

## Emergence of Noncontextuality under Quantum Darwinism

Roberto D. Baldijão<sup>1,\*</sup>, Rafael Wagner<sup>2,3,4</sup>, Cristhiano Duarte<sup>5,6,7</sup>, Bárbara Amaral,<sup>3</sup>  
and Marcelo Terra Cunha<sup>8</sup>

<sup>1</sup>*Institute of Physics “Gleb Wataghin”, University of Campinas—UNICAMP, Campinas, São Paulo 13083-859, Brazil*

<sup>2</sup>*International Iberian Nanotechnology Laboratory (INL), Avenida Mestre José Veiga, Braga 4715-330, Portugal*

<sup>3</sup>*Department of Mathematical Physics, Institute of Physics, University of São Paulo, R. do Matão 1371, São Paulo, São Paulo 05508-090, Brazil*

<sup>4</sup>*Centro de Física, Universidade do Minho, Braga 4710-057, Portugal*

<sup>5</sup>*Wigner Research Center for Physics, Budapest H-1121, Hungary*

<sup>6</sup>*School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, United Kingdom*

<sup>7</sup>*International Institute of Physics, Federal University of Rio Grande do Norte, Natal 59070-405, Brazil*

<sup>8</sup>*Instituto de Matemática, Estatística e Computação Científica, Universidade Estadual de Campinas, Cidade Universitária Zeferino Vaz, 651, Sérgio Buarque de Holanda, Campinas, São Paulo 13083059, Brazil*

 (Received 20 April 2021; revised 7 August 2021; accepted 23 August 2021; published 27 September 2021)

Quantum Darwinism proposes that the proliferation of redundant information plays a major role in the emergence of objectivity out of the quantum world. Is this kind of objectivity necessarily classical? We show that if one takes Spekkens’s notion of noncontextuality as the notion of classicality and the approach of Brandão, Piani, and Horodecki to quantum Darwinism, the answer to the above question is “yes,” if the environment encodes the proliferated information sufficiently well. Moreover, we propose a threshold on this encoding, above which one can unambiguously say that classical objectivity has emerged under quantum Darwinism.

DOI: [10.1103/PRXQuantum.2.030351](https://doi.org/10.1103/PRXQuantum.2.030351)

### I. INTRODUCTION

Quantum theory is a normative set of rules the explanatory power of which finds no counterpart in classical theories [1–3]. As a framework adapted to deal with probabilistic descriptions of microscopical phenomena, quantum theory generalizes, and includes, the whole set of classical probability theories [4]. Remarkably, the generalization and explanatory capacity of quantum theory come together with a long list of extremely useful practical applications [5] and philosophical implications that challenge our classical understanding of physical reality [6,7].

Amongst the many challenges that quantum theory presents, two of them still merit attention and a proper connection: first, an unambiguous definition of what is classical and what is quantum; and, second, the emergence, out of a deeper microscopical quantum world, of the objective

reality with which we are familiar. Arguably, an important candidate for the former is *noncontextuality*, as it expresses an important ingredient of the reasoning underlying classical probability theories—operational equivalences reflect and imply ontological equivalences [8–11].

To investigate the emergence of an objective (commonly agreed upon) reality, we consider that the environment plays a crucial role in mediating interactions between observers and the observed system. Within this paradigm, *quantum Darwinism* (QD) is undoubtedly one of the most physically appealing notions, seeking to explain why distinct and independent observers usually obtain the same information regarding the system with which they are interacting [12–14]. In other words, quantum Darwinism helps to explain the emergence of objectivity. Here, we consider the approach to quantum Darwinism due to Brandão, Piani, and Horodecki [15], which is independent of specific aspects of the system-environment interaction and is thus capable of showing the generic emergence of objectivity. Moreover, it encapsulates some essential features of other important processes related to the emergence of objectivity under the quantum realm [12,16,17].

Can we say that the objectivity obtained in Darwinist processes necessarily implies classicality? Considering

\*Corresponding author. [rdaldi@ifi.unicamp.br](mailto:rdaldi@ifi.unicamp.br)

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

solely the original argument of Brandão *et al.*, this is certainly not clear: as we shall see, no restriction is made on the information upon which the observers agree. Moreover, a recent result reinforces the idea that agreement—classical or not—may be an important defining aspect of *quantum* theory itself, making this question even more subtle [18]. In this work, we show that whenever “enough objectivity” emerges due to a Darwinist process, observers can construct noncontextual models explaining the statistics of their experiments. Therefore, the novel aspect we put forward is that noncontextuality emerges out of (sufficiently successful) quantum Darwinism or, in other words, that a precise notion of classicality emerges from this notion of objectivity. Moreover, we discuss how even if quantum Darwinism does not hold, contextuality might still fade out in an *environment-as-a-witness* dynamics, thus being a more sensitive feature. Finally, we propose a quantitative condition for the emergence of *classical objectivity* under quantum Darwinism. Surprisingly, our proofs can also be of interest to quantum state discrimination, as we show that a high success probability on this task can point to a geometric structure of the states being discriminated—they must be vertices of a simplex in the space of density matrices.

## II. BACKGROUND

### A. Environment as a witness and quantum Darwinism

Quantum Darwinism aims to explain how independent observers may obtain the same information about a central quantum system, despite the odd features of quantum theory that defy a classical notion of objectivity. For instance, if quantum theory may be applied to arbitrary systems and a cat’s probability of staying alive or dying are both non-null, why do we all see the same outcome (hopefully alive)? Ultimately, providing an answer to this question explains the emergence of an objective world out of the quantum realm—where the notion of objectivity comes from the common experience of potential independent observers [19]. In this section, we concisely describe the results of Brandão *et al.* [15] that are of importance to this work.

Initially proposed in Ref. [12], quantum Darwinism recognizes the active role of a fragmented environment in the emergence of objectivity, as dictated by the so-called “environment-as-a-witness” paradigm. Broadly speaking, potential observers extract information about the system of interest by *interacting with a portion of its environment*—rather than interacting with the system itself [20].

Let us look into the environment-as-a-witness paradigm in more detail (see Fig. 1). The central system of interest,  $A$ , interacts with its environment,  $\mathcal{E}$ , composed of  $N$  subsystems,  $B_1, \dots, B_N$ . We focus on the information that the environment has about  $A$ , represented by a completely positive and trace-preserving (CPTP) map  $\Phi : \mathcal{D}(\mathcal{H}_A) \rightarrow$

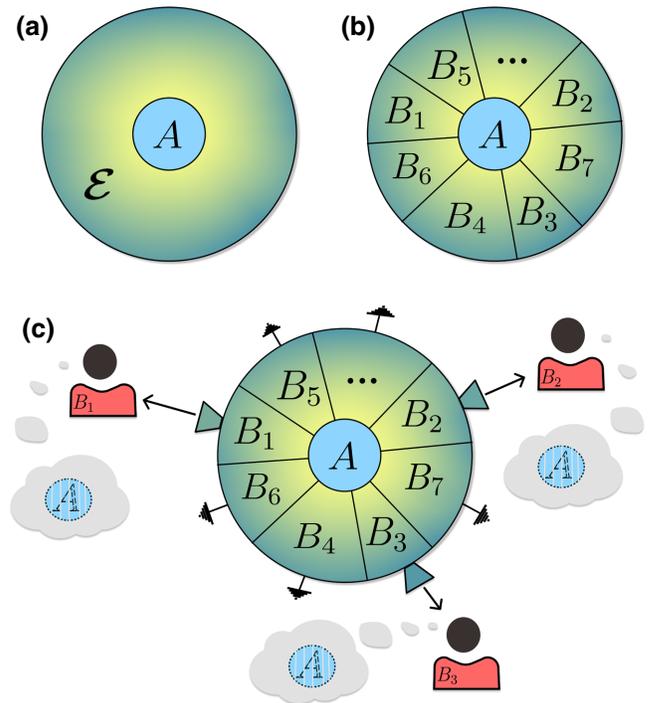


FIG. 1. Different approaches to open systems. (a) The decoherence paradigm. The focus is on the central system  $A$ . The environment is treated as a monolithic system, usually to be traced out after the interaction. (b) The environment split into fractions. The environment is more accurately described as a composition of several subsystems. (c) Environment-as-a-witness dynamics and the quantum Darwinism process. Under this paradigm, the environment is promoted to an information channel, spreading redundant information about system  $A$  to independent observers. In the example of the figure, the number of portions in  $S_t$  is  $t = 3$  and  $B_{S_t} = \{B_1, B_2, B_3\}$ . Theorem 4 provides us with a sufficient condition for such redundant information to be noncontextual, i.e., classical.

$\mathcal{D}(\mathcal{H}_\mathcal{E})$ , where  $\mathcal{D}(\mathcal{H}_A)$  represents the density matrices acting on the Hilbert space  $\mathcal{H}_A$  of the central system and analogously for  $\mathcal{D}(\mathcal{H}_\mathcal{E})$ . Now, the actual dynamics of interest, which we call  $\text{EW}_t$  dynamics (short for environment-as-a-witness  $t$  dynamics), consider only a portion of such an environment and are defined as follows.

**Definition 1** ( $\text{EW}_t$  dynamics). *Consider a dynamics  $\Phi : \mathcal{D}(\mathcal{H}_A) \rightarrow \mathcal{D}(\mathcal{H}_\mathcal{E})$ , from a central system to a multipartite environment,  $A$  and  $\mathcal{E}$ , respectively. Let  $S_t \subset \{1, \dots, N\}$  be a set of labels describing  $t$  portions of the environment  $\mathcal{E}$ . An **EW $_t$  dynamics** for the subset  $B_{S_t} := \{B_j\}_{j \in S_t}$  is the CPTP map  $\Phi^{B_{S_t}} := \text{Tr}_{\mathcal{E} \setminus B_{S_t}} \circ \Phi$ .*

Tracing out  $\mathcal{E} \setminus B_{S_t}$ , an  $\text{EW}_t$  dynamics selects exactly those fractions of the environment on which one wants to center attention. Brandão *et al.* [15] show that this dynamics already implies a certain kind of objectivity,

called objectivity of observables, in which a pointer positive operator-valued measure (POVM) [21] is selected by the interaction.

**Theorem 1** (Adapted Theorem 2 of Ref. [15], objectivity of observables). *Consider an  $EW_t$  dynamics  $\Phi^{B_{S_t}}$ . If the total environment is big enough compared to  $S_t$  ( $N \gg t$ ), then there exists a POVM  $\{\tilde{E}_k\}$  acting on  $\mathcal{D}(\mathcal{H}_A)$  such that, for most choices of  $S_t$ ,*

$$\Phi^{B_{S_t}}(\rho^A) \approx \sum_k \text{Tr}[\tilde{E}_k \rho^A] \sigma_k^{B_{S_t}}, \quad (1)$$

where  $\sigma_k^{B_{S_t}} \in \mathcal{D}(\bigotimes_{j \in S_t} \mathcal{H}_{B_j})$  and  $\{\tilde{E}_k\}_k$  is independent of  $B_{S_t}$ .

A consequence of this theorem is that, when its conditions are met, the environmental state  $\Phi^{B_{S_t}}(\rho^A)$  approximately encodes information regarding the probability distribution  $\tilde{p}_k := \text{Tr}[\tilde{E}_k \rho^A]$ . (Note, however, that nothing is said on how good this encoding is.) The most important feature of this result is that the POVM  $\{\tilde{E}_k\}_k$  does not depend on the elements of  $B_{S_t}$ . The emergence of an objective reality comes from this fact, i.e., that the observable (possibly) monitored by the environment is essentially independent of which portion of the  $t$  subsystems of the environment one ends up with.

From now on, we only consider  $S_t$  to be portions of the environment such that the approximation in Eq. (1) is valid. Moreover, we assume an infinite-sized environment, as is usual in open-quantum-systems approaches, such that the approximation can be substituted by an equality (for details, and for a relaxation of this assumption, see Appendix A). Without loss of generality, we also assume that each  $\tilde{E}_k \neq 0$ .

With these assumptions, we see that any  $EW_t$  dynamics,  $\Phi^{B_{S_t}}$ , is effectively a measure-and-prepare map, defined by the pairs  $(\tilde{E}_k, \sigma_k^{B_{S_t}})_k$ , where  $\{\tilde{E}_k\}_k$  forms a POVM acting on the central system with non-null elements and  $\{\sigma_k^{B_{S_t}}\}_k$  are states of  $B_{S_t}$  (as such, these states may depend on the portion of the environment on which we are focusing).

The objectivity of observables obtained through  $EW_t$  dynamics tells us that different observers might gather information regarding the same observable,  $\{\tilde{E}_k\}_k$ , (almost) independently of the choice of  $B_{S_t} \subset \mathcal{E}$ . However, objectivity of observables does not solve adequately the emergence of an objective reality, as a crucial feature of the chosen notion of objectivity is that independent observers should agree on which *outcome* they see. In other words, they should all be able to infer the same label  $k$  upon measuring their share of the environment—and this is not a consequence of objectivity of observables. Brandão *et al.* [15] tell us that this happens when the full Darwinist process occurs, i.e., when information about  $\{\tilde{E}_k\}_k$  is indeed well encoded on the environmental subsystems.

