

Nonlocality without entanglement: Quantum theory and beyond

Some Sankar Bhattacharya,¹ Sutapa Saha,² Tamal Guha,² and Manik Banik³

¹*Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong*

²*Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata-700108, India*

³*School of Physics, Indian Institute of Science and Educational Research (IISER), Thiruvananthapuram, Vithura, Kerala 695551, India*



(Received 2 September 2019; accepted 19 February 2020; published 18 March 2020)

Quantum nonlocality without entanglement (Q-NWE) captures nonlocal behavior of multipartite product states as they may entail global operation for optimal decoding of the classical information encoded in the state ensemble that allows *local* preparation. In this Rapid Communication we show that the phenomena of NWE is not specific to quantum theory only, but rather a class of generalized probabilistic theories that can exhibit such behavior. In fact, several manifestations of NWE, e.g., asymmetric local discrimination, suboptimal local discrimination, the notion of separable but locally unimplementable measurements arise generically in operational theories other than quantum theory. We propose a framework to compare the strength of NWE in different theories and show that such behavior in quantum theory is limited, suggesting a specific topological feature of quantum theory, namely, the continuity of state space structure. Our work adds the erstwhile missing foundational appeal to the study of NWE phenomena along with its information-theoretic relevance.

DOI: [10.1103/PhysRevResearch.2.012068](https://doi.org/10.1103/PhysRevResearch.2.012068)

Introduction. One of the most counterintuitive aspects of quantum theory is its *nonlocal* behavior. Bell, in his seminal result, showed that entangled states of composite quantum systems can result in correlations that do not allow any *local realistic* explanation [1] (see [2,3] for reviews on Bell nonlocality). Such correlations, however, are not available in its most strengthened form [4]. This limited behavior of Bell nonlocality has later been axiomatized in deriving quantum theory [5] and subsequently motivates several computational and information-theoretic primitives that aim to single out the correlations allowed in the physical world [6–14].

Nonlocal behaviors of quantum theory, however, do not always necessitate entanglement. In a pioneering work, Bennett *et al.* recognized that multipartite quantum systems can exhibit nonlocal properties involving only product states in a way fundamentally different from Bell nonlocality [15]. In particular, they constructed sets of product states that cannot be exactly discriminated using local operations and classical communication (LOCC) while mutual orthogonality assures their perfect global discrimination. The authors coined the phrase “quantum nonlocality without entanglement” (Q-NWE) for this phenomenon as the sets allow local preparation (with some preshared strategy) but prohibit perfect local discrimination. In other words, global measurement can be more efficient than LOCC for extracting classical information encoded in a locally prepared ensemble of quantum states. At this point, it should be noted that this particular phenomenon is different from the other “nonlocality without entanglement”

aspect of quantum theory as established through the seminal Aharonov-Bohm (AB) effect [16]. The AB effect shows that, while electromagnetic scalar and vector potentials are mere calculational aids in classical mechanics, in quantum mechanics they are an essential part of the formalism: a charged quantum particle can respond to electromagnetic potentials without ever passing through an electromagnetic field. The AB effect has a nonlocal aspect as it underlines the necessity of describing quantum systems via gauge-dependent quantities rather than local forces, which cannot account for abrupt changes in modular velocity [17]. However, in our work the term “nonlocality without entanglement” should be understood in the sense of Bennett *et al.* [15].

Local indistinguishability has later been identified as a crucial primitive for a number of distributed quantum protocols, namely, quantum data hiding [18–20] and quantum secret sharing [21–23]; and it underlies the nonzero gap between single-shot and multishot classical capacities of noisy quantum channels [24]. On the foundational part, the recent Pusey-Barrett-Rudolph theorem uses such a nonlocal feature of nonorthogonal product states to establish a ψ -*ontic* nature of quantum wave function [25].

In the present work, we ask the question, what quantum feature is indeed captured in quantum nonlocality without entanglement? More particularly we look for whether this phenomenon is specific to quantum theory or is it possible to devise generalized probabilistic models other than quantum mechanics that also manifest similar behavior. Quite interestingly, we show that the nonlocality without entanglement phenomenon is not specific to quantum theory; it is indeed possible in a broader class of generalized probabilistic theories. Recall that indistinguishability of pure states in quantum theory can be thought of as an artifact of Hilbert space structure as it entails nonorthogonality among the states; and the impossibility of perfect local discrimination of an orthogonal product

