experimental generalized contextuality with single-photon qubits, Optica **4**, 966 (2017).

[7] A. Zhang, H. Xu, J. Xie, H. Zhang, B. J. Smith, M. S. Kim, and L. Zhang, Experimental test of contextuality in quantum and classical systems, Phys. Rev. Lett. **122**, 080401 (2019).

[8] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories, PRX Quantum **2**, 020302 (2021).

[9] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi, N. Spagnolo, R. Osellame, E. F. Galvão, and F. Sciarrino, Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor, Science Advances **9**, eadj4249 (2023).

[10] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation Contextuality Powers Parity-Oblivious Multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).

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

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

[13] S. A. Yadavalli and R. Kunjwal, Contextuality in entanglement-assisted classical communication, Quantum **6**, 839 (2022).

[14] A. M. Fonseca, V. P. Rossi, R. D. Baldijão, J. H. Selby, and A. B. Sainz, Robustness of contextuality under different types of noise as quantifiers for parity-oblivious multiplexing tasks, Phys. Rev. A **111**, 022217 (2025).

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

[16] J. Shin, D. Ha, and Y. Kwon, Quantum contextual advantage depending on nonzero prior probabilities in state discrimination of mixed qubit states, Entropy **23**, 10.3390/e23121583 (2021).

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

[18] S. Mukherjee, S. Naonit, and A. K. Pan, Discriminating three mirror-symmetric states with a restricted contextual advantage, Phys. Rev. A **106**, 012216 (2022).

[19] C. Roch i Carceller and J. B. Brask, Contextuality witness inspired by optimal state discrimination, Phys. Rev. A **109**, 032202 (2024).

[20] K. Flatt and J. Bae, Unifying contextual advantages in state discrimination, arXiv:2511.04100 (2025).

[21] Y. Yīng, D. Schmid, and R. W. Spekkens, A contextual advantage for conclusive exclusion: repurposing the pusey-barrett-rudolph construction, arXiv:2512.04173 (2025).

[22] R. Wagner, A. Camillini, and E. F. Galvao, Coherence and contextuality in a mach-zehnder interferometer, Quantum **8**, 1240 (2024).

[23] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, Aspects of the phenomenology of interference that are genuinely nonclassical, Phys. Rev. A **108**, 022207 (2023).

[24] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, What is nonclassical about uncertainty relations?, Phys. Rev. Lett. **129**, 240401 (2022).

[25] M. Lostaglio and G. Senno, Contextual advantage for state-dependent cloning, Quantum **4**, 258 (2020).

[26] A. Tavakoli and R. Uola, Measurement incompatibility and steering are necessary and sufficient for operational contextuality, Phys. Rev. Res. **2**, 013011 (2020).

[27] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Contextuality without incompatibility, Phys. Rev. Lett. **130**, 230201 (2023).

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

[29] M. Lostaglio, Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response, Phys. Rev. Lett. **125**, 230603 (2020).

[30] N. E. Comar, D. Cius, L. F. Santos, R. Wagner, and B. Amaral, Contextuality in anomalous heat flow, PRX Quantum **6**, 030359 (2025).

[31] M. F. Pusey, Anomalous weak values are proofs of contextuality, Phys. Rev. Lett. **113**, 200401 (2014).

[32] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, Emergence of noncontextuality under quantum darwinism, PRX Quantum **2**, 030351 (2021).

[33] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110 (2007).

[34] R. W. Spekkens, Quasi-quantization: Classical statistical theories with an epistemic restriction, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht,

2016) pp. 83–135.

[35] J. J. Wallman and S. D. Bartlett, Non-negative subtheories and quasiprobability representations of qubits, Phys. Rev. A **85** (2012).

[36] P. Lillystone, J. J. Wallman, and J. Emerson, Contextuality and the single-qubit stabilizer subtheory, Phys. Rev. Lett. **122**, 140405 (2019).

[37] R. Wagner, A. Camillini, and E. F. Galvão, Coherence and contextuality in a Mach-Zehnder interferometer, Quantum **8**, 1240 (2024).

[38] J. C. Garcia-Escartin and P. Chamorro-Posada, swap test and hong-ou-mandel effect are equivalent, Phys. Rev. A **87**, 052330 (2013).

The toy theory is the unique largest noncontextual theory satisfying A_1^3-symmetry

Daniel J. Brod
Instituto de Fisica, Universidade Federal Fluminense, 24210-346 Niteroi, RJ, Brazil

Lorenzo Catani
International Iberian Nanotechnology Laboratory (INL),
Av. Mestre Jose Veiga s/n, 4715-330 Braga, Portugal

Robert W. Spekkens
Perimeter Institute for Theoretical Physics, 31 Caroline St. N, Waterloo, Ontario, N2L 2Y5, Canada

In this work we show that, for single systems, the toy theory developed in [Spekkens, PRA 75, 032110 (2007)] is essentially the unique largest theory with a tomographically complete set of three two-outcome measurements that satisfies A_1^3-symmetry and that admits of a noncontextual ontological model. The A_1^3-symmetry assumption is satisfied by many commonly studied theories, such as qubit quantum theory and its stabilizer subtheory, and has been adopted in several previous works. This result implies that, in order to establish contextuality proofs for quantum phenomena (in theories satisfying A_1^3-symmetry) it suffices to analyse them in the toy theory. As an application, we investigate contextual features of the SWAP test phenomenology. In particular, we show that, in the toy theory (and therefore in any noncontextual theory satisfying A_1^3-symmetry), any two-particle test that is passed by the tensor product of two identical pure states with high probability must also be passed by the tensor product of two orthogonal states with high probability. This is in stark contrast with what occurs in quantum theory, where there is a gap between the output probabilities associated to the tensor product of two identical and the tensor product of two orthogonal input states. Consequently, the quantum phenomenology of the SWAP test is a witness of contextuality.

CONTENTS

I. INTRODUCTION

In order to understand in which respects quantum theory truly resists classical explanation, it is crucial to define a well-motivated, rigorous, experimentally testable, and practically useful notion of classical explainability. Generalized noncontextuality has been shown to provide such a notion [1]. It refers to the requirement that a theory admits of a generalized-noncontextual ontological model. This means that experimental procedures that are operationally equivalent are represented identically in the ontological model of the theory. The ontological model constitutes a formal framework for providing a realist explanation of the theory's predictions [2]. The assumption of noncontextuality is motivated by a methodological principle inspired by Leibniz's principle of the identity of indiscernibles [3], or, equivalently, by the principle of no operational fine-tuning [4]. The violation of noncontextuality – which we refer to as *contextuality* – has been subjected to experimental tests [5–9] and has been shown to constitute a resource in information-processing tasks [10–14].

Despite its importance, determining which aspects of a given quantum phenomenon witness contextuality is often nontrivial. Results have been established in state discrimination [15–21], interference [22, 23], uncertainty relations [24], cloning [25], measurement compatibility [26, 27], coherence [22, 28], metrology [29], thermodynamics [29, 30], weak values [31], and quantum Darwinism [32]. However, although this list is extensive, it is still outpaced by the number of quantum phenomena for which such analysis has not been implemented. Moreover, there is currently no systematic method to address the question in general. Each new case typically requires the development of a dedicated proof.

In this work, we describe a novel method for deriving proofs of contextuality in theories with a tomographically complete set of three two-outcome measurements satisfying the assumption of A_1^3-symmetry. Specifically, we show (Theorem II.1) that, under these assumptions, the only noncontextual theory is the toy theory of [33], modulo a certain gauge freedom in the ontological model that we refer to as *junk ontology*, and which can be ruled out by appealing to the same methodological principle of Leibnizianity that motivates noncontextuality. Thus, for experiments in which A_1^3-symmetry is tested and tomographic completeness assumed, if the corresponding quantum phenomenology can be reproduced within the toy theory, then it admits of a noncontextual ontological model; otherwise, it constitutes a proof of contextuality. There is no need to examine all possible noncontextual ontological models to determine whether they can reproduce the phenomenology under consideration.

An operational theory satisfies A_1^3-symmetry with respect to three two-outcome measurements on a single system, which we denote by X, Y, Z, if, for every state in the theory, there exist seven other states with the same X-, Y-, and Z-predictabilities, such that the resulting set of eight states spans the eight possible combinations of expectation values $\langle X\rangle$, $\langle Y\rangle$, and $\langle Z\rangle$. Many operational theories satisfy this condition, including qubit quantum theory, its stabilizer subtheory, and gbit theory. The assumption, which implies a nontrivial set of operational equivalences among single system states, was originally employed in [24] and [23].

