Measurement incompatibility and steering are necessary and sufficient for operational contextuality

Armin Tavakoli[*] and Roope Uola[*]

Département de Physique Appliquée, Université de Genève, CH-1211 Genève, Switzerland

Contextuality is a signature of operational nonclassicality in the outcome statistics of an experiment. This notion of nonclassicality applies to a breadth of physical phenomena. Here, we establish its relation to two fundamental nonclassical entities in quantum theory; measurement incompatibility and steering. We show that each is necessary and sufficient the failure of operational contextuality. We exploit the established connection to contextuality to provide a novel approach to problems in measurement incompatibility and steering.

I. INTRODUCTION

The nonclassical nature of quantum theory has a variety of different manifestations. On the one hand, quantum theory postulates theoretical entities with properties that lack a counterpart in classical physics. On the other hand, the nonclassicality of quantum theory is also present on the observable level, i.e., in the outcome statistics of experiments. Evidently, if an experiment takes the reality of the quantum formalism for granted, every nonclassical entity of quantum theory can be experimentally detected. However, if the assumption of nature being quantum is dropped, the outcome statistics can frequently be reproduced with some classical model. Matters become more interesting when the nonclassicality of outcome statistics can be operationally determined in the spirit of device independence [1], that is, in experiments that demonstrate nonclassicality while making weak assumptions on the underlying physical nature.

The strongest form of operational inference is encountered in tests of Bell inequalities [2]. These experiments statistically analyze the correlations between the outcomes of measurements performed in space-like separated events. If the correlations violate a Bell inequality, it follows that the outcome statistics cannot be explained by *any* classical (local hidden variable) theory. Famously, by sharing entangled states and performing incompatible measurements that together steer the remote partner system, quantum theory can violate these inequalities and therefore provide an unequivocal demonstration of nonclassicality [3,4]. Surprisingly however, not all incompatible measurements, nor all steerable ensembles, enable Bell inequality violations [5–7]. This motivates the question: Is nonclassicality at the level of theoretical entities both necessary and sufficient for some form of operational nonclassicality?

We focus on two fundamental theoretical entities in quantum theory: the incompatibility of quantum measurements and the ability to steer another system by local measurements and classical communication (a feature of quantum theory originating from Schrödinger's remarks [8] on the Einstein-Podolsky-Rosen paradox [9]). These two features of quantum theory have been thoroughly researched, see, e.g., Refs. [10–13] and Refs. [14,15] respectively. We show that both measurement incompatibility and steering admit a generally valid one-to-one connection with a family of physical tasks which in turn correspond to tests of operational *contextuality*.[1] Contextuality in quantum theory has for long been researched in its own interest and is closely related to, e.g., advantages in quantum computation [19–22], advantages in particular communication tasks [23–26], quantum zero-error communication [27], and anomalous weak values [28].

The established general connection between the theoretical entities of measurement incompatibility and steering on the one hand and operational contextuality on the other, enables us to approach relevant problems in the former ones using tools originally developed for the latter. We exploit this to present a family of noncontextuality inequalities and provide numerical evidence that these are necessary and sufficient conditions for certifying the incompatibility of any set of binary qubit observables, and that they also constitute optimal tests of the steerability of a pair of qubits in a singlet state subject to noisy environments. Moreover, since our task-oriented characterization of measurement incompatibility and steering makes reasonably weak assumptions on the characterization of physical devices, such applications make possible more stringent experimental certificates of all incompatible measurements and steerable states via experimental proofs of contextuality.

---

[*]These authors have contributed equally to this work.

---

[1]As originally introduced by Bell, Kochen, and Spekker [16,17], contextuality is a property of projective measurements in quantum theory. However, the concept has seen a generalization that applies on the level of ontological models, and therefore to general operational theories used to model outcome statistics [18].

II. CONTEXTUALITY

Contextuality in an operational theory [18] is developed within the framework of ontological models [29]. An ontological model describes a preparation procedure **P** by an ontic state (hidden variable) λ with a distribution $p(\lambda|\mathbf{P})$. When a measurement procedure **M** is applied, the ontological model determines the probability of an outcome b with some response function $p(b|\mathbf{M}, \lambda)$. Therefore, the outcome statistics reads

$$p(b|\mathbf{P}, \mathbf{M}) = \sum_\lambda p(\lambda|\mathbf{P})p(b|\mathbf{M}, \lambda). \qquad (1)$$

Furthermore, ontological models are linear in the sense that convex combinations of preparation and measurement procedures are represented by convex sums of the relevant ontic state distributions and response functions respectively. See Ref. [18] for a discussion of this property.