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

basis, there, results in a *separable measurement* that cannot be implemented locally. The generalized probability theory (GPT) framework admits a broader mathematical setup that includes quantum and classical theory as special examples. It can incorporate the notion of indistinguishable pure states without invoking the much constrained Hilbert space structure of quantum mechanics [26]. Here we show that this general framework can also exhibit a nonzero gap in optimal success probabilities of product states discrimination under local and global protocols, i.e., it evinces the NWE phenomenon. In fact, we find that several aspects of NWE, as observed in quantum theory, are indeed feasible in this generalized framework. For instance, it is possible to have a set of bipartite product states that allows perfect local discrimination when one of the parties starts the protocol, whereas the protocol fails for the other party, a fact already known in quantum theory [27,28]. In the GPT framework, we then constructed sets of product states that cannot be perfectly discriminated by any local protocol, whereas a global measurement serves the purpose exactly. Furthermore, all the effects constituting the perfect discrimination measurement are product effects ensuring the existence of separable but locally unimplementable measurement in the GPT framework. This mimics the NWE phenomena as established in the seminal work of Bennett *et al.* [15]. We propose a methodology to compare the strength of NWE in different theories and subsequently show that such behavior in quantum theory is limited in nature. Importantly, it turns out that this limited behavior of NWE in quantum theory is deeply linked with one of its topological features. More particularly the limited NWE is caused due to the continuity of state space structure in quantum theory which assures the existence of a continuous reversible transformation on a system between any two pure states [29]. Our present work thus adds a deep foundational appeal to the study of NWE phenomena. In the following we start with a brief discussion on the GPT framework.

Generalized probabilistic theory. This mathematical framework is broad enough to describe all possible probabilistic theories that use the notion of states to yield the outcome probabilities of measurements (see the Supplemental Material [30], and Refs. [29,31–36] for details of this framework). In a GPT, a system $S_{ys} \equiv (\Omega, \mathcal{E})$ is associated with a set of states Ω and a set of effects \mathcal{E} . Typically, Ω is considered to be a compact convex set embedded in the positive convex cone V_+ of some real vector space V while \mathcal{E} is embedded in the cone V_+^* which is dual to the state cone V_+ . An effect $e \in \mathcal{E}$ corresponds to a linear functional on Ω that maps each state $\omega \in \Omega$ onto a probability $p(e|\omega)$, representing successful filter of the effect e on the state ω . A collection of effects $\{e_i\}_i$ forms a measurement whenever $\sum_i e_i = u$, with u being the unit effect such that $p(u|\omega) = 1, \forall \omega \in \Omega$. A preparation or state ω thus specifies outcome probabilities for all measurements that can be performed on it. Given two systems $S_{ys}(A) \equiv (\Omega^A, \mathcal{E}^A)$ and $S_{ys}(B) \equiv (\Omega^B, \mathcal{E}^B)$ the theory also delineates their composition. The composite system $S_{ys}(AB) \equiv (\Omega^{AB}, \mathcal{E}^{AB})$ satisfies some natural conditions, such as no-signaling and local tomography [29,37], that narrow down the possibilities of such compositions and assure that Ω^{AB} is embedded in the positive cone of $V^A \otimes V^B$ [38–40].

Hilbert space description of quantum theory lies within this framework. The state of the system is represented by a density operator $\rho \in \mathcal{D}(\mathcal{H})$ [41] while effects correspond to positive semidefinite operators on \mathcal{H} , where \mathcal{H} is the Hilbert space associated with the system. The outcome probabilities follow the Born (trace) rule. The Hilbert space of a composite system consisting of the subsystems A_i 's is given by $\bigotimes_i \mathcal{H}_{A_i}$, where \mathcal{H}_{A_i} is the Hilbert space of the i th subsystems.

Here we recall another class of GPTs, namely, polygonal models $\mathbb{P}(n) \equiv [\Omega(n), \mathcal{E}(n)]$ [42]. The state spaces $\Omega(n)$ for elementary systems are regular polygons with n vertices. The states and effects are represented by vectors in \mathbb{R}^3 and $p(e|\omega)$ is given by usual Euclidean inner product. For a fixed n , $\Omega(n)$ is the convex hull of n pure states $\{\omega_i\}_{i=0}^{n-1}$ with $\omega_i := [r_n \cos(2\pi i/n), r_n \sin(2\pi i/n), 1]^T \in \mathbb{R}^3$; T denotes transpose and $r_n := \sqrt{\sec(\pi/n)}$. The unit effect is given by $u := (0, 0, 1)^T$. The set $\mathcal{E}(n)$ of all possible measurement effects consists of a convex hull of zero effect, unit effect, and the extremal effects $\{e_i, \bar{e}_i\}_{i=0}^{n-1}$, where $e_i := \frac{1}{2}[r_n \cos[(2i+1)\pi/n], r_n \sin[(2i+1)\pi/n], 1]^T$ for even n and $e_i := \frac{1}{1+r_n}[r_n \cos(2\pi i/n), r_n \sin(2\pi i/n), 1]^T$ for odd n and $\bar{e} := u - e$. This class of models has attracted considerable interest in the recent past [43–46].

