Eur. Phys. J. D (2019) 73: 207
https://doi.org/10.1140/epjd/e2019-90452-2

THE EUROPEAN PHYSICAL JOURNAL D

Regular Article

Hierarchical axioms for quantum mechanics*

S. Aravinda[1], Anirban Pathak[2,a], and R. Srikanth[3]

[1] Department of Physics, Indian Institute of Technology Madras, Chennai 600036, India
[2] Jaypee Institute of Information Technology, A10, Sector 62, Noida 201309, India
[3] Poornaprajna Institute of Scientific Research, Bangalore, Karnataka, India

Received 31 August 2018 / Received in final form 9 August 2019
Published online 24 September 2019

Abstract. The origin of nonclassicality in quantum mechanics (QM) has been investigated recently by a number of authors with a view to identifying axioms that would single out quantum mechanics as a special theory within a broader framework such as convex operational theories. In these studies, the axioms tend to be logically unconnected in the sense that no specific ordering of the axioms is implied. Here, we identify a hierarchy of five nonclassical features that separate QM from a classical theory. By hierarchy is meant an axiomatic scheme where the succeeding axioms can be regarded as superstructure built on top of the structure provided by the preceding axioms. In a sense, the latter are necessary, but not sufficient, for the succeeding axioms. In our scheme, the axioms briefly are: (Q1) incompatibility and uncertainty; (Q2) contextuality; (Q3) entanglement; (Q4) nonlocality and (Q5) indistinguishability of identical particles. Such a hierarchy isn't obvious when viewed from within the quantum mechanical framework, but, from the perspective of generalized probability theories (GPTs), relevant toy GPTs are introduced at each layer when useful to illustrate the action of the nonclassical features associated with the particular layer.

1 Introduction

What exactly makes quantum mechanics (QM) nonclassical? This question has been answered in different ways in quantum optics, in quantum information and in the foundations of QM [1]. For example, in quantum optics, a state is considered to be nonclassical if its Glauber-Sudarshan P function [2] cannot be described as classical probability distribution function [3], i.e., it takes negative values. Other quasi-probability distributions that are used similarly to characterize nonclassical properties of light include the Wigner distribution $W(x,p)$ and the Husimi Q distribution [4].

With the advent of quantum information theory, the study of nature of quantum correlations and of the question of what makes quantum mechanics (QM) special within a larger class of operational probability theories has been investigated recently by various authors. In quantum information theory, we associate nonclassicality with bi-partite or multi-partite *quantum* correlations that correspond to nonlocality [5], entanglement [6,7], the weaker condition of non-vanishing discord [8] and Einstein-Podolsky-Rosen steering.

Quantum correlations are nonclassical when nonlocal, in that they can violate Bell-type inequalities [9,10] that classical correlations cannot. However, the converse is not true as there exists nonclassical states which are local. In what follows, this point will be illustrated through a proposed hierarchical structure of the axioms of QM. Correlation inequalities for temporal situation can be proposed based on the assumption of realism and non-invasiveness [11].

The assumptions behind the derivation of a Bell-type inequality are localism and realism. A classical theory is necessarily local and realist. Consequently, a violation of Bell's inequality essentially implies non-classicality. Likewise, as a classical theory is necessarily non-contextual and realist, a violation of a contextuality inequality would also imply non-classicality, but as before the converse is not true.

In the context of multi-partite systems, the relation between local properties and its non-local nature has been extensively studied by various authors [12–19]. In the inverse direction, bounds on nonlocality in nonsignaling theories have been derived by assumptions about monopartite system properties like uncertainty [20] and incompatibility [21]. In reference [16], it's shown that any theory which cannot be ascribed a simplex state space structure, with pure states being one-shot distinguishable, has a no-cloning theorem. Monopartite nonclassical systems has been considered in the ontological framework by

* Contribution to the Topical Issue "Quantum Correlations", edited by Marco Genovese, Vahid Karimipour, Sergei Kulik, and Olivier Pfister.

[a] e-mail: anirban.pathak@gmail.com

Spekkens in reference [22,23]. An argument for the classicality for discrete physical theories satisfying an information theoretic axiom is presented in reference [24].

In this work, we identify five basic elements that separate QM from classical physics: (Q1) incompatibility and uncertainty; (Q2) contextuality; (Q3) quantum entanglement; (Q4) nonlocality; (Q5) indistinguishability of identical particle. A toy theory is associated with each axioms, and we list the relevant information theoretical tasks that can be achievable in those theories.

In the foundations of quantum mechanics, one often studies nonclassical features in the framework of generalized probability theories (GPTs) [12,25–27], with the aim to identify the minimal set of axioms to guarantee the nonclassical properties in QM. In this approach, QM, classical theory and a set of other nonclassical probabilistic theories can be considered as special cases in a framework of GPTs.