Let us make the above claim more precise. Consider a setting with several observers (call them all Bob). For simplicity, assume that we have  $t$  Bobs, each receiving one subsystem of  $B_{S_t}$ , and let us label each Bob by the element  $B_j \in B_{S_t}$ . Each Bob will be the end of a reduced dynamics,  $\Phi^{B_j} := \text{Tr}_{B_{S_t} \setminus B_j} \circ \Phi^{B_{S_t}}$ , so Bob  $B_j$  will hold a state of the form  $\sum_k \tilde{p}_k \sigma_k^{B_j}$ , where  $\sigma_k^{B_j} = \text{Tr}_{B_{S_t} \setminus B_j}[\sigma_k^{B_{S_t}}]$ . Now, the hypotheses that the different subsystems encode well enough the information regarding the selected observable, as presumed by quantum Darwinism [22], introduces an important requirement into the landscape of Ref. [15]: sufficient distinguishability of the states  $(\sigma_k^{B_j})_k$  for each  $B_j$ . We thus define the following.

**Definition 2** (QD $_\eta$  process). *Consider an  $EW_t$  dynamics  $\Phi^{B_{S_t}} := (\tilde{E}_k, \sigma_k^{B_{S_t}})_k$  and the  $t$  partial traces  $\Phi^{B_j}$ . Define the quantity*

$$p_{\text{guess}}[(\tilde{p}_k, \sigma_k^{B_j})_k] := \max_{\{F_k^{B_j}\}_k} \sum_k \tilde{p}_k \text{Tr}[F_k^{B_j} \sigma_k^{B_j}], \quad (2)$$

where  $\tilde{p}_k = \text{Tr}[\tilde{E}_k \rho^A]$ ,  $\sigma_k^{B_j} = \text{Tr}_{B_{S_t} \setminus B_j}[\sigma_k^{B_{S_t}}]$  and  $\{F_k^{B_j}\}_k$  are local POVMs for each Bob. A **quantum Darwinism process with distinguishability**  $\eta$ , QD $_\eta$ , is said to occur when there exists an  $0 \leq \eta \leq 1$  such that, for all  $B_j \in B_{S_t}$ ,

$$\min_{\rho^A} p_{\text{guess}}[(\tilde{p}_k, \sigma_k^{B_j})_k] \geq \eta. \quad (3)$$

The quantity  $p_{\text{guess}}$  tells us how well one can discriminate between states  $\{\sigma_k^{B_j}\}_k$ , depending on  $\rho^A$  (recall that  $\tilde{p}_k$  is a function of  $\rho^A$ ). Thus, QD $_\eta$  sets a lower bound on the distinguishability of the  $\{\sigma_k^{B_j}\}_k$ , to all  $B_j$  and regardless of the initial state of the central system. From now on, when we write QD $_\eta$ , we assume that  $\eta$  is the maximal value for which Inequality (3) holds.

Finally, Brandão *et al.* [15] tell us that a Darwinist process essentially leads to all Bobs seeing the same outcome.

**Proposition 1** (Adapted Proposition 3 of Ref. [15], objectivity of outcomes). *Consider the probability that each Bob, by performing a local measurement  $\{F_k^{B_j}\}_k$ , gets the same outcome  $k$ , provided that the encoded label is indeed  $k$ , i.e.,*

$$\min_{\rho^A} \sum_k \tilde{p}_k \text{Tr} \left[ \bigotimes_{j \in S_t} F_k^{B_j} \sigma_k^{S_t} \right]. \quad (4)$$

For any  $\text{QD}_\eta$  process, there exist local POVMs  $\{\tilde{F}_k^{B_j}\}_k$  for each Bob such that

$$\min_{\rho^A} \sum_k \tilde{D}_k \text{Tr} \left[ \bigotimes_{j \in S_k} \tilde{F}_k^{B_j} \sigma^{S_k} \right] \geq 1 - 6t\delta^{1/4}, \quad (5)$$

where  $\delta = 1 - \eta$ .

In other words, the better the distribution  $\text{Tr}[\tilde{E}_k \rho^A]$  is encoded on the states  $\{\sigma_k^{B_j}\}$ , the higher is the probability that all Bobs will see the same outcome (provided that they implement the correct measurements). Note that whether or not the Bobs will see the same outcome is not the key aspect here; rather, what is important is that if they try to obtain the information correctly, they will agree. Roughly speaking, this means that the higher the  $\eta$  in a  $\text{QD}_\eta$  process, the higher is the “amount” of emergent objectivity. We thus have a continuous emergence of objectivity under this approach to quantum Darwinism. In the (often considered) limit  $\eta \rightarrow 1$  of perfect encoding, nearly perfect objectivity is obtained. Whether this condition of “high  $\eta$ ” is generic or not among physically motivated interactions is still an open question. Even though it is a natural assumption to take in the environment-as-a-witness paradigm, it is not known whether generic interactions will reach high distinguishability or not. Since this is vital to the Darwinism program, we hope that future work will provide such an answer.

The emergence of objectivity under quantum Darwinism in the framework of Brandão *et al.* is a remarkable result. It comes, nevertheless, with a small drawback, as there are no restrictions on the nature of  $\{\tilde{E}_k\}_k$  [15]—in principle,  $\{\tilde{E}_k\}_k$  can even represent an informationally complete POVM. This is incompatible with the idea of a “classical observable” selected by the interaction, as usually thought of within the Darwinist program [22]. The possibility of having informationally complete POVMs spreading over the environment means that essentially *all information* (including *quantum* information) contained in  $\rho^A$  can be stored in the environment.

Can we still say that some sort of classicality arises in those cases? Here, we argue that, even if we allow for  $\{\tilde{E}_k\}_k$  to be an informationally complete POVM, there is still a precise notion of the emergence of classicality in some  $\text{QD}_\eta$  processes: all the Bobs can construct a noncontextual ontological model for the statistics of their measurements on every  $\Phi^{B_j}(\rho^A)$ .

## B. Spekkens’ contextuality

The conception of classicality that we adopt is Spekkens’s notion of noncontextuality [10] for prepare-and-measure scenarios. This notion captures essential features of classical theories and thus encapsulates several

criteria associated with classicality in different research areas. For instance, Spekkens’s noncontextuality is equivalent to the existence of positive quasiprobability distributions [9], which is a widely used classicality criteria in quantum optics [23,24]. In the framework of generalized probabilistic theories [25,26], in which physical principles can be studied without resorting to the quantum or classical formalisms, a noncontextual theory must be embedded in classical probability theory [27]. Moreover, Spekkens’s noncontextuality also encapsulates Gaussian quantum mechanics [28] and subsumes the well-known notions of Kochen-Specker noncontextuality [29–32] and local causality [33]. In this section, we briefly describe this idea of classicality.

Spekkens’ framework is operational and allows for a representation of laboratory procedures—a list  $\mathcal{P}$  of preparations, a list  $\mathcal{M}$  of measurements, and a set of effects  $\{b|M\}_b$  for each measurement  $M$  in  $\mathcal{M}$ . In particular, it is assumed that convex combinations of preparations in  $\mathcal{P}$  also result in valid preparations. In general, there will be equivalences among procedures of the same type: a preparation  $P_1 \in \mathcal{P}$  is equivalent to  $P_2 \in \mathcal{P}$  in a prepare-and-measure scenario when, for all *conceivable* measurement procedures,  $p(b|M, P_1) = p(b|M, P_2)$ . That is, one cannot distinguish those preparation procedures from the probabilities arising on *any* measurement—or, equivalently, for a subset of measurements called tomographically complete. The above relation defines equivalence classes of preparations, each class denoted by  $[P]$ . For instance, in quantum theory, a density matrix could label an equivalence class: all the alternative ways to produce a certain  $\rho \in \mathcal{D}(\mathcal{H})$  will always lead to the same statistics. Similarly, we can define the equivalences for events of measurements: they are equivalent when no *conceivable* preparation leads to different probabilities. In quantum theory, these can be labeled by the elements of POVMs. We denote the sets of all equivalences of each type of procedure as  $\text{Equiv}\{\mathcal{P}\}$  and  $\text{Equiv}\{\mathcal{M}\}$ .

As an attempt to classically explain how the probabilities arise within the operational theory, one might try to construct an ontological model. Ontological models in prepare-and-measure scenarios comprise of three ingredients: a measurable space  $(\Lambda, \Sigma)$ ; a linear map  $\mu : \mathcal{P} \mapsto (\mu_P)_{P \in \mathcal{P}}$ , where  $\mu_P(\lambda)$  is a probability measure over  $\Lambda$ ; and another linear map  $\xi : (b|M) \mapsto \xi_M(b|\lambda)$ , such that each measurement  $M \in \mathcal{M}$  is mapped to a set of response functions  $\{\xi_M(b|\lambda)\}_b$  satisfying  $\xi_M(b|\lambda) \geq 0$ ,  $\sum_b \xi_M(b|\lambda) = 1 \forall \lambda$ . Linearity of the maps defined above implies that convex mixtures of procedures should be mapped to the appropriate convex mixtures of the corresponding ontic objects. Finally, the probabilities of the operational theory must be reproduced by ontological models via  $p(b|M, P) = \sum_\lambda \mu_P(\lambda) \xi_M(b|\lambda)$  [7]. Up to here, any operational theory will accept explanations via some ontological model. However, the important aspect is

whether it is possible to recover operational probabilities via a *noncontextual ontological model* [10].

**Definition 3** (Noncontextual ontological models). *Consider an operational theory with procedures  $\mathcal{P}$ ,  $\mathcal{M}$  and equivalences  $\text{Equiv}\{\mathcal{P}\}, \text{Equiv}\{\mathcal{M}\}$ . An ontological model recovering the probabilities of the operational theory is said to be **preparation noncontextual** if*

$$P \in [P'] \implies \mu_P = \mu_{P'}. \quad (6)$$

*Analogously, the model is said to be **measurement noncontextual** if*

$$b|M \in [b'|M'] \implies \xi_M(b|\lambda) = \xi_{M'}(b'|\lambda) \quad \forall \lambda \in \Lambda. \quad (7)$$

*If both conditions are satisfied, the ontological model is said to be **universally noncontextual**, or noncontextual in brief.*

This definition says that procedures that are operationally indistinguishable should be represented by the same entities on the ontological model. The essence of this definition of noncontextuality is the so-called principle of the identity of indiscernibles, due to Leibniz [8, 10]. Roughly, it says that two empirically indistinguishable things should be ontologically described as *identical*. The methodology implied by (generalizations of) this principle has been important in building classical theories in the past; for instance, it has played an important role in the birth of relativity [8].

Even though noncontextuality is a natural feature to demand of physical theories, quantum theory is known to be a “contextual” operational theory: there is no noncontextual ontological model reproducing all of its statistics [10, 34]. This remarkable nonclassical feature of quantum theory has important practical consequences: contextuality is a useful resource, as it provides genuine quantum advantage over classical computation. Indeed, contextuality is the crucial resource in information processing [35–40], and it provides advantages for quantum state discrimination [41], quantum cloning [42], quantum metrology [43], and quantum machine learning [44]. It also constitutes a necessary resource for models of restricted quantum computation [45–48], capable of providing robust numerical bounds of advantage [49], not only for deterministic but also for probabilistic restricted quantum computation.

In summary, Spekkens’ notion of noncontextuality is a mathematically well-defined and amply used notion of classicality. In this sense, it is fair to assume that the emergence of generalized noncontextuality in an intrinsically contextual theory can be naturally understood as a classical limit—with a very precise meaning. In what follows, we show a deep connection between classicality in these terms and quantum Darwinism objectivity.

### III. RESULTS

#### A. Emergence of noncontextuality

To approach the question of the emergence of noncontextuality, as a proxy for the emergence of classicality in the considered dynamics, we need to describe the scenario that we care about—we need to say what are the procedures and operational equivalences, thus establishing what a noncontextual model must respect. As demanded by the environment-as-a-witness paradigm, we focus on Bob’s perspectives: *to each Bob*, we must say what the different preparations, measurements, and equivalences are. Measurement procedures are straightforward: each Bob implements a measurement procedure  $M^{B_j} \in \mathcal{M}^{B_j}$  on the system that he receives. Therefore, the equivalence classes are identified with the elements of POVM at each Bob’s side.