In quantum theory, preparations are represented by density matrices ρ and measurements are positive operator-valued measures (POVMs) $M = \{M_b\}$, i.e., $M_b \geq 0$ and $\sum_b M_b = \mathbb{1}$. Outcome statistics is given by the Born rule $p(b|\rho, M) = \mathrm{tr}\,[\rho M_b]$. A state (measurement) can be realized in as many ways as it can be decomposed into mixtures of other states (decomposed into element-wise mixtures or coarse-graining of other measurements). Different ways of preparing the same state (performing the same measurement) are called contexts for ρ (M). An ontological model is said to be preparation noncontextual if the ontic state distribution is independent of the context, i.e., if $p(\lambda|\mathbf{P}) = p(\lambda|\rho)$. Similarly, an ontological model is said to be measurement noncontextual if the response functions are context independent, i.e., if $p(b|\mathbf{M}, \lambda) = p(b|M, \lambda)$. These notions embody the idea that if two laboratory procedures are indistinguishable, then they are also indistinguishable on the level of ontic states. We remark that to ensure that two procedures truly are indistinguishable, one needs to be able to perform measurements (prepare states) that span the measurement (state) space. In contrast, if outcome statistics cannot be reproduced with any preparation (measurement) noncontextual model, it is said to be preparation (measurement) contextual. See Ref. [18] for a detailed discussion of operational contextuality.

III. MEASUREMENT INCOMPATIBILITY

Measurement incompatibility [10–13] is the impossibility of jointly measuring a set of (at least two) POVMs by employing only a single measurement and classical postprocessing of its outcomes. More precisely, let $\{A_{a|x}\}$ be a set of POVMs, with a labeling the outcome and x labeling the measurement. The set is called compatible (jointly measurable) if there exists a POVM $\{G_\lambda\}$ which allows us to recover the set $\{A_{a|x}\}$ via some postprocessing probability distribution $p(a|x, \lambda)$:

$$A_{a|x} = \sum_\lambda p(a|x, \lambda)G_\lambda. \qquad (2)$$

If such a model does not exist, the set $\{A_{a|x}\}$ is called incompatible (not jointly measurable). This extends the textbook concept of commutativity in the sense that mutually commuting POVMs are jointly measurable, but the converse does not hold in general. The converse holds, however, for textbook observables, i.e., projective measurements. It is worth noting that joint measurability can be characterized as the existence of a common Naimark dilation in which the projective measurements commute.

IV. STEERING

Steering [7] is a qualitative property of some entangled quantum states regarding the set of ensembles that can be remotely prepared with local measurements and classical communication. Specifically, one considers a pair of entangled systems in state ρ and performs a set of measurements $\{A_{a|x}\}$ on the first system. Given the choice of x, this renders the second system in the state $\rho_{a|x} = \mathrm{tr}_A\,[A_{a|x} \otimes \mathbb{1}\rho]/\mathrm{tr}\,[A_{a|x} \otimes \mathbb{1}\rho]$ with probability $p(a|x) = \mathrm{tr}\,[A_{a|x} \otimes \mathbb{1}\rho]$. It is important to underline the fact that classical communication is necessary for the steered party to be able to distinguish between different local states $\rho_{a|x}$. These local states can be effectively described with a set of unnormalized states (called an assemblage) $\{\sigma_{a|x}\}$ where $\sigma_{a|x} = \mathrm{tr}_A\,[A_{a|x} \otimes \mathbb{1}\rho]$. Such assemblages are no-signaling, i.e., $\sum_a \sigma_{a|x} = \sum_a \sigma_{a|x'}$. In this work all assemblages are assumed to be no-signaling. We remark that the Gisin-Hughston-Josza-Wootters theorem [30,31] ensures that every assemblage can be prepared by a distant party's local measurements (supported by classical communication) on a properly chosen entangled state. An assemblage is said to be unsteerable if it admits a so-called local hidden state model. Such models use (a, x) as information toward a postprocessing $p(a|x, \lambda)$ of a set of local states ρ_λ appearing with probability $p(\lambda)$ to explain the assemblage $\{\sigma_{a|x}\}$. Hence, if the state is unsteerable, it can be written as

$$\sigma_{a|x} = \sum_\lambda p(\lambda)p(a|x, \lambda)\rho_\lambda. \qquad (3)$$

If no local hidden state model is possible, the assemblage is called steerable.

V. MAIN RESULTS

We begin by proving a one-to-one relation between measurement incompatibility and preparation contextuality.

Theorem 1. A set of measurements is compatible if and only if their statistics admit a preparation noncontextual model for all states.