A theory that satisfies A_1^3-symmetry is the toy theory of [33][1], an example of a classical statistical theory with an epistemic restriction—namely, a specific constraint on the allowed probability distributions over the underlying phase-space points. A more formal treatment of the toy theory than that of [33], appealing to symplectic structure and applicable to both discrete and continuous-variable systems, can be found in [34]. The toy theory demonstrates that many phenomena previously regarded as genuinely nonclassical features of quantum theory can in fact be reproduced within such a noncontextual theory, and it has inspired a substantial body of subsequent research [35–46].

We apply the result obtained above to identify contextual aspects of the SWAP test phenomenology. The SWAP test is a central primitive in quantum information processing: it provides a way to compare two quantum states by estimating their overlap (equivalently, by testing their symmetry/antisymmetry properties), and it underlies a variety of state-comparison and overlap-estimation tasks.

Assuming A_1^3-symmetry with respect to three two-outcome measurements forming a tomographically complete set, we first establish the ideal-case result (Theorem III.1), in which we consider any two-particle test that is passed by the tensor product of two identical pure states with probability 1. We show that, in the toy theory, and therefore in any noncontextual theory, such a test must be the trivial test, i.e., the one that is passed with probability 1 by all products of two pure states. We then prove a robust extension of this result (Theorem III.2): any two-particle test that is passed by the tensor product of two identical pure states with high probability must also be passed by a product of two orthogonal states with high probability in the toy theory, and therefore in any noncontextual theory.

---

[1] More precisely, it is the convexified version of the toy theory of [33], i.e., the version in which any state that is a convex mixture of valid pure states is itself a valid state. We will refer to this as "the toy theory".

The noncontextual predictions of both the ideal and robust cases differ from those of the SWAP test in quantum theory, where there is a gap between the output probabilities for the cases of identical and orthogonal states. This gap allows us to identify the SWAP test phenomenology as a witness of contextuality. Since this phenomenology already appears in the stabilizer subtheory, our result constitutes a new proof of contextuality for the two-qubit stabilizer theory.

The remainder of the article is structured as follows. In Section II we prove our main result for a single system: the toy theory is the largest noncontextual theory in this symmetry class. The composite-system case is treated in Appendix A, where we establish the result on the separable and maximally entangled sectors and explain why the general (partially entangled) case remains open. In Section III we apply this result to show that the phenomenology of the SWAP test witnesses contextuality. In Section IV we conclude and outline avenues for future research. We also include Appendix B, where we show that the same noncontextual SWAP-test phenomenology arises in the 8-state model of Wallman and Bartlett [47], which is preparation- and measurement-noncontextual for the single-system case but transformation-contextual [48].

II. UNIQUENESS OF THE TOY THEORY

In this section we prove the following theorem.

Theorem II.1. *For a single system, the toy theory of [33] is the unique largest theory that admits of a noncontextual ontological model and satisfies A_1^3-symmetry with respect to three two-outcome measurements forming a tomographically complete set; that is, every such theory is a subtheory of the toy theory, up to the gauge freedom (junk ontology) described below. The underlying ontological model is, in turn, unique up to this gauge freedom.*

A. A_1^3-symmetry

Before proving the result, let us clarify what we mean by A_1^3-symmetry and how we impose it, given it is an important assumption behind our key result.

We follow the definition of [24], which states that a theory satisfies A_1^3-symmetry relative to the three observables X, Y, and Z if, for every state ρ_1 in the theory, there exist other 7 states $\{\rho_2, \rho_3 \dots \rho_8\}$ in the theory which agree with ρ_1 on the absolute values $|\langle X\rangle|, |\langle Y\rangle|$, and $|\langle Z\rangle|$, but which correspond to all 8 choices of different signs for these quantities. That is, we have, for example

$$\langle X\rangle_{\rho_1} = \langle X\rangle_{\rho_2}\,, \tag{1}$$
$$\langle Y\rangle_{\rho_1} = \langle Y\rangle_{\rho_2}\,, \tag{2}$$
$$\langle Z\rangle_{\rho_1} = -\langle Z\rangle_{\rho_2}\,, \tag{3}$$

and so on. Geometrically, if we represent the states as points with coordinates $\{\langle X\rangle, \langle Y\rangle, \langle Z\rangle\}$ (i.e., in Bloch-sphere coordinates), then the 8 points $\{\rho_1, \dots, \rho_8\}$ lie in the 8 vertices of a rectangular prism with faces parallel to the XY, XZ and YZ planes.

Since we assume that the X, Y and Z measurements are tomographically complete, this symmetry leads to several operational equivalences, such as

$$\frac{1}{2}(\rho_1 + \rho_8) = \frac{1}{4}(\rho_1 + \rho_4 + \rho_6 + \rho_7).$$

In this work, we impose this symmetry in a different manner that turns out to be more convenient, but is ultimately equivalent. For any model, we posit the existence of a set of Pauli-like transformations, which act on the expectation values of the Pauli observables as expected. For instance, for any state ρ, we have that applying X to ρ leaves the value of $\langle X\rangle$ invariant but flips the signs of $\langle Y\rangle$ and $\langle Z\rangle$, with the obvious extensions to Y and Z transformations. We also consider a transformation, which we call R, which acts as an "inversion about the origin" in these coordinates, namely flipping the sign of the three expectations above.

It is then clear to see that the set of 8 states $\{\rho_1, \dots, \rho_8\}$ are actually the orbit of ρ_1 under combinations of the three Pauli operators and the inversion. The operational equivalences can now be written as

$$\frac{1}{2}\left(\rho + R[\rho]\right) = \frac{1}{4}\left(\rho + X[\rho] + Y[\rho] + Z[\rho]\right), \tag{4}$$

which is the form that we will use. We emphasize that, even though we are describing them in terms of a set of transformations, the expression above should be interpreted as the ontological equivalence between two preparations (and subsequently, when imposed at the ontological level, as a consequence of preparation noncontextuality).

B. Proof of Theorem II.1 – states

We now prove that any theory admitting of a noncontextual ontological model for a single system which satisfies A_1^3-symmetry with respect to the three Pauli measurements is the toy theory [33], up to a spurious unobservable degree of freedom (i.e., junk ontology).

Consider an operational theory with a tomographically complete set of three measurements $\{X, Y, Z\}$, each with two outcomes ± 1, and consider an ontological model of this theory. Without loss of generality, we may assume that the ontic state space can be taken to consist of 8 ontic states, corresponding to all deterministic assignments of outcomes to the three measurements.[2] Equivalently, we represent the ontic states by the standard basis vectors of $\mathbb{R}^8$, namely

$$\{(1,0,0,0,0,0,0,0),\ (0,1,0,0,0,0,0,0),\ \ldots,\ (0,0,0,0,0,0,0,1)\}. \tag{5}$$

Without loss of generality we can identify the three effects corresponding to the Pauli measurements

$$\xi_{+1|X} = (1,1,0,0,1,1,0,0) \tag{6}$$

$$\xi_{+1|Y} = (1,0,1,0,1,0,1,0) \tag{7}$$

$$\xi_{+1|Z} = (1,1,1,1,0,0,0,0) \tag{8}$$

with the other three effects defined such that

$$\xi_{+1|A} + \xi_{-1|A} = \vec{u},$$

for $A = \{X, Y, Z\}$ and where $\vec{u}$ is the unit effect. We define the corresponding three expectation values on some state ρ as

$$\langle A \rangle = (\xi_{+1|A} - \xi_{-1|A}) \cdot \rho.$$

Finally, we also define the X, Y and Z transformations in this model as the following permutations:

$$X : (16)(25)(38)(74), \tag{9}$$

$$Y : (17)(28)(35)(46), \tag{10}$$

$$Z : (14)(23)(58)(67). \tag{11}$$

These Pauli-like transformations have been defined by the property that the X should preserve the expectation of the X observable and flip the sign of the Y and Z observables, and so on.

Consider now an arbitrary vector in this ontological model

$$\vec{\rho} = (a, b, c, d, e, f, g, h), \tag{12}$$

with normalization $h = 1 - a - b - c - d - e - f - g$. Define the Pauli-twirled state obtained from it as follows

$$\mathcal{P}[\rho] = \frac{1}{4}\left(X[\rho] + Y[\rho] + Z[\rho] + \rho\right). \tag{13}$$

It is easy to check that, for this state, $\langle X \rangle = \langle Y \rangle = \langle Z \rangle = 0$. Indeed, in quantum theory and in the toy theory, this preparation corresponds to the maximally mixed state for any ρ. In our model, at this stage, this preparation corresponds to the following state

$$\mathcal{P}[\rho] = \tfrac{1}{4}(A, 1-A, 1-A, A, 1-A, A, A, 1-A) =: v_A,$$

for $A = a + d + f + g \in [0, 1]$.