A more subtle subject in our case is defining preparations and equivalences thereof. Suppose, for the sake of simplicity, that there is an experimenter (say, Alice) preparing the quantum states of system  $A$ , according to a set of possible preparations  $\mathcal{P}$ ; then, the equivalence class of  $P \in \mathcal{P}$  would be  $[\rho_P^A]$  (whether there *really is* an experimenter is not of importance to this work—the procedures can be understood as any specific sequence of events “preparing” system  $A$ ). However, each Bob will *infer* information about system  $A$  by making measurements on his share of the environment of  $A$ . Therefore, from the perspective of Bob  $B_j$ , every preparation of system  $A$  is always followed by a fixed transformation  $T_j$ , which we can describe as: interact  $A$  with a large number of systems in a predefined way and discard all subsystems except  $B_j \in B_{S_t}$ . Thus, the preparations arriving at  $B_j$  are effectively  $\mathcal{P}'_j := T_j(\mathcal{P})$  (see Fig. 2). From Bob’s standpoint, the equivalence classes for preparations must then be defined by the density matrix available for Bob, i.e.,  $\Phi^{B_j}(\rho_P^A)$ . In sum, the scenario of each  $B_j$  is composed of preparations  $\mathcal{P}'_j$  with equivalences  $[\Phi^{B_j}(\rho^A)]$ ,

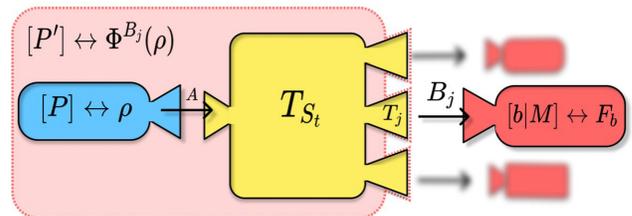


FIG. 2. The scenario from the perspective of Bob  $B_j$ : preparation  $P \in \mathcal{P}$  of the central system is followed by a fixed transformation (i.e., interaction with the environment, always prepared in the same way), leading to an effective preparation  $P' = T_j(\mathcal{P})$ . The nontraced-out portion of the environment is given by the labels  $S_t$  (here,  $t = 3$ ). For Bob,  $[P'] \leftrightarrow \Phi^{B_j}(\rho^A)$  and  $[b|M] \leftrightarrow F_b$ .

while measurement procedures are the local implementations  $\mathcal{M}^{B_j}$ , with equivalence classes for effects represented by elements of POVMs.

Now, we have set the stage to present our results. Below, we show a mathematical condition on the encoding states  $\{\sigma_k^{B_j}\}_k$  of an  $\text{EW}_t$  dynamics that, if satisfied for some  $B_j \in S_t$ , ensures that there exists a noncontextual ontological model recovering the statistics for those Bobs.

**Lemma 2** (Affinely independent  $\{\sigma_k^{B_j}\}_k$  implies noncontextuality). *Consider an  $\text{EW}_t$  dynamics, with an infinite environment, defined by the pairs  $(\tilde{E}_k, \sigma_k^{S_t})_k$ . If, for a given  $B_j \in B_{S_t}$ , the environmental systems  $\{\sigma_k^{B_j}\}_k$  form an affinely independent set of matrices on  $\mathcal{D}(\mathcal{H}_{B_j})$ , then there exists a noncontextual ontological model for the distribution about  $A$  available to  $B_j$ . This holds true for any measurement  $\{F_b^{B_j}\}$  implemented by Bob  $B_j$ .*

*Proof.* We prove this result by constructing such a noncontextual ontological model. As mentioned above, the equivalence classes for the “preparations” arriving at  $B_j$  are defined by the quantum states  $\sigma_{P'}^{B_j} = \sum \text{Tr}[\tilde{E}_k \rho_P^A] \sigma_k^{B_j}$ . Affine independence of  $\{\sigma_k^{B_j}\}_k$  implies the bijection

$$\sigma_{P'_1}^{B_j} = \sigma_{P'_2}^{B_j} \iff \text{Tr}[\tilde{E}_k \rho_{P'_1}^A] = \text{Tr}[\tilde{E}_k \rho_{P'_2}^A] \quad \forall k. \quad (8)$$

(Such a bijection can also be seen by the fact that affine independence of  $\{\sigma_k^{B_j}\}_k$  implies that  $\text{ConvHull}[\{\sigma_k^{B_j}\}_k]$  is a  $(k_{\max} - 1)$ -simplex on  $\mathcal{D}(\mathcal{H}_{B_j})$ ; thus, every point has a unique convex decomposition in terms of the  $k_{\max}$  vertices). Therefore, the equivalences can be ultimately labeled by the distribution  $(\text{Tr}[\tilde{E}_k \rho_P^A])_k$ . Now, consider the following ontological model:

$$\Lambda := \{k\}_k; \quad (9a)$$

$$\mu_{P'}(k) := \text{Tr}[\tilde{E}_k \rho_P^A]; \quad (9b)$$

$$\xi_M(b|k) := \text{Tr}[F_b^{B_j} \sigma_k]. \quad (9c)$$

This indeed respects the equivalence classes of the scenario of  $B_j$ :  $\mu_{P'}$  depends uniquely on the distribution  $\text{Tr}[\tilde{E}_k \rho_P^A]$ , which labels each class  $[P']$ , while  $\xi_M$  depends only on  $F_b^{B_j}$ , the label of each  $[b|M]$ . It remains to show that this ontological model recovers the statistics. Since the probabilities are obtained via a measure-and-prepare channel defined by  $(\tilde{E}_k, \sigma_k)_k$ , we have  $p(b|M, P') = \sum_k \text{Tr}[\tilde{E}_k \rho_P^A] \text{Tr}[F_b \sigma_k] = \sum_k \mu_{[P']}(k) \xi_{[M]}(b|k)$ . ■

**Remark.** *The above proof has interesting connections to characterizations of noncontextuality in the framework of generalized probabilistic theories [27,50]. In Appendix C,*

*we make this connection more explicit for the interested reader.*

Lemma 2 shows how noncontextuality may *emerge* in an  $\text{EW}_t$  dynamics: generally, there is no noncontextual model for the scenario with full quantum probabilities, with all possible  $\mathcal{P}$  and their equivalences  $\text{Equiv}\{\mathcal{P}\}$ ; however, the  $\text{EW}_t$  dynamics maps  $[\rho^A] \mapsto [\Phi^{B_j}(\rho^A)]$ , enabling a noncontextual explanation. Nonetheless, the condition of affine independence of  $\{\sigma_k\}_k$  is not clearly motivated within quantum Darwinism. In other words, up to this point, Lemma 2 is merely a mathematical condition signaling the existence of noncontextual models within  $\text{EW}_t$  dynamics. Therefore, a connection between emergence of noncontextuality and emergence of objectivity under quantum Darwinism remains open.

By looking at the extreme case in which  $\{p_{\text{guess}}[(\sigma_k^{B_j}, \tilde{p}_k)_k] = 1\}$  for some Bob, we find an indication of such a connection: affine independence of  $\{\sigma_k^{B_j}\}_k$  is necessary for perfect encoding of the distribution  $(\text{Tr}[\tilde{E}_k \rho^A])_k$ . Indeed, if the distribution  $(\text{Tr}[\tilde{E}_k \rho^A])_k$  is perfectly encoded in the states  $\{\sigma_k^{B_j}\}_k$ , these states must be deterministically distinguishable; in turn, these states must have orthogonal support, thus being affinely independent. Moving away from the extreme case, we can obtain affine independence even if the discrimination is not perfect but is good enough. This is summarized in the following lemma—its proof deferred to Appendix B 1. Remarkably, this proof can be of interest to quantum state discrimination: if the probability of success is high enough, we obtain information about the geometrical disposition of  $\{\sigma_k^{B_j}\}_k$ , namely, that they form vertices of a  $(k_{\max} - 1)$ -simplex in  $\mathcal{D}(\mathcal{H}_{B_j})$ .

**Lemma 3** (Operational signature implying affine independence). *Consider an  $\text{EW}_t$  dynamics defined by  $(\tilde{E}_k, \sigma_k^{B_j})_k$ . Then, there exists a bound  $\hat{P}[(\tilde{E}_k)_k]$  such that, if  $p_{\text{guess}}[(\tilde{p}_k, \sigma_k^{B_j})] > \hat{P}[(\tilde{E}_k)_k] \forall \rho^A$  (recall that  $\tilde{p}_k$  depends on the state  $\rho^A$  of the central system), the set of states  $\{\sigma_k^{B_j}\}_k$  must be affinely independent.*

**Remark.** *For the condition above to be valid, it is necessary that  $\dim(\mathcal{H}_{B_j}) \geq k_{\max}$ .*

In Appendix B 1, we construct the bound  $\hat{P}$ , which is generally not trivial, i.e., there are cases where  $\hat{P} < 1$ . Therefore, Lemma 3 tells us that if Bob  $B_j$  can recover well enough the statistics of  $(\text{Tr}[\tilde{E}_k \rho^A])_k$  through measurements on his system (i.e.,  $p_{\text{guess}} > \hat{P}$  for that Bob),  $\{\sigma_k^{B_j}\}$  are affinely independent. Thus, by Lemma 2, he can construct a noncontextual model for such statistics. Now, if such good distinguishability also holds for every Bob  $B_j \in S_t$ , a process of  $\text{QD}_\eta$  occurs, with  $\eta > \hat{P}$ . In this case, while

the results of Brandão *et al.* [15] put a lower bound on the probability that all Bobs will see the same outcome when making the correct measurements, the above results imply that *every Bob can explain his results using a noncontextual model*. This is the summary of the following theorem, our main result.

**Theorem 4** (Main result: emergence of noncontextuality under  $\text{QD}_\eta$ ). *Suppose that the conditions for  $\text{QD}_\eta$  are met with  $\eta > \hat{P}[(\tilde{E}_k)_k]$ . Then, each Bob can construct a noncontextual ontological model for each respective scenario.*

*Proof.* If the conditions are met, it must hold that

$$\min_{\rho} p_{\text{guess}}[(\sigma_k^{B_j})_k, (\text{Tr}[\tilde{E}_k \rho])_k] > \hat{P}[(\tilde{E}_k)_k] \quad \forall B_j \in S_t, \quad (10)$$

which, in turn, implies that

$$p_{\text{guess}}[(\sigma_k^{B_j})_k, (\text{Tr}[\tilde{E}_k \rho])_k] > \hat{P}[(\tilde{E}_k)_k] \quad \forall \rho^A \in \mathcal{D}(\mathcal{H}_A) \quad \text{and} \quad B_j \in S_t. \quad (11)$$

Therefore, by Lemma 3 and Eq. (10), the states  $\{\sigma_k^{B_j}\}$  must be affinely independent, for every  $B_j \in B_{S_t}$ . Then, Lemma 2 guarantees that every Bob will be able to construct a noncontextual ontological model for his scenario. ■

### B. A cutoff for classical objectivity in QD

The emergence of objectivity in  $\text{QD}_\eta$  processes has a continuous form: the closer  $\eta$  is to 1, the “more objectivity” arises due to Darwinism; only in the idealized limit  $\eta = 1$  is “perfect objectivity” obtained. This leads to problems, as there is no cutoff on  $\eta$  defining whether or not objectivity has emerged—and all we can do without a cutoff is to say that some process may lead to more objectivity than some other. That is, whenever  $\eta \neq 1$ , one may not be able to say, safely, that “objectivity has emerged.” If we decide to choose an arbitrary bound for  $\eta$  close enough to 1, above which we are confident to say that “objectivity has emerged,” we also face some issues.

First, the definition of objectivity itself becomes subjective, since what one considers “close enough to 1” is certainly not objective. Second, and most importantly, we want any cutoff signaling objectivity to ensure emergence of noncontextuality as well, as this provides a well-motivated and broad notion of classicality. Our results tell us that sufficiently high values of  $\eta$  ensure that this is true—but how high  $\eta$  should be depends on the monitored observable  $\{E_k\}_k$ . In fact, for any arbitrary fixed bound  $\eta$  one chooses, there is always a central system and specific dynamics in which  $\eta < \hat{P}[(\tilde{E}_k)_k]$  (for an example

of this behavior on complete decoherence dynamics, see Appendix B 2). In these cases, we have no guarantee that noncontextuality emerges, even though  $\eta$  would be “close enough to 1” according to an arbitrary choice.

To avoid those problems, and in view of our results, we propose  $\hat{P}$  to be such a cutoff for classical objectivity. This proposal gives a bound to talk about objectivity that is neither subjective nor signals objectivity without noncontextuality.

**Definition 4** (Emergence of classical objectivity under  $\text{QD}_\eta$ ). *Consider a process  $\text{QD}_\eta$ , with observable  $\{\tilde{E}_k\}_k$ . We say that **classical objectivity** has emerged under this process if  $\eta > \hat{P}[(\tilde{E}_k)_k]$ .*

As said above, setting a nonsubjective bound for objectivity is crucial: we are now able to say whether or not classical objectivity has emerged in  $\text{QD}_\eta$  processes depending on  $\eta$ . By taking  $\hat{P}$  as such a cutoff, we are sure that noncontextuality has emerged and the bound *depends on the monitored observable*,  $\{\tilde{E}_k\}_k$ . This makes sense, as insensitivity to the specifics of the interaction is rather artificial. In other words, central systems with different dimensions and different dynamics, leading to selection of different pointer observables, *should* impact the decision on whether or not classical objectivity has arisen. This is also the case in other approaches to quantum Darwinism, where mutual information between the system and a fragment of the environment is compared to the von Neumann entropy of the system, which depends on the initial state of  $A$  and on the dynamics [12].