Proof. Assume that the set of POVMs $\{A_{a|x}\}$ when applied to any quantum state ρ returns outcome statistics that is preparation noncontextual. We denote as $\mathcal{P}_\rho$ the set of preparation procedures (contexts) in which ρ can be prepared. Then, using the label x to denote the measurement procedure, it holds that

$$\forall\,\mathbf{P} \in \mathcal{P}_\rho : p(a|x, \mathbf{P}) = \sum_\lambda p(\lambda|\rho)p(a|x, \lambda). \qquad (4)$$

For each λ, the object $p(\lambda|\rho)$ is a convexity-preserving map from the space of quantum states to the interval [0,1]. The Riesz representation theorem [13,32] asserts that such maps can be written as an inner product $p(\lambda|\rho) = \mathrm{tr}\,[G_\lambda \rho]$ for some unique operator $0 \leq G_\lambda \leq \mathbb{1}$. Moreover, since $\forall\,\rho: \sum_\lambda p(\lambda|\rho) = 1$, it follows that $\sum_\lambda G_\lambda = \mathbb{1}$. Inserting this into

Eq. (4), we have

$$\forall\, \mathbf{P} \in \mathcal{P}_\rho : p(a|x, \mathbf{P}) = \sum_\lambda p(a|x, \lambda)\, \mathrm{tr}\,[G_\lambda \rho]. \qquad (5)$$

We have recovered the outcome statistics obtained from measuring ρ with a compatible set of POVMs.

Conversely, assume that $\{A_{a|x}\}$ is a compatible set of POVMs. Then, the statistics obtained from measuring any state ρ prepared with a procedure $\mathbf{P}$ is given by Eq. (5). By defining $p(\lambda|\rho) = \mathrm{tr}\,[G_\lambda \rho]$, we recover the definition of outcome statistics being preparation noncontextual. ∎

It is interesting to note that tests of preparation contextuality can be formulated as communication tasks between two separated parties, in which the receiver is kept oblivious about parts of the sender's input [23–25]. Such obliviousness corresponds to different contexts for the states. From Theorem 1, we can therefore infer the following corollary:

Corollary. Every set of incompatible measurements enables a quantum-over-classical advantage in an oblivious communication task.

We remark that the advantages of all incompatible sets of measurements have recently been shown in various measurement-device-independent communication tasks [33–36].

In a spirit similar to that of Theorem 1, we prove a one-to-one relation between steering and measurement contextuality.

Theorem 2. An assemblage is unsteerable if and only if its statistics admits a preparation and measurement noncontextual model for all measurements.

Proof. Assume that the assemblage $\{\sigma_{a|x}\}$ when measured with any POVM M returns outcome statistics that is measurement noncontextual. We denote the set of measurement procedures (contexts) in which M can be realized by $\mathcal{M}_M$. Due to assemblages being no-signaling, we have that $p(a, b|x, \mathbf{M}) = p(b|a, x, \mathbf{M})p(a|x)$ and that

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M})p(a|x)$$
$$= p(a|x)\sum_\lambda p(\lambda|a, x)p(b|M, \lambda), \qquad (6)$$

where (a, x) labels the preparation procedure. For every λ, the object $p(b|M, \lambda)$ is a map from the space of POVMs to the space of probability distributions. Such maps are characterized by the works of Gleason [37] and Busch [38]. The Gleason-Busch theorem asserts that $p(b|M, \lambda) = \mathrm{tr}\,[\rho_\lambda M_b]$ for some unique state ρ_λ. Inserting this into Eq. (6), we have

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M})p(a|x)$$
$$= p(a|x)\sum_\lambda p(\lambda|a, x)\, \mathrm{tr}\,[\rho_\lambda M_b]. \qquad (7)$$

Using Bayes' rule and the fact that x and λ are independent,[2] one straightforwardly finds that $p(a|x)p(\lambda|a, x) = p(\lambda)p(a|x, \lambda)$. Inserting this in (7), we recover the outcome
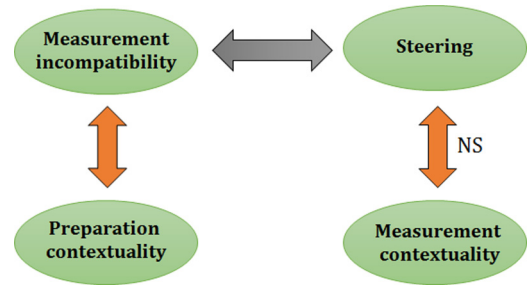

FIG. 1. Orange arrows illustrate Theorems 1 and 2. Measurement incompatibility and steering each enable a proof of a form of contextuality provided that one possesses a proper catalyst preparation or measurement procedure. Conversely, having observed preparation contextuality, one infers measurement incompatibility. Similarly, having observed measurement contextuality in a bipartite no-signaling scenario, one infers steerability of the shared state. The grey arrow indicates the previously known relation that a set of measurements is incompatible if and only if it enables steering with a proper catalyst state [39,40]. It is worth noting that this connection can also be seen as a mapping between the problems in measurement incompatibility and steerability [41].

statistics obtained from applying M to an unsteerable assemblage (3).