We now invoke the assumption of preparation noncontextuality, and demand that the maximally mixed state is *unique*. This breaks down our model into a class of distinct models, each corresponding to a different choice of A. Let us begin by considering the simplest case, where $A = 1$.

---

[2] If the ontic state space contains more than these 8 ontic states, we can coarse-grain it into 8 disjoint regions labelled by the induced deterministic assignments to (X, Y, Z). By tomographic completeness, this coarse-graining preserves all operational statistics, so no generality is lost by restricting attention to these 8 states.

1. The case $A = 1$.

If we choose the model corresponding to $A = 1$, its maximally mixed state is given simply by

$$v_1 = \frac{1}{4}(1, 0, 0, 1, 0, 1, 1, 0),$$

and the arbitrary state ρ from Eq. (12) becomes

$$\rho = (1 - d - f - g, b, c, d, e, f, g, -b - c - e).$$

The fact that ρ must only have nonnegative entries implies $b = c = e = 0$, and thus

$$\rho = (1 - d - f - g, 0, 0, d, 0, f, g, 0).$$

These correspond to arbitrary convex combinations of the following four states

$$\begin{gathered}(1, 0, 0, 0, 0, 0, 0, 0)\\(0, 0, 0, 1, 0, 0, 0, 0)\\(0, 0, 0, 0, 0, 1, 0, 0)\\(0, 0, 0, 0, 0, 0, 1, 0)\end{gathered}$$

If we compute the expectations of the three Pauli operators for these four states — i.e., their "Bloch sphere coordinates", we find that they lie at the vertices of a tetrahedron

$$\{(-1, 1, -1), (1, -1, -1), (-1, -1, 1), (1, 1, 1)\}.$$

We now want to impose that the theory satisfies A_1^3 symmetry. This effectively corresponds to demanding that the equal mixture of a state ρ and the one with diametrically opposed Bloch sphere coordinates (let us call it $\rho^\perp$) is the maximally mixed state, following Eq. (4). However, there is no obvious transformation that generates $\rho^\perp$ given an arbitrary ρ. Therefore, we instead flip the logic of the A_1^3 symmetry, and rather than demanding that ρ and $\rho^\perp$ mix to v_1, we will use this as a *definition* of $\rho^\perp$:

$$\rho^\perp = 2v_1 - \rho.$$

If we now impose the constraint that $\rho^\perp$ is a valid state (i.e., all entries in the range $[0, 1]$) whenever ρ is a valid state we obtain a set of inequalities relating the free variables d, f, g. This simple set of constraints defines a polytope, and it is not hard to show that this polytope is the convex hull of the six extremal points:

$$\begin{gathered}\tfrac{1}{2}(0, 0, 0, 0, 0, 1, 1, 0)\\\tfrac{1}{2}(0, 0, 0, 1, 0, 0, 1, 0)\\\tfrac{1}{2}(0, 0, 0, 1, 0, 1, 0, 0)\\\tfrac{1}{2}(1, 0, 0, 0, 0, 0, 1, 0)\\\tfrac{1}{2}(1, 0, 0, 0, 0, 1, 0, 0)\\\tfrac{1}{2}(1, 0, 0, 1, 0, 0, 0, 0)\end{gathered}$$

Note that, if we simply ignore the entries 2, 3,5 and 8 in each of the above vectors, they correspond exactly to the 6 epistemic states in the toy theory of [33].

We now need to consider the set of allowed transformations in this model. A priori, all permutations of the ontic states should be valid. However, note that there are permutations that do not preserve the maximally mixed state. For example, if we define the reversal operation by

$$R : (18)(27)(36)(45).$$

then we get

$$R[v_1] = \frac{1}{4}(0, 1, 1, 0, 1, 0, 0, 1).$$

If we demand that the maximally mixed state is invariant under any transformation in the model, this means that we must discard R as a valid transformation. Alternatively, notice that the statistics of $R[v_1]$ is again that of the maximally mixed state, and thus its presence would contradict the assumption that the maximally mixed state is unique. From these considerations it follows that, in order to preserve the uniqueness of the maximally mixed state, the set of allowed transformations in this model must correspond to only permutations within entries 1, 4, 6 and 7, which correspond exactly to the set permutations allows in the toy theory.

In conclusion, the state space of the resulting model uses only 4 of the 8 ontic states from the original model, which is preserved by all allowed transformations. If we were to simply disregard vectors entries 2, 3, 5, and 8, in all states, transformations and effects, the resulting model would be exactly the one corresponding to the toy theory.

Clearly the same reasoning applies in the case where $A = 0$, which would also correspond exactly to the toy theory, but now using ontic states 2, 3, 5 and 8.

2. *The case* $A \neq 0, \frac{1}{2}, 1$.

Let us now consider a more general case, where we choose any value of $A \in [0, 1/2) \cup (1/2, 1]$. The case $A = 1/2$ has some extra subtleties which we delay to the next section.

For this choice, the maximally mixed state is now given by

$$v_A = \tfrac{1}{4}(A, 1-A, 1-A, A, 1-A, A, A, 1-A).$$

We now apply the same set of constraint as before. First, we discard from the model all transformations that mix entries 1,4,6,7 with entries 2,3,5,8. The reason is the same as before: these permutations (e.g., the reversal R), when applied to v_A, generate a state that is distinct from it but that also has maximally mixed statistics, thus would violate the constraint that the maximally mixed state is unique.

Next, we impose A_1^3 symmetry, in the form of demanding that $\rho^\perp = 2v_A - \rho$ is a valid state whenever ρ is a valid state. The resulting inequalities again define a polytope, whose vertices are characterized by 36 distinct states. This can be understood as a natural consequence of the previous result, as follows. The state v_A is a convex combination of states v_1 and v_0. In the cases where $A = 0$ and $A = 1$ we showed, in the previous section, that we recover the toy theory using sets of ontic states $\{1, 4, 6, 7\}$ and $\{2, 3, 5, 8\}$, respectively. The states that we obtain for general A are then all 36 convex combinations, with weights $\{A, 1-A\}$, of the 6 states in the $\{1, 4, 6, 7\}$ sector and the 6 states in the $\{2, 3, 5, 8\}$ sector. They include states of the type:

$$\{\tfrac{1}{2}(A, 1-A, 1-A, A, 0, 0, 0, 0), \tfrac{1}{2}(A, 1-A, 0, A, 1-A, 0, 0, 0) \ldots\}$$

However, we now argue that the resulting model is preparation contextual. To see why, we separate the states in two classes. There are 6 of them, which we call “pure” states, which have the Pauli statistics that corresponds to pure stabilizer states in quantum theory, or the epistemic states of the toy theory. In these cases, we have a convex combinations of two “copies” of the toy theory in the two sectors, but they are perfectly correlated, in the sense that for both sectors the states have the same statistics. For example, the state

$$\tfrac{1}{2}(A, 1-A, 1-A, A, 0, 0, 0, 0)$$

has $\langle X \rangle = 0 = \langle Y \rangle$ and $\langle Z \rangle = 1$, for any A. Crucially, these states (and their convex combinations) are already sufficient to generate any set of Pauli statistics compatible with the toy theory or stabilizer quantum theory.

The remaining 30 states, which we call “pseudo-pure” states, correspond to choices where, in each sector, we have a state with a different statistics. For example, the state

$$\tfrac{1}{2}(A, 1-A, 0, A, 1-A, 0, 0, 0)$$

is such that $\langle Y \rangle = 0$, $\langle X \rangle = A$ and $\langle Z \rangle = 1 - A$. Now note that the exact same Pauli statistics can be obtained by a convex combination of pure states:

$$\tfrac{A}{2}(A, 1-A, 1-A, A, 0, 0, 0, 0) + \tfrac{1-A}{2}(A, 1-A, 0, 0, 1-A, A, 0, 0).$$

Since we assume that Pauli measurements are tomographically complete, we have two distinct preparations that correspond to the same state. Therefore, in order for our model to include the 6 states corresponding to perfect predictability for all three Pauli measurements, that is, the 6 pure states, preparation nonconextuality implies that we must discard all the 30 pseudo-pure states from the model.

Similarly, the only set of allowed transformations in this model should correspond to those which take allowed states into allowed states. This means that we can apply arbitrary permutations between ontic states $\{1, 4, 6, 7\}$, but they must always be accompanied by a corresponding permutation in sector $\{2, 3, 5, 8\}$ such that every pure state is taken into another pure state, and not into a pseudo-pure one.

After imposing all described restrictions on the set of allowed states and operations, we are left again with a model that corresponds effectively to the toy theory. Every state is fully determined by its entries on the $\{1, 4, 6, 7\}$ sector, and every operation is also fully determined by how it permutes the elements in this sector—the corresponding states and transformations in the $\{2, 3, 5, 8\}$ sector must always be fixed in order to maintain the model with the convex hull of the 6 pure states.

