

# Operational nonlocality

R. Srikanth<sup>1,\*</sup>

<sup>1</sup>*Poornaprajna Institute of Scientific Research, Bengaluru 560 080, India*

An operational concept of locality whose quantum violation is indicated independently of any other assumption(s) seems to be lacking in the quantum foundations literature so far. Bell’s theorem only shows that quantum correlations violate the conjunction of the ontological assumptions of localism and determinism. Taking a cue from computational complexity theory, here we define such a concept of locality in terms of a class of operational decision problems, and propose that Einstein-Podolsky-Rosen (EPR) steering in its basic asymmetric formulation gives a specific realization of this class. Various nonclassical (convex) operational theories, including quantum mechanics, are shown to be nonlocal according to this operational criterion. We discuss several ramifications arising from this result. It indicates a basic difference between quantum no-signaling and relativistic no-signaling, and suggests an information theoretic derivation of the former from other basic principles in the convex framework. It elucidates the connection between two different bounds on nonlocality, due to Oppenheim and Wehner (2010) and Banik et al. (2015), thereby highlighting the interplay and contrast between uncertainty and measurement incompatibility. Our result provides an argument supporting the realist interpretation of the quantum state, and helps clarify why Spekkens’ toy model (2007) features steering but not Bell nonlocality.

*Introduction.* Is there an operational criterion by which a measurement on particle  $A$  can be said to disturb a distant quantum-correlated particle  $B$ , independently of other assumptions? While quantum information processing tasks such as remote state preparation [1, 2], quantum teleportation [3] and remote steering [4] are suggestive of such a disturbance (one that, in the last case, Schrödinger found disconcerting [5]), yet by virtue of intrinsic randomness in measurement outcomes, which enforces no-signaling, the reduced state of  $B$  is unaffected by local actions at  $A$ . This obfuscates the issue of whether the remote disturbance of the state of  $B$  is indeed an operational phenomenon (i.e., one that refers to operational quantities only), or merely an influence felt in a HV model reproducing the experimental observations, or yet again, just an epistemic update conditioned on Alice’s measurement. It has been the subject of some debate (cf. [6]) since the historic paper by Einstein, Podolsky and Rosen [7].

In this context, Bell’s theorem [8] only shows that quantum correlations violate the conjunction of the ontological assumptions of localism and determinism, and implies nothing about localism by itself. The theorem can be formulated as the complementarity between randomness ( $I$ ) and signaling ( $S$ ) resources required for simulating a violation of the CHSH inequality [9], given by:

$$S + 2I \geq C, \quad (1)$$

where  $C$  is the average communication cost in bits [10, 11] (generalizing earlier results reported in [12, 13]).

In a deterministic HV model ( $I = 0$ ), or even certain predictively superior ( $0 < I < \frac{C}{2}$ ) HV models, nonlocal correlations ( $C > 0$ ) imply a disturbance at the HV level ( $S > 0$ ). The operational implication of Bell’s theorem is this: since special relativity and the structure of QM impose signal locality ( $S = 0$ ), given nonlocal cor-

relations ( $C > 0$ ), from Eq. (1), we infer non-vanishing operational randomness, i.e., unpredictability ( $I > 0$ )—a conclusion reached differently in [6].

As far as we know, there is thus far in the literature, no operational concept of locality that captures the notion of disturbance referenced in the question posed at the start of this Section. In this article, we propose an affirmative response to the question. Taking a leaf from computational complexity theory’s book, we identify this concept with a class of operationally decidable problems, and propose that quantum steering realizes this class. Conceptual ramifications of our result include making a distinction between quantum and relativistic no-signaling; and making a case for no-signaling being a derived—rather than basic—principle in the framework of generalized probability theories (GPTs) [14, 15], or convex operational theories (Appendix I). Quantum mechanics (QM), formulated as an operational theory, is a special case in this framework. We also show that our result indicates a rich interplay between uncertainty [16], measurement incompatibility [17], steering, nonlocality [18] and no-signaling. Our starting point takes a cue from computer science.

*Detectable vs. verifiable disturbance.* Two major computational complexity classes are  $\mathbf{P}$ , the set of all decision problems that are quickly (i.e., in time that scales polynomially with problem input size) solvable, and  $\mathbf{NP}$ , the set of problems that are quickly verifiable (against a certificate). Obviously,  $\mathbf{P} \subseteq \mathbf{NP}$ , but whether the containment is strict is the  $\mathbf{P}$  versus  $\mathbf{NP}$  problem, a major open problem in computer science [19].

Here we consider an analogous use of the concept of verification against a certificate to capture the idea that although the putative remote disturbance is not detectable (thanks to no-signaling), it may be checkable *a posteriori* in an operational way. In the context of

bipartite correlations, Bob’s measurement is considered to be the analogue of polynomial-time computation, and his guessing of Alice’s measurement choice with better-than-random probability based on his outcome, as the analogue of problem solving. The analogues of decision problems are operational theories, and those of problem instances are individual (bi-partite) states.

Accordingly, **Det**— the class of signaling theories, i.e., ones wherein Bob can (with better than random probability) unilaterally determine Alice’s input— is analogous to **P** (or **BPP**, but this distinction isn’t significant here). The analog of **NP** would be the class of theories wherein Bob can operationally verify (through some procedure  $\mathcal{P}$ ) Alice’s action against a certificate subsequently issued by her, which (class) we designate **Ver**. Clearly, **Det**  $\subseteq$  **Ver**.

If (over many trials) a state  $\omega_{AB}$  can pass the verification test  $\mathcal{P}$  for some measurement setting(s), then  $\omega_{AB}$  is deemed *operationally nonlocal*. A theory that features any such state belongs to **Ver**. Since remote disturbance is trivial to operationally verify in signaling theories, our first concern is to identify  $\mathcal{P}$  suitable for “**Ver**-complete theories”— i.e., non-signaling theories in **Ver**. In what follows, GPTs are implicitly assumed to be non-signaling, unless explicitly stated to be otherwise.

Theories in  $\overline{\mathbf{Ver}}$ , the complement of **Ver** in the space of all GPTs, are, by definition, operationally local (or, “Einstein local”). They correspond to the classical (and strongest) notion of locality. By contrast, theories in  $\overline{\mathbf{Det}} - \overline{\mathbf{Ver}}$  are local in a weaker sense. They violate operational locality and correspond to the quantum or nonclassical notion of locality.

In general, theories in **Det** will lack a tensor product structure, and be higher-dimensional than their non-signaling counterparts in **Ver**. For example, in the context of theories of finite-input-finite-output bipartite correlations  $P(ab|xy)$ , where  $x \in \mathcal{X}$  and  $a \in \mathcal{A}$  (resp.,  $y \in \mathcal{Y}$  and  $b \in \mathcal{B}$ ) are the inputs and outputs of Alice (resp., Bob), a theory in **Det** is larger dimensional than its non-signaling counterpart by (Appendix II):

$$C_{\text{nosig}} = |\mathcal{X}|(|\mathcal{Y}| - 1)(|\mathcal{A}| - 1) + |\mathcal{Y}|(|\mathcal{X}| - 1)(|\mathcal{B}| - 1),$$

the number of independent no-signaling constraints.

Being signaling and hence incompatible with special relativity, theories in **Det** will be considered unphysical. However, as would be clear at the end of this Article, it is advantageous to widen the GPT framework to include them, in order to better understand what principles single out QM as a special non-signaling theory in Nature.