A composite system allows the possibility of a state $\omega_{AB} \in \Omega^{AB}$ that cannot be prepared as a statistical mixture of some product states, i.e., $\omega_{AB} \neq \sum_i p_i \omega_A^i \otimes \omega_B^i$ with $\{p_i\}_i$ being a probability distribution. Such states are called entangled states. Entangled effects are defined similarly. Whenever such entangled states and/or entangled effects are invoked in a GPT, they must satisfy the basic self-consistency condition—any valid composition of systems, states, effects, and their transformations should produce non-negative conditional probabilities. However, mathematically several self-consistent compositions of the elementary system are possible while considering a multipartite system. For instance, bipartite composition of square-bit model $\mathbb{P}(4)$ allows four such nontrivial compositions: (i) Popescu-Rohrlich model, (ii) hypersignaling (HS) model, (iii) hybrid model, and (iv) frozen model [47]. Even in quantum mechanics, different, in fact infinite, self-consistent compositions are possible [38–40,48]. Among these only the quantum composite state space $\mathcal{D}(\mathcal{H}^{\otimes n})$ possesses the property of *self-duality* [49], which in the GPT framework has recently been derived from a computational primitive [50]. With this prelude we now move to the main part of this work.

Nonlocality without entanglement. This particular phenomenon is related to the multipartite state discrimination problem under local protocols. In the GPT framework the task can be defined as follows. Suppose an n -partite state chosen randomly from an ensemble $\{p_i, \omega_{A_1 \dots A_n}^i\}_{i=1}^k$ is distributed among n number of spatially separated parties who know the ensemble but not the exact state and aim to identify it given one copy of the system. However, there are severe restrictions on their action—they can only perform operations on their respective parts of the composite system and can classically communicate with each other. In general, the ensemble can consist of entangled states. But while studying the NWE phenomena we will consider ensembles of product states only. Before proceeding further, let us discuss a bit more about operations which are locally implementable or otherwise. In

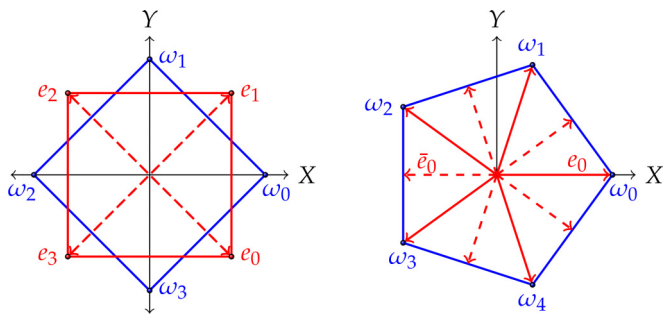


FIG. 1. Polygonal models $\mathbb{P}(4)$ (left) and $\mathbb{P}(5)$ (right). Normalized state plane (at $z = 1$) is depicted. Pentagon model is self-dual, whereas squit is not. In both models ω_i 's represent pure states. In the squit model there are only two extremal measurements, $\mathcal{M}_i \equiv \{e_i, e_{i+1}|e_i + e_{i+1} = u\}$ with $i \in \{0, 1\}$, whereas the other has five extremal measurements, $\mathcal{M}_i \equiv \{e_i, \bar{e}_i|e_i + \bar{e}_i = u\}$ with $i \in \{0, \dots, 4\}$. In both models e_i 's are ray extremal effects, whereas \bar{e}_i 's are extremal effects of $\mathcal{E}(5)$ but not ray extremal.

the following we define separable measurements in the GPT framework.

Definition 1. Consider an n -partite system $S_{\text{sy}}(A_1 \dots A_n) \equiv (\Omega^{A_1 \dots A_n}, \mathcal{E}^{A_1 \dots A_n})$. A measurement $\mathcal{M} \equiv \{e_{A_1 \dots A_n}^i | e_{A_1 \dots A_n}^i \in \mathcal{E}^{A_1 \dots A_n} \forall i \text{ and } \sum_i e_{A_1 \dots A_n}^i = u_{A_1 \dots A_n}\}$ is called separable if $e_{A_1 \dots A_n}^i = e_{A_1}^i \otimes \dots \otimes e_{A_n}^i$ for all i , where $e_X^i \in \mathcal{E}^X$ with $S_{\text{sy}}(X) \equiv (\Omega, \mathcal{E}^X)$ being the subsystems for $X \in \{A_1, \dots, A_n\}$.

Quite surprisingly, all separable quantum measurements are not locally implementable, i.e., cannot be realized by LOCC [15]. Generally a LOCC protocol consists of multi-round steps that make its mathematical characterization hard in quantum theory [51]. During such a protocol when some party performs a measurement on her/his subsystem and obtains some outcome, then naturally the question arises how the given state gets updated. As noted in Ref. [52], any valid postmeasurement update rule in a GPT must satisfy some basic consistency requirements imposed through Bayesian reasoning. While the von Neumann–Lüders rule in quantum theory is a consistent update rule, there is no natural way in an arbitrary GPT to come up with such a rule. Despite this we now show that several features of the local state discrimination problem, as observed in quantum theory, have similar manifestations in GPT.

We start with the example of asymmetric local discrimination. For composite quantum systems there exist orthogonal product states that can be perfectly discriminated locally if and only if one of the parties starts the protocol [27,28]. For instance, the set $\{|00\rangle_{AB}, |01\rangle_{AB}, |1+\rangle_{AB}, |1-\rangle_{AB}\} \subset \mathbb{C}_A^2 \otimes \mathbb{C}_B^2$ is perfectly discriminable when Alice starts the protocol, but not the other way around. Similar is possible in other generalized probabilistic models.