Conversely, if the assemblage has a local hidden state model, then for every POVM the outcome statistics reads

$$\forall\, \mathbf{M} \in \mathcal{M}_M : p(b|a, x, \mathbf{M})$$
$$= \frac{1}{p(a|x)}\sum_\lambda p(\lambda)p(a|x, \lambda)\, \mathrm{tr}\,[\rho_\lambda M_b]. \qquad (8)$$

From Bayes' rule and the independence of x and λ, we have that $p(\lambda)p(a|x, \lambda)/p(a|x) = p(\lambda|a, x)$. Note that said independence implies preparation noncontextuality. Inserted into Eq. (8) we find the outcome statistics obtained in a measurement noncontextual model. ∎

We have illustrated the theorems in Fig. 1. Notice that Theorems 1 and 2 give a characterization of the ontic variables using quantum theory. Whereas this characterization is relevant for noncontextual models covering all states or measurements, it would be interesting to see whether such characterization exists in the case of fragments of quantum theory, i.e., for noncontextual models covering subsets of states and measurements.

Also, it is worth noting that a number of works have (in different ways) shown that outcome statistics that violate a Bell inequality is proof of preparation contextuality [18,23,24,42]. In Appendix A, we note that this fact follows immediately from ontological models and the no-signaling principle (see also Ref. [43] for similar results).[3]

VI. NONCONTEXTUALITY INEQUALITIES FOR QUBIT MEASUREMENT INCOMPATIBILITY AND STEERING

We proceed to use the established connection to contextuality to address two relevant problems in measurement

---

[2]The independence of x and λ follows from the fact that λ cannot carry information about the oblivious variable x (the obliviousness comes from no-signaling), i.e., the assumption of preparation noncontextuality. See, e.g., Refs. [23,25] for an elaboration.

[3]This result was shown originally in the unpublished note Ref. [44].

incompatibility and steering: (i) Can one find a preparation noncontextuality inequality whose violation is both necessary and sufficient for certifying measurement incompatibility for interesting families of measurements? Note that despite Theorem 1 this is a nontrivial matter since any *single* test of preparation contextuality is only a sufficient condition for measurement incompatibility (and the full characterization of all tests of preparation contextuality is a demanding problem). (ii) Can one find a measurement noncontextuality inequality (in a no-signaling scenario) that for an interesting class of states optimally certifies their steerability? In analogy to the previous, this is nontrivial since any single test of measurement contextuality in a no-signaling scenario is only a sufficient condition for steering.

To answer these questions, we present a family of correlation inequalities (parametrized by an integer $n \geqslant 2$) inspired by the works of Refs. [25,45,46]. Consider a Bell-like (no-signaling) experiment in which two separated observers, Alice and Bob, share parts of a physical system. Alice (Bob) performs measurements labeled by her (his) uniformly random input $x \in \{0, 1\}^{n-1}$ ($y \in \{1, \ldots, n\}$). The outcome is denoted by $a \in \{0, 1\}$ ($b \in \{0, 1\}$). Alice's measurement procedures are constrained by operational equivalences. That is, her outcome statistics always upholds suitable indistinguishability relations which enable us to consider the statistics of different contexts for her measurements. Specifically, for every bit string $r \in \{0, 1\}^n$, we require that the measurement procedures $\mathbf{M}_{r,0}$ and $\mathbf{M}_{r,1}$ corresponding to a uniform mixing of all (a, x) satisfying $r \cdot \bar{x} = 0$ and $r \cdot \bar{x} = 1$ respectively [where $\bar{x} = (a, x + a)$], are indistinguishable from each other. In quantum theory, this means that

$$\sum_{a,x|r\cdot\bar{x}=0} M_{a|x} = \sum_{a,x|r\cdot\bar{x}=1} M_{a|x}. \quad (9)$$

Note that whenever r has an even number of ones, this condition is always satisfied since $M_{0|x} + M_{1|x} = \mathbb{1}$. For odd strings r, Eq. (9) is a nontrivial constraint. Now, let Alice and Bob play a game in which they aim to maximize the probability of finding $a + b = \bar{x}_y \mod 2$. When Alice is considered the sender of Bob's remotely prepared local states, we can consider the scenario as a test of preparation contextuality. In contrast, when Bob is considered the sender of Alice's remotely prepared local states, we can consider the scenario as a test of preparation and measurement contextuality. In the case of either being noncontextual, the average success probability is bounded by

$$\mathcal{A}_n \equiv \frac{1}{n2^{n-1}} \sum_{x,y} p(a + b = \bar{x}_y | x, y) \leqslant \frac{n+1}{2n}. \quad (10)$$

The proof of this result is a simple modification of the arguments presented in Ref. [25] and is discussed in Appendix B. A violation of the inequality (10) means that Bob's measurements (which are unconstrained) are incompatible (by Theorem 1) and that Alice's local assemblage (prepared by Bob) is steerable (by Theorem 2). We now study the usefulness of the inequality (10) for certifying qubit measurement incompatibility and two-qubit steerability.