*Operational nonlocality.* Alice and Bob live in a world whose physical laws are governed by GPT  $\theta$ . They have access to well characterized state preparations and measurements, and cooperatively implement the following steering-inspired realization of the verification test  $\mathcal{P}$ . Other possible realizations are discussed later.

They agree on two dichotomic incompatible measurements of Bob,  $y = y_0$  and  $y = y_1$ , with a single-system

uncertainty relation. Let  $q(y, \omega)$  denote the probability  $\max_b P(b|y, \omega)$  for an arbitrary (mixed) state  $\omega$  with Bob. A nontrivial uncertainty exists if, for any given state  $\omega$

$$q(y_0) + q(y_1) \leq v_{\text{loc}}^*. \quad (2)$$

and  $v_{\text{loc}}^* < 2$ . As an example, in QM, let  $y_0 \equiv \sigma_X$  and  $y_1 \equiv \sigma_Z$ . Then,  $v_{\text{loc}}^* = 1 + \frac{1}{\sqrt{2}} \approx 1.7$ , with the optimal states attaining this bound being eigenstates of  $\frac{1}{\sqrt{2}}(\sigma_X \pm \sigma_Z)$ . Suppose Alice prepares and sends one of these states. If Bob announces one of  $y_0$  and  $y_1$  randomly, then she can predict Bob’s outcome with probability  $\frac{v_{\text{loc}}^*}{2}$ .

Instead, suppose Alice prepares a joint state  $\sum_{\lambda} \mathbf{p}(\lambda) \omega_{AB}^{\lambda}$ , with  $\sum_{\lambda} \mathbf{p}(\lambda) = 1$ , and sends particle  $B$  to Bob. After Bob announces  $y_j$ , Alice measures a corresponding  $x_j$  on particle  $A$ . To prove that her action remotely disturbed Bob’s state, Alice sends Bob a classical certificate predicting Bob’s outcome  $b_j$ . Then, Bob measures  $y$  and checks her claim. Over many runs, he determines the uncertainty conditioned on her certificate.

The assemblage of unnormalized states of Bob  $\{\tilde{\omega}_B^{a|x}\}$  (the tilde indicating non-normalization) produced if Alice measures  $x$  obtaining outcome  $a$  has the general decomposition

$$\tilde{\omega}_B^{a|x} = \sum_{\lambda} \mathbf{p}(\lambda) p(a|x, \lambda) \omega_B^{a|x, \lambda}, \quad (3)$$

where  $p(a|x, \lambda) = [e^{a|x} \otimes u_B](\omega_{AB}^{\lambda})$ ,  $e^{a|x}$  is Alice’s effect, and  $u_B$  is the identity operation on Bob’s particle. If  $\omega_{AB}^{\lambda}$  has the product form  $\omega_A^{\lambda} \otimes \omega_B^{\lambda}$ , Eq. (3) reduces to:

$$\tilde{\omega}_B^{a|x} = \sum_{\lambda} \mathbf{p}(\lambda) p(a|x, \lambda) \omega_B^{\lambda}, \quad (4)$$

where states  $\omega_B^{\lambda}$  are classically correlated with  $A$ , with probability  $p(a|x, \lambda)$  reducing to  $e^{a|x}(\omega_A^{\lambda})$ . In contrast to Eq. (3), the form Eq. (4) defines an *unsteerable* assemblage of local hidden states at  $B$  [20] in the context of GPTs [21]. In this case, conditioning on outcome ( $a|x$ ) doesn’t reduce Alice’s uncertainty about Bob’s outcomes. Therefore, the conditioned uncertainty relation

$$q(y_0|x_0) + q(y_1|x_1) \leq v_{\text{loc}}^*, \quad (5)$$

which reduces to the uncertainty relation (2), must hold. Here,  $q(y_j|x_j) \equiv \sum_{\lambda, a_j} \mathbf{p}(\lambda) p(a_j|x_j, \lambda) q(y_j, \omega_B^{a_j|x_j, \lambda})$ . Thus, a violation of inequality Eq. (5) can arise only from a steerable (from Alice to Bob) assemblage of form Eq. (3). By convexity, to produce a violation,  $\lambda$  may be fixed to be some optimal state.

The following purely operational facts demonstrate a kind of nonlocality without invoking assumptions about a HV model. Let  $t_{\text{prep}}$  be the time when Alice prepares in some unspecified way the state of particle  $B$ , and  $t_{\text{rec}}$  the time when she receives Bob’s message on his chosen

$y$ . (a) If Alice prepared the state of  $B$  before knowing  $y$ , then  $B$  is constrained by the bound of Eq. (5), i.e.,

$$t_{\text{prep}} \leq t_{\text{rec}} \implies \sum_j q(y_j|x_j) \leq v_{\text{loc}}^*. \quad (6)$$

(b) The violation of Eq. (5) would imply that Alice prepared Bob's state *after* knowing  $y$ :

$$\sum_j q(y_j|x_j) > v_{\text{loc}}^* \implies t_{\text{prep}} > t_{\text{rec}}, \quad (7)$$

with no assumptions made concerning Alice's preparation method. (c) If Alice prepared Bob's state by measuring  $A$  on (several copies of a given) state  $\omega_{AB}$ , and particles  $A$  and  $B$  are sufficiently far from each other, and further Eq. (5) is found violated, then it follows that she prepared his state from afar—implying the operational nonlocality of  $\omega_{AB}$ .

In other words, violation of Eq. (5) would represent evidence of Alice's remote disturbance of  $B$  purely at the operational level. By contrast, Bell nonlocality only entails an *ontic* disturbance in a class of HV models. Operational locality of  $\omega_{AB}$  is identified with the satisfaction of inequality (5) for all possible settings. Any other uncertainty-based steering inequality than Eq. (5) can equally well be used for this argument.

As a quantum realization of operational nonlocality, let  $\omega_{AB}$  be the quantum singlet state  $\frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ , and Bob announces one of  $y_0 = \sigma_X$  and  $y_1 = \sigma_Z$ . Alice measures in the announced basis  $y$ , and is able to predict his outcome with complete certainty, leading to a violation of Eq. (5), with the lhs being 2, whilst the rhs  $v_{\text{loc}}^* \approx 1.7$ .

To violate Eq. (5), Alice's observables  $x = x_0$  and  $x = x_1$  must be mutually incompatible. Suppose, to the contrary, they are jointly measurable. Let the corresponding outcomes be  $a_0$  and  $a_1$ . By definition of joint measurability [22, 23], the outcome function  $p(a|x, \lambda)$  should be derivable as the marginal statistics of a joint probability distribution (JD) over  $x_0$  and  $x_1$ . In this case, Eq. (3) becomes:

$$\tilde{\omega}_B^{a_i|x_i} = \sum_{\lambda} \mathfrak{p}(\lambda) \sum_{a_i \oplus 1} p(a_0, a_1|x_0, x_1, \lambda) \omega_B^{a_i|x_i, \lambda}. \quad (8)$$

where  $\oplus$  signifies addition modulo 2. No-signaling implies that  $\sum_{a_0} \tilde{\omega}_B^{a_0|x_0} = \sum_{a_1} \tilde{\omega}_B^{a_1|x_1}$ . It follows from Eq. (8) that  $\sum_{a_0, a_1, \lambda} p(a_0, a_1|x_0, x_1, \lambda) (\omega_B^{a_0|x_0, \lambda} - \omega_B^{a_1|x_1, \lambda}) = 0$  for all states  $\omega_{AB}$ . This can hold true in general only if for either  $a_i$ ,  $\omega_B^{a_i|x_i, \lambda} = \omega_B^\lambda$ , i.e., the assemblage Eq. (8) takes the unsteerable form Eq. (4). In the inverse, if the  $a_j$ 's are incompatible, then clearly in general the unsteerable form Eq. (8) holds (cf. [21, 24–26]). It is straightforward to extend this argument to more than two pairs of observables.