Moreover, since  $\eta > \hat{P}$  is not a necessary condition for noncontextuality to have emerged but a sufficient condition obeyed with sufficiently high  $p_{\text{guess}}$  for each Bob, it sets a good threshold that will not be artificially close to zero—this could happen if we only asked for emergence of noncontextuality, as affinely independent states could be almost indistinguishable. Indeed, the perfect decoherence example shows that  $\hat{P}$  can be fairly close to 1, setting a high lower bound on the probability that all Bobs will see the same outcome when making the correct measurements. Therefore, the above definition proposes a new and physically well motivated way to treat the emergence of Darwinist objectivity within the framework of Brandão *et al.* [15].

### C. The case of state-spectrum broadcasting

Before we conclude, let us mention some interesting byproduct of Theorem 4: noncontextuality necessarily emerges under the process of state-spectrum broadcasting [16]. This process has been proposed as an alternative way to reach objectivity from the interaction of quantum systems. In such a process, the central system suffers full decoherence and its spectrum is perfectly broadcast to

the environment. As we show in Appendix B, it can be described as a special form of the  $\text{QD}_\eta$  process adapted to deal with a postinteraction state of the central system  $A$  (for more details, see Appendix B 2). It demands full distinguishability, so we have  $\eta = 1$  and full decoherence implies  $\tilde{E}_k = |k\rangle\langle k|$ , with  $\{|k\rangle\langle k|\}_k$  a basis of  $\mathcal{H}_A$ . This form of  $\tilde{E}_k$  leads to a  $\hat{P} < 1 = \eta$ . By applying Theorem 4, we obtain the following.

**Corollary 5** (Emergence of noncontextuality under state-spectrum broadcasting). *If the  $\text{EW}_t$  dynamics leads to the occurrence of a state-spectrum broadcasting process for arbitrary initial states  $\rho^A$ , all the Bobs can construct a noncontextual ontological model to their statistics.*

*Proof.* As we mentioned above, if a state-spectrum broadcasting process occurs, we have  $\eta = 1$  and  $\tilde{E}_k$  is a projection into a basis of  $\mathcal{H}_A$ . We prove in Appendix B that, for such a kind of observable  $\tilde{E}_k$ , the bound  $\hat{P}$  gives  $\hat{P}[\{|k\rangle\langle k|\}_k] = 1 - 1/2\dim(\mathcal{H}_A)$ . Since this is smaller than  $\eta = 1$ , the corollary follows. For more details, see Appendix B 2. ■

#### IV. CONCLUSION

This paper shows that a particular expression of nonclassicality may emerge from quantum Darwinism processes. More precisely, we have proved that noncontextuality is implied for every observer in  $B_{S_i}$  in quantum Darwinism processes  $\text{QD}_\eta$  with an appropriate distinguishability rate  $\eta > \hat{P}$ , the cutoff  $\hat{P}$  depending only on the monitored POVM  $\{\tilde{E}_k\}_k$ .

This result is crucially important, as it shows that  $\text{QD}_\eta$  processes with high  $\eta$ , besides leading to objectivity (i.e., a high probability of independent observers seeing the same outcome when implementing the adequate measurements), also implies that each of them can adopt noncontextual explanations for their aggregated statistics. Moreover, they can construct noncontextual models with some aspects in common: the ontic space  $\Lambda$  and distributions  $\mu_{P'}$ , as defined in Eq. (9), are independent of  $B_j$ . Such noncontextual explanations may provide a classical view underpinning the probabilities that each Bob sees, obviating the need for any nonclassical (in particular, quantum) description of the central system. Interestingly, these conclusions also hold true for state-spectrum broadcasting, an alternative process for the emergence of objectivity in the quantum realm.

Another important byproduct of our results is the fact that one can use  $\hat{P}$  as a cutoff for the emergence of classical objectivity in  $\text{QD}_\eta$  processes. This brings a new perspective to the Darwinist objectivity proposed by Brandão *et al.* [15], in which the specifics of the ongoing dynamics set a bound for  $\eta$ , above which objectivity and noncontextuality

have emerged. Indeed, Definition 4 suggests that the emergence of noncontextuality should also be considered as an important benchmark for emergent classicality within Darwinist processes.

Remarkably, even if Eq. (3) is only obeyed for small values of  $\eta$ , so that objectivity under a Darwinist process essentially does not occur, noncontextuality can still emerge, as long as  $\{\sigma_k^{B_j}\}_k$  are affinely independent. Indeed, even almost indistinguishable states can be affinely independent; in such cases, it is fair to say that objectivity will not emerge, whereas the probabilities can still be explained by noncontextual ontological models. In this sense, our findings reinforce the idea that the environment-as-a-witness paradigm alone already represents a step toward emergent classicality. One can also see this by using the minimum dimension of the Bobs' systems as a figure of merit: on one hand, for noncontextuality to emerge due to an  $\text{EW}_t$  dynamics together with affine independence of  $\{\sigma_k\}_k$ , one must have  $\dim(\mathcal{H}_{B_j})^2 \geq k_{\max}$ ; on the other hand, our bound identifying that the states  $\{\sigma_k^{B_j}\}_k$  are “almost orthogonal” among  $\{\sigma_k^{B_j}\}_k$  (which implies affine independence) implies  $\dim(\mathcal{H}_{B_j}) \geq k_{\max}$  (see the remark after Lemma 3). Therefore, a portion of the environment that cannot carry as much classical information about  $A$  as required by Darwinism (to lead to “enough” objectivity of outcomes) can still lead to noncontextuality.

Interesting perspectives also arise from this work. A first question, similar to the problem of how often interactions lead to  $\eta \approx 1$ , is if generic interactions lead to the condition  $\eta > \hat{P}$ . Second, it would be interesting to understand whether strong quantum Darwinism [17]—defined as a set of conditions on the mutual information and discord of system-environment states—also leads to noncontextuality. Third, it would be interesting to consider what is the impact of some special interactions on this emergence of classicality, such as those leading to non-Markovianity, the impact of which on Darwinist processes is a subject of debate [51,52]. Fourth, analyzing how to leverage the results of this work to the generalized probabilistic theories framework would certainly be of interest, as well as possible connections with state discrimination in this framework. Indeed, this could provide a way to understand the emergence of noncontextuality due to dynamics without resorting to the quantum formalism.

In sum, this work is strongly based on the description of the environment-as-a-witness paradigm and the  $\text{QD}$  process as given by Brandão *et al.* [15]—which shows the generic impacts of these dynamics. Therefore, our results also point to the generic emergence of noncontextuality under the same paradigm, which is undeniably a good approximation of our experiences as classical observers [14,20]. Thus, we may conclude that our noncontextual experience as observers, which is naturally embodied in

classical physics, may emerge from this indirect interaction with the systems that surround us. This work also generates new reasons to deem as classical the features emerging under  $\text{QD}_\eta$  processes, even if the POVM pointer  $\{\tilde{E}_k\}_k$  is informationally complete.

## ACKNOWLEDGMENTS

We thank Renan Cunha for important discussions about relations to minimum-error state discrimination. We are all grateful to the reviewers for important suggestions that helped to improve the presentation and deepen the discussions; in particular, for pointing out the possibility of using  $\hat{P}$  as a cutoff. R.D.B., R.W., B.A., and M.T.C. are grateful to the Brazilian agencies National Council for Scientific and Technological Development (CNPq) and Coordination for the Improvement of Higher Education Personnel (CAPES) for financial support. This work is part of the Brazilian National Institute for Science and Technology of Quantum Information (INCT-IQ). R.W. also acknowledges financial support by Portuguese funding institution, the Fundação para Ciência e Tecnologia (FCT), through scholarship No. SFRH/BD/151199/2021. R.D.B. acknowledges funding from Grant No. 2016/24162-8 from the São Paulo Research Foundation (FAPESP). This research was supported by Grant No. FQXi-RFP-IPW-1905 from the Foundational Questions Institute and Fetzer Franklin Fund, a donor advised fund of Silicon Valley Community Foundation. This work was also supported by the National Research, Development and Innovation Office of Hungary (NKFIH) through the Quantum Information National Laboratory of Hungary and through Grant No. FK 135220.

## APPENDIX A: FORMAL PRESENTATION AND DISCUSSION OF THE RESULTS OF BRANDÃO *ET AL.*

### 1. The results of Brandão *et al.*

In this subsection, for the benefit of the reader, we first present one of the results of Brandão *et al.* in rigorous form and then discuss quickly the consequences of our assumption of infinite subsystems in the environment.

Recall the  $\text{EW}_t$  dynamics: there is an environment with  $N$  subsystems,  $\{B_1, \dots, B_N\}$ , and we want to focus on a portion with  $t$  individual subsystems,  $B_{S_t}$ , of this environment. Theorem 1 of the main text, proved in Ref. [15] and responsible for objectivity of observables in this dynamics, can be rigorously put as follows.

**Theorem 6** (Theorem 2 in Ref. [15]). *Let  $\Phi^{S_t} : \mathcal{D}(\mathcal{H}_A) \rightarrow \mathcal{D}(\otimes_{j \in S_t} \mathcal{H}_{B_j})$  be an  $\text{EW}_t$  dynamics, where  $S_t \subset \{1, \dots, N\}$ . For every  $0 < \delta < 1$ , there exists a POVM  $\{\tilde{E}_k\}_k$  such that for more than a  $(1 - \delta)$  fraction*

*of the subsets  $S_t$ ,*

$$\left\| \Phi^{S_t} - \Phi_{\text{obs}}^{S_t} \right\|_{\diamond} \leq \left( \frac{27 \ln(2) d_A^6 \log(d_A) t}{N \delta^3} \right)^{1/3}, \quad (\text{A1})$$

*with  $d_A \equiv \dim(\mathcal{H}_A)$ , and where  $\Phi_{\text{obs}}^{S_t}$  is a measure-and-prepare map with respect to the family of states  $(\sigma_k^{S_t})_k$ , meaning that for all  $\rho \in \mathcal{D}(\mathcal{H}_A)$ ,*

$$\Phi_{\text{obs}}^{S_t}(\rho) = \sum_k \text{Tr}[\tilde{E}_k \rho] \sigma_k^{S_t}. \quad (\text{A2})$$

In the above theorem,  $\delta$  quantifies how many fractions of  $t$  portions of the environment the approximation will hold. From the upper bound given above, it is possible to see that if  $\delta$  is close enough to 1, our hypothesis of  $S_t$  being one such fraction is not too restrictive. Next, if this value  $\delta$  is fixed (or even is increasing slowly enough with  $N$ ), when we take the limit  $N \rightarrow \infty$  we have  $\Phi^{S_t} = \Phi_{\text{obs}}^{S_t}$ , as considered in the main text. Since environments are usually considered to have a very large quantity of subsystems, this hypothesis is well motivated.

### 2. Relaxing the assumption of infinite environment

In this subsection, we relax the assumption of infinite-sized environments, allowing for a deviation  $\Phi^{B_{S_t}} \neq \Phi_{\text{obs}}^{B_{S_t}}$ . We show that as long as we consider a finite set of procedures, outcomes, and equivalences, objectivity of observables hardly constrains contextuality, as measured by the quantifier based on the  $l_1$  distance [53].

In what follows, we assume that the number of procedures and outcomes are finite, as well as  $\{\text{Equiv}(\mathcal{P}'), \{\text{Equiv}(\mathcal{M}^{B_j})\}$ —which matches the features of any realist experiment. This means that the effective prepare-and-measure contextuality scenario for each Bob  $B_j$  can be defined as the tuple

$$\mathfrak{S}_j \equiv [T_j(\mathcal{P}), \mathcal{M}^{B_j}, \mathcal{O}_{\mathcal{M}^{B_j}}, \text{Equiv}[T_j(\mathcal{P})], \text{Equiv}(\mathcal{M}^{B_j})],$$

where these sets are as defined in the main text and where the set  $\mathcal{O}_{\mathcal{M}^{B_j}}$  is the set of outcome labels of Bob  $B_j$ . Given that all these sets are finite, they define prepare-and-measure scenarios as in Ref. [33]. Let  $C(\mathfrak{S}_j)$  represent the set of all possible behaviors,  $p \equiv [p(b|M, P')]_{b \in \mathcal{O}_{\mathcal{M}^{B_j}}, M \in \mathcal{M}^{B_j}, P' \in \mathcal{P}'}$  of  $\mathfrak{S}_j$ ; in the case of finitely many procedures, equivalences, and outcomes, this forms a polytope [33]. We write the inner set of noncontextual behaviors (also a polytope in the finite scenario) as  $\text{NC}(\mathfrak{S}_j)$ . We consider the resource-theory framework constructed for these scenarios [53,54], in particular, the

definition of the  $l_1$  contextuality distance.

$$\mathbf{d}(p) := \min_{q \in \text{NC}(\mathfrak{S}_j)} \max_{\substack{M \in \mathcal{M}^{B_j} \\ P' \in \mathcal{P}'}} \sum_b |p(b|M, P') - q(b|M, P')|.$$

For any particular noncontextual behavior  $q_*$ ,

$$\mathbf{d}(p) \leq \max_{\substack{M \in \mathcal{M}^{B_j} \\ P' \in \mathcal{P}'}} \sum_b |p(b|M, P') - q_*(b|M, P')|. \quad (\text{A3})$$

Recall that, in particular, when  $(\sigma_k)_k^{B_j}$  constitutes a family of affinely independent states, we have that  $q_* = \left( \text{Tr}[\Phi_{\text{obs}}^{B_j}(\rho) F_b^M] \right)_{b \in \mathcal{O}_{\mathcal{M}^{B_j}}, M \in \mathcal{M}^{B_j}, P' \in \mathcal{P}'}$  is a noncontextual behavior (Lemma 2 in the main text), where we are using POVM notation  $\{F_b^M\}_b = M$ . Since  $\mathcal{P}' = T_j(\mathcal{P})$ , we can use the labels for states in  $\mathcal{P}$ , meaning that Alice prepares quantum realizations as the states  $\rho_P$  for any  $P \in \mathcal{P}$  that are later transformed before reaching Bob  $B_j$ . If  $p$  is any quantum behavior in this scenario, we must have that

$$\begin{aligned} \mathbf{d}(p) &\leq \max_{\substack{M \in \mathcal{M}^{B_j} \\ P \in \mathcal{P}}} \sum_b \left| \text{Tr}[F_b^M \Phi^{B_j}(\rho_P)] - \text{Tr}[F_b^M \Phi_{\text{obs}}^{B_j}(\rho_P)] \right| \\ &= \max_{\substack{M \in \mathcal{M}^{B_j} \\ P \in \mathcal{P}}} \sum_b \left| \text{Tr}[F_b^M (\Phi^{B_j} - \Phi_{\text{obs}}^{B_j})(\rho_P)] \right| \\ &= \max_{\substack{M \in \mathcal{M}^{B_j} \\ P \in \mathcal{P}}} \sum_b \left| \text{Tr}[F_b^M \mathcal{R}_j(\rho_P)] \right|, \end{aligned}$$

where  $\mathcal{R}_j := \Phi^{B_j} - \Phi_{\text{obs}}^{B_j}$  for brevity. Note that  $\mathcal{R}_j(\rho_P)$  is a self-adjoint operator for any  $\rho_P$  in the domain. To proceed, recall the following inequality [55].