For the case of $n = 2$ the inequality (10) reduces to the Clauser-Horne-Shimony-Holt Bell inequality [4] for which it

is known that all pairs of incompatible measurements enable a violation [47]. For $n > 2$ (specifically studying $n = 3, \ldots, 7$) we have numerically obtained support (10 000 examples for each n) for the following conjecture:

Conjecture 1. Every set of n incompatible two-outcome qubit measurements enable a proof of preparation contextuality by a violation of the inequality (10).

In Appendix C, we describe the numerical procedure employed to motivate this conjecture.

Consider now the case of steering. For simplicity, let Alice and Bob share the noisy singlet state $\rho_v = v|\psi^-\rangle\langle\psi^-| + (1 - v)\mathbb{1}/2$, where $|\psi^-\rangle = (|01\rangle - |10\rangle)/\sqrt{2}$ for some visibility $v \in [0, 1]$. What is the critical value of $v = v_n$ so that Bob can steer Alice using n projective measurements? Although this question is well studied (see, e.g., Refs. [48–50]) an analytical formula is lacking. However, Ref. [48] presented nearly matching upper and lower bounds on v_n for $n = 2, \ldots, 13$ and $n = 2, \ldots, 5$ respectively. Using our inequality (10), we have numerically implemented alternating convex searches to find an upper bound on the critical v_n (below which we can no longer find a quantum violation). This returns

$$v_2 = 0.7071, \quad v_3 = 0.5774, \quad v_4 = 0.5547,$$
$$v_5 = 0.5422, \quad v_6 = 0.5270, \quad v_7 = 0.5234. \quad (11)$$

Interestingly, these numbers coincide precisely with those presented in Ref. [48] (up to the number of decimals presented in Ref. [48]). This motivates the conjecture

Conjecture 2. The inequality (10) is a tight steering inequality for the noisy singlet state under n projective measurements.

Finding a conclusive proof of Conjectures 1 and 2 would be interesting. We remark that although the above considerations are straightforwardly analyzed with a computer, the criterion (10) can be treated in a fully analytical manner.

VII. DISCUSSION

We have shown that every set of incompatible measurements and every steerable assemblage can be operationally certified as nonclassical in a test of operational contextuality, and that the latter also implies the formers. A direct consequence is that problems of joint measurability and steering can be viewed through the lens of contextuality, as we illustrated through our conjectures. In this sense, our results bridge the two research directions of quantum measurements and quantum steering with the line of research focused on quantum contextuality.

Moreover, since tests of operational contextuality only rely on weak characterization of the experimental devices [51], our results can also be considered as semi-device-independent certificates of measurement incompatibility and steering. Naturally, fully device-independent certificates are found by violating a Bell inequality. However, in addition to such tests being experimentally demanding, it is importantly also the case that not all incompatible measurements nor all steerable ensembles violate any Bell inequality [5–7]. This makes tests of operational contextuality relevant for practical considerations when no fully device-independent certificate is either possible or known.

APPENDIX A: BELL NONLOCALITY IMPLIES PREPARATION CONTEXTUALITY

We give a simple argument for every probability distribution that violates a Bell inequality also being a proof of preparation contextuality (see also Ref. [43]). We show this immediately from ontological models supplemented with the no-signaling principle encountered in Bell inequality tests.

To see this, we write a general ontological model for a Bell experiment as

$$p(a, b|x, y) = \sum_\lambda p(a|x, y)p(\lambda|a, x)p(b|y, \lambda). \tag{A1}$$

If we also impose no-signaling, then Alice's local marginals are independent of Bob's input. Therefore,

$$p(a, b|x, y) = \sum_\lambda p(a|x)p(\lambda|a, x)p(b|y, \lambda). \tag{A2}$$

Bayes' rule together with the independence of x and λ gives that $p(a|x)p(\lambda|a, x) = p(\lambda)p(a|x, \lambda)$. Inserting this into Eq. (A2), we obtain

$$p(a, b|x, y) = \sum_\lambda p(\lambda)p(a|x, \lambda)p(b|y, \lambda). \tag{A3}$$

This is a local hidden variable model, i.e., the notion of classicality in Bell inequality tests. The assumption of preparation noncontextuality is enforced due to the assignment of the same ontic-state distribution for the preparation procedures corresponding to the remotely prepared state on Bob's side when averaged over Alice's outcomes, i.e., the principle of no-signaling. Therefore, whenever $p(a, b|x, y)$ has no local hidden variable model, it also has no preparation noncontextual model.