We now return to the big picture. We started from the assumptions of preparation noncontexuality, tomographic completeness of the Pauli measurements and A_1^3 symmetry. By demanding that the maximally mixed state is unique, the original model broke down into a class of distinct models parameterized by some $A \in [0, 1]$. However, for every value of A (except for $1/2$, which we will return to shortly) the resulting model is completely isomorphic to the one associated with the toy theory in every way, and the Pauli measurements are insensitive to the value of A. Therefore, we can interpret the possible choices of A as a form of gauge freedom, concluding therefore that the toy theory is the unique theory admitting of a noncontextual model compatible with this set of assumptions, up to unobservable "junk ontology". Notice that the removal of junk ontology can be justified by appealing to the same credential underlying noncontextuality, namely the principle of Leibnizianity. In its contrapositive form, this principle states that ontological distinctness must correspond to empirical distinguishability, which precisely prohibits junk ontology.

3. *The case $A = 1/2$*

We now consider the more subtle case where $A = 1/2$. First, notice that the arguments from the previous section can be repeated in the case $A = 1/2$, and would lead to the same conclusions. However, in this case there is one step that would not be well justified, namely the removal of certain permutations from the model. To see this, recall that the maximally mixed state now is

$$v_{1/2} = \tfrac{1}{8}(1, 1, 1, 1, 1, 1, 1, 1).$$

This state is invariant under a larger set of transformations than v_A. For example, we can define

$$P : (2431)(5687), \tag{14}$$
$$H : (12)(36)(45)(78), \tag{15}$$
$$R : (18)(27)(36)(45), \tag{16}$$

where H and P are analogues of the Hadamard and Phase operations from quantum theory, and we encountered R before as a Bloch-sphere inversion (i.e., which flips the sign on all three Pauli expectations). There is no good *a priori* reason to exclude them from our model. However, it is now easy to see that the state

$$\frac{1}{2}\left(\rho + R[\rho]\right)$$

has again the same Pauli statistics of the maximally mixed state. Once more invoking the assumption of preparation noncontextuality, we demand that

$$\frac{1}{2}\left(\rho + R[\rho]\right) = v_{1/2}.$$

Since $v_{1/2} = \mathcal{P}[\rho]$, it is clear that the expression above is nothing more than a manifestation of A_1^3−symmetry. By imposing this constraint, the resulting model corresponds to a polytope whose vertices are given by the six states

$$\tfrac{1}{4}(0, 0, 0, 0, 1, 1, 1, 1)$$
$$\tfrac{1}{4}(1, 1, 1, 1, 0, 0, 0, 0)$$
$$\tfrac{1}{4}(0, 0, 1, 1, 0, 0, 1, 1)$$
$$\tfrac{1}{4}(1, 1, 0, 0, 1, 1, 0, 0)$$
$$\tfrac{1}{4}(1, 0, 1, 0, 1, 0, 1, 0)$$
$$\tfrac{1}{4}(0, 1, 0, 1, 0, 1, 0, 1)$$

which are exactly the states in the 8-state model of Wallman and Bartlett [47].

There is a subtly different route to the 8-state model, in which one defines $\rho^\perp := 2v_{1/2} - \rho$, analogously to the previous subsection. This construction yields two *a priori* distinct states, $R[\rho]$ and $\rho^\perp$, that nevertheless have identical operational statistics. If one additionally imposes $R[\rho] = \rho^\perp$ on the grounds of preparation noncontextuality, then one again recovers the 8-state model. The key difference from the analysis of the previous subsection is that here one does not need to exclude permutations such as R from the model: the availability of R, together with the preceding argument, bypasses the issue of pseudopure states and leads directly to the 8-state model.

However, from our perspective, the standard 8-state model should still be excluded from our conclusions, since it is transformation contextual, as shown in [48], precisely due to the existence of transformations that permute between the sectors $\{1,4,6,7\}$ and $\{2,3,5,8\}$. Therefore, when $A = 1/2$ we find that either (i) the model coincides with the 8-state model of [47], includes transformations such as R, and is contextual, or (ii) it is a subtheory of the 8-state model of [47] in which the allowed transformations are restricted to permutations within each sector, in which case it is simply another instance of the toy theory up to a gauge transformation, as described in the previous section.

Since the SWAP test fits within a prepare-and-measure paradigm, one might worry that excluding the Wallman–Bartlett 8-state model on the grounds that it is "only" transformation contextual is too strong. In appendix B we show that, in the SWAP-test setting, one recovers the same conclusion as for the toy theory even if the 8-state model is contextual. Along the way, we also show that any preparation noncontextual fragment of the two-particle 8-state model admits no nontrivial entangled states.

C. Proof of Theorem II.1 – effects

In this subsection, the goal is to complete our reconstruction of Spekkens' single-particle toy theory by obtaining also its effect space as a consequence of A_1^3 symmetry. Our starting point is the ontic state space spanned by 4 states described in the previous section. In other words, we assume it has already been established that the underlying ontological model that explains our system satisfies A_1^3 symmetry and preparation noncontextuality, and that we have already discarded any junk ontology. The goal now is to prove that the largest possible effect space compatible with the underlying ontological model and with A_1^3 symmetry also coincides with that of the toy theory. Here the four surviving ontic states from the previous section are relabelled $1,\dots,4$, and the components $e_1,\dots,e_4$ are the response probabilities of the effect on them.

We begin by considering an unconstrained set of effects represented in the underlying ontological model, i.e., we assume 4-dimensional effect vectors

$$\vec{e} = (e_1, e_2, e_3, e_4) \tag{17}$$

subject only to the constraints that $0 \le e_i \le 1$ for $i = 1 \dots 4$. We remark that this is not fully unconstrained in the GPT sense [49], where the effects are only required to satisfy $0 \le \vec{e}\cdot\vec{p} \le 1$ for all states $\vec{p}$ in the state space. If we use this GPT definition to construct all effects compatible with the state space of the toy theory, we would obtain effects such as $(2,0,0,0)$, which give sensible results over all epistemic states of the toy theory but do not make sense in the underlying ontological model. Instead, by starting with $\vec{e}$ subject only to $0 \le e_i \le 1$ we are allowing all possible effects compatible with the ontological model, whilst making no reference to the allowed epistemic states.

Given an arbitrary effect $\vec{e}$ as in Eq. (17), we can map it to a more convenient set of coordinates as

$$\vec{\epsilon} := (\epsilon_x, \epsilon_y, \epsilon_z, t) = \frac{1}{4}(e_1 - e_2 + e_3 - e_4, e_1 - e_2 - e_3 + e_4, e_1 + e_2 - e_3 - e_4, e_1 + e_2 + e_3 + e_4). \tag{18}$$

The first three entries of $\vec{\epsilon}$ correspond to the "Bloch sphere" coordinates of the effect vectors, whereas t is the probability associated to the maximally mixed epistemic state $(1,1,1,1)/4$. We could have described the states ρ in the previous section in terms of these same coordinates, where r_i would have corresponded to the expectation of the corresponding Pauli observable in state ρ, and the last coordinate would be trivially equal to 1 due to normalization of the state vectors.

As in the previous section, we now demand that the effects space satisfies A_1^3 symmetry in terms of its Bloch sphere coordinates. Concretely this means that, for any effect vector $\vec{\epsilon}$, there should exist 7 other effects vectors obtained by flipping the sign of the three first entries of $\vec{\epsilon}$. Computationally, the most straightforward way to obtain the largest effect space compatible with this symmetry is to (i) map all 16 extreme effects defined by Eq. (17) to the corresponding $\vec{\epsilon}$ coordinates, (ii) compute the 4-dimensional convex hull of these 16 extreme effects, encoded as a set of linear inequalities, (iii) compute the corresponding inequalities defining the other 7 polytopes defined by the orbit of these 16 effect vectors under A_1^3 symmetry, (iv) map all inequalities back in terms of the coordinates in the original ontic state space, and (v) solve the vertex ennumeration problem for the resulting polytope. By construction, that is the largest polytope which is a subset of the unconstrained polytope and that is compatible with A_1^3 symmetry. Performing these steps, the extreme effects of the resulting polytope are given by the two trivial effects (the zero effect and the unit effect)

$$(0,0,0,0),(1,1,1,1),$$

and the six toy theory effects

$$(1,1,0,0),(1,0,1,0),(1,0,0,1),(0,1,1,0),(0,1,0,1),(0,0,1,1).$$

This concludes the proof that A_1^3 symmetry is sufficient to reconstruct the toy theory (together with noncontextuality, here embodied in the assumption of the underlying ontological model). We reiterate that this argument did not use any information about the epistemic state space of the toy theory.