**Lemma 7** (Hölder's inequality). *Let  $A, B$  be any  $n \times n$  complex matrices. Then,*

$$|\text{Tr} A^\dagger B| \leq (\text{Tr}|A|^l)^{1/l} (\text{Tr}|B|^s)^{1/s} \quad (\text{A4})$$

such that  $1 \leq l, s \leq \infty$  with  $1/l + 1/s = 1$ .

Here, we fix the notation  $|A| := \sqrt{A^\dagger A}$ , while for scalars  $x$  the same symbol  $|x|$  means the absolute value. From Hölder's inequality, we have that

$$|\text{Tr}[F_b^M \mathcal{R}_j(\rho_P)]| \leq \sqrt{\text{Tr}[|\mathcal{R}_j(\rho_P)|^2]} \sqrt{\text{Tr}[|F_b^M|^2]}$$

and we obtain

$$\begin{aligned} \mathbf{d}(p) &\leq \max_{\substack{M \in \mathcal{M}^{B_j} \\ P \in \mathcal{P}}} \sum_b \left| \text{Tr}[F_b^M \mathcal{R}_j(\rho_P)] \right| \\ &\leq \max_{\substack{M \in \mathcal{M}^{B_j} \\ P \in \mathcal{P}}} \sum_b \sqrt{\text{Tr}[|\mathcal{R}_j(\rho_P)|^2]} \sqrt{\text{Tr}[|F_b^M|^2]} \\ &= \max_{M \in \mathcal{M}^{B_j}} \sum_b \sqrt{\text{Tr}[|F_b^M|^2]} \max_{P \in \mathcal{P}} \sqrt{\text{Tr}[|\mathcal{R}_j(\rho_P)|^2]} \\ &= C \max_{P \in \mathcal{P}} \sqrt{\text{Tr}[|\mathcal{R}_j(\rho_P)|^2]}, \end{aligned}$$

where  $C := \max_{M \in \mathcal{M}^{B_j}} \sum_b \sqrt{\text{Tr}[|F_b^M|^2]}$  is a nonzero bounded constant for any set  $\mathcal{M}^{B_j}$ . For informationally complete (IC) POVMs,  $0 < C \leq d_{B_j}^3$ . In general,  $\{0 < C \leq d_{B_j} \#\mathcal{O}_{\mathcal{M}^{B_j}}\}$ , the dimension of Bob's system  $d_{B_j}$  and the number of outcomes. We can further improve this inequality with the following lemma.

**Lemma 8.** *Let  $A \in \mathcal{B}(\mathcal{H})$ , with  $\dim(\mathcal{H}) < \infty$ . Then,*

*Proof.* Calling  $G = A^\dagger A$ , we have  $\sqrt{G} = \{\sum_{\lambda \in \sigma(G)} \sqrt{\lambda} P_\lambda\} = G$ , where  $\sigma(G)$  is the spectrum of  $G$  and  $P_\lambda$  are spectral projectors. Hence,  $\sqrt{\text{Tr}|A|^2} = \sqrt{\sum_{\lambda \in \sigma(G)} \lambda} \leq \sum_{\lambda \in \sigma(G)} \sqrt{\lambda} = \sum_{\lambda \in \sigma(\sqrt{G})} \lambda = \text{Tr}|A|$  since  $\sqrt{G} = |A|$ . ■

Using this lemma, we show that

$$\begin{aligned} \mathbf{d}(p) &\leq C \max_{P \in \mathcal{P}} \sqrt{\text{Tr}[|\mathcal{R}_j(\rho_P)|^2]} \\ &\stackrel{\text{Lemma 8}}{\leq} C \max_{P \in \mathcal{P}} \text{Tr}[|\mathcal{R}_j(\rho_P)|] \end{aligned}$$

and therefore

$$\begin{aligned} \mathbf{d}(p) &\leq C \max_{\rho \in \mathcal{D}(\mathcal{H}_A)} \text{Tr}[|\mathcal{R}_j(\rho)|] \\ &= C \max_{\rho \in \mathcal{D}(\mathcal{H}_A)} \left\| (\Phi^{B_j} - \Phi_{\text{obs}}^{B_j})(\rho) \right\|_1 \\ &= C \max_{\rho \in \mathcal{D}(\mathcal{H}_A)} \left\| \frac{1}{d_A} \text{id}_A(\mathbb{1}_A) \otimes ((\Phi^{B_j} - \Phi_{\text{obs}}^{B_j})(\rho)) \right\|_1, \end{aligned}$$

where the last equality comes from the fact that the singular values of a matrix that is a tensor product is simply the product of the singular values. Our notation is  $\mathbb{1}_A \in \mathcal{D}(\mathcal{H}_A)$  and  $\text{id}_A : \mathcal{D}(\mathcal{H}_A) \rightarrow \mathcal{D}(\mathcal{H}_A)$ , with  $\text{id}_A(\rho) = \rho$ . Continuing with the manipulations, we finally find,

defining  $\mathcal{H}_A^{\otimes 2} \equiv \mathcal{H}_A \otimes \mathcal{H}_A$ ,

$$\begin{aligned} \mathbf{d}(p) &\leq \frac{C}{d_A} \max_{\mathbb{1}_A \otimes \rho \in \mathcal{D}(\mathcal{H}_A^{\otimes 2})} \left\| \text{id}_A \otimes \left( \Phi^{B_j} - \Phi_{\text{obs}}^{B_j} \right) (\mathbb{1}_A \otimes \rho) \right\|_1 \\ &\leq \frac{C}{d_A} \max_{\sigma \in \mathcal{D}(\mathcal{H}_A^{\otimes 2})} \left\| \text{id}_A \otimes \left( \Phi^{B_j} - \Phi_{\text{obs}}^{B_j} \right) (\sigma) \right\|_1 \\ &= \frac{C}{d_A} \left\| \Phi^{B_j} - \Phi_{\text{obs}}^{B_j} \right\|_{\diamond}. \end{aligned}$$

The relation found above tells us that contextuality, as quantified by the  $l_1$  distance, is bounded by the distance from  $\Phi^{B_{S_t}}$  to  $\Phi_{\text{obs}}^{B_{S_t}}$ , as measured by the diamond norm. Now, the theorem of Brandão *et al.* reproduced in the previous subsection tells us that this distance is bounded, showing that the EW<sub>t</sub> dynamics bounds contextuality even in the case of a finite-sized environment.

As an important instance, let Bob  $B_j$  perform informationally complete POVMs, such that  $C \leq d_{B_j}^3$ . Our conclusion is that for a sufficiently large number  $N$  with respect to  $d_{B_j}$ ,  $d_A$ ,  $t$ ,  $\delta^{-1}$ , the behavior  $p$  will be noncontextual up to precision in the estimate of  $\mathbf{d}(p)$ , since  $\mathbf{d}(p) \leq \left( 27 \ln(2) d_A^3 d_{B_j}^3 \log(d_A) t / N \delta^3 \right)^{1/3}$ .

## APPENDIX B: PROOFS OF THE RESULTS

The first problem to be treated here is as follows: we have a measure-and-prepare map defined by the pairs  $(\tilde{E}_k, \sigma_k)_k$  and we wish to detect affine independence of the states  $(\sigma_k)_k$  in an operational task. What we do in the next subsection is to consider that, for each initial state of system  $A$ ,  $\rho^A \in \mathcal{D}(\mathcal{H}_A)$ , we have an instance of a minimum-error state discrimination task. Indeed, each  $\rho^A$  leads to an *a priori* distribution  $\tilde{p}_k := \text{Tr}[\tilde{E}_k \rho^A]$  for the states to be discriminated,  $\{\sigma_k\}_k$ . This leads us to a bound  $\hat{P}$  such that, if  $p_{\text{guess}}[\tilde{p}_k, \sigma_k]_k > \hat{P}$  for all  $\rho^A$ , the states  $\{\sigma_k\}_k$  must be affinely independent.

### 1. Proof of Lemma 3

Before proving our results, let us give a simple mathematical argument of why there should be a bound on the guessing probability separating the affinely independent and dependent cases.

Define the matrix  $(W)_{ik} := \text{Tr}[F_i \sigma_k]$ , where  $\{F_i\}_i$  is a POVM that best distinguishes among the states  $\{\sigma_k\}_k$ . If the states are perfectly distinguishable—in which case they must have disjoint support, thus being affinely independent—we have  $W = \mathbb{1}_{k_{\text{max}} \times k_{\text{max}}}$ . This implies that  $\det(W) = 1$ .

However, in the case where  $\{\sigma_k\}_k$  is affinely dependent, the matrix  $W$  must have a column that is linearly dependent on the others. This implies that  $\det(W) = 0$ . From the continuity of the determinant, we see that there exists a

ball  $b$  around  $\mathbb{1}$  such that  $V \in b \implies \det(V) > 0$ ; thus the underlying states defining the entries of  $V$ ,  $(\sigma_k)_k$ , must be affinely independent.

With this general intuition, we follow to the specifics of our proof. First, we define our distinguishability bound,  $\hat{P}$ , able to detect affine independence when  $p_{\text{guess}} > \hat{P}$ .

**Definition 5** (Distinguishability bound). *Consider a measure-and-prepare channel defined by  $(\tilde{E}_k, \sigma_k)_k$ , with  $\tilde{E}_k \neq 0$  for all  $k$ . Then, there will be states  $\rho^A \in \mathcal{D}(\mathcal{H}_A)$  such that  $\text{Tr}[\tilde{E}_k \rho^A] \neq 0$  for all  $k$ . Denote the set of such states by  $\mathcal{S}$ . Now, assume without loss of generality that  $\text{Tr}[\tilde{E}_1 \rho^A] \geq \text{Tr}[\tilde{E}_2 \rho^A] \geq \dots \geq \text{Tr}[\tilde{E}_{k_{\text{max}}} \rho^A] > 0$  (otherwise, relabel  $\{\tilde{E}_k\}$  so that it does). We define the **distinguishability bound**  $\hat{P}$  as*

$$\begin{aligned} \hat{P}[(\tilde{E}_k)_k] &:= \min_{\rho^A \in \mathcal{S}} \left[ \sum_{k=1}^{k_{\text{max}}-1} \text{Tr}[\tilde{E}_k \rho^A] + \frac{\text{Tr}[\tilde{E}_{k_{\text{max}}} \rho^A]}{2} \right] \\ &= 1 - \frac{1}{2} \max_{\rho^A \in \mathcal{S}} \text{Tr}[\tilde{E}_{k_{\text{max}}} \rho^A]. \end{aligned} \quad (\text{B1})$$

Note that, indeed,  $\mathcal{S}$  is not empty: since  $\tilde{E}_k \neq 0$  for all  $k$ , at least states of the form  $(1-a)\rho^A + a(\mathbb{1}/d_A)$ , with  $a \neq 0$ , belong to this set. In what follows and for simplicity, we always consider that  $(\tilde{p}_k)_k = (\text{Tr}[\tilde{E}_k \rho^A])_k$  is nonincreasing, so that  $\tilde{p}_{k_{\text{max}}}$  is always the smallest value of the distribution  $(\tilde{p}_k)_k$  generated by a state  $\rho^A \in \mathcal{S}$ , as in Definition 5. There is no loss of generality here, since equivalent proofs can be written by relabeling and/or changing  $\tilde{p}_{k_{\text{max}}}$  to  $\min_k[(\tilde{p}_k)_k]$ . We take advantage of the fact that if  $\{\sigma_k\}_k$  are affinely dependent, they *do not* form vertices of a  $(k_{\text{max}} - 1)$ -simplex, thus interior points will have nonunique convex decompositions in terms of these states and this will make it impossible to have  $p_{\text{guess}} > \hat{P}$  for all initial states  $\rho^A$ . We actually consider a stronger consequence of affine dependence, captured by Carathéodory's theorem: we can always describe interior points using convex combinations of  $\{\sigma_k\}_k$  with some null coefficients (see Fig. 3).