It may be noted that, in this context, by QM is meant the operational formulation of QM in terms of measurements, probabilities and correlations as would be observed in a laboratory experiment, and states are considered to be the lists of probabilities of outcomes, with state spaces being the convex set of such operational states. This formulation avoids terminology such as Hilbert spaces or phase that can't be directly observed. Note that, mathematically, quantum mechanics can be viewed as a non-classical probability calculus based on a non-classical propositional logic, such that quantum states are measures on a suitably defined non-Boolean (non-distributive), orthocomplemented lattice.

Other approaches to reconstruct quantum mechanics include identifying the set of quantum correlations [28] and characterizing QM in terms of information theoretic axioms [29]. For a review and analysis of various approaches of such reconstructions, see Grinbaum [30]. All these efforts represent valid ways to understand the mathematical structure of quantum mechanics, and are mutually insightful.

Yet, none of these axiomatizations of quantum mechanics is hierarchical. That is, in them the axioms appear with no logical connection to each other. In the hierarchical scheme that we propose, in contrast, the preceding axioms are necessary to provide the structure required for the later axioms, but are not sufficient to imply the later axioms, these being logically independent. Thus, our hierarchical approach allows us, informally speaking, to visualize Nature as starting with a simple theory, and then by stepwise addition of more axioms, creating the more complex theory that is QM. This physically inspired visualization is not possible in the non-hierarchical schemes of axiomatization. In that sense, our approach provides a complementary approach to understand the structure of QM.

One usually associates nonclassicality of QM with features like fundamental randomness [31], Heisenberg uncertainty, monogamy of nonlocal correlations [32], privacy of nonlocal correlations [33], and the impossibility of perfect cloning [34,35]. However, to the best of our knowledge, till now no hierarchical understanding of quantum nonclassicality is known, such that certain features are understood to be built on top of the others. This is of course because within QM, the features are mathematically inter-dependent, with no obvious ordering discernible. Here, we address this issue, by presenting a hierarchy of quantum features, that are inspired by GPT considerations.

There are certain details to which we will return elsewhere: underlying the hierarchy is the assumption of linearity, which is not specific to quantum mechanics, and exists already in classical mechanics; the hierarchical transition from (Q2) to (Q3) assumed a tensor product structure between the state space, which itself is not a nonclassical feature; (Q3) to (Q4) must include classical notions of entanglement and quantum steering. Let us clarify that our aim here is not to provide a complete axiomatic structure, rather what we offer here is a hierarchically axiomatic structure inspired by quantum information theory. The idea of the hierarchy is that later axioms sit on top of the structure provided by the earlier axioms. Such an arrangement is apparently not possible, working within the framework of standard QM. Hence we do it via the framework of GPTs.

The rest of the paper is organized as follows. In Section 3, we discuss the axiom (Q1), its relevance in quantum cryptography and show that it's possible to design a scheme of quantum key distribution in a local toy theory that allows (Q1), but does not allow other axioms. Similarly, in Section 4–7, we elucidate the axioms (Q2)–(Q5), and the toy theories valid in each layer. Quantum infromation processing tasks that can be performed in a layer is also discussed. Finally, the paper is concluded in Section 8.

2 Framework

In any laboratory experiment the basic elements are the preparation of physical system, its manipulation and finally measuring the system. A GPT can be viewed as a framework in which operationally available data, namely experimental observations based on state preparations, manipulations and measurements, form the basis of the theory. Thus, concepts such as "phase" or "Hilbert space", which are not directly observable, are avoided in preference to observed experimental probabilities. Mathematically, a physical system is represented by its state space Ω, which is a convex set of allowed states ω. A pure state ω represents an operationally equivalent class of preparations. Any measurement X_i can be represented as the set of effects (yes/no measurements) $e_i : e(\omega_i) \to \{0, 1\}$, such that $\sum_i e_i = u$, where u is the "unit effect" defined by $u(\omega) = 1$ for all allowed states ω. Here, $e(\omega)$ represents the probability that the "yes" event corresponding to the effect e happens under the measurement, when the system is in the state ω. A physical theory is characterized by a probability rule that determines the outcome statistics of any measurement on any state.

3 Incompatibility and uncertainty

Single systems constitute the simplest system to study nonclassicality. Multipartite structure can be built upon it by assuming further assumptions related to the compositional structure. For single systems, two fundamental facets of nonclassicality are uncertainty and

measurement incompatibility. The former postulates the existence of observables whose values cannot be jointly specified to arbitrary accuracy. The latter postulates the existence of observables whose values cannot be jointly measured in the GPT. Complementarity refers to the smallest value of mixing or "unsharpness" for which two observables will be compatible, in the sense that they can be obtained as marginals of a master observable [36].

For projective measurements in QM, the incompatibility of two observables coincides with non-commutativity and uncertainty. In an arbitrary GPT, incompatible observables may lack uncertainty, an example being generalized bit (gbit) or generalized dit (gdit) theory [37]. In this work, we observe that it's convenient to treat both uncertainty and incompatibility as existing on the same, basic level of nonclassicality, which we have designated Q1. In this connection, we note that both uncertainty and incompatibility have been separately implicated in bounding nonlocality in a GPT, in references [20] and [21], respectively.