D. Composite systems

This completes the proof of Theorem II.1 for a single system: its state space, its effect space, and its allowed transformations all coincide with those of the toy theory, up to junk ontology. The extension to two particles is more delicate. As shown in Appendix A, the same set of single-particle constraints recovers the toy theory exactly on the separable and maximally entangled sectors, but does not by itself rule out a family of spurious "bad" states with partial entanglement; we discuss there the additional principles that would close this gap. We stress that our application to the SWAP test (Sec. III) does not require resolving this issue: it relies only on the single-particle result, the product structure of the composite ontic space, and the freedom to take the relevant effects to be arbitrary response functions.

III. CONTEXTUAL ASPECTS OF THE SWAP TEST

In this Section we leverage Theorem II.1 to provide a noncontextual bound for the phenomenology of the SWAP test.

Theorem III.1. *Consider any noncontextual model satisfying A_1^3-symmetry with respect to a tomographically complete set of three two-outcomes observables. Suppose there is a two-particle test that is passed with probability 1 by all states of the type $\phi\otimes\phi$. Then it is the trivial test, and thus is also passed with probability 1 by all states $\psi\otimes\phi$, for $\psi\neq\phi$.*

Notice that in (stabilizer) quantum theory, which does satisfy A_1^3-symmetry, the conclusion of the theorem does not hold, as evidenced by the SWAP-test, which is passed with probability 1 for all symmetric states but passed with probability 1/2 by any state $|\psi\rangle|\phi\rangle$ for $\langle\phi|\psi\rangle=0$. Therefore, this consists of a new proof of contextuality of the stabilizer theory of 2 qubits.

We can also prove a noise-robust version of the result, stated as follows.

Theorem III.2 (Noise-robust version)**.** *Consider any noncontextual model satisfying A_1^3-symmetry with respect to a tomographically complete set of three two-outcomes observables. If there exists a test such that all six pairs of pure states $\psi\otimes\psi$ pass with probability $p\geq 3/4+\delta$, then any pair of states $\psi\otimes\phi$, with $\psi\neq\phi$, must pass the same test with probability $p\geq 1/2+2\delta$.*

We recover theorem III.1 when $\delta=1/4$. This version of the Theorem is noise-robust in the following sense. Suppose we wish to experimentally implement a proof of contextuality via the SWAP test. This involves two steps: (i) checking that all states $\psi\otimes\psi$ pass a particular test with probability 1, and (ii) checking that some state $\psi\otimes\phi$ passes it with smaller probability, say 1/2. However, neither outcome will perfectly match the theoretically predictions. The robust version then tells us that some noise is acceptable on both estimates of probabilities while still providing a proof of contextuality. Below we prove only Theorem III.2, since Theorem III.1 immediately follows.

Proof of Theorem III.2. By Theorem II.1, each single system in a noncontextual model satisfying A_1^3-symmetry is, up to junk ontology, a toy bit, with the four-element ontic space of the toy theory. We assume that the ontic state space of the two-particle system is the Cartesian product of the single-particle ones, so that it consists of the $4 \times 4 = 16$ cells of two toy bits. We make no assumption on which two-particle states are allowed and, following the standard treatment of effects in an ontological model, we take the relevant test to be an arbitrary response function on these 16 cells, i.e., a vector with entries in $[0, 1]$. The argument below uses only the ontic representations of product states together with this effect; in particular, it does not rely on any characterization of the full two-particle state space (for which see Appendix A).

The main step of the result relies on the following important fact: in the toy theory for two toy bits, the collection of all states of the type $\psi \otimes \psi$, where ψ runs over all 6 pure states, spans the full two-particle ontic state space. More precisely, we can write the uniform mixture over all symmetric states (i.e., $\psi \otimes \psi$) as

$$\mathbf{p}_I = \frac{1}{24}\begin{pmatrix} 1 & 1 & 1 & 3 \\ 1 & 1 & 3 & 1 \\ 1 & 3 & 1 & 1 \\ 3 & 1 & 1 & 1 \end{pmatrix} \tag{19}$$

where the 4×4 grid corresponds to the ontic state space of two toy bits, following [33], and the numbers correspond to the relative weight associated to each ontic state. The uniform mixture over all non-identical states (i.e., $\psi \otimes \phi$, for $\psi \neq \phi$) can be written as

$$\mathbf{p}_D = \frac{1}{120}\begin{pmatrix} 8 & 8 & 8 & 6 \\ 8 & 8 & 6 & 8 \\ 8 & 6 & 8 & 8 \\ 6 & 8 & 8 & 8 \end{pmatrix} \tag{20}$$

(as a sanity check we can immediately verify that $\mathbf{p}_I + \mathbf{p}_D$ is the maximally mixed state). We also need the uniform mixture over all orthogonal states (i.e., where $\phi.\psi = 0$):

$$\mathbf{p}_O = \frac{1}{24}\begin{pmatrix} 2 & 2 & 2 & 0 \\ 2 & 2 & 0 & 2 \\ 2 & 0 & 2 & 2 \\ 0 & 2 & 2 & 2 \end{pmatrix} \tag{21}$$

Suppose now that there is some effect vector in this two-particle ontic space given by

$$\mathbf{e}_{=}\begin{pmatrix} e_{11} & e_{12} & e_{13} & e_{14} \\ e_{21} & e_{22} & e_{23} & e_{24} \\ e_{31} & e_{32} & e_{33} & e_{34} \\ e_{41} & e_{42} & e_{43} & e_{44} \end{pmatrix}, \tag{22}$$

for which $\mathbf{p}_I \cdot \mathbf{e} \geq C$ for some constant C. This means that

$$3\sum_i e_{ii} + \sum_{i\neq j} e_{ij} \geq 24C.$$

It then follows that the probability assigned to the $\mathbf{p}_D$ is

$$\mathbf{p}_D \cdot \mathbf{e} = \frac{1}{120}\left(6\sum_i e_{ii} + 8\sum_{i\neq j} e_{ij}\right) \geq \frac{8}{5}C - \frac{3}{5}.$$

Similarly, the probability assigned to $\mathbf{p}_O$ is

$$\mathbf{p}_O \cdot \mathbf{e} = \frac{1}{12}\sum_{i\neq j} e_{ij} \geq 2C - 1.$$

Meanwhile, in the ideal *quantum* case, the 6 stabilizer product states of the form $|\psi\rangle|\phi\rangle$ where $\langle\phi|\psi\rangle = 0$ pass the SWAP test with probability $1/2$, while other product states where $\langle\phi|\psi'\rangle = 1/\sqrt{2}$ (of which there are 24), pass it with probability $3/4$. Consequently, the quantum-mechanical predictions are

$$\mathbf{p}_D \cdot \mathbf{e} = \frac{6}{30}\frac{1}{2} + \frac{24}{30}\frac{3}{4} = \frac{7}{10}, \tag{23}$$

$$\mathbf{p}_O \cdot \mathbf{e} = \frac{1}{2}. \tag{24}$$

This means that, as long as C is not too small, we expect to observe a quantum violation of the noncontextual bound. How much noise can we tolerate? In order to witness contextuality with a state $\mathbf{p}_I$ we require

$$\frac{8}{5}C - \frac{3}{5} > \frac{7}{10}$$

which means

$$C > 0.8125,$$

whereas to witness it with a state $\mathbf{p}_O$ it suffices to have

$$2C - 1 > \frac{1}{2}$$

which means

$$C > \frac{3}{4}.$$

Clearly the latter case is more noise robust. By setting $C = \frac{3}{4} + \delta$ we conclude the proof of Theorem III.2. By setting $C = 1$ we obtain the particular case of Theorem III.1. $\square$

IV. CONCLUSION AND FUTURE DIRECTIONS

In this work we identified a regime in which generalized noncontextuality rigidly constrains the space of admissible operational theories. Assuming A_1^3-symmetry with respect to a tomographically complete triple of two-outcome measurements, we proved that the toy theory is essentially the largest theory that admits of a noncontextual ontological model, with this model unique up to junk ontology. In particular, once operational equivalences implied by A_1^3-symmetry are imposed at the ontological level, the single-system state space and allowed transformations collapse (modulo this gauge freedom) to those of the toy theory; for composite systems, the same argument fixes the toy theory on the separable and maximally entangled sectors, while the general case remains open (Appendix A). Conceptually, this elevates the toy theory to the canonical representative of noncontextual theories in this symmetry class.