Lemma 1. Consider the four states, $\$(4) := \{\omega_0 \otimes \omega_0, \omega_0 \otimes \omega_3, \omega_1 \otimes \omega_0, \omega_2 \otimes \omega_1\}$ of the composite system $\mathbb{P}(4)^{\otimes 2}$, where all $\omega_i \in \Omega(4)$ (see Fig. 1). This set can be perfectly discriminated locally if and only if Alice starts the protocol (proof provided in [30]).

As already mentioned, for $\mathbb{P}(4)^{\otimes 2}$ four different self-consistent nontrivial compositions are possible along with the

trivial minimal composition $\mathbb{P}(4)_{\text{min}}^{\otimes 2} \equiv [\Omega(4)_{\text{min}}^{\otimes 2}, \mathcal{E}(4)_{\text{min}}^{\otimes 2}]$, where both the state space $\Omega(4)_{\text{min}}^{\otimes 2}$ and the effect space $\mathcal{E}(4)_{\text{min}}^{\otimes 4}$ contain only product states and product effects, respectively. Among these, the minimal composition and the HS composition [47] are Bell local models as they contain no entangled state. However, Lemma 1 holds true in all of these five models. We now consider a more exotic phenomenon of NWE. In their classic paper [15], Bennett *et al.* provided an example of orthogonal product bases for $(\mathbb{C}^3)^{\otimes 2}$ and $(\mathbb{C}^2)^{\otimes 3}$ Hilbert spaces that cannot be perfectly discriminated under LOCC operations. However, global separable measurements perfectly discriminate the states. To obtain a similar manifestation in GPT we consider the tripartite system $\mathbb{P}(5)^{\otimes 3}$. Likewise $\mathbb{P}(4)^{\otimes 2}$, here also it will be interesting to find out all possible self-consistent compositions. However, the minimal composition $\mathbb{P}(5)_{\text{min}}^{\otimes 3} \equiv [\Omega(5)_{\text{min}}^{\otimes 3}, \mathcal{E}(5)_{\text{min}}^{\otimes 3}]$ will suffice for our purpose.

Theorem 1. Consider the set of states $\$(5) \equiv \{\phi_1 := \omega_{000}, \phi_2 := \omega_{222}, \phi_3 := \omega_{102}, \phi_4 := \omega_{402}, \phi_5 := \omega_{021}, \phi_6 := \omega_{024}, \phi_7 := \omega_{210}, \phi_8 := \omega_{240}\} \subset \Omega(5)_{\text{min}}^{\otimes 3}$, where $\omega_{klm} := \omega_k \otimes \omega_l \otimes \omega_m$ with each $\omega_i \in \Omega(5)$. This set is perfectly discriminable under global operation, whereas no local protocol can discriminate them exactly.

Proof. Symmetry of the states in $\$(5)$ assures that any one of the parties (say Alice) can start the local discrimination protocol. Suppose that Alice performs the extremal measurement $\mathcal{M}_0 \equiv \{e_0, \bar{e}_0\}$ from such five possible choices $\{\mathcal{M}_i\}_{i=0}^5$ (see Fig. 1). Since postmeasurement update is not well defined in polygonal models, a measurement outcome should either exactly identify the given state or it should conclusively eliminate some possibilities. Alice's outcome e_0 ensures that the given state is none of $\{\phi_2, \phi_7, \phi_8\}$. Similarly the outcome \bar{e}_0 excludes $\{\phi_1, \phi_5, \phi_6\}$. Having this information, Bob and/or Charlie (in any possible order) perform suitable measurements on their respective subsystems. At this step, one of the outcomes corresponding to the best chosen measurement conclusively eliminates two states. Thus the last party has to identify the state from the remaining three states, which is impossible (a flow chart of the protocol is provided in the Supplemental Material). If Alice measures $\mathcal{M}_1 \equiv \{e_1, \bar{e}_1\}$, her outcome e_1 (\bar{e}_1) eliminates only one state ϕ_4 (ϕ_3) making the protocol less effective. It is not hard to see that whichever measurement Alice starts with, no perfect discrimination is possible.