A subtlety here is that the JD  $p(a_0, a_1|x_0, x_1, \lambda)$  may exist even though the joint measurement isn't part of the

operational theory. An example is Spekkens' toy theory [27]. This peculiarity due to the non-convexity of the theory (see Appendix III). In this case, we deem the measurements to be incompatible, but “meta-compatible” (Appendix IV). Such a theory supports EPR steering, despite being local.

In a non-signaling theory, Alice's ability to predict Bob's outcome is unaffected even if Bob measures before announcing  $y$ . Then, by an argument similar to that used above to show the incompatibility of the  $x_j$ 's, one can show that  $y_0$  and  $y_1$  must be mutually incompatible to violate Eq. (5) (cf. [28]).

The above observations suggest that operational nonlocality is mathematically equivalent to steerability, although the corresponding protocols have different objectives. (In the EPR steering scenario, performing unknown-to-Bob measurements, Alice aims to convince Bob, who has full control over his quantum measurements, that her state is entangled with his.) Indeed, Eq. (5) can be considered as a steering inequality, analogous to entropic or fine-grained steering inequalities [29–32], but based on a different quantification of uncertainty and applicable to any operational theory, not just QM. The above mentioned equivalence indeed holds good in QM, and therefore, within QM, operational nonlocality is strictly weaker than Bell nonlocality. However, the equivalence fails in the space of all non-signaling GPTs. In particular, PR box world [33, 34], which allows perfect steering, as evidenced by its maximal Bell nonlocality [35], lacks operational nonlocality because pure qubits lack uncertainty, i.e.,  $v_{\text{loc}}^* = 1$  (Appendix III).

Class **Ver** can be realized as the set of theories that are operationally nonlocal in the sense of allowing the violation of Eq. (5) for some state(s). Evidently, QM belongs to **Ver**-complete and so does Spekkens' (local) toy theory. But PR box world, for the reason stated above, doesn't, though this may not be the case in other realization of **Ver**. See Appendix III for further discussion on the status of operational nonlocality in various theories.

*Uncertainty, incompatibility, steering and nonlocality.* Our realization of operational nonlocality based on uncertainty fails to distinguish distinct theories having the same level of uncertainty bound  $v_{\text{loc}}$ , despite their distinct steering capabilities, e.g., classical theory and PR box world. By contrast, Bell nonlocality obviously distinguishes these two theories.

Indeed, using a bound on the violation of the CHSH inequality from measurement incompatibility [36], and a relationship between uncertainty and incompatibility for a family of GPTs (Appendices V, VI), we obtain the Bell-CHSH inequality

$$|\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle| \leq 4\varsigma(v_{\text{loc}}^* - 1), \quad (9)$$

featuring a different bound in terms of uncertainty and steering strength  $\varsigma$ , than that obtained in [35] in terms of fine-grained uncertainty and steering. This contrasts

with the case of incompatibility, which can bound Bell nonlocality by itself [37, 38]. The reason is that whereas incompatibility directly relates to the (in)existence of a joint probability distribution over all inputs, uncertainty requires to be supplemented by steering strength (Appendix VI), as suggested by our above discussion.

In Eq. (9),  $\frac{1}{2} \leq \zeta \leq 1$ , with the extremes representing classical and PR box theories, respectively. Putting  $v_{\text{loc}}^* = 2$  and  $1 + \frac{1}{\sqrt{2}}$  in Eq. (9), with  $\zeta = 1$ , we obtain the Bell-CHSH inequality for PR boxworld and QM, respectively, whilst putting  $v_{\text{loc}}^* = 2$  with  $\zeta = \frac{1}{2}$ , we obtain that for classical theory.

*Reality and relativity.* In a quantum violation of inequality Eq. (5), Alice’s measurement and her remote-preparation of Bob’s state are operationally well defined and evidently spacelike-separated. This can be experimentally tested, but is already implicit in loophole-free tests of quantum steering such as [39]. Yet, there can’t be any operational mechanism causally linking these two events, QM being non-signaling. On the other hand, there is an asymmetry and natural causal ordering in the experiment—namely, that Alice’s measurement results in Bob’s particle’s remote preparation, rather than the other way round. (By contrast, in a Bell test, which is symmetric, the case for such intrinsic causal ordering is less compelling.) The only operational element available encompassing both events in spacetime is the probability field corresponding to the quantum wave function,  $\psi(x)$ . Minimalistically, we attribute to this field itself the capacity to act as a kind of atemporal causal matrix (cf. [40]). The quantum state is arguably real to possess such causal efficacy. A similar attribution of reality can be made for states in any GPT in **Ver**-complete.

In light of operational nonlocality, it is natural to ask why QM, a theory in **Ver**, isn’t found in **Det**. To clarify this question, we note that in complexity theory, there are sound mathematical grounds to believe that intractable problems exist, and therefore that  $\mathbf{P} \neq \mathbf{NP}$  [19]. In other words, there seems to be a computational barrier separating **P** and **NP**. If QM lay outside **Ver** (like classical theory), then the above query wouldn’t be well motivated. To rephrase the question: is there a natural barrier that precludes theories in **Ver** from living in **Det**?

An obvious response would be to invoke the special relativistic prohibition on superluminal signaling. But, the fact is that quantum and classical relativistic non-signaling are fundamentally different. The former corresponds to **Det** – **Ver** and is a consequence of the tensor product structure of multipartite quantum systems, whereas the latter corresponds to **Ver** and is an axiom of spacetime geometry. Indeed, even non-relativistic QM is non-signaling. In other words, one would like to look for an *information theoretic* basis for quantum no-signaling in the GPT framework. This would potentially provide a more natural justification for the **Ver/Det** barrier.

An argument, that we elaborate elsewhere, is the fol-

lowing. It is known that no-signaling can be used to derive various no-go theorems for quantum cloning [41], state discrimination [42], etc. Here, we wish to invert this argument in the context of GPTs. Let us consider a simple 2-input-2-output scenario that illustrates the basic intuition. Correlation  $P(a, b|x, y)$  is nonlocal iff the probability  $p_{\text{success}}$  to satisfy the CHSH condition  $a + b = xy \bmod 2$  exceeds  $\frac{3}{4}$ , with  $x, y, a, b \in \{0, 1\}$ . Assume that Alice and Bob share a PR box state  $|PR\rangle_{AB}$ . If superluminal signaling were possible, then Bob could access Alice’s input  $x$  and prepare a second qubit  $B'$  in the state  $|b' = a + xy\rangle_{B'}$ , with  $a$  obtained via pre-shared randomness. Thus, particle  $B'$  would be a clone of  $B$ , which would contravene the local no-cloning rule for qubits.