A direct consequence is a systematic route to contextuality proofs for phenomena within this regime: rather than searching over all noncontextual ontological models, it suffices to analyse the phenomenon in the toy theory. We illustrated this by studying the SWAP-test phenomenology. Using the fact that, in the toy theory, the collection of all symmetric product states $\psi \otimes \psi$ span the entire two-system ontic state space, we derived a strong restriction on any noncontextual behaviour: in the ideal case, any two-particle test passed with probability 1 by all $\psi \otimes \psi$ must be the trivial test, hence passed with probability 1 by all pairs of pure states; and in the noise-robust case, if $\psi \otimes \psi$ passes with probability at least $3/4 + \delta$, then any other pair must pass with probability at least $1/2 + 2\delta$. These noncontextual predictions are incompatible with the SWAP test in (stabilizer) quantum theory, where a certain gap persists between identical and orthogonal inputs. This also yields a new proof of contextuality for the two-qubit stabilizer subtheory.

Our uniqueness theorem implies that, in any setting where A_1^3-symmetry can be certified (and the measurements are tomographically complete), noncontextual predictions can be read off directly from the toy theory, thereby providing an easier and more systematic route to demonstrating contextual/noncontextual separations. In an updated version of this work, we plan to show how a range of existing contextuality proofs can be reproduced via our result, including those found in the contexts of state discrimination [15], conclusive exclusion [21], uncertainty relations [24], state-dependent cloning [25], and two-state overlap [50]. Beyond re-deriving known separations, it would be particularly interesting to use the toy-theory reduction as a search tool for identifying new contextuality witnesses in phenomena that have not yet been analysed.

A second direction for future research concerns our contextuality witness based on the SWAP test phenomenology. Since the SWAP test has been shown to be operationally equivalent to the Hong–Ou–Mandel effect [51], our results suggest a possible route to studying contextual aspects of bosonic indistinguishability. More broadly, it would be valuable to understand to what extent the present approach can be extended beyond the assumption of A_1^3-symmetry.

V. BIBLIOGRAPHY

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

[2] N. Harrigan and R. W. Spekkens, Einstein, Incompleteness, and the Epistemic View of Quantum States, Foundations of Physics **40**, 125 (2010).

[3] R. W. Spekkens, The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein, arXiv:1909.04628 (2019).

[4] L. Catani and M. Leifer, A mathematical framework for operational fine tunings, Quantum **7**, 948 (2023).

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

[6] X. Zhan, E. G. Cavalcanti, J. Li, Z. Bian, Y. Zhang, H. M. Wiseman, and P. Xue, Experimental generalized contextuality with single-photon qubits, Optica **4**, 966 (2017).

[7] A. Zhang, H. Xu, J. Xie, H. Zhang, B. J. Smith, M. S. Kim, and L. Zhang, Experimental test of contextuality in quantum and classical systems, Phys. Rev. Lett. **122**, 080401 (2019).

[8] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories, PRX Quantum **2**, 020302 (2021).

[9] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi, N. Spagnolo, R. Osellame, E. F. Galvão, and F. Sciarrino, Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor, Science Advances **9**, eadj4249 (2023).

[10] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, Preparation Contextuality Powers Parity-Oblivious Multiplexing, Phys. Rev. Lett. **102**, 010401 (2009).

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

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

[13] S. A. Yadavalli and R. Kunjwal, Contextuality in entanglement-assisted classical communication, Quantum **6**, 839 (2022).

[14] A. M. Fonseca, V. P. Rossi, R. D. Baldijão, J. H. Selby, and A. B. Sainz, Robustness of contextuality under different types of noise as quantifiers for parity-oblivious multiplexing tasks, Phys. Rev. A **111**, 022217 (2025).

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

[16] J. Shin, D. Ha, and Y. Kwon, Quantum contextual advantage depending on nonzero prior probabilities in state discrimination of mixed qubit states, Entropy **23**, 10.3390/e23121583 (2021).

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

[18] S. Mukherjee, S. Naonit, and A. K. Pan, Discriminating three mirror-symmetric states with a restricted contextual advantage, Phys. Rev. A **106**, 012216 (2022).

[19] C. Roch i Carceller and J. B. Brask, Contextuality witness inspired by optimal state discrimination, Phys. Rev. A **109**, 032202 (2024).

[20] K. Flatt and J. Bae, Unifying contextual advantages in state discrimination, arXiv:2511.04100 (2025).

[21] Y. Yīng, D. Schmid, and R. W. Spekkens, A contextual advantage for conclusive exclusion: repurposing the pusey-barrett-rudolph construction, arXiv:2512.04173 (2025).

[22] R. Wagner, A. Camillini, and E. F. Galvao, Coherence and contextuality in a mach-zehnder interferometer, Quantum **8**, 1240 (2024).

[23] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, Aspects of the phenomenology of interference that are genuinely nonclassical, Phys. Rev. A **108**, 022207 (2023).

[24] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, What is nonclassical about uncertainty relations?, Phys. Rev. Lett. **129**, 240401 (2022).

[25] M. Lostaglio and G. Senno, Contextual advantage for state-dependent cloning, Quantum **4**, 258 (2020).

[26] A. Tavakoli and R. Uola, Measurement incompatibility and steering are necessary and sufficient for operational contextuality, Phys. Rev. Res. **2**, 013011 (2020).

[27] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Contextuality without incompatibility, Phys. Rev. Lett. **130**, 230201 (2023).

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

[29] M. Lostaglio, Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response, Phys. Rev. Lett. **125**, 230603 (2020).

[30] N. E. Comar, D. Cius, L. F. Santos, R. Wagner, and B. Amaral, Contextuality in anomalous heat flow, PRX Quantum **6**, 030359 (2025).

[31] M. F. Pusey, Anomalous weak values are proofs of contextuality, Phys. Rev. Lett. **113**, 200401 (2014).

[32] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, Emergence of noncontextuality under quantum darwinism, PRX Quantum **2**, 030351 (2021).

[33] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110 (2007).

[34] R. W. Spekkens, Quasi-quantization: Classical statistical theories with an epistemic restriction, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht,

2016) pp. 83–135.
[35] L. Catani and D. E. Browne, Spekkens' toy model and its relationship with stabiliser quantum mechanics, New J. Phys. **96**, 052112 (2017).
[36] L. Catani and D. E. Browne, State-injection schemes of quantum computation in spekkens' toy theory, Phys. Rev. A **98**, 052108 (2018).
[37] S. J. van Enk, A toy model for quantum mechanics, Found. Phys. **37**, 1447 (2007).
[38] M. F. Pusey, Stabilizer notation for spekkens'toy theory, Found. Phys. **42**, 688 (2012).
[39] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Reconstruction of gaussian quantum mechanics from liouville mechanics with an epistemic restriction, Phys. Rev. A **86**, 012103 (2012).
[40] J.-Å. Larsson, A contextual extension of spekkens' toy model, AIP Conf. Proc. **1424**, 211 (2012).
[41] P. Blasiak, Quantum cube: A toy model of a qubit, Phys. Lett. A **377**, 847 (2013).
[42] L. Disilvestro and D. Markham, Quantum protocols within Spekkens' toy model, Phys. Rev. A **95**, 052324 (2017).
[43] L. Hausmann, N. Nurgalieva, and L. del Rio, A consolidating review of spekkens' toy theory, arXiv:2105.03277 10.48550/arXiv.2105.03277 (2021).
[44] W. F. Braasch Jr. and W. K. Wootters, A quantum prediction as a collection of epistemically restricted classical predictions, Quantum **6**, 659 (2022).
[45] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, Why interference phenomena do not capture the essence of quantum theory, Quantum **7**, 1119 (2023).
[46] L. Hausmann, N. Nurgalieva, and L. d. Rio, Toys can't play: physical agents in spekkens' theory, New Journal of Physics **25**, 023018 (2023).
[47] J. J. Wallman and S. D. Bartlett, Non-negative subtheories and quasiprobability representations of qubits, Phys. Rev. A **85** (2012).
[48] P. Lillystone, J. J. Wallman, and J. Emerson, Contextuality and the single-qubit stabilizer subtheory, Phys. Rev. Lett. **122**, 140405 (2019).
[49] M. Plávala, General probabilistic theories: An introduction, Physics Reports **1033**, 1 (2023).
[50] R. Wagner, A. Camillini, and E. F. Galvão, Coherence and contextuality in a Mach-Zehnder interferometer, Quantum **8**, 1240 (2024).
[51] J. C. Garcia-Escartin and P. Chamorro-Posada, swap test and hong-ou-mandel effect are equivalent, Phys. Rev. A **87**, 052330 (2013).

Appendix A: Uniqueness for two particles

In this appendix we discuss the extension of Theorem II.1 to two particles. We show that the natural two-particle analogues of the single-particle constraints recover the toy theory exactly on the separable and maximally entangled sectors, and we explain why they do not, on their own, pin it down for general (partially entangled) states. None of the open issues below affects the SWAP-test application of Sec. III, which uses only the single-particle result, the product structure of the composite ontic space, and the freedom to take the relevant effects to be arbitrary response functions.