APPENDIX B: NONCONTEXTUAL BOUND

In the main text, we considered a scenario in which separated parties Alice and Bob share a state and perform local measurements with binary outcomes $a, b \in \{0, 1\}$. Alice's measurement settings are labeled by a bit string $x \in \{0, 1\}^{n-1}$ and Bob's measurement settings are labeled by $y \in \{1, \ldots, n\}$. Alice and Bob aim to satisfy the relation $a + b = \bar{x}_y \mod 2$ where $\bar{x} = (a, a + x)$ is an n-bit string. The notation $\bar{x}_y$ labels the yth bit in the string $\bar{x}$. Their average success probability is

$$\mathcal{A}_n \equiv \frac{1}{n2^{n-1}} \sum_{x,y} p(a + b = \bar{x}_y|x, y). \tag{B1}$$

Alice and Bob are restricted by two constraints. First, they obey the no-signaling principle. This means that the preparations of Alice on Bob's side (denoted $\mathbf{P}_{a,x}$), effectively achieved by a local measurement on her system, realize the same preparation in different contexts. That is, the following

operational equivalences hold: $\sum_a \mathbf{P}_{a,x} \sim \sum_a \mathbf{P}_{a,x'}$. The analogy holds in the other direction, i.e., by the preparations of Bob on Alice's side achieved by him locally measuring his system. Second, Alice's measurements are required to uphold certain operational equivalences. In quantum theory, these are written

$$\sum_{a,x|r\cdot\bar{x}=0} M_{a|x} = \sum_{a,x|r\cdot\bar{x}=1} M_{a|x}, \tag{B2}$$

for every n-bit string $r \in \{0, 1\}^n$ with at least two instances of $'1'$. For clarity, we give as an example the case of $n = 3$. There exists eight three-bit strings of which four have at least two instances of $'1'$. Those are $r = 011$, $r = 101$, $r = 110$, and $r = 111$. For each r we have the relation in Eq. (B2). In the case of, for example, $r = 011$ we find

$$M_{0|00} + M_{1|00} + M_{0|11} + M_{1|11}$$
$$= M_{0|01} + M_{1|01} + M_{0|10} + M_{1|10}. \tag{B3}$$

However, this is trivially satisfied since $\forall x : M_{0|x} + M_{1|x} = \mathbb{1}$. Similarly, one finds that the constraint (B2) is trivial also for $r = 101$ and $r = 110$. However, for $r = 111$ we obtain

$$M_{0|00} + M_{0|11} + M_{1|01} + M_{1|10}$$
$$= M_{0|01} + M_{0|10} + M_{1|00} + M_{1|11}, \tag{B4}$$

which is a nontrivial constraint.

Imagine now that instead of performing local measurements on a shared state, Alice directly prepares the would-have-been post-measurement states of Bob's system [labeled by the pair (a, x)] and sends them to Bob, who then measures the system and records $b \in \{0, 1\}$. This represents a prepare-and-measure scenario in which Alice has 2^n inputs (a, x) with some prior distribution $p(a, x) = p(a|x)/2^{n-1}$. Alice's preparations are required to satisfy the operational equivalence which in quantum theory reads

$$\forall r : \sum_{a,x|r\cdot\bar{x}=0} p(a|x)\rho_{a,x} = \sum_{a,x|r\cdot\bar{x}=1} p(a|x)\rho_{a,x}. \tag{B5}$$

Notice first that in the original scenario, every assemblage prepared by Alice on Bob's side can also be directly sent in this prepare-and-measure model; simply define $\rho_{a,x} = \text{tr}_A [M_{a|x} \otimes \mathbb{1}\rho]/\text{tr} [M_{a|x} \otimes \mathbb{1}\rho]$, and the prior distribution as $p(a|x) = \text{tr} [M_{a|x} \otimes \mathbb{1}\rho]$. Conversely, every ensemble that Alice can communicate to Bob in the prepare-and-measure scenario can also be realized in the original scenario via local measurements on an entangled state and classical communication. This follows from the Gisin-Hughston-Josza-Wootters theorem [30,31] and the fact that Eq. (B5) enforces a no-signaling-like preparation ensemble.

In Ref. [25] it was shown that when $p(a|x) = 1/2$, the considered prepare-and-measure scenario serves as the following test of preparation contextuality: the inequality

$$\frac{1}{n2^{n-1}} \sum_{a,x,y} p(a|x)p(b = (a, x)_y|a, x, y) \leqslant \frac{n+1}{2n} \tag{B6}$$

holds for every preparation noncontextual model. Moreover, it is a trivial modification of the arguments of Ref. [25] to show that the same bound holds regardless of the prior distribution

$p(a|x)$. Therefore, due to the connection between the prepare-and-measure scenario and the original scenario, it also holds that

$$\mathcal{A}_n \leqslant \frac{n+1}{2n} \tag{B7}$$

in a preparation noncontextual model, in which we view Alice as effectively preparing the local states of Bob.

Moreover, in the original scenario, we can equally well consider Bob as the effective sender of Alice's local states. If we impose measurement noncontextuality, the response function of Alice takes no regard of the different contexts of her measurement (related to r). Note that preparation noncontextuality is still present due to Alice and Bob being no-signaling.