For perfect discrimination, consider the set of effects $\{E_1 := e_{000}, E_2 := \bar{e}_{000}, E_3 := e_{10}, \otimes \bar{e}_0, E_4 := \bar{e}_1 \otimes e_0 \otimes \bar{e}_0, E_5 := e_0 \otimes \bar{e}_0 \otimes e_1, E_6 := e_0 \otimes \bar{e}_{01}, E_7 := \bar{e}_0 \otimes e_{10}, E_8 := \bar{e}_{01} \otimes e_0\} \subset \mathcal{E}(5)_{\text{min}}^{\otimes 3}$, where $e_{ijk} := e_i \otimes e_j \otimes e_k$ and $\bar{e}_{lmn} := \bar{e}_l \otimes \bar{e}_m \otimes \bar{e}_n$, with each $e_r, \bar{e}_s \in \mathcal{E}(5)$. They form a measurement as $\sum_{i=1}^8 E_i = u(5)^{\otimes 3}$. Straightforward calculation also yields $p(E_i|\phi_j) = \delta_{ij}$ ensuring perfect discrimination of the set $\$(5)$. This concludes the proof. ■

The state discriminable measurement in Theorem 1 is a separable measurement (see Definition 1). Since no local protocol can perfectly identify the states, this particular separable measurement cannot be implemented locally. Similar construction is also possible in other higher gonal models. Please see the Supplemental Material for explicit constructions in hexagonal and heptagonal models. Regarding the squit model, however, we have the following observation.

Observation 1. Construction of nonlocal product states as of Theorem 1 is not possible in $\mathbb{P}(4)^{\otimes 3}$.

This follows from the fact that an elementary system of squit does not allow any pair of indistinguishable pure states. In other words, local indistinguishability among pure states is necessary for the existence of nonlocal product states of Theorem 1.

So far we have studied different aspects of NWE in the broader mathematical framework of GPTs. Naturally the question arises how to compare the strength of NWE in different GPTs. To do so, first note that the elementary systems considered in different GPTs must be of the *same type*. Recall that the phenomena of NWE fundamentally demonstrates a difference between local and global operations in extracting classical information encoded in product states. Thus to be on a similar footing, different such systems must have the same “classical information storage capacity,” a notion recently studied for quantum systems in Ref. [53] and generalized for GPTs in Ref. [47] by the name “signaling dimension.” The concept can be understood with the following communication scenario. Given two finite alphabets $\mathcal{X} = \{x\}$ and $\mathcal{Y} = \{y\}$ of cardinalities m and n , respectively, let $\mathcal{P}_{S_{ys}}^{m \rightarrow n}$ denote the convex set of all m -input/ n -output conditional probability distributions generated by transmitting an elementary system S_{ys} from a sender to a receiver who may have preshared randomness. In such a scenario signaling dimension is defined as follows.

Definition 2. (Dall’Arno *et al.* [47]) The signaling dimension of a system S_{ys} , denoted by $\kappa(S_{ys})$, is defined as the smallest integer d such that $\mathcal{P}_{S_{ys}}^{m \rightarrow n} \subseteq \mathcal{P}_{C_d}^{m \rightarrow n}$, for all m and n .

Here, $\mathcal{P}_{C_d}^{m \rightarrow n}$ denotes the set of all m -input/ n -output conditional probability distributions obtained by means of a d -dimensional classical noiseless channel and shared random data. Suppose that S_{ys}^1 and S_{ys}^2 are elementary systems of two different theories having identical signaling dimension and $(S_{ys}^i)^{\otimes n}$ is their n -partite self-consistent composition, with $i \in \{1, 2\}$. Consider now sets of product states (having the same cardinality) in both these theories that cannot be distinguished locally while respective global measurements can discriminate the states perfectly. The quantity $\Delta[i] := 1 - P_L[i]$ amounts to the gap between global and local success probabilities in discriminating the states, where $P_L[i]$ is optimized under all local protocols allowed in the i th theory. If it turns out that $\Delta[1] > \Delta[2]$, then we can say Theory 1 depicts stronger NWE in comparison to Theory 2. In quantum theory both for the systems $(\mathbb{C}^3)^{\otimes 2}$ and $(\mathbb{C}^2)^{\otimes 3}$ we have two different sets of product states with cardinality 8 that exhibit NWE phenomena [15]. However, the example of $(\mathbb{C}^2)^{\otimes 3}$ is comparable with that of Theorem 1, since both the elementary systems have same signaling dimension and both examples consider tripartite composite systems. What follows next is the comparison between the strength of NWE in these two examples.

Theorem 2. $\Delta[\text{QT}] \leq \frac{1}{8}(4 - \sqrt{10}) < \frac{1}{8} = \Delta[\text{pentagon}]$.

Proof of the theorem is provided in the Supplemental Material [30]. Here it is worth mentioning that for the pentagon model local success probability is optimized over all possible local protocols which consists of one-way protocols only. The corresponding quantum value is also obtained under

one-way LOCC protocol. More general local protocol (consisting of two-way LOCC) may further decrease the value of $\Delta[\text{QT}]$ which in general is difficult to optimize [54]. From the structure of constructions, arguably it follows that other polygonal models also have $\Delta = 1/8$. Furthermore, we note that instead of the uniform prior distribution of the states $\{\phi_i\}_{i=1}^8$ if one considers a biased prior distribution (a more feasible situation for the experimental purpose), then also limited NWE behavior of quantum theory can be established (see [30]). Importantly, the continuity of the quantum state space plays the crucial role in the resulting limited NWE behavior compared to the discrete polygonal models.