We begin by assuming that the ontic state space for two copies of the system is given by the cartesian product of two copies the ontic state space of a single system, which we take to be the 4-dimensional space of the toy theory. By consistency, we admit the existence of all 36 pure product states, e.g.,

$$\frac{1}{4}\begin{pmatrix}1&1&0&0\\1&1&0&0\\0&0&0&0\\0&0&0&0\end{pmatrix} \tag{A1}$$

corresponding to the case where both systems are in the state $(1/2,1/2,0,0)$, and so on. The state corresponding to an equal mixture of all pairs of pure product states is the two-particle maximally mixed state:

$$\frac{1}{36}\sum_{\psi,\phi}\psi\otimes\phi=\frac{1}{16}\begin{pmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{pmatrix}. \tag{A2}$$

Let us now construct the entangled states available in this model. As before, we start with an arbitrary state ρ

given by

$$\begin{pmatrix} a & b & c & d \\ e & f & g & h \\ j & k & l & m \\ n & o & p & q \end{pmatrix}. \tag{A3}$$

We now impose preparation noncontextuality by demanding that ρ satisfies the following identities (satisfied by all perfectly-entangled states in quantum theory)

$$\mathcal{P}_1[\rho] = \frac{1}{36}\sum_{\psi,\phi} \psi \otimes \phi,$$

$$\mathcal{P}_2[\rho] = \frac{1}{36}\sum_{\psi,\phi} \psi \otimes \phi,$$

where $\mathcal{P}_i$ is the Pauli twirl from Eq. (13) applied to particle i. These relations imply a series of constraints[3] between the matrix elements of ρ:

$$n + p + o + q = j + k + l + m = \frac{1}{4},$$

$$e + f + g + h = d + h + m + q = \frac{1}{4},$$

$$c + g + l + p = b + f + k + o = \frac{1}{4},$$

$$f + g + h + k + l + m + o + p + q = a + \frac{1}{2}.$$

The set of states compatible with these constraints and the demand that every matrix element of ρ is in $[0, 1]$ forms a polytope, which is the convex hull of 24 extreme points, given by

$$\frac{1}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \frac{1}{4}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \frac{1}{4}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \frac{1}{4}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \ldots,$$

which corresponds exactly to the 24 perfectly entangled states in the toy theory [33].

Two remarks are in order. First, the constraints above amount to demanding that both single-particle marginals of ρ are maximally mixed. The extreme points of the polytope of 4×4 nonnegative matrices with uniform marginals are precisely the 24 maximally entangled toy states. This argument therefore characterises only the *maximally entangled sector*: it constrains neither the product states (whose marginals are pure) nor, more importantly, the partially entangled states (whose marginals are neither pure nor maximally mixed). Together with the 36 product states admitted above and their convex combinations, we thus recover the toy theory exactly on the separable and maximally entangled sectors.

Second, the general case is genuinely open. If, instead of fixing the marginals, one imposes a two-particle generalisation of A_1^3-symmetry together with the available Pauli, Clifford, and (positive but not completely positive) transpose constraints, vertex enumeration returns the 60 toy-theory pure states (36 product and 24 entangled) together with a family of 192 spurious "bad" states. These correspond to non-positive semidefinite Hermitian operators, are supported on 6 of the 16 ontic cells, and have squared Bloch radius $|\vec{r}|^2 = 5/3$ rather than the value 3 shared by all toy-theory and quantum pure states. We have explored several approaches to remove these bad states, but the question is still open. We leave a fully operational two-particle uniqueness statement to future work.

Appendix B: SWAP test in the 8-state model

In this appendix we show that not only there exist a separation between the SWAP-test phenomenology in contextual theories and in the toy theory (and hence in any noncontextual theory satisfying A_1^3-symmetry with respect to a

[3] Incidentally, exactly the same set of constraints also follows if we assume, instead, that ρ reduces to the single-particle maximally-mixed state when marginalized over either particle.

tomographically complete triple of measurements) but also with respect to the the 8-state model of Wallman and Bartlett [47]. The key ingredient is to impose uniqueness of the ontic representation of the maximally mixed state of two systems $\mathbb{I}/4$, in particular under the operational equivalences $\frac{1}{4}\big(\rho+X\rho X+Y\rho Y+Z\rho Z\big)=\frac{\mathbb{I}}{4}$ and $\frac{1}{2}\rho+\frac{1}{2}\,\mathrm{R}_i(\rho)=\frac{\mathbb{I}}{4}$, where R_i denotes the transformation that flips the signs of the X,Y,Z expectation values of the ith subsystem. This transformation is allowed in the 8-state model, but it is not valid in the toy theory (nor in quantum theory, where it would not correspond to a completely positive map). For this reason, it is natural to invoke it only in the 8-state model.

More precisely, our proof proceeds as follows. We first note that the argument does not apply immediately, since the $\psi\otimes\psi$ states in the 8-state model do not span the entire ontic state space. In particular, they have no support on the main diagonal (i.e., on a singlet-like state, where the two particles are always perfectly anticorrelated). This is perhaps unsurprising: the model is known to be transformation-contextual already at the single-system level (via Lillystone's work [48]), and one can show that it becomes preparation-contextual as well if this singlet-like state is admitted, simply by applying a Choi–Jamiołkowski version of Lillystone's proof.

Our goal is therefore to argue that, under the same set of constraints imposed at the single-system level, the 8-state model admits no nontrivial entangled states. We start from an arbitrary two-particle state ρ and impose the following constraints:

(i) We impose that ρ reduces to the maximally mixed state upon marginalization over either subsystem. Importantly, the maximally mixed state is fixed by our choice of single-particle model. In particular, if a single particle is described by the 8-state model, then the two-particle maximally mixed state is the 8×8 matrix with all entries equal to $1/64$, obtained as the uniform mixture over all product states of the form $\psi\otimes\phi$.

(ii) We impose that the state is invariant under $U\otimes U$ and under the SWAP transformation. This step is not meant to characterize the full set of two-particle states in the 8-state model, but rather to reduce the search space for the effects of interest. Since the SWAP test is invariant under these operations (i.e., it depends only on the two-particle overlap), it is natural to restrict attention to effects satisfying the same symmetry (alternatively, one could state a theorem for fragments of quantum theory/the 8-state model whose effects satisfy it). Note that this concerns $U\otimes U$ symmetry: transformations of the form $U\otimes\mathbb{I}$, for example, do affect the statistics.

(iii) We impose the ontological equivalence corresponding to the operational equivalence that we call the Pauli twirl: $\frac{1}{4}\big(\rho+X_1[\rho]+Y_1[\rho]+Z_1[\rho]\big)=\frac{\mathbb{I}}{4}$. Note that this is applied to a single subsystem (though it does not matter which). This operational equivalence is also satisfied by all 24 perfectly entangled stabilizer states, both in quantum theory and in the toy theory. Here we impose it on the grounds that, for any state ρ satisfying (i) and (ii), the statistics of its Pauli-twirled version coincide with those of the maximally mixed state (i.e., all single- and two-qubit Pauli expectation values vanish). Notice how in this case the choice of maximally mixed state is fixed by our single-particle model.

(iv) We impose the ontological equivalence corresponding to the operational equivalence $\frac{1}{2}\rho+\frac{1}{2}\,\mathrm{R}_i(\rho)=\frac{\mathbb{I}}{4}$. This equivalence does not make sense in quantum theory or in the toy theory because R_i is not a valid operation in those theories: it is not completely positive in quantum theory, and it is not a permutation in the toy theory. However, there is no reason not to include this transformation in the 8-state model. Furthermore, as in (iii), one can verify that the resulting state has the statistics of the maximally mixed state.

We now begin by applying conditions (i) and (ii). In particular, demanding invariance under $U\otimes U$ for the single-qubit Pauli gates reduces the set of extreme points in the convex hull of entangled states to 64. However, this set still exhibits contextuality: it contains distinct states that nevertheless agree on the statistics of all two-qubit Pauli measurements (which form a tomographically complete set). If we restrict the effect space to the convex hull of these 64 extreme points together with all product stabilizer effects, then we recover the ideal version of our result. We also obtain a robustness curve similar to that of the toy theory, though not identical.

Figure 1 shows four curves, corresponding to the different sets of requirements imposed, as explained in the caption. When we also impose (ii) and (iii), we obtain the desired result. Interestingly, in this case we also find that the only states satisfying all constraints are convex mixtures of product states of the form $\psi\otimes\psi$ and $\psi\otimes\psi^{\perp}$. Figure 2 compares the curves for the two-system toy theory (black) and for the noncontextual 8-state model (blue), illustrating the expected behaviour of the robustified theorem. The region of interest is the top-right corner,[4] and we see that the two theories agree there. This establishes the desired result.