More generally, let spaces  $\Omega_A$  and  $\Omega_B$  of GPT  $\theta$  be non-simplicial (and thus have nonclassical features such as no-cloning [43–46]), and joint space  $\Omega_{AB}$  contain elements outside the minimal tensor product. The above argument suggests that if theory  $\theta$  admits signaling, then it would lack a natural trace operation on  $\Omega_{AB}$  to recover the subsystem nonclassicality.

*Conclusions.* An operational concept of locality stronger than signal locality is formulated in terms of a class of operational decision problems, and EPR steering is proposed to realize it. This possibility is particularly tied to its asymmetric character, which Bell nonlocality lacks. Quantum mechanics, and some other operational theories, are shown to be nonlocal according to this operational criterion.

As regards practical demonstration, evidently experimental tests of uncertainty reduction via steering can be adapted to a test of operational nonlocality. Other phenomena, such as those mentioned in the Introduction as evocative of remote disturbance, could be used to propose other versions of operational nonlocality. Also, the argument for the *insecurity* of quantum bit commitment [47–49] can presumably be reconstructed as a criterion for operational locality, indeed one that, unlike the present one, identifies PR box world as operationally nonlocal.

Steering corresponds to strong correlations between incompatible observables of two particles, leading to the violation of the single-system uncertainty relation. Einstein *et al.* [7] believed that this violation indicated QM to be incomplete, since the alternative would be a non-local influence that they (wrongly) deemed contradicted by relativity. Here, by grounding this nonlocality in an operational setting, we find this alternative inevitable. In a sense, perhaps we have simply only drawn attention to “an elephant in the room”.

The author thanks S. Aravinda, U. Shrikant and N. Vinod for discussions, and AMEF, Bengaluru and DST-SERB, Govt. of India (project EMR/2016/004019) for financial support.

- 
- \* srik@poornaprajna.org
- [1] Arun K. Pati, “Minimum classical bit for remote preparation and measurement of a qubit,” *Phys. Rev. A* **63**, 014302 (2000).
  - [2] Charles H Bennett, Patrick Hayden, Debbie W Leung, Peter W Shor, and Andreas Winter, “Remote preparation of quantum states,” *IEEE Transactions on Information Theory* **51**, 56–74 (2005).
  - [3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, “Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels,” *Phys. Rev. Lett.* **70**, 1895–1899 (1993).
  - [4] Daniel Cavalcanti and Paul Skrzypczyk, “Quantum steering: a review with focus on semidefinite programming,” *Reports on Progress in Physics* **80**, 024001 (2016).
  - [5] E Schrödinger, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 31 (1935) pp. 555–563.
  - [6] E. Cavalcanti and H. Wiseman, “Bell nonlocality, signal locality and unpredictability (or what Bohr could have told einstein at solvay had he known about bell experiments),” *Found. Phys.* **42**, 1329 (2012).
  - [7] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?” *Phys. Rev.* **47**, 777–780 (1935).
  - [8] John S. Bell, “On the problem of hidden variables in quantum mechanics,” *Rev. Mod. Phys.* **38**, 447–452 (1966).
  - [9] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt, “Proposed experiment to test local hidden-variable theories,” *Phys. Rev. Lett.* **23**, 880–884 (1969).
  - [10] S. Aravinda and R. Srikanth, “Complementarity between signalling and local indeterminacy in quantum nonlocal correlations,” *Quant. Inf. Comp.* **15**, 308–315 (2015).
  - [11] S Aravinda and R Srikanth, “Extending quantum mechanics entails extending special relativity,” *J. Phys. A: Math. Theor.* **49**, 205302 (2016).
  - [12] M. J. W. Hall, “Complementary contributions of indeterminism and signaling to quantum correlations,” *Phys. Rev. A* **82**, 062117 (2010).
  - [13] G. Kar, MD. R. Gazi, M. Banik, S. Das, A. Rai, and S. Kunkri, “A complementary relation between classical bits and randomness in local part in the simulating singlet state,” *Journal of Physics A: Mathematical and Theoretical* **44**, 152002 (2011).
  - [14] Peter Janotta and Haye Hinrichsen, “Generalized probability theories: what determines the structure of quantum theory?” *Journal of Physics A: Mathematical and Theoretical* **47**, 323001 (2014).
  - [15] Lucien Hardy, “Reformulating and reconstructing quantum theory,” arXiv preprint arXiv:1104.2066 (2011).
  - [16] S. Wehner and A. Winter, “Entropic uncertainty relations - a survey,” *NJP* **12**, 025009 (2010).
  - [17] Paul Busch, “Indeterminacy relations and simultaneous measurements in quantum theory,” *International Journal of Theoretical Physics* **24**, 63–92 (1985).
  - [18] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, “Bell nonlocality,” *Rev. Mod. Phys.* **86**, 419–478 (2014).
  - [19] Scott Aaronson, *Quantum computing since Democritus* (Cambridge University Press, 2013).
  - [20] 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).
  - [21] Manik Banik, “Measurement incompatibility and Schrödinger-Einstein-Podolsky-Rosen steering in a class of probabilistic theories,” *Journal of Mathematical Physics* **56**, 052101 (2015).
  - [22] Paul Busch and Pekka J. Lahti, “On various joint measurements of position and momentum observables in quantum theory,” *Phys. Rev. D* **29**, 1634–1646 (1984).
  - [23] Paul Busch, Teiko Heinosaari, Jussi Schultz, and Neil Stevens, “Comparing the degrees of incompatibility inherent in probabilistic physical theories,” *EPL (Europhysics Letters)* **103**, 10002 (2013).
  - [24] Marco Túlio Quintino, Tamás Vértesi, and Nicolas Brunner, “Joint measurability, einstein-podolsky-rosen steering, and bell nonlocality,” *Phys. Rev. Lett.* **113**, 160402 (2014).
  - [25] Roope Uola, Tobias Moroder, and Otfried Gühne, “Joint measurability of generalized measurements implies classicality,” *Phys. Rev. Lett.* **113**, 160403 (2014).
  - [26] H. S. Karthik, A. R. Usha Devi, and A. K. Rajagopal, “Joint measurability, steering, and entropic uncertainty,” *Phys. Rev. A* **91**, 012115 (2015).
  - [27] R. W. Spekkens, “Evidence for the epistemic view of quantum states: A toy theory,” *Phys. Rev. A* **75**, 032110 (2007).
  - [28] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M Renes, and Renato Renner, “The uncertainty principle in the presence of quantum memory,” *Nature Physics* **6**, 659 (2010).
  - [29] S. P. Walborn, A. Salles, R. M. Gomes, F. Toscano, and P. H. Souto Ribeiro, “Revealing hidden einstein-podolsky-rosen nonlocality,” *Phys. Rev. Lett.* **106**, 130402 (2011).
  - [30] James Schneeloch, Curtis J. Broadbent, Stephen P. Walborn, Eric G. Cavalcanti, and John C. Howell, “Einstein-podolsky-rosen steering inequalities from entropic uncertainty relations,” *Phys. Rev. A* **87**, 062103 (2013).
  - [31] Tanumoy Pramanik, Marc Kaplan, and A. S. Majumdar, “Fine-grained einstein-podolsky-rosen-steering inequalities,” *Phys. Rev. A* **90**, 050305 (2014).
  - [32] Priyanka Chowdhury, Tanumoy Pramanik, A. S. Majumdar, and G. S. Agarwal, “Einstein-podolsky-rosen steering using quantum correlations in non-gaussian entangled states,” *Phys. Rev. A* **89**, 012104 (2014).
  - [33] Jonathan Barrett and Stefano Pironio, “Popescu-Rohrlich correlations as a unit of nonlocality,” *Phys. Rev. Lett.* **95**, 140401 (2005).
  - [34] Jonathan Barrett, “Information processing in generalized probabilistic theories,” *Phys. Rev. A* **75**, 032304 (2007).
  - [35] Jonathan Oppenheim and Stephanie Wehner, “The uncertainty principle determines the nonlocality of quantum mechanics,” *Science* **330**, 1072–1074 (2010).
  - [36] Manik Banik, Md. Rajjak Gazi, Sibasish Ghosh, and Guruprasad Kar, “Degree of complementarity determines the nonlocality in quantum mechanics,” *Phys. Rev. A* **87**, 052125 (2013).
  - [37] Wonmin Son, Erika Andersson, Stephen M. Barnett, and M. S. Kim, “Joint measurements and Bell inequalities,” *Phys. Rev. A* **72**, 052116 (2005).
  - [38] Guruprasad Kar, Sibasish Ghosh, Sujit K Choudhary,