One may ask for the GPT analog of the nonlocal product bases in $(\mathbb{C}^3)^{\otimes 2}$ [15]. We have a negative impression at this point that such analogy will not be possible in bipartite composition of the polygonal models. This is due to the fact that all the polygonal models have signaling dimension 2, whereas that of the qutrit system is 3. Of course a rigorous mathematical proof of this intuition will be worth investigating. Such a result will generalize Theorem 4 of Ref. [28] in the GPT framework.

Discussion. In this work we study the nonlocality without entanglement phenomena in the broader mathematical framework of generalized probabilistic theories. We show that this particular nonclassical behavior in quantum theory is limited as compared to other GPT models. This, in a sense, is similar to the fact of limited Bell nonlocality in quantum theory as observed by Rohrlich and Popescu in their seminal work [5]. In fact, Rohrlich-Popescu proposed to axiomatize this limited Bell nonlocal behavior (along with relativistic causality) to derive quantum theory. Subsequently, several information and physical principles have been proposed to explain limited Bell nonlocality in quantum theory [6–14], and its connection with other quantum features has also been established [55–61]. In our work, we observe that the limited nonlocality without entanglement feature in quantum theory is owing to the continuity of the quantum state space structure which is presumed in axiomatic derivation of quantum theory either directly [29], or invoked through other assumptions such as “reversible transformation” [62] or “purification” [63]. However, the present work shows that the same feature can be obtained as an emerging fact if we presume the limited NWE as a fundamental characteristic of the theory. It therefore welcomes novel information/physical primitive(s) to explain this limited NWE behavior in quantum theory.

Our work also motivates further research to study other exotic features of NWE phenomena in the GPT framework. For instance, the phenomena of NWE in quantum theory was first anticipated by Peres and Wootters [64]. They conjectured that LOCC measurements are suboptimal for discrimination of a specific set of nonorthogonal product states (see also [65]). More recently, the authors in Ref. [66] revisited the classic problem of Peres and Wootters and proved that their conjecture is indeed true. A similar example in the GPT framework is yet to be obtained. On the other hand, multipartite generalization of NWE phenomena have been studied very recently [67–69]. Similar study in the GPT framework may provide further insight into the structure of quantum theory.

Acknowledgments. This work is supported by the National Natural Science Foundation of China through Grant No. 11675136, the Foundational Questions Institute through Grant No. FQXi-RFP3-1325, the Hong Kong Research Grant Council through Grant No. 17300918, and the John Templeton Foundation through Grant No. 60609, Quantum Causal Structures. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. Part of the work has been done when M.B. was a DST INSPIRE Faculty at

S. N. Bose National Center for Basic Sciences. M.B. acknowledges support through an INSPIRE-faculty fellowship by the Department of Science and Technology, Government of India. We gratefully acknowledge discussions with Guruprasad Kar, Saronath Halder, Amit Mukherjee, Arup Roy, Mir Alimuddin, and Sumit Rout. We would like to thank Amit Mukherjee for his help with Fig. 2 in the Supplemental Material and Saronath Halder for pointing out Ref. [54]. We gratefully acknowledge the useful suggestions by Giulio Chiribella.