APPENDIX C: NUMERICAL EVIDENCE IN SUPPORT OF CONJECTURE 1

The numerical evidence behind Conjecture 1 was obtained as follows. We used the prepare-and-measure variant (discussed in Appendix B, based on Ref. [25]) for the numerics. We sample a set of n random two-outcome qubit POVMs $\mathcal{M} = \{B_{b|y}\}_{y=1}^n$. The sampling is done by using the Bloch sphere parametrization of the most general two-outcome qubit measurement, i.e.,

$$\forall\, y : B_{0|y} = \frac{\alpha_y \mathbb{1} + \eta_y \vec{n}_y \cdot \vec{\sigma}}{2}, \tag{C1}$$

$$B_{1|y} = \frac{(2 - \alpha_y)\mathbb{1} - \eta_y \vec{n}_y \cdot \vec{\sigma}}{2} \tag{C2}$$

for some random unit vectors $\vec{n}_y$, some random numbers $\eta_y \in [0, 1]$, and some random numbers $\eta_y \leqslant \alpha_y \leqslant 2 - \eta_y$.

For the sampled $\mathcal{M}$, we evaluate the largest possible value of the witness $\mathcal{A}_n$ via a semidefinite program optimizing over the state ensemble of Alice. This returns the maximal value of $\mathcal{A}_n(\mathcal{M})$ attainable with $\mathcal{M}$. We denote the optimal ensemble returned by the semidefinite program by $\mathcal{P}$. Provided that $\mathcal{A}_n(\mathcal{M})$ violates the noncontextuality inequality (in its prepare-and-measure form), we construct new measurements $B'_{b|y} = v B_{b|y} + (1 - v)\mathbb{1}/2$ where $v \in [0, 1]$. We write $\mathcal{M}' = \{B'_{b|y}\}$. For the states $\mathcal{P}$ we have that

$$\mathcal{A}_n(\mathcal{M}', \mathcal{P}) = v\mathcal{A}_n(\mathcal{M}) + (1 - v)\mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P}). \tag{C3}$$

We choose the value of v for which $\mathcal{A}_n(\mathcal{M}', \mathcal{P})$ saturates the noncontextual bound, i.e.,

$$v = \frac{\mathcal{C}_n - \mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P})}{\mathcal{A}_n(\mathcal{M}) - \mathcal{A}_n(\{\mathbb{1}/2\}, \mathcal{P})}, \tag{C4}$$

where $\mathcal{C}_n = (n + 1)/(2n)$ is the noncontextual bound. Then, via a semidefinite program, we check whether $\mathcal{M}'$ is jointly measurable. Evidently, any perturbation of v to the positive renders $\mathcal{M}'$ incompatible since it implies a violation of the preparation noncontextuality inequality. We have repeated the procedure 10 000 times]postselected on the cases in which $\mathcal{A}_n(\mathcal{M})$ constitutes a proof of preparation contextuality] for $n = 3, 4, 5, 6, 7$ respectively. Without exception, we have found that $\mathcal{M}'$ is jointly measurable.

[1] S. Pironio, V. Scarani, and T. Vidick, Focus on device independent quantum information, New J. Phys. **18**, 100202 (2016).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics **1**, 195 (1964).

[4] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).

[5] E. Bene and T. Vértesi, Measurement incompatibility does not give rise to Bell violation in general, New J. Phys. **20**, 013021 (2018).

[6] F. Hirsch, M. T. Quintino, and N. Brunner, Quantum measurement incompatibility does not imply Bell nonlocality, Phys. Rev. A **97**, 012129 (2018).

[7] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox, Phys. Rev. Lett. **98**, 140402 (2007).

[8] E. Schrödinger, Discussion of probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[9] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete? Phys. Rev. **47**, 777 (1935).

[10] T. Heinosaari, T. Miyadera, and M. Ziman, An invitation to quantum incompatibility, J. Phys. A: Math. Theor. **49**, 123001 (2016).

[11] P. Busch, P. J. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement* (Springer, New York, 1996).

[12] P. Busch, M. Grabowski, and P. J. Lahti, *Operational Quantum Physics* (Springer, New York, 1995).

[13] P. Busch, P. Lahti, J-P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer, New York, 2016).

[14] D. Cavalcanti and P. Skrzypczyk, Quantum steering: A review with focus on semidefinite programming, Rep. Prog. Phys. **80**, 024001 (2017).

[15] R. Uola, A. C. S. Costa, H. C. Nguyen, and O. Gühne, Quantum steering, arXiv:1903.06663.

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

[17] S. Kochen and E. P. Specker, The problem of hidden variables in quantum mechanics, J. Math. Mech. **17**, 59 (1967).

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

[19] R. Raussendorf, Contextuality in measurement-based quantum computation, Phys. Rev. A **88**, 022322 (2013).