**Theorem 9** (Carathéodory, adapted from version of Ref. [56]). *Given any affine space  $E$  of dimension  $n$ , for any (nonvoid) family  $f = \{\sigma_k\}_{k=1}^{k_{\text{max}}}$  in  $E$ , the set  $\text{ConvHull}[f]$  is equal to the set of convex combinations of families of  $n + 1$  points of  $f$ .*

Let us interpret this result in our terms. The affine space  $E := \text{AffineHull}[\{\sigma_k\}_k]$  will have dimension  $n \leq k_{\text{max}} - 1$ , where equality is reached if and only if  $\{\sigma_k\}_k$  is an affinely independent set. Carathéodory's theorem thus says that any point  $\sigma \in \text{ConvHull}[\{\sigma_k\}_{k=1}^{k_{\text{max}}}]$  can be written as a convex combination of at most  $n + 1$  points of  $\{\sigma_k\}_k$ .

Now, suppose that  $\{\sigma_k\}_k$  *does not* form an affinely independent set. Then, the dimension of the affine space  $E$  is

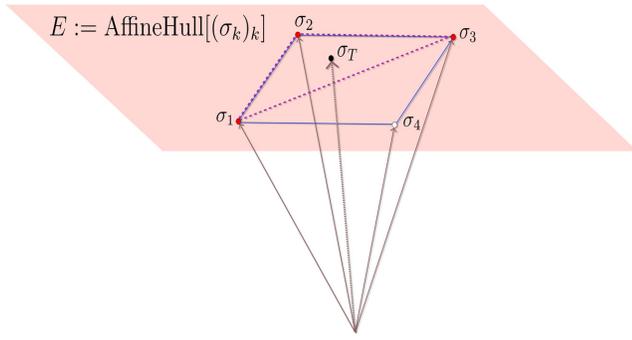


FIG. 3. An example of Caratheodory's theorem: an example of an affinely dependent set with four states ( $k_{\max} = 4$ ). Their affine hull  $E$  forms a plane; thus  $\dim(E) = 2$  and  $\text{ConvHull}[\{\sigma_k\}_k]$  is a quadrilateral contained in  $E$ . An interior point,  $\sigma_T$ , can be written as a convex combination of  $2 + 1 = 3$  vertices (as shown by the red vertices and the purple dashed triangle). Thus, we can write  $\sigma_T = \sum_k q_k \sigma_k$  with  $q_4 = 0$ .

strictly less than  $k_{\max} - 1$ , i.e.,  $\dim(E) \leq k_{\max} - 2$ . From Carathéodory's theorem, any point of  $\text{ConvHull}[\{\sigma_k\}_k]$  can thus be written as a convex combination of  $k_{\max} - 1$  vertices. In particular, consider a density matrix  $\sigma_T = \sum_k \tilde{p}_k \sigma_k$  with  $\tilde{p}_k > 0$  for all  $k$ ; we can be sure that there exists another set of convex coefficients,  $\{q_k\}$ , such that  $q_k = 0$  for at least one  $k$  [there may be more zeros if  $\dim(E) < k_{\max} - 2$ ]. We formalize this as follows.

**Corollary 10.** *Consider any state of the form  $\sigma_T = \sum_k \tilde{p}_k \sigma_k$  with  $\tilde{p}_k > 0$  for  $k \in \{1, \dots, k_{\max}\}$ . Then, if  $(\sigma_k)_k$  is an affinely dependent set, there exists a set of convex coefficients  $\{q_k\}_k$ , with  $q_k = 0$  for at least one value of  $k$ , such that  $\sigma_T = \sum_k q_k \sigma_k$ .*

The comparison of  $\{q_k\}$  and  $\{\tilde{p}_k\}$  has important consequences for us. We might quantify it via “the” statistical distance.

**Remark 1.** *Assume  $(\sigma_k)_k$  to be an affinely dependent set. Consider a state  $\sigma_T = \sum_k \tilde{p}_k \sigma_k = \sum_k q_k \sigma_k$  with  $\tilde{p}_1 \geq \tilde{p}_2 \dots \geq \tilde{p}_{k_{\max}} > 0$  and  $q_k = 0$  for some  $k$ . Define the statistical distance between  $\{q_k\}_k$  and  $\{\tilde{p}_k\}_k$  as*

$$D(\{q_k\}) = \frac{1}{2} \sum_k |q_k - \tilde{p}_k|. \quad (\text{B2})$$

Then,  $D(\{\tilde{p}_k\}, \{q_k\}) \geq \tilde{p}_{k_{\max}}$ .

Let us justify this remark. First, define the set  $K := \{k | q_k \geq \tilde{p}_k\}$  and suppose that  $q_{k_{\max}} = 0$ . Now, note that we

can rewrite the statistical distance

$$\begin{aligned} D(\{q_k\}, \{\tilde{p}_k\}) &= \frac{1}{2} \sum_k |q_k - \tilde{p}_k| \\ &= \frac{1}{2} \left( \sum_{k \in K} (q_k - \tilde{p}_k) + \sum_{k \notin K} (\tilde{p}_k - q_k) \right) \\ &= \sum_{k \in K} (q_k - \tilde{p}_k). \end{aligned}$$

The last step uses the fact that  $\sum_{k \notin K} q_k = 1 - \sum_{k \in K} q_k$  and similarly for  $\tilde{p}_k$ , so we can substitute the second summation by another over  $k \in K$  and the last equation follows. Now, we assume that  $\tilde{p}_{k_{\max}} > 0$ , while  $q_{k_{\max}} = 0$ , so  $k_{\max} \notin K$ . Thus,

$$\begin{aligned} D(\{q_k\}, \{\tilde{p}_k\}) &= \sum_{k \in K} (q_k - \tilde{p}_k) \\ &\geq \sum_{k \neq k_{\max}} (q_k - \tilde{p}_k) \\ &= 1 - \sum_{k \neq k_{\max}} \tilde{p}_k = \tilde{p}_{k_{\max}}. \end{aligned}$$

The inequality comes from the fact that we might have included some values of  $k \notin K$ , which can only decrease the summation. Now, suppose that  $q_{k_{\max}} \neq 0$ ; we know by Corollary 10 that there will be at least one value of  $k$ , call it  $k^*$ , such that  $q_{k^*} = 0$ . Then, by repeating the above calculation, we find that  $D(\{q_k\}, \{\tilde{p}_k\}) > \tilde{p}_{k^*} \geq \tilde{p}_{k_{\max}}$ , justifying our remark.

Finally, we have the ingredients to prove our lemma.

**Lemma 11** (Violation of  $p_{\text{guess}} \leq \hat{P}$  for all  $\rho^A$  implies  $(\sigma_k)_k$  is an affinely independent set). *Consider an environment-as-a-witness dynamics defined by the measure-and-prepare channel  $(\tilde{E}_k, \sigma_k)_k$ , with  $\tilde{E}_k \neq 0$  for all  $k$ . Consider the associated distinguishability bound  $\hat{P}[(\tilde{E}_k)_k]$  as per Definition 5. Then, if for all  $\rho^A \in \mathcal{D}(\mathcal{H}_A)$  the inequality*

$$p_{\text{guess}}[\text{Tr}[\tilde{E}_k \rho^A], \sigma_k] > \hat{P}[(\tilde{E}_k)_k] \quad (\text{B3})$$

holds, the states  $(\sigma_k)_k$  are affinely independent.

*Proof.* Fix  $\rho^A = \bar{\rho}$ , where  $\bar{\rho}$  is a state attaining the minimum in the definition of the bound  $\hat{P}$ . Since  $\bar{\rho} \in \mathcal{S}$ ,  $\tilde{p}_k$  does not increase with  $k$  and  $\tilde{p}_{k_{\max}} > 0$ . Let us denote the POVM that maximizes the probability of guessing (for the specific  $\rho^A \in \mathcal{S}$ ) as  $\{F_k^*\}_k$ . Now, if  $\text{Tr}[F_b^* \sigma_b] \leq \frac{1}{2}$  for some

value  $b$ , the bound cannot be violated. Indeed,

$$p_{\text{guess}}[(\tilde{p}_k, \sigma_k)_k] = \sum_k \tilde{p}_k \text{Tr}[F_k^* \sigma_k] \quad (\text{B4})$$

$$\begin{aligned} &\leq \sum_{k \neq b} \tilde{p}_k \overbrace{\text{Tr}[F_k^* \sigma_k]}^{\leq 1} + \frac{\tilde{p}_b}{2} \\ &\leq \sum_{k \neq b} \tilde{p}_k + \frac{\tilde{p}_b}{2} \leq \sum_{k \neq k_{\max}} \tilde{p}_k + \frac{\tilde{p}_{k_{\max}}}{2} = \hat{P} \end{aligned} \quad (\text{B5})$$

[remember that  $\tilde{p}_{k_{\max}}$  is the minimum of  $(\tilde{p}_k)$ , hence the last inequality. By considering the appropriate  $\{F_k^*\}_k$  for each  $\rho^A$ , all but the last step above are valid for all  $\rho^A \in \mathcal{S}$ . The last equality is only valid for states attaining the minimum at the definition of the bound]. So, if  $\text{Tr}[F_b^* \sigma_b] \leq 1/2$  for some  $b$ , we can see that  $p_{\text{guess}} \leq \hat{P}$  at least for  $\bar{\rho}$  as the initial state. Therefore, one needs  $\text{Tr}[E_k^* \sigma_k] > 1/2$  for all  $k$  so that  $p_{\text{guess}} > \hat{P}$  might be possible for all  $\rho^A$ .

Now, we assume that  $\{\sigma_k\}_k$  is an affinely dependent set, which leads to a contradiction with Eq. (B3). We still consider  $\rho^A = \bar{\rho} \in \mathcal{S}$  and show that if the effective average state at Bob's side ( $\sigma_T = \sum_k \tilde{p}_k \sigma_k$  with  $\tilde{p}_k > 0$  for all  $k$ ) admits a convex decomposition  $\sigma_T = \sum_k q_k \sigma_k$  with  $q_k = 0$  for some  $k$ ,  $p_{\text{guess}} < \hat{P}$  must be obeyed for  $\bar{\rho}$ .

The probability of error in the discrimination is defined by  $P_{\text{err}} = \sum_k \sum_{b \neq k} \tilde{p}_b \text{Tr}[F_k^* \sigma_b]$ . Since  $\sum_{b \neq k} \tilde{p}_b \sigma_b = \sigma_T - \tilde{p}_k \sigma_k$ , we can rewrite  $P_{\text{err}}$  as follows:

$$P_{\text{err}} = \sum_k \overbrace{\text{Tr}[F_k^* (\sigma_T - \tilde{p}_k \sigma_k)]}^{\geq 0} \quad (\text{B6})$$

$$\begin{aligned} &\geq \sum_{k \in K} \text{Tr}[F_k^* (\sigma_T - \tilde{p}_k \sigma_k)] \\ &= \sum_{k \in K} \text{Tr} \left[ F_k^* \left( \sum_i q_i \sigma_i - \tilde{p}_k \sigma_k \right) \right] \end{aligned} \quad (\text{B7})$$

$$\geq \sum_{k \in K} \text{Tr}[F_k^* (q_k \sigma_k - \tilde{p}_k \sigma_k)]. \quad (\text{B8})$$

In the first inequality, we use the fact that the summand is positive, so restricting  $k \in K$  (recall,  $K := \{k | q_k \geq \tilde{p}_k\}$ ) can only decrease the summation. After substituting  $\sigma_T = \sum_i q_i \sigma_i$  (second equation), we use the fact that  $\text{Tr}[F_k \sigma_k] \leq \text{Tr}[F \sigma_T]$  for any POVM element  $F$ . Therefore, we obtain

$$\begin{aligned} P_{\text{err}} &\geq \sum_{k \in K} (q_k - \tilde{p}_k) \overbrace{\text{Tr}[F_k^* \sigma_k]}^{> 1/2} > \frac{1}{2} \sum_{k \in K} (q_k - \tilde{p}_k) \\ &= \frac{1}{2} D(\{\tilde{p}_k\}, \{q_k\}) \geq \frac{p_{k_{\max}}}{2}, \end{aligned} \quad (\text{B9})$$

where we use Remark 1 to replace  $D(\{q_k\}, \{\tilde{p}_k\})$ .