[4] Recall that the desired statement is of the form: if $\psi\otimes\psi$ passes a test with probability $p>X$, then $\psi\otimes\psi^{\perp}$ must pass it with probability $p>Y$. In the SWAP test in quantum theory, however, $\psi\otimes\psi$ passes with probability 1 while $\psi\otimes\psi^{\perp}$ passes with probability 1/2. Therefore, as long as $Y>1/2$, there is a gap between the quantum value and the lowest value achievable by a noncontextual model. Accordingly, the plot is only relevant in the region $Y>1/2$.

Let us conclude by considering an additional requirement:

(v) We impose the ontological equivalence corresponding to the operational equivalence that we call the *Clifford twirl*: $\frac{1}{4}\big(H\rho H + HX\rho HX + HY\rho HY + HZ\rho HZ\big) = \frac{\mathbb{1}}{4}$. This operational equivalence is satisfied by all 24 perfectly entangled stabilizer states in quantum theory, but not in the toy theory, where the Hadamard gate H does not correspond to a valid permutation. As in conditions (iii) and (iv), applying this twirl to ρ yields statistics identical to those of the maximally mixed state.

Imposing requirement (v) in place of (iv), together with (i), (ii), and (iii), yields the same blue curve in Fig. 2, consistent with the desired result.

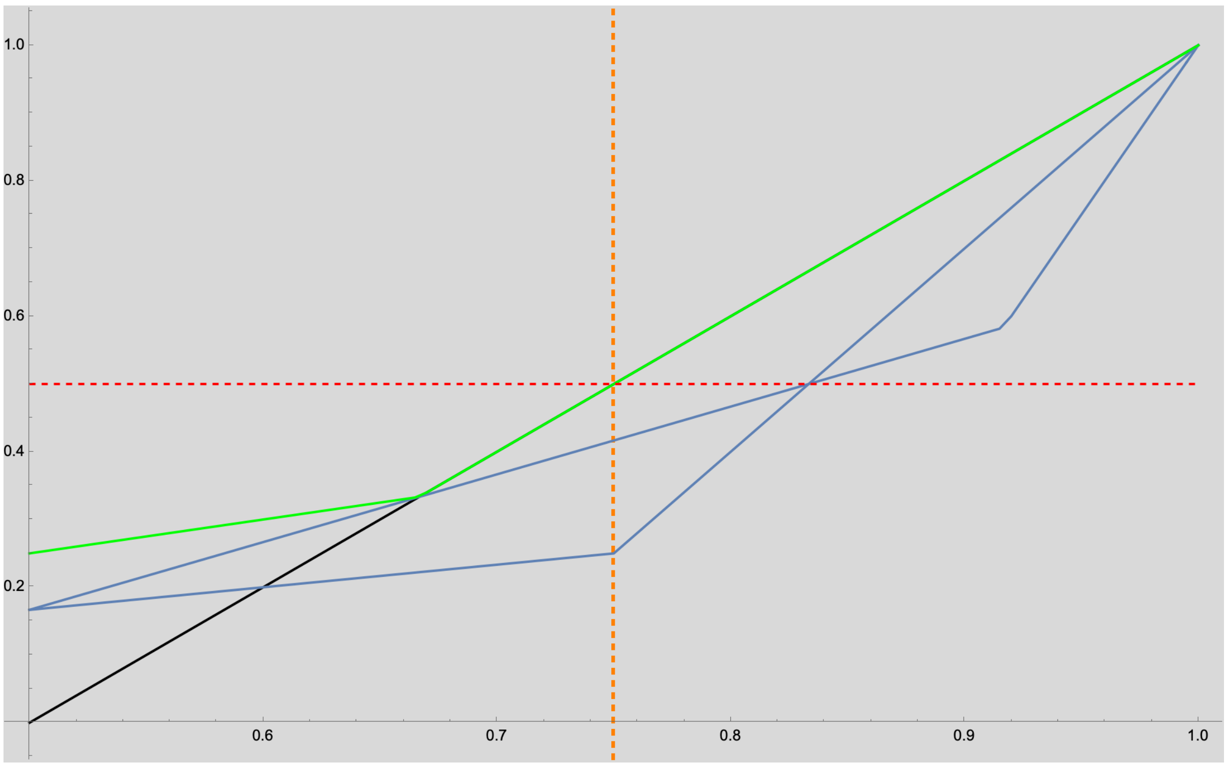

FIG. 1: In the x axis we plot the probability that $\psi \otimes \psi$ passes the test, and in the y axis the probability that $\psi \otimes \phi$ passes the test. The point $(1, 1)$ corresponds to the ideal (exact) version of the result. The black curve is associated with the toy theory. The blue curves arise when we impose only (i), (ii), and (iii), or only (i), (ii), and (iv), respectively. In both cases the desired result still holds, but with weaker robustness than in the toy theory. The green curve arises when we impose (i), (ii), (iii), and (iv), or alternatively (i), (ii), (iii), and (v). Since the result is relevant only in the top-right corner of the plot, we conclude that in this region the green curve agrees with the toy-theory curve (see also Fig. 2).

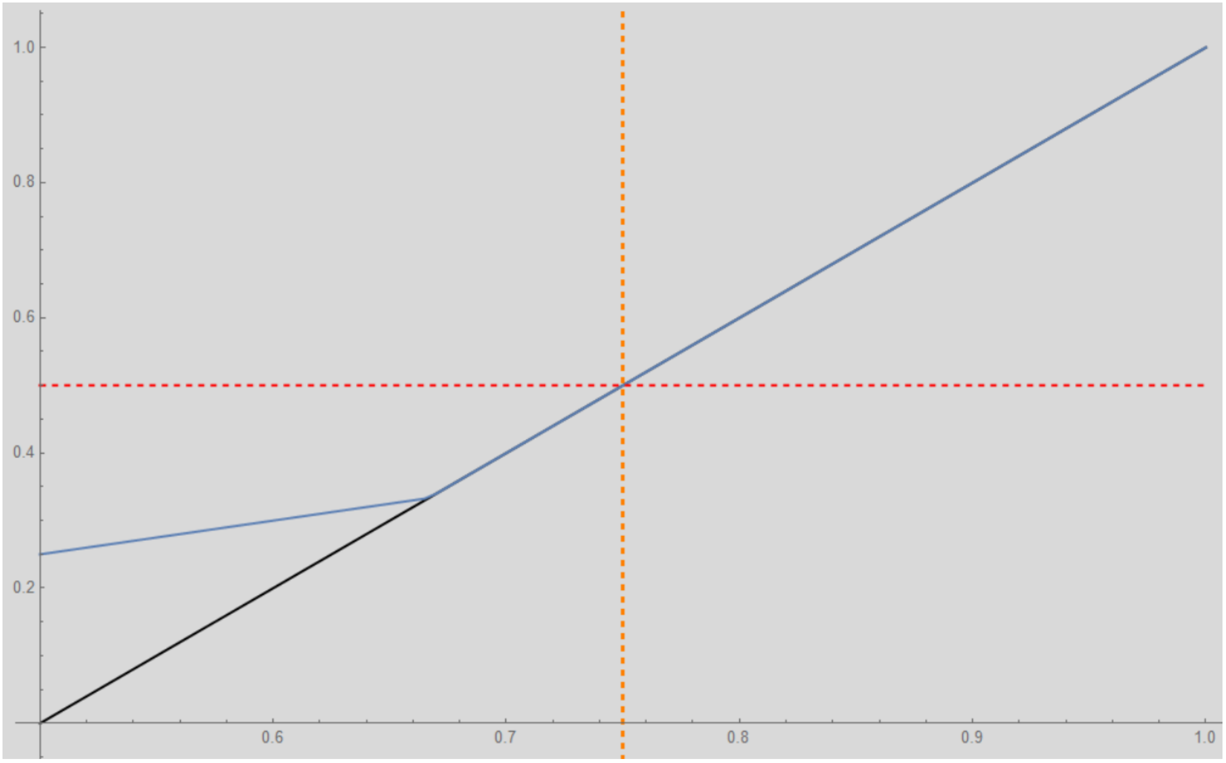

FIG. 2: In the x axis we plot the probability that $\psi \otimes \psi$ passes the test, and in the y axis the probability that $\psi \otimes \phi$ passes the test. The point $(1, 1)$ corresponds to the ideal (exact) version of the result. The black curve is associated with the toy theory. Since we are only interested in the top-right corner of the plot, the toy-theory curve agrees there with the blue curve, which arises when we impose (i), (ii), (iii), and (iv), or alternatively (i), (ii), (iii), and (v).