-
- [1] J. S. Bell, On the Einstein Podolsky Rosen paradox, *Physics* **1**, 195 (1964); On the problem of hidden variables in quantum mechanics, *Rev. Mod. Phys.* **38**, 447 (1966).
- [2] N. David Mermin, Hidden variables and the two theorems of John Bell, *Rev. Mod. Phys.* **65**, 803 (1993).
- [3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, *Rev. Mod. Phys.* **86**, 419 (2014).
- [4] B. S. Cirel'son, Quantum generalizations of Bell's inequality, *Lett. Math. Phys.* **4**, 93 (1980).
- [5] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, *Found. Phys.* **24**, 379 (1994); S. Popescu, Nonlocality beyond quantum mechanics, *Nat. Phys.* **10**, 264 (2014).
- [6] W. van Dam, Implausible consequences of superstrong nonlocality, *Nat. Comput.* **12**, 9 (2013).
- [7] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on Nonlocality in Any World in Which Communication Complexity is Not Trivial, *Phys. Rev. Lett.* **96**, 250401 (2006).
- [8] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum Nonlocality and Beyond: Limits from Nonlocal Computation, *Phys. Rev. Lett.* **99**, 180502 (2007).
- [9] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, *Nature (London)* **461**, 1101 (2009).
- [10] M. Navascués and H. Wunderlich, A glance beyond the quantum model, *Proc. R. Soc. London, Ser. A* **466**, 881 (2009).
- [11] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, *Nat. Commun.* **4**, 2263 (2013).
- [12] A. Cabello, Simple Explanation of the Quantum Violation of a Fundamental Inequality, *Phys. Rev. Lett.* **110**, 060402 (2013).
- [13] A. Acín, T. Fritz, A. Leverrier, and A. Belén Sainz, A combinatorial approach to nonlocality and contextuality, *Commun. Math. Phys.* **334**, 533 (2015).
- [14] J. Henson, Bounding Quantum Contextuality with Lack of Third-Order Interference, *Phys. Rev. Lett.* **114**, 220403 (2015).
- [15] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, and W. K. Wootters, Quantum nonlocality without entanglement, *Phys. Rev. A* **59**, 1070 (1999).
- [16] Y. Aharonov and D. Bohm, Significance of electromagnetic potentials in the quantum theory, *Phys. Rev.* **115**, 485 (1959).
- [17] Y. Aharonov, E. Cohen, and D. Rohrlich, Nonlocality of the Aharonov-Bohm effect, *Phys. Rev. A* **93**, 042110 (2016).
- [18] B. M. Terhal, D. P. DiVincenzo, and D. W. Leung, Hiding Bits in Bell States, *Phys. Rev. Lett.* **86**, 5807 (2001).
- [19] T. Eggeling and R. F. Werner, Hiding Classical Data in Multipartite Quantum States, *Phys. Rev. Lett.* **89**, 097905 (2002).
- [20] W. Matthews, S. Wehner, and A. Winter, Distinguishability of quantum states under restricted families of measurements with an application to quantum data hiding, *Commun. Math. Phys.* **291**, 813 (2009).
- [21] D. Markham and B. C. Sanders, Graph states for quantum secret sharing, *Phys. Rev. A* **78**, 042309 (2008).
- [22] R. Rahaman and M. G. Parker, Quantum scheme for secret sharing based on local distinguishability, *Phys. Rev. A* **91**, 022330 (2015).
- [23] J. Wang, L. Li, H. Peng, and Y. Yang, Quantum-secret-sharing scheme based on local distinguishability of orthogonal multipartite entangled states, *Phys. Rev. A* **95**, 022320 (2017).
- [24] C. A. Fuchs, Just two nonorthogonal quantum states, in *Quantum Communication, Computing, and Measurement 2*, edited by P. Kumar, G. M. D'Ariano, and O. Hirota (Springer, Boston, 2002).
- [25] M. F. Pusey, J. Barrett, and T. Rudolph, On the reality of the quantum state, *Nat. Phys.* **8**, 475 (2012).
- [26] M. Banik, S. Saha, T. Guha, S. Agrawal, S. S. Bhattacharya, A. Roy, and A. S. Majumdar, Constraining the state space in any physical theory with the principle of information symmetry, *Phys. Rev. A* **100**, 060101(R) (2019).
- [27] B. Groisman and L. Vaidman, Nonlocal variables with product states eigenstates, *J. Phys. A: Math. Gen.* **34**, 6881 (2001).
- [28] J. Walgate and L. Hardy, Nonlocality, Asymmetry, and Distinguishing Bipartite States, *Phys. Rev. Lett.* **89**, 147901 (2002).
- [29] L. Hardy, Quantum theory from five reasonable axioms, [arXiv:quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
- [30] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevResearch.2.012068> for detailed discussion of GPT and few technical results mentioned in the paper.
- [31] J. Barrett, Information processing in generalized probabilistic theories, *Phys. Rev. A* **75**, 032304 (2007).
- [32] G. Chiribella, G. Mauro D'Ariano, and P. Perinotti, Probabilistic theories with purification, *Phys. Rev. A* **81**, 062348 (2010).
- [33] H. Barnum and A. Wilce, Information processing in convex operational theories, *Electron. Notes Theor. Comput. Sci.* **270**, 3 (2011).
- [34] D. A. Yopp and R. D. Hill, Extremals and exposed faces of the cone of positive maps, *Linear Multilinear Algebra* **53**, 167 (2007).
- [35] G. Kimura and K. Nuida, On affine maps on non-compact convex sets and some characterizations of finite-dimensional solid ellipsoids, *J. Geom. Phys.* **86**, 1 (2014).