and Manik Banik, “Role of measurement incompatibility and uncertainty in determining nonlocality,” *Mathematics* **4**, 52 (2016).

- [39] Bernhard Wittmann, Sven Ramelow, Fabian Steinlechner, Nathan K Langford, Nicolas Brunner, Howard M Wiseman, Rupert Ursin, and Anton Zeilinger, “Loophole-free einstein–podolsky–rosen experiment via quantum steering,” *New Journal of Physics* **14**, 053030 (2012).
- [40] Nicolas Gisin, “Quantum nonlocality: how does nature do it?” *Science* **326**, 1357–1358 (2009).
- [41] N Gisin, “Quantum cloning without signaling,” *Physics Letters A* **242**, 1–3 (1998).
- [42] Sarah Croke, Erika Andersson, and Stephen M. Barnett, “No-signaling bound on quantum state discrimination,” *Phys. Rev. A* **77**, 012113 (2008).
- [43] Lucien Hardy and David D Song, “No signalling and probabilistic quantum cloning,” *Physics Letters A* **259**, 331–333 (1999).
- [44] Martin Plávala, “All measurements in a probabilistic theory are compatible if and only if the state space is a simplex,” *Physical Review A* **94**, 042108 (2016).
- [45] S Aravinda, R Srikanth, and Anirban Pathak, “On the origin of nonclassicality in single systems,” *Journal of Physics A: Mathematical and Theoretical* **50**, 465303 (2017).
- [46] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce, “Cloning and broadcasting in generic probabilistic theories,” arXiv preprint quant-ph/0611295 (2006).
- [47] D. Mayers, “Unconditionally secure quantum bit commitment is impossible,” *Phys. Rev. Lett.* **78**, 3414–3417 (1997).
- [48] Hoi-Kwong Lo and H.F. Chau, “Why quantum bit commitment and ideal quantum coin tossing are impossible,” *Physica D: Nonlinear Phenomena* **120**, 177–187 (1998).
- [49] G. M. D’Ariano, P. Perinotti, D. M. Schlingemann, and R. F. Werner, “A short impossibility proof of quantum bit commitment,” *Phys. Lett. A* **377**, 1076–1087 (2013).
- [50] Arthur Fine, “Hidden variables, joint probability, and the bell inequalities,” *Phys. Rev. Lett.* **48**, 291–295 (1982).
- [51] Liang-Liang Sun, Sixia Yu, and Zeng-Bing Chen, “Uncertainty-complementarity balance as a general constraint on non-locality,” arXiv preprint arXiv:1808.06416 (2018).
- [52] Neil Stevens and Paul Busch, “Steering, incompatibility, and Bell-inequality violations in a class of probabilistic theories,” *Phys. Rev. A* **89**, 022123 (2014).

## I. Generalized probability theories (GPTs)

In the framework of GPTs, set  $\Omega$  of states  $\omega$  of a theory is a convex subset of a real vector space  $V$ . A valid measurement or observable is any affine functional that maps any state  $\omega \in \Omega$  to a probability distribution over measurement outcomes (usually taken to be finite in number). The unnormalized states of the theory form a convex positive cone living in  $V$ .

For a bipartite system satisfying the assumptions of no-signaling and tomographic locality [15], the set  $\Omega_{AB}$  of all bipartite states lies between (according to a natu-

ral ordering) the minimal tensor product  $\Omega_A \otimes_{\min} \Omega_B$ , which denotes the collection of all possible separable states (having the form  $\sum_j p_j \omega_A^j \otimes \omega_B^j$ , where  $p_j$  is a probability distribution) and the maximal tensor product  $\Omega_A \otimes_{\max} \Omega_B$ .

## II. No-signaling

Suppose Alice’s and Bob’s correlations are described by the conditional probability distribution  $P(a, b|x, y)$ , where  $(a, x)$  (resp.,  $(b, y)$ ) are the (output, input) pair of Alice (resp., Bob), and  $a, b, x, y$  are drawn respectively from the sets  $\mathcal{A}, \mathcal{B}, \mathcal{X}, \mathcal{Y}$ .

Signal locality requires that

$$\forall_{y \neq y'} P(a|x, y) = P(a|x, y') \equiv P(a|x), \quad (10a)$$

$$\forall_{x \neq x'} P(b|x, y) = P(b|x', y) \equiv P(b|y), \quad (10b)$$

for any given state shared between Alice and Bob.

The dimension of a non-signaling theory (i.e., the number of independent real numbers required to describe an arbitrary mixed state in the theory) is:

$$D_{\text{nosig}} = |\mathcal{X}| \cdot |\mathcal{Y}| (|\mathcal{A}| - 1)(|\mathcal{B}| - 1) + |\mathcal{X}| (|\mathcal{A}| - 1) + |\mathcal{Y}| (|\mathcal{B}| - 1) \quad (11)$$

which can be derived as follows. The last two terms in the RHS of Eq. (11) come from the marginal probability distributions  $P(a|x)$  and  $P(b|y)$  in Eq. (10), whilst the first term in the RHS comes from the fact that for each of the  $|\mathcal{X}| \cdot |\mathcal{Y}|$  pairs of two-party inputs, there are  $(|\mathcal{A}| - 1)(|\mathcal{B}| - 1)$  independent output pairs.

In Eq. (10), consider the top equation (10a). For each input  $x = X$ , given output  $a$ , there are  $|\mathcal{Y}| - 1$  independent constraints by varying  $y$ , and then  $|\mathcal{A}| - 1$  independent values to set  $a$  (minus 1 for normalization). This gives a total of  $|\mathcal{X}| (|\mathcal{Y}| - 1)(|\mathcal{A}| - 1)$  no-signaling constraints. Repeating the similar exercise for Eq. (10b), in all there are

$$C_{\text{nosig}} = |\mathcal{X}| (|\mathcal{Y}| - 1)(|\mathcal{A}| - 1) + |\mathcal{Y}| (|\mathcal{X}| - 1)(|\mathcal{B}| - 1)$$

constraints (see main text).