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

[21] N. Delfosse, P. A. Guerin, J. Bian, and R. Raussendorf, Wigner Function Negativity and Contextuality in Quantum Computation on Rebits, Phys. Rev. X **5**, 021003 (2015).

[22] J. Bermejo-Vega, N. Delfosse, D. E. Browne, C. Okay, and R. Raussendorf, Contextuality as a Resource for Models of Quantum Computation with Qubits, Phys. Rev. Lett. **119**, 120505 (2017).

[23] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, Communication Games Reveal Preparation Contextuality, Phys. Rev. Lett. **119**, 220402 (2017).

[24] D. Saha, and A. Chaturvedi, Preparation contextuality: The ground of quantum communication advantage? Phys. Rev. A **100**, 022108 (2019).

[25] 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).

[26] D. Schmid and R. W. Spekkens, Contextual Advantage for State Discrimination, Phys. Rev. X **8**, 011015 (2018).

[27] T. S. Cubitt, D. Leung, W. Matthews, and A. Winter, Improving Zero-Error Classical Communication with Entanglement, Phys. Rev. Lett. **104**, 230503 (2010).

[28] M. F. Pusey, Anomalous Weak Values Are Proofs of Contextuality, Phys. Rev. Lett. **113**, 200401 (2014).

[29] N. Harrigan and R. W. Spekkens, Einstein, incompleteness, and the epistemic view of quantum states, Found. Phys. **40**, 125 (2010).

[30] N. Gisin, Stochastic quantum dynamics and relativity, Helv. Phys. Acta **62**, 363 (1989).

[31] L. P. Hughston, R. Jozsa, and W. K. Wootters, A complete classification of quantum ensembles having a given density matrix, Phys. Lett. A **183**, 14 (1993).

[32] T. Heinosaari and M. Ziman, *The Mathematical Language of Quantum Theory* (Cambridge University, New York, 2011).

[33] P. Skrzypczyk, I. Šupić, and D. Cavalcanti, All Sets of Incompatible Measurements give an Advantage in Quantum State Discrimination, Phys. Rev. Lett. **122**, 130403 (2019).

[34] R. Uola, T. Kraft, J. Shang, X-D. Yu, and O. Gühne, Quantifying Quantum Resources with Conic Programming, Phys. Rev. Lett. **122**, 130404 (2019).

[35] C. Carmeli, T. Heinosaari, and A. Toigo, Quantum Incompatibility Witnesses, Phys. Rev. Lett. **122**, 130402 (2019).

[36] L. Guerini, M. T. Quintino, and L. Aolita, Distributed sampling, quantum communication witnesses, and measurement incompatibility, Phys. Rev. A **100**, 042308 (2019).

[37] A. M. Gleason, Measures on the closed subspaces of a Hilbert space, J. Math. Mech. **6**, 885 (1957).

[38] P. Busch, Quantum States and Generalized Observables: A Simple Proof of Gleason's Theorem, Phys. Rev. Lett. **91**, 120403 (2003).

[39] M. T. Quintino, T. Vértesi, and N. Brunner, Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality, Phys. Rev. Lett. **113**, 160402 (2014).

[40] R. Uola, T. Moroder, and O. Gühne, Joint Measurability of Generalized Measurements Implies Classicality, Phys. Rev. Lett. **113**, 160403 (2014).

[41] R. Uola, C. Budroni, O. Gühne, and J-P. Pellonpää, One-to-One Mapping between Steering and Joint Measurability Problems, Phys. Rev. Lett. **115**, 230402 (2015).

[42] M. S. Leifer and O. J. E. Maroney, Maximally Epistemic Interpretations of the Quantum State and Contextuality, Phys. Rev. Lett. **110**, 120401 (2013).

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

[44] J. Barrett (unpublished).

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

[46] A. Tavakoli, B. Marques, M. Pawłowski, and M. Bourennane, Spatial versus sequential correlations for random access coding, Phys. Rev. A **93**, 032336 (2016).

[47] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Measurements Incompatible in Quantum Theory Cannot Be Measured Jointly in Any Other No-Signaling Theory, Phys. Rev. Lett. **103**, 230402 (2009).

[48] J. Bavaresco, M. T. Quintino, L. Guerini, T. O. Maciel, D. Cavalcanti, and M. T. Cunha, Most incompatible measurements for robust steering tests, Phys. Rev. A **96**, 022110 (2017).

[49] R. Uola, K. Luoma, T. Moroder, and T. Heinosaari, Adaptive strategy for joint measurements, Phys. Rev. A **94**, 022109 (2016).

[50] S. J. Jones and H. M. Wiseman, Nonlocality of a single photon: Paths to an Einstein-Podolsky-Rosen-steering experiment, Phys. Rev. A **84**, 012110 (2010).

[51] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unphysical idealizations, Nat. Commun. **7**, 11780 (2016).