- [36] A. B. Sainz, Y. Guryanova, A. Acín, and M. Navascués, Almost-Quantum Correlations Violate the No-Restriction Hypothesis, *Phys. Rev. Lett.* **120**, 200402 (2018).
- [37] L. Hardy and W. K. Wootters, Limited holism and real-vector-space quantum theory, *Found. Phys.* **42**, 454 (2012).
- [38] I. Namioka and R. R. Phelps, Tensor products of compact convex sets, *Pac. J. Math.* **31**, 469 (1969).
- [39] A. Wilce, Tensor products in generalized measure theory, *Int. J. Theor. Phys.* **31**, 1915 (1992).
- [40] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Generalized No-Broadcasting Theorem, *Phys. Rev. Lett.* **99**, 240501 (2007).
- [41] $\mathcal{D}(\mathcal{H})$ denotes the collection of all positive semidefinite operators acting on \mathcal{H} and having unit trace. $\mathcal{D}(\mathcal{H})$ is isomorphic to a convex set embedded in the positive cone of \mathbb{R}^{d^2-1} , where d is the dimension of \mathcal{H} .
- [42] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, Limits on nonlocal correlations from the structure of the local state space, *New J. Phys.* **13**, 063024 (2011).
- [43] P. Janotta and R. Lal, Generalized probabilistic theories without the no-restriction hypothesis, *Phys. Rev. A* **87**, 052131 (2013).
- [44] S. Massar and M. K. Patra, Information and communication in polygon theories, *Phys. Rev. A* **89**, 052124 (2014).
- [45] S. W. Al-Safi and J. Richens, Reversibility and the structure of the local state space, *New J. Phys.* **17**, 123001 (2015).
- [46] S. S. Bhattacharya, S. Saha, T. Guha, S. Halder, and M. Banik, Supremacy of quantum theory over supra-quantum models of communication, [arXiv:1806.09474](https://arxiv.org/abs/1806.09474).
- [47] M. Dall'Arno, S. Brandsen, A. Tosini, F. Buscemi, and V. Vedral, No-Hypersignaling Principle, *Phys. Rev. Lett.* **119**, 020401 (2017).
- [48] H. Arai, Y. Yoshida, and M. Hayashi, Perfect discrimination of non-orthogonal separable pure states on bipartite system in general probabilistic theory, *J. Phys. A: Math. Theor.* **52**, 465304 (2019).
- [49] G. Birkhoff and J. Von Neumann, The logic of quantum mechanics, *Ann. Math.* **37**, 823 (1936).
- [50] M. P. Müller and C. Ududec, Structure of Reversible Computation Determines the Self-Duality of Quantum Theory, *Phys. Rev. Lett.* **108**, 130401 (2012).
- [51] E. Chitambar, D. Leung, L. Mancinska, M. Ozols, and A. Winter, Everything you always wanted to know about LOCC (but were afraid to ask), *Commun. Math. Phys.* **328**, 303 (2014).
- [52] G. Chiribella, A. Cabello, M. Kleinmann, and M. P. Müller, General Bayesian theories and the emergence of the exclusivity principle, [arXiv:1901.11412](https://arxiv.org/abs/1901.11412).
- [53] P. E. Frenkel and M. Weiner, Classical information storage in an n -level quantum system, *Commun. Math. Phys.* **340**, 563 (2015).
- [54] S. Croke and S. M. Barnett, Difficulty of distinguishing product states locally, *Phys. Rev. A* **95**, 012337 (2017).
- [55] J. Oppenheim and S. Wehner, The uncertainty principle determines the non-locality of quantum mechanics, *Science* **330**, 1072 (2010).
- [56] M. Banik, M. R. Gazi, S. Ghosh, and G. Kar, Degree of complementarity determines the nonlocality in quantum mechanics, *Phys. Rev. A* **87**, 052125 (2013).
- [57] G. Kar, S. Ghosh, S. K. Choudhary, and M. Banik, Role of measurement incompatibility and uncertainty in determining nonlocality, *Mathematics* **4**, 52 (2016).
- [58] G. Kar and M. Banik, Several foundational and information theoretic implications of Bell's theorem, *Int. J. Quantum Inf.* **14**, 1640027 (2016).
- [59] N. Stevens and P. Busch, Steering, incompatibility, and Bell-inequality violations in a class of probabilistic theories, *Phys. Rev. A* **89**, 022123 (2014).
- [60] M. Banik, S. S. Bhattacharya, A. Mukherjee, A. Roy, A. Ambainis, and A. Rai, Limited preparation contextuality in quantum theory and its relation to the Cirel'son bound, *Phys. Rev. A* **92**, 030103(R) (2015).
- [61] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, Parity oblivious d -level random access codes and class of noncontextuality inequalities, *Quant. Inf. Process.* **18**, 111 (2019).
- [62] L. Masanes and M. P. Müller, A derivation of quantum theory from physical requirements, *New J. Phys.* **13**, 063001 (2011).
- [63] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Informational derivation of quantum theory, *Phys. Rev. A* **84**, 012311 (2011).
- [64] A. Peres and W. K. Wootters, Optimal Detection of Quantum Information, *Phys. Rev. Lett.* **66**, 1119 (1991).
- [65] S. Massar and S. Popescu, Optimal Extraction of Information from Finite Quantum Ensembles, *Phys. Rev. Lett.* **74**, 1259 (1995).
- [66] E. Chitambar and M.-H. Hsieh, Revisiting the optimal detection of quantum information, *Phys. Rev. A* **88**, 020302(R) (2013).
- [67] S. Halder, M. Banik, S. Agrawal, and S. Bandyopadhyay, Strong Quantum Nonlocality without Entanglement, *Phys. Rev. Lett.* **122**, 040403 (2019).
- [68] S. Rout, A. G. Maity, A. Mukherjee, S. Halder, and M. Banik, Genuinely nonlocal product bases: Classification and entanglement assisted discrimination, *Phys. Rev. A* **100**, 032321 (2019).
- [69] S. Rout, A. G. Maity, A. Mukherjee, S. Halder, and M. Banik, Local state discrimination and ordering of multipartite entangled states, [arXiv:1910.14308](https://arxiv.org/abs/1910.14308).