For theories in **Det**, relaxing the no-signaling constraints (10), one obtains

$$D_{\text{prob}} = D_{\text{nosig}} + C_{\text{nosig}} = |\mathcal{X}| \cdot |\mathcal{Y}| (|\mathcal{A}| \cdot |\mathcal{B}| - 1),$$

the full dimensionality of a probability polytope.

## III. Status of operational locality in various theories

We survey various nonclassical operational theories, besides quantum mechanics, concerning their operationally nonlocal behavior.

*Classical theory:* is operationally local since  $v_{\text{loc}}^* = 1$  for any pair of observables, implying that Eq. (5) can never be violated.

*Quantum mechanics:* is operationally nonlocal, as already shown in the main text.

*Spekkens' toy theory* [27] is operationally nonlocal. The state space of a single system is characterized by four ontic states, labelled 1, 2, 3 and 4. The only single-system pure states are the ‘‘eigenstates’’ of three dichotomic and mutually unbiased measurements,  $\sigma_{X;\text{Sp}} \equiv \{1\vee 2, 3\vee 4\}$ ,  $\sigma_{Y;\text{Sp}} \equiv \{1\vee 3, 2\vee 4\}$  and  $\sigma_{Z;\text{Sp}} \equiv \{1\vee 4, 2\vee 3\}$ .

Without loss of generality, suppose  $b \in \{\sigma_{X;\text{Sp}}, \sigma_{Z;\text{Sp}}\}$ . If the measured state is an eigenstate, its outcome can be deterministically predicted, whereas if it is not, then it can be predicted only half the time. Thus,  $v_{\text{loc}}^* = \frac{3}{2}$ . On the other hand, the lhs of Eq. (5) evaluates to 2 because the theory admits perfect steering (see below), which is related to the incompatibility of any pair of its measurements.

Bipartite entangled states in the toy theory have the form

$$(a \wedge e) \vee (b \wedge f) \vee (c \wedge g) \vee (d \wedge h), \quad (12)$$

where  $a, b, c, d, e, g, h \in \{1, 2, 3, 4\}$  such that  $a \neq b \neq c \neq d$  are ontic states of the first particle, and  $e \neq f \neq g \neq h$  are ontic states of the second particle. To see that the theory allows steering, consider the entangled state with  $a = e = 1, b = f = 2, c = g = 3$  and  $d = h = 4$ . Measuring in  $\sigma_{X;\text{Sp}}$  (resp.,  $\sigma_{Z;\text{Sp}}$ ) basis, Alice collapses Bob's state into the corresponding eigenstate of  $\sigma_{X;\text{Sp}}$  (resp.,  $\sigma_{Z;\text{Sp}}$ ). There is no transgression of no-signaling since the uniform mixture of  $1 \vee 2$  and  $3 \vee 4$  equals that of  $1 \vee 3$  and  $2 \vee 4$ , which is the fully mixed state in the theory. This perfect steering behavior implies that the lhs of Eq. (5) evaluates to 2.

However, the theory doesn't admit states that are Bell-nonlocal, because any pair of measurements, although incompatible in the operational theory, are ‘‘meta-compatible’’, i.e., the outcome statistics admits a joint distribution (JD); see Appendix IV. Compatibility in the operational theory is thwarted because the required master observable is not part of the theory's set of allowed measurements. This peculiarity has to do with the non-convexity of the theory (arbitrary convex combinations of pure states are not part of the state space  $\Omega$ ).

*Generalized local theory:* Single system states are *gdits*, characterized by  $d$  fiducial measurements with  $k$  outcomes each [34]. The state space is the convex hull of the  $|\Omega| = k^d$  pure states of single systems, which correspond to deterministic outcomes for each fiducial measurement. The dimension  $\dim(\Omega)$  of the system is  $d(k-1)$ , the number of parameters required to describe  $d$  probability distributions. Nonclassicality arises from the fact that  $\Omega$  is not a simplex, noting that  $|\Omega| - \dim(\Omega) > 1$ . The joint space of a multi-partite system is the minimal

tensor product  $\otimes_{\text{min}}$ , namely the set of convex combinations of the direct product of two single-system states. Therefore, the states are unsteerable and consequently operationally local.

*Boxworld:* The single-system states are two-dimensional *gdits*, namely *gdits*. By convention,  $x, y, a, b \in \{0, 1\}$ . Pure bipartite entangled states are PR boxes  $P(x, y|a, b)$ , characterized by  $x \oplus y = a \cdot b$ , which ensures the maximal violation of the CHSH inequality, while  $P(0|a) = P(0|b) = \frac{1}{2}$ , ensuring no-signaling [33]. Clearly, the lhs of Eq. (5) evaluates to 2, the maximal possible value. Yet, the theory is not operationally nonlocal, because *gdits* are maximally certain, i.e.,  $v_{\text{loc}}^* = 2$ .

In particular, *gdits* (see Figure 1) take simultaneous deterministic values for  $\sigma_{X;g}$  (corresponding to  $x, y = 0$ ) and  $\sigma_{Z;g}$  (corresponding to  $x, y = 1$ ), the *gdit* analogues of Pauli  $\sigma_X$  and  $\sigma_Z$ . The state space  $\Omega_g$  is the convex hull of four pure points, denoted  $(0, 0), (0, 1), (1, 0)$  and  $(1, 1)$ , where the first (resp., second) coordinate represents the probability to get outcome 0 if  $\sigma_{X;g}$  (resp.,  $\sigma_{Z;g}$ ) is measured. Thus,  $v_{\text{loc}}^* = 2$ .

The state space is non-simplicial, and hence features measurement disturbance [45], so that only one of the two measurement values can be read out, while the other is maximally disturbed. Therefore, the *gdit* pure states in the theory can't be prepared by direct measurement, but instead by measuring one of two particles in a PR box state, which prepares the partner particle in a *gdit* pure state.

#### IV. Complementarity in Spekkens' toy theory

Consider two observables in Spekkens' toy theory, say  $\sigma_{X;\text{Sp}}$  and  $\sigma_{Z;\text{Sp}}$ , the analogues of Pauli  $\sigma_X$  and  $\sigma_Z$ . We construct the joint measurement  $M$  as a ‘‘master effect’’ such that  $\sum_j M[j, k]$  reproduces  $\sigma_{Z;\text{Sp}}[k]$ , the effect that corresponds to outcome  $k$  on measuring  $\sigma_{Z;\text{Sp}}$ , and similarly  $\sum_k M[j, k] = \sigma_{X;\text{Sp}}[j]$ . Let  $x^{\pm 1}$  be the vector representing the pure (maximum information) states of measurement  $\sigma_{X;\text{Sp}}$ , and  $z^{\pm 1}$  those of measurement  $\sigma_{Z;\text{Sp}}$ , and so on. Now,

$$\begin{aligned} M[++] \cdot x^+ + M[+-] \cdot x^+ &= 1 \\ M[-+] \cdot x^+ + M[--] \cdot x^+ &= 0 \\ M[++] \cdot x^+ + M[-+] \cdot x^+ &= 0.5 \\ M[+-] \cdot x^+ + M[--] \cdot x^+ &= 0.5 \end{aligned} \quad (13)$$

from which it follows that

$$\begin{aligned} M[++] \cdot x^+ &= M[+-] \cdot x^+ = 0.5 \\ M[-+] \cdot x^+ &= M[--] \cdot x^+ = 0. \end{aligned} \quad (14)$$