We can thus conclude that if  $(\sigma_k)_k$  is an affinely dependent set,  $P_{\text{err}} > \tilde{p}_{k_{\max}}/2$  for states attaining the minimum on the definition of the bound. Equivalently,  $p_{\text{guess}} < \sum_{k \neq k_{\max}} \tilde{p}_k + \tilde{p}_{k_{\max}}/2$  for such states, thus being impossible to obey the inequality in Eq. (B3). By contradiction, we arrive at the claim. (Note that  $p_{\text{guess}}$  is not bounded by the condition of affine independence, since it can get as high as 1, when  $\{\sigma_k\}_k$  are perfectly distinguishable.) ■

The above proof may be of interest for minimum-error state discrimination: as mentioned above, for each  $\rho^A$ , we have an instance of this task. Moreover, the attention we pay to states in the set  $\mathcal{S}$  is very natural in minimum-error state discrimination [57]. Indeed, in those tasks, one is interested in distinguishing states  $\{\sigma_k\}$ , chosen according to a fixed prior  $(\tilde{p}_k)$  where, usually,  $\tilde{p}_k > 0$  for all  $k$  is assumed. Otherwise, one could just consider a smaller subset of the states  $\{\sigma_k\}$  that have a nonzero chance of being selected.

Adapting to this task, our proof shows that if  $p_{\text{guess}} > 1 - \min_k \tilde{p}_k/2$ , the states  $\{\sigma_k\}_k$  must be affinely independent, thus forming a  $(k_{\max} - 1)$ -simplex in  $\mathcal{D}(\mathcal{H})$ . In other words, by observing a high value of  $p_{\text{guess}}$ , we obtain information on the geometrical disposition of the states  $\{\sigma_k\}_k$ .

## 2. Perfect decoherence and state-spectrum broadcasting cases

### a. Dynamics with perfect decoherence

Let us look into how our bound behaves in a simple and important example of an  $\text{EW}_t$  dynamics: the case in which  $\{\tilde{E}_k\}$  is a rank-1 projective pointer observable, projecting onto a basis of  $\mathcal{H}_A$ . This is the case of a perfect decoherence process, which is important for ideal cases of quantum Darwinism [14], thus forming a milestone in such processes.

In this case, as mentioned above,  $\tilde{E}_k = |k\rangle \langle k|$ , and information regarding projection on this basis,  $\{|k\rangle \langle k|\}_k$ , may be available to the environmental states  $\sigma^{B_{S_t}}$ . Initial states  $\rho^A$ , having the same population on this basis and differing only on the phases, will lead to the same distribution  $\text{Tr}[\tilde{E}_k \rho^A] = \text{Tr}[|k\rangle \langle k| \rho^A]$ —as mandated by an ideal decoherence process. Therefore, the ontological model constructed in the main text [Eq. (8)] encapsulates this idea:  $\mu_{P'}$  is the same for all initial states that decohere to the same final state. What would our bound  $\hat{P}$  be?

To answer this, we need to look at the set of states  $\mathcal{S}$  in Definition 5, which contains the states having a nonzero component on all of the pointer basis states. In other words, the states in  $\mathcal{S}$  must obey  $\tilde{p}_k = \text{Tr}[|k\rangle \langle k| \rho] > 0$  for all  $k$ . Among those, we need to find the one with the highest possible value for the smaller  $\tilde{p}_k$ . It is possible to see that such a state is the maximally mixed state:  $\text{Tr}[|k\rangle \langle k| \rho^A] = 1/\dim(\mathcal{H}_A)$  for all  $k$ . Then, in the case of

perfect decoherence with a pointer *basis* being selected, if

$$p_{\text{guess}}[(\text{Tr}[|k\rangle\langle k| \rho^A], \sigma_k)_k] > 1 - \frac{1}{2\dim(\mathcal{H}_A)} \quad \forall \rho^A, \quad (\text{B10})$$

we are sure that  $\{\sigma_k\}_k$  are affinely independent states.

As we can see, for complete decoherence of systems with different dimensions, we have a different bound  $\hat{P}$ —the higher the dimension, the higher is  $\hat{P}$ . This shows that choosing an arbitrary bound  $\bar{\eta}$  to  $\text{QD}_{\bar{\eta}}$  processes as a threshold for objectivity can lead to problems, even though it might look “close enough to 1.” Indeed, suppose that one fixes  $\bar{\eta} < 1$ . Now, for complete decoherence of a system with dimension  $\dim(\mathcal{H}_A) > 1/2(1 - \bar{\eta})$ , the bound  $\hat{P}$  obeys  $\hat{P} > \bar{\eta}$ .

Another interesting consequence of this particular case is that we can use it to show the emergence of noncontextuality in an alternative process for objectivity in the quantum realm, namely, state-spectrum broadcasting [16].

### b. State-spectrum broadcasting

The process of state-spectrum broadcasting was first proposed in Ref. [16], imposing a specific form for the final state of the central system and a fragment of its environment as the reason for objectivity. It was suggested as a necessary and sufficient condition for emergence of objectivity in the quantum realm. Later, however, it was proven to be too restrictive, providing sufficient, but not necessary, conditions [17]. Here, we adapt the Brandão *et al.* approach to dealing with state-spectrum broadcasting and show that noncontextuality also emerges under this process.

Consider the whole system-environment dynamics  $\Phi^{AE}$  as a CPTP map  $\mathcal{D}(\mathcal{H}_A) \mapsto \mathcal{D}(\mathcal{H}_{A\mathcal{E}})$ . Note that this is similar to the  $\Phi$  map we consider in Definition 1 but with the difference that we are not discarding the state of the central system after the interaction. Now, as in the case of  $\text{EW}_t$  dynamics, let us discard all of the environment except a small fragment with  $r$  subsystems [58],  $B_{S_r}$ , i.e.,  $\Phi^{AB_{S_r}} := \text{Tr}_{\mathcal{E} \setminus B_{S_r}} \circ \Phi^{AE}$ . Again, the essential difference between  $\Phi^{AB_{S_r}}$  and  $\Phi^{B_{S_r}}$  is the explicit presence of  $A$ . Depending on the final state induced by the  $\Phi^{AB_{S_r}}$  dynamics, we may arrive at a state-spectrum broadcasting process.

**Definition 6** (State-spectrum broadcasting). *Consider the map  $\Phi^{AE}$  describing the interaction of system  $A$  and its environment  $\mathcal{E}$ , composed of  $N$  subsystems. Now, consider the dynamics  $\Phi^{AB_{S_r}}$ , obtained by tracing out the whole environment except for  $r$  subsystems denoted by  $B_{S_r}$ . A state-spectrum broadcasting process has occurred if the final joint state of  $A$  and  $B_{S_r}$ ,  $\rho^{AB_{S_r}} := \Phi^{AB_{S_r}}(\rho^A)$ , takes the*

following form:

$$\rho^{AB_{S_r}} = \sum \tilde{p}_k |k\rangle\langle k| \otimes \bigotimes_{j \in S_r} \sigma_k^{B_j}, \quad (\text{B11})$$

$\tilde{p}_k := \text{Tr}[|k\rangle\langle k| \rho^A]$  being the probability distribution arising due to perfect decoherence and all  $\{\sigma_k^{B_j}\}_k$  having disjoint supports, i.e.,  $\sigma_k^{B_j} \sigma_{k'}^{B_j} = 0$  whenever  $k \neq k'$ .

In the above definition, the name of the process becomes clear: the spectrum of the state of the system after decoherence (with respect to a certain pointer basis),  $(\tilde{p}_k)_k$ , is perfectly broadcast to the environment. The proposal that such a structure for the joint state  $\rho^{AB_{S_r}}$  provides sufficient conditions for objectivity comes from the following. First, the system undergoes a complete decoherence process, having no decoherence-free subspace left. This means that it can be considered as a classical mixture of orthogonal states. Second, the condition of disjoint supports for  $\{\sigma_k^{B_j}\}_k$  for every  $j \in B_{S_r}$  means that the distribution  $\tilde{p}_k$  is perfectly broadcast to the environmental subsystems  $B_j \in B_{S_r}$ . That is, this condition corresponds to the  $\eta = 1$  case for quantum Darwinism! Thus, if one accepts a condition  $\eta \approx 1$  leading to objectivity under the approach to quantum Darwinism of Brandão *et al.*, one is forced to accept state-spectrum broadcasting as reaching objectivity as well. Finally, the additional presence of the (completely decohered) state of the central system ensures that one can directly probe it, thus allowing for confirmation of the information broadcast to the environment (as long as one can probe it using the  $\{|k\rangle\langle k|\}_k$  measurement).

Note that we can describe a state-spectrum broadcasting process as a special and more restrictive case of Darwinism in the Brandão *et al.* formalism. All we need to do is to adapt our interpretation, by considering a postinteraction state of  $A$  in addition to an environmental portion,  $B_{S_r}$ . In other words, we can choose a special  $B_{S_t}$  portion in our definition of Darwinist process:  $B_{S_t} = A \cup B_{S_r}$ , where  $B_{S_t}$  now also contains the final state of  $A$ . This is not problematic, as there is no need to demand all of  $B_j$  to be different from  $A$ . Indeed, all that is required by the  $\text{EW}_t$  dynamics is that  $\Phi$  maps  $\mathcal{D}(\mathcal{H}_A)$  to  $\mathcal{D}(\mathcal{H}_{B_1} \otimes \dots \otimes \mathcal{H}_{B_N})$ , with  $N$  large, and we should trace out all but the  $B_{S_t}$  subsystems in the end—and one of the  $B_j \in B_{S_t}$  can certainly represent the state of  $A$  after the interaction [59]. In the main text, for the sake of simplicity, we consider  $B_{S_t}$  as being composed of only environmental subsystems.

Therefore, we can see state-spectrum broadcasting as a special kind of  $\text{QD}_{\bar{\eta}}$  process, in which (i)  $\eta = 1$ , (ii) perfect decoherence has occurred ( $\tilde{E}_k = |k\rangle\langle k|$  for some pointer basis  $|k\rangle$  of the Hilbert space  $\mathcal{H}_A$ ), and (iii)  $A \in B_{S_t}$  and  $\sigma_k^A = |k\rangle\langle k|$ . Indeed, the state in Eq. (B10) is already in a measure-and-prepare form, with objective observable  $\tilde{E}_k = |k\rangle\langle k|$  and conditionally prepared

states  $|k\rangle\langle k| \otimes \bigotimes_{j \in S_i} \sigma_k^{B_j}$ . Those states *have disjoint support*, thus being perfectly distinguishable, i.e.,  $\eta = 1$ . From the example of perfect decoherence, we know that  $\hat{P}[(|k\rangle\langle k|)_k] = 1 - 1/2\dim(\mathcal{H}_A)$ , which means that  $\{\eta (= 1) > \hat{P}\}$ . Since we are dealing with a specific form of a  $\text{QD}_\eta$  process, we can apply our Theorem 4 and conclude that, under a process of state-spectrum broadcasting, non-contextuality emerges. We state this formally as follows.

**Corollary** (State-spectrum broadcasting—restatement). *If the interaction leads to the occurrence of state-spectrum broadcasting process for arbitrary initial states  $\rho^A$ , all the Bobs can construct a noncontextual ontological model for their statistics.*

*Proof.* As discussed above, we have  $\eta = 1$  and  $\tilde{E}_k = |k\rangle\langle k|$  for all  $k$ , with  $|k\rangle\langle k|$  a basis of  $\mathcal{H}_A$ . This last condition implies that  $\hat{P} = 1 - 1/2\dim(\mathcal{H}_A)$ , as seen in Appendix B 2 a. If the interaction is such that a state-spectrum broadcasting process occurs regardless of the initial state of the system, we have a  $\text{QD}_\eta$  process satisfying  $\eta > \hat{P}$  and, by Theorem 4, all of the Bobs  $B_j \in B_{S_i}$  can construct a noncontextual ontological model for their statistics. Interestingly, one of the Bobs will be receiving the central system  $A$  after the interaction, in state  $\sigma^A = \sum \tilde{p}_k |k\rangle\langle k|$ . The fact that this is a state of  $A$  and not an environmental subsystem makes no difference to our results. ■

### APPENDIX C: CONNECTION TO THE CHARACTERIZATION OF NONCONTEXTUALITY IN THE GENERALIZED PROBABILISTIC THEORIES FRAMEWORK

In Lemma 2, we show that noncontextuality emerges for those Bobs in which the encoding states  $\{\sigma_k^{B_j}\}_k$  are affinely independent. The condition of affine independence of these states can be linked to the characterization of noncontextuality in the generalized probabilistic theories (GPT) framework, as formalized in Refs. [27,50]. For the benefit of readers familiar with such characterization, in this appendix, we concisely demonstrate this connection for the case  $k_{\max} = \dim(\mathcal{H}_{B_j})^2$ . In what follows, we use the notation  $n := \dim(\mathcal{H}_{B_j})$ .

As mentioned in the proof of Lemma 2, affine independence of  $\{\sigma_k^{B_j}\}_k$  implies that  $\text{ConvHull}[\{\sigma_k^{B_j}\}_k]$  is a  $(k_{\max} - 1)$ -simplex, denoted by  $\Delta_{k_{\max}}$ . Thus, the effective states arriving at Bob  $B_j$ ,  $\sum \tilde{p}_k \sigma_k^{B_j}$ , can be identified with points on that simplex. In the case on which we focus here, i.e.,  $k_{\max} = n^2$ , this simplex is written  $\Delta_{n^2}$ . Since these are quantum states on Bob's subsystem,  $\Delta_{n^2} \subset \mathcal{D}(\mathcal{H}_{B_j})$ .