Proceeding thus, one finds

$$\begin{aligned}
M[++] \cdot x^- &= M[+-] \cdot x^- = 0 \\
M[-+] \cdot x^- &= M[--] \cdot x^- = 0.5 \\
M[+-] \cdot z^+ &= M[--] \cdot z^+ = 0 \\
M[++] \cdot z^+ &= M[-+] \cdot z^+ = 0.5 \\
M[++] \cdot z^- &= M[-+] \cdot z^- = 0 \\
M[+-] \cdot z^- &= M[--] \cdot z^- = 0.5.
\end{aligned} \tag{15}$$

The nonclassicality turns up in the fact that the extreme points are not linearly independent. In particular,

$$x^+ + x^- = z^+ + z^-. \tag{16}$$

It may be verified that each of the components  $M[j, k]$  determined in Eqs. (14) and (15) are consistent with Eq. (16). Thus, in this toy theory,  $\sigma_{X;Sp}$  and  $\sigma_{Z;Sp}$ , and by a similar argument any pair of observables admits a JD. Therefore, although the pairs are incompatible in the operational theory, they are “meta-compatible”, i.e., compatible in an underlying HV model. Meta-compatibility entails that a joint probability distribution exists for all measurement outcomes in the two-party Bell-CHSH scenario. Therefore, by Fine’s theorem [50], the correlations must be local.

## V. Uncertainty and steering strength bound on Bell nonlocality

The interplay of uncertainty and steering brought out by our result casts light on the bound on Bell nonlocality from (fine-grained) uncertainty and steering in a non-signaling theory [35]. (In a signaling theory, the signal directly demonstrates nonlocality, even without uncertainty and steering.) Further comments concerning Bell nonlocality are in order here.

Like in operational nonlocality, incompatibility is necessary in Bell nonlocality (although the specific settings for maximal violation can be different). In the two-input-two-output case with outputs  $\pm 1$ , suppose  $y_0$  and  $y_1$  are compatible. Measuring  $\tilde{\omega}_B^{a_0|x_0}$  (resp.,  $\tilde{\omega}_B^{a_1|x_1}$ ) in bases  $y_0$  and  $y_1$  yields JD  $P(b_0, b_1, a_0 | y_0, y_1, x_0)$  (resp.,  $P(b_0, b_1, a_1 | y_0, y_1, x_1)$ ). One can then construct JD  $P(a_0, a_1, b_0, b_1 | x_0, x_1, y_0, y_1)$  given by

$$\frac{P(b_0, b_1, a_0 | y_0, y_1, x_0)P(b_0, b_1, a_1 | y_0, y_1, x_1)}{P(b_0, b_1 | y_0, y_1)}, \tag{17}$$

which reproduces the observed JD’s by tracing over  $a_0$  or  $a_1$ . By Fine’s theorem [50], such a correlation can’t violate a Bell inequality, and thus must satisfy the CHSH locality condition

$$|\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle| \leq 2, \tag{18}$$

which we obtained from only arguments related to compatibility, rather than local-realism (cf. [37]). Eq. (18) is maximally violated for  $a \in \{\frac{\sigma_X + \sigma_Z}{\sqrt{2}}\}$  and  $b \in \{\sigma_X, \sigma_Z\}$ , with  $\omega_{AB}$  being the singlet state. On the other hand, we saw that maximal qubit violation of Eq. (5) requires setting  $a = b$ .

In the context of GPTs, incompatibility can be quantified in terms of an “unsharpness” parameter,  $\kappa$ : given dichotomic observable  $O$ , its unsharp version is defined to be  $O^{(\kappa)} = \kappa O + (1 - \kappa)\frac{I}{2}$ , where  $(0 < \kappa \leq 1)$  and  $I$  is the uniform distribution over two inputs [23]. It follows that the expectation value  $\langle a_j b_k^{(\kappa)} \rangle = \kappa \langle a_j b_k \rangle$ . The degree of compatibility is the maximum  $\kappa$  such that  $b_0^{(\kappa)}$  and  $b_1^{(\kappa)}$  are jointly measurable. Let  $\kappa_{\text{opt}}$  denote such a maximum, optimized over all pairs of measurements. It follows that the CHSH inequality Eq. (18), with Alice and Bob measuring  $a_j$  and  $b_k^{(\kappa_{\text{opt}})}$ , takes the form

$$|\langle a_0 b_0 \rangle + \langle a_0 b_1 \rangle + \langle a_1 b_0 \rangle - \langle a_1 b_1 \rangle| \leq \frac{2}{\kappa_{\text{opt}}}, \tag{19}$$

an incompatibility-based bound on nonlocality [36].

To relate compatibility  $\kappa_{\text{opt}}$  to uncertainty, we observe that roughly speaking, among GPTs with the same number of inputs and outputs, a more uncertain theory tends to feature more compatibility (also, cf. [51]). For example, in the 2-input-2-output case, PR boxworld features the least compatibility ( $\kappa_{\text{opt}} = \frac{1}{2}$ ) and zero uncertainty ( $v_{\text{loc}}^* = 1$ ), whereas in Spekkens theory all pairs of measurements are meta-compatible (as noted earlier) and uncertainty is maximal (all measurements are mutually unbiased). QM occupies an intermediate position here.

This pattern may be understood as follows. Among these GPTs, gdit theory can be considered as providing an ontological model, with states and measurements in the other theories, having greater uncertainty, being considered as (ontologically) more smeared versions of those in gdit theory (see Figure 1, Appendix VI). Thus, measurements in a more uncertain theory require lesser fuzzification to attain joint measurability.

For example, consider a family of 2-input-2-output GPTs  $\theta_\tau$  such that  $\lambda^\tau + \mu^\tau \leq 1$  ( $\tau \geq 1$ ) determines the parameter region  $\Delta$  where observables  $\sigma_{X;\tau}^{(\lambda)}$  and  $\sigma_{Z;\tau}^{(\mu)}$  (the  $\theta_\tau$  analogues of  $\sigma_X$  and  $\sigma_Z$ , with unsharpness parameters  $\lambda$  and  $\mu$ , respectively) are jointly measurable. Larger  $\tau$  represents a greater area of  $\Delta$  and hence [23] greater compatibility. We set  $\kappa_{\text{opt}}$  to be the maximum  $\lambda$  such that  $\lambda = \mu$ , giving  $\kappa_{\text{opt}} = 2^{-(1/\tau)}$  for theory  $\theta_\tau$ , and the Tsirelson bound in Eq. (19) becomes  $2 \cdot 2^{\frac{1}{\tau}}$ . Here,  $\tau = 1$  and  $2$  yield the bound for PR boxworld and QM, respectively [52]. For theories  $\theta_\tau$  and their more classical counterparts, we find (Appendix VI):

$$\kappa_{\text{opt}} = \frac{\beta}{2(v_{\text{loc}}^* - 1)}, \tag{20}$$



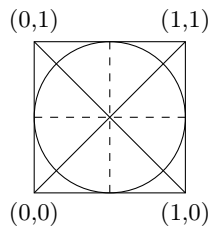


FIG. 1. State space of gdit theory, fragment of QM (with  $b_0 = \sigma_X$  and  $b_1 = \sigma_Z$ ) and of Spekkens' toy theory: the vertices of the solid square represent the four extreme states of the gdit theory, denoted by  $(P(+1|b_0), P(+1|b_1))$ . Qubits are represented by the inscribed circle (see Figure 2 in [34]). The dashed lines represent Spekkens' theory (Appendices III and IV). The point where the diagonal lines intersect the circle, which represent the quantum states with maximum uncertainty, are eigenstates of  $(\sigma_X \pm \sigma_Z)/\sqrt{2}$ . All these theories are nonclassical in that the state space is non-simplicial [44, 45], i.e., their extreme points aren't linearly independent. Note that gdit theory can be considered as providing an ontological model for the other theories.