The states  $\mathcal{D}(\mathcal{H}_{B_j})$ , as well as the restricted  $\Delta_{n^2}$ , live in (a hyperplane of) the real vector space of  $n \times n$  Hermitian matrices,  $\mathbb{H}_n$ . The set of elements of POVM,

which will assign the probabilities to each outcome of all possible measurements that Bob  $B_j$  can perform, can be identified with the compact convex set  $\mathcal{D}^*(\mathcal{H}_{B_j}) := \{F_b \in \mathbb{H}_n | \text{Tr}[F_b \rho] \in [0, 1] \forall \rho \in \mathcal{D}(\mathcal{H}_{B_j})\}$ . We also consider the analogously defined set of Hermitian matrices having inner product with elements of the simplex between 0 and 1, i.e.,  $\Delta_{n^2}^* := \{F'_b \in \mathbb{H}_n | \text{Tr}[F'_b \rho] \in [0, 1] \forall \rho \in \Delta_{n^2}\}$ . Since  $\Delta_{n^2} \subset \mathcal{D}(\mathcal{H}_{B_j})$ , the “dual” set  $\Delta_{n^2}^*$  must be larger than  $\mathcal{D}^*(\mathcal{H}_{B_j})$ , i.e.,  $\mathcal{D}^*(\mathcal{H}_{B_j}) \subset \Delta_{n^2}^*$ .

Now, consider a generalized probabilistic theory that is the result of the following procedure: start with quantum theory (as a GPT) and then impose the restriction  $\mathcal{D}(\mathcal{H}_{B_j}) \rightarrow \Delta_{n^2}$  on Bob's states due to the dynamics—let us call this GPT the “EWQuantum” theory. Under this perspective, we can view  $\Delta_{n^2}$  as the normalized state space of “EWQuantum” and  $\mathcal{D}^*$  as the (restricted) set of effects of such a theory. Under this perspective, the fact that the normalized state space is a simplex,  $\Delta_{n^2}$ , and that the effects are a subset of the dual,  $\mathcal{D}^* \subset \Delta_{n^2}^*$ , imply that EWQuantum “fits inside” classical probability theory! Formally, this means that this GPT is *simplex-embeddable*, which is an equivalent condition for noncontextuality of the underlying operational theory in prepare-and-measure scenarios, as proven in Ref. [27].

In Ref. [50], noncontextuality in GPTs is characterized in a similar way but with the additional condition that the classical theory into which the GPT “fits” should have the same dimension as the GPT. This is exactly what happens here, since the classical probability theory (as a GPT) with the set of normalized states  $\Delta_{n^2}$  and effects  $\Delta_{n^2}^*$  lives in a real vector space of dimension  $n^2$ , which is the same as  $\dim(\mathbb{H}_n)$ . Therefore, the conditions for noncontextuality according to Ref. [50] are also satisfied.

For cases in which  $k_{\max} < n^2$ , the situation is not so straightforward (note that there is no need to consider  $k_{\max} > n^2$ , as it is impossible to have as many affinely independent states). We can still think of the restricted state space  $\Delta \subset \mathcal{D}(\mathcal{H}_{B_j})$  but now  $\mathbb{H}_n$  is strictly larger than  $\text{span}(\Delta_{k_{\max}})$ . This leads to some technical problems, e.g.,  $\Delta_{k_{\max}}^*$  would not be a compact set on  $\mathbb{H}_n$ . There are a few possibilities to try to dodge these problems; for instance, one could see the impact of considering nontomographically complete GPTs, as introduced in Ref. [60]. This, however, is beyond the scope of this work and will be addressed in the future.

- 
- [1] J. S. Bell, On the problem of hidden variables in quantum mechanics, *Rev. Mod. Phys.* **38**, 447 (1966).
  - [2] J. S. Bell, On the Einstein Podolsky Rosen paradox, *Phys. Phys. Fizika* **1**, 195 (1964).
  - [3] D. Jennings and M. Leifer, No return to classical reality, *Contemp. Phys.* **57**, 60 (2016).

- [4] K. T. Goh, J. M. K. Kaniewski, E. Wolfe, T. Vértesi, X. Wu, Y. Cai, Y.-C. Liang, and V. Scarani, Geometry of the set of quantum correlations, *Phys. Rev. A* **97**, 022104 (2018).
- [5] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 419 (2014).
- [6] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, *Nat. Phys.* **8**, 475 (2012).
- [7] M. S. Leifer, Is the quantum state real? An extended review of  $\Psi$ -ontology theorems, *Quanta* **3**, 67 (2014).
- [8] R. W. Spekkens, The ontological identity of empirical indiscernibles: Leibniz’s methodological principle and its significance in the work of Einstein, [arXiv:1909.04628](https://arxiv.org/abs/1909.04628) [physics.hist-ph] (2019).
- [9] R. W. Spekkens, Negativity and Contextuality Are Equivalent Notions of Nonclassicality, *Phys. Rev. Lett.* **101**, 020401 (2008).
- [10] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, *Phys. Rev. A* **71**, 052108 (2005).
- [11] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, A structure theorem for generalized-noncontextual ontological models, [arXiv:2005.07161](https://arxiv.org/abs/2005.07161) [quant-ph] (2020).
- [12] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, *Rev. Mod. Phys.* **75**, 715 (2003).
- [13] W. H. Zurek, Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? *Phys. Rev. D* **24**, 1516 (1981).
- [14] W. H. Zurek, Relative states and the environment: Einselection, enviance, quantum Darwinism, and the existential interpretation, [arXiv:0707.2832](https://arxiv.org/abs/0707.2832) [quant-ph] (2007).
- [15] F. G. S. L. Brandão, M. Piani, and P. Horodecki, Generic emergence of classical features in quantum Darwinism, *Nat. Commun.* **6**, 7908 (2015).
- [16] R. Horodecki, J. K. Korbicz, and P. Horodecki, Quantum origins of objectivity, *Phys. Rev. A* **91**, 032122 (2015).
- [17] T. P. Le and A. Olaya-Castro, Strong Quantum Darwinism and Strong Independence are Equivalent to Spectrum Broadcast Structure, *Phys. Rev. Lett.* **122**, 010403 (2019).
- [18] P. Contreras-Tejada, G. Scarpa, A. M. Kubicki, A. Brandenburger, and P. L. Mura, Agreement between observers: A physical principle? [arXiv:2102.08966](https://arxiv.org/abs/2102.08966) [quant-ph] (2021).
- [19] Given the dependence on the observers’ experience, this notion has also been called “intersubjectivity” rather than objectivity [61]. Nevertheless, we will use the term “objectivity,” as often used in the quantum Darwinism literature.
- [20] H. Ollivier, D. Poulin, and W. H. Zurek, Environment as a witness: Selective proliferation of information and emergence of objectivity in a quantum universe, *Phys. Rev. A* **72**, 042112 (2005).
- [21] C. Bény, Unsharp pointer observables and the structure of decoherence, [arXiv:0802.0685](https://arxiv.org/abs/0802.0685) [quant-ph] (2008).
- [22] M. Zwolak and W. H. Zurek, Redundancy of einselected information in quantum Darwinism: The irrelevance of irrelevant environment bits, *Phys. Rev. A* **95**, 030101 (2017).
- [23] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).
- [24] E. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.* **40**, 749 (1932).
- [25] J. Barrett, Information processing in generalized probabilistic theories, *Phys. Rev. A* **75**, 032304 (2007).
- [26] P. Janotta and H. Hinrichsen, Generalized probability theories: What determines the structure of quantum theory? *J. Phys. A: Math. Theor.* **47**, 323001 (2014).
- [27] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, Characterization of noncontextuality in the framework of generalized probabilistic theories, *PRX Quantum* **2**, 010331 (2021).
- [28] 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).
- [29] R. Kunjwal and R. W. Spekkens, From the Kochen-Specker Theorem to Noncontextuality Inequalities without Assuming Determinism, *Phys. Rev. Lett.* **115**, 110403 (2015).
- [30] R. Kunjwal and R. W. Spekkens, From statistical proofs of the Kochen-Specker theorem to noise-robust noncontextuality inequalities, *Phys. Rev. A* **97**, 052110 (2018).
- [31] R. Kunjwal, Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements, *Quantum* **3**, 184 (2019).
- [32] R. Kunjwal, Hypergraph framework for irreducible non-contextuality inequalities from logical proofs of the Kochen-Specker theorem, *Quantum* **4**, 219 (2020).
- [33] D. Schmid, R. W. Spekkens, and E. Wolfe, All the non-contextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences, *Phys. Rev. A* **97**, 062103 (2018).
- [34] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, *J. Math. Mech.* **17**, 59 (1967).
- [35] 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).
- [36] D. Saha and A. Chaturvedi, Preparation contextuality as an essential feature underlying quantum communication advantage, *Phys. Rev. A* **100**, 022108 (2019).
- [37] S. Ghorai and A. Pan, Optimal quantum preparation contextuality in an n-bit parity-oblivious multiplexing task, *Phys. Rev. A* **98**, 032110 (2018).
- [38] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, Parity oblivious  $d$ -level random access codes and class of noncontextuality inequalities, *Quantum Inf. Process.* **18**, 1 (2019).
- [39] D. Saha, P. Horodecki, and M. Pawłowski, State independent contextuality advances one-way communication, *New J. Phys.* **21**, 093057 (2019).
- [40] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, Optimal bounds for parity-oblivious random access codes, *New J. Phys.* **18**, 045003 (2016).
- [41] D. Schmid and R. W. Spekkens, Contextual Advantage for State Discrimination, *Phys. Rev. X* **8**, 011015 (2018).
- [42] M. Lostaglio and G. Senno, Contextual advantage for state-dependent cloning, *Quantum* **4**, 258 (2020).
- [43] M. Lostaglio, Certifying Quantum Signatures in Thermodynamics and Metrology via Contextuality of Quantum Linear Response, *Phys. Rev. Lett.* **125**, 230603 (2020).

- [44] X. Gao, E. R. Anschuetz, S.-T. Wang, J. I. Cirac, and M. D. Lukin, Enhancing Generative Models via Quantum Correlations, [arXiv:2101.08354](#) (2021).
- [45] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey, The only noncontextual model of the stabilizer subtheory is Gross's, [arXiv:2101.06263](#) (2021).
- [46] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Contextuality supplies the “magic” for quantum computation, *Nature* **510**, 351 (2014).
- [47] R. Raussendorf, Contextuality in measurement-based quantum computation, *Phys. Rev. A* **88**, 022322 (2013).
- [48] S. Mansfield and E. Kashefi, Quantum Advantage from Sequential-Transformation Contextuality, *Phys. Rev. Lett.* **121**, 230401 (2018).
- [49] S. Abramsky, R. S. Barbosa, and S. Mansfield, Contextual Fraction as a Measure of Contextuality, *Phys. Rev. Lett.* **119**, 050504 (2017).
- [50] F. Shahandeh, Contextuality of general probabilistic theories, *PRX Quantum* **2**, 010330 (2021).
- [51] F. Galve, R. Zambrini, and S. Maniscalco, Non-Markovianity hinders quantum Darwinism, *Sci. Rep.* **6**, 19607 (2016).
- [52] S. M. Oliveira, A. L. de Paula, and R. C. Drumond, Quantum Darwinism and non-Markovianity in a model of quantum harmonic oscillators, *Phys. Rev. A* **100**, 052110 (2019).
- [53] C. Duarte and B. Amaral, Resource theory of contextuality for arbitrary prepare-and-measure experiments, *J. Math. Phys.* **59**, 062202 (2018).
- [54] R. Wagner, R. D. Baldijão, A. Tezzin, and B. Amaral, Using a resource theoretic perspective to witness and engineer quantum generalized contextuality for prepare-and-measure scenarios, [arXiv:2102.10469](#) [quant-ph] (2021).
- [55] B. Baumgartner, An inequality for the trace of matrix products, using absolute values, [arXiv:1106.6189](#) [math-ph] (2011).
- [56] J. H. Gallier, Properties of convex sets: A glimpse, Lecture Notes.
- [57] J. A. Bergou, U. Herzog, and M. Hillery, in *Quantum State Estimation*, edited by M. Paris and J. Řeháček (Springer-Verlag, Berlin, 2004), p. 417.
- [58] The use of the label  $r$  for the number of subsystems, instead of  $t$ , will be clear soon.
- [59] Actually, this is explicitly mentioned in Ref. [15].
- [60] V. Gitton and M. P. Woods, Solvable criterion for the contextuality of any prepare-and-measure scenario, [arXiv:2003.06426](#) [quant-ph] (2021).
- [61] P. Mironowicz, J. K. Korbicz, and P. Horodecki, Monitoring of the Process of System Information Broadcasting in Time, *Phys. Rev. Lett.* **118**, 150501 (2017).