with the allowed range being  $\frac{1}{2} \leq \kappa_{\text{opt}} \leq 1$ . Eq. (20) shows that more compatible theories are more uncertain, for fixed  $\beta$  (see above). Here, family parameter  $\beta$  ( $\in [1, 2]$ ) interpolates between the family  $\theta_\tau$  (characterized by a non-simplicial state space  $\Omega$ ;  $\beta = 1$ ) and classical theory (simplicial  $\Omega$ ;  $\beta = 2$ ), and the range of  $v_{\text{loc}}^*$  is determined by the above allowed range of  $\kappa_{\text{opt}}$ . Importantly, unlike incompatibility, uncertainty must be combined with parameter  $\beta$  to capture nonclassicality: for example,  $v_{\text{loc}}^* = 1$  can refer to both classical theory ( $\kappa_{\text{opt}} = 1$ ) and gbit theory ( $\kappa_{\text{opt}} = \frac{1}{2}$ ).

Using Eq. (20) in Eq. (19), we obtain the inequality (9), where we have set  $\varsigma \equiv \frac{1}{\beta}$ . By virtue of our earlier result linking incompatibility and steering,  $\varsigma$  represents the degree of steering, i.e., greater  $\varsigma$  corresponds to lesser classicality and more steering.

## VI. Incompatibility and uncertainty for PR boxworld, QM and a family of theories

Figure 1 depicts the idea that gbit theory can be considered as an ontological model for a family of other 2-input-2-output theories with non-vanishing uncertainty. Essentially, we regard observables and states in these theories as the smeared or fuzzified versions (at the ontological level) of their gbit counterparts. The family of GPTs  $\theta_\tau$  is a simple quantitative illustration of this idea.

Let  $\sigma_{X;g}$  and  $\sigma_{Z;g}$  be the analogues of Pauli  $\sigma_X$  and  $\sigma_Z$  in gbit theory. The corresponding observables in theory  $\theta_\tau$  are denoted  $\sigma_{X;\tau}$  and  $\sigma_{Z;\tau}$ , and their unsharp versions by  $\sigma_{X;\tau}^{(\lambda)}$  and  $\sigma_{Z;\tau}^{(\mu)}$ . Theory  $\theta_\tau$  is characterized by region  $\Delta$  of joint measurability (in the  $(\lambda, \mu)$ -parameter space)

of  $\sigma_{X;\tau}^{(\lambda)}$  and  $\sigma_{Z;\tau}^{(\mu)}$  given by

$$\lambda^\tau + \mu^\tau \leq 1, \quad (21)$$

where  $0 < \lambda, \mu \leq 1$  and  $\sigma_{X;\tau}$ . The larger is  $\tau$ , the larger is the area of  $\Delta$ , and correspondingly more compatibility in the theory, according to the criterion proposed in [23]. A measure of compatibility, which can be identified with  $\kappa_{\text{opt}}$ , is obtained by setting  $\lambda = \mu$  such that Eq. (21) is saturated. Accordingly, we find:

$$\kappa_{\text{opt}} = 2^{-1/\tau}, \quad (22)$$

which, for gdit theory, assumes the minimum value

$$\kappa_{\text{opt}}^{\text{gdit}} = \frac{1}{2}, \quad (23)$$

setting  $\tau = 1$ .

Suppose  $\sigma_{X;\tau} = \alpha\sigma_{X;g} + (1 - \alpha)(I/2)$ , by the (ontological) smearing of the corresponding gbit observable, where  $\alpha$  is the ontological unsharpness parameter. Smearing  $\sigma_{X;\tau}$  at the operational level, we obtain  $\mu\sigma_{X;\tau} + (1 - \mu)\sigma_{X;\tau} = \mu\alpha\sigma_{X;g} + (1 - \mu\alpha)(I/2)$ . Therefore the parameter  $\mu\alpha$ , which is like the effective smearing at the ontological gbit level, must satisfy

$$\mu\alpha = \frac{1}{2}, \quad (24)$$

in view of Eq. (23). Letting  $\mu \equiv \kappa_{\text{opt}}$  in Eq. (22), we find:

$$\alpha = 2^{-1+(1/\tau)}, \quad (25)$$

which represents the required ontological smearing with respect to gdit theory, to reproduce a  $\theta_\tau$  observable.

We can now estimate uncertainty parameter  $v_{\text{loc}}^*$  in  $\theta_\tau$  as follows. Suppose that the action of  $\sigma_{X;\tau}$  with respect to some state in this GPT is captured by the above smearing, but, conservatively speaking,  $\sigma_{Z;\tau}$  is not, so that the output of  $\sigma_{Z;\tau}$  on this state is deterministic as in gbit theory. Thus:

$$\begin{aligned} v_{\text{loc}}^* &= 1 + \alpha \\ &= 1 + 2^{-1+(1/\tau)} \\ &= 1 + \frac{1}{2\kappa_{\text{opt}}}, \end{aligned} \quad (26)$$

as the relation between uncertainty and incompatibility in the  $\theta_\tau$  GPT family. Cases  $\tau = 1$  and  $\tau = 2$  give the known bounds for gdit theory and QM [52].

Classical theory, like PR boxworld, lacks uncertainty, but unlike in the latter, the state space is a simplex. Thus, it does not belong to the above  $\theta_\tau$  family and cannot be derived from by an ontological fuzzification of gbit theory. This is reflected, for example, in the fact that setting  $\kappa_{\text{opt}} := 1$  in Eq. (26) doesn't lead to full certainty. Classical theory can be incorporated into this scheme, by

letting space  $\Omega^{\text{classical}}$  be obtained from gdit space  $\Omega^{\text{gbit}}$  by “simplexifying” the latter. This is a bijective map that “inflates” the outer square in Figure 1 to a tetrahedron (cf. [45]). For our present purpose, we can capture this process by a parameter  $\beta$  ( $1 \leq \beta \leq 2$ ) that modifies Eq. (26) to

$$v_{\text{loc}}^* = 1 + \frac{\beta}{2\kappa_{\text{opt}}}, \quad (27)$$

from which Eq. (20) follows. Here,  $\beta = 1$  corresponds

to the  $\theta_\tau$  family of nonclassical theories derived from gdit theory, whilst  $\beta = 2$  corresponds to classical theory. Thus,  $\beta$  is like a family parameter interpolating between classical theory and the  $\theta_\tau$  family. In Eq. (27), both gdit theory ( $\kappa_{\text{opt}} = \frac{1}{2}, \beta = 1$ , i.e., minimal compatibility and non-simplicial  $\Omega$ ) and classical theory ( $\kappa_{\text{opt}} = \frac{1}{2}, \beta = 1$ , or maximum compatibility and simplicial  $\Omega$ ) are seen to correspond to vanishing uncertainty.