No-cloning, superposition, indistinguishability, measurement disturbance, uncertainty in measurement outcomes, randomness associated with the measurement, etc. are other important nonclassical features associated with single systems, besides incompatibility. In reference [37], the present authors have showed that all these properties are interrelated and associated with the non-simplicial structure of the state space. Thus, in our hierarchical axiomatics, incompatibility of the observables forms the lowest layer.

Let X_1 and X_2 be two observables with n outcomes. The observables X_1 and X_2 are jointly measurable or both the observables are compatible if there exists a joint observable X_{12} such that the statistics of both the observables X_1 and X_2 can be obtained by marginalizing X_{12} : $M_{X_1}(x_1|\omega) = \sum_{x_2} M_{X_{12}}(x_1, x_2|\omega)$ and $M_{X_2}(x_2|\omega) = \sum_{x_1} M_{X_{12}}(x_1, x_2|\omega), \forall \omega \in \Omega$.

The cryptographic power of incompatibility, without any other nonclassical feature being assumed, can be illustrated via the BB84-like key distribution protocol, which we call "local key distribution" (LKD).

LKD works as follows. Alice and Bob each have two copies of the key to a strongbox. Alice opens the box and leaves a random bit $\kappa \in \{0, 1\}$ for Bob to read, and then locks the box. Bob comes along later, open the box with his key, and receives his message.

In the real world, a classical implementation of LKD is not unconditionally secure, since classical laws do not preclude that an eavesdropper can break open the box, read the secret bit κ and then rebuild the box, without Alice and Bob knowing about it.

Consider the two-input-two-output operational theory $\mathcal{T}$, with two dichotomic measurements $\mathtt{X}$ and $\mathtt{Z}$. The pure states of the theory $\mathcal{T}$, which forms the state space Σ, are:

$$\begin{aligned} \psi_{\mathtt{X}}^{+} &\equiv (1, 0 \mid \tfrac{1}{2}, \tfrac{1}{2}) \\ \psi_{\mathtt{X}}^{-} &\equiv (0, 1 \mid \tfrac{1}{2}, \tfrac{1}{2}) \\ \psi_{\mathtt{Z}}^{+} &\equiv (\tfrac{1}{2}, \tfrac{1}{2} \mid 1, 0) \\ \psi_{\mathtt{Z}}^{-} &\equiv (\tfrac{1}{2}, \tfrac{1}{2} \mid 0, 1), \end{aligned} \tag{1}$$

where $\psi \equiv (P(0|X), P(1|X) \mid P(0|X), P(1|X))$ represents the state with $0 \leq P(m|M) \leq 1, \sum_m P(m|M) = 1$, and $P(m|M)$ is the probability of getting outcome m upon measuring M.

The nonsimpliciality condition corresponds to the operational indistinguishability of two mixtures which are prepared differently from pure states. For example,

$$\frac{1}{2}(\psi_{\mathtt{X}}^{+} + \psi_{\mathtt{X}}^{-}) = \frac{1}{2}(\psi_{\mathtt{Z}}^{+} + \psi_{\mathtt{Z}}^{-}) = \left(\frac{1}{2}, \frac{1}{2} \mid \frac{1}{2}, \frac{1}{2}\right). \tag{2}$$

Thus, an equal weight convex combination of pure states $\psi_{\mathtt{X}}^{+}$ and $\psi_{\mathtt{X}}^{-}$ and of pure states $\psi_{\mathtt{Z}}^{+}$ and $\psi_{\mathtt{Z}}^{-}$ are indistinguishable. A theory $\mathcal{T}$ whose state space manifests such non-simpliciality is nonclassical.

Now consider the following protocol implemented in the nonclassical but noncontextual theory given by (1). It is the LKD version of the Bennett-Brassard 1984 (BB84) quantum key distribution (QKD) protocol [38]; and we may refer to it as **BB84 LKD** [39]

1. Alice randomly prepares n particles in one of the four states $|\psi_X^{\pm}\rangle, |\psi_Z^{\pm}\rangle$ given by (1) by measuring X or Z on each particle. She transmits them to Bob.
2. Bob measures the particles randomly in the basis X or Z. He notes the outcomes τ.
3. On the key string so extracted, Alice and Bob publicly discuss to retain only those outcomes where their bases agree; this forms their raw key string;
4. They agree on certain coordinates and announce the outcomes on those coordinates. If too many of them are mismatched, they deem the protocol round secure and abort the round.
5. Else Alice and Bob proceed to classically extract a secure, smaller key from the remaining bits.

Security of the above protocol originates from the fact that although Eve can deterministically extract the encoded bit by measurement if she measures in the right basis, she will produce measurement disturbance if she gets the basis wrong, which can be detected in Step (3).

Suppose Eve implements such an intercept-resend attack, by measuring X or Z on m gbits from a total of n particles transmitted. She will be able to extract $I(A{:}E) = \frac{m}{2}$ bits of information on average, where $I(A{:}E) = H(A) - H(A|E)$ is the mutual information, with $H(A)$ being the entropy and $H(A|E)$ the conditional entropy. Let $f \equiv \frac{m}{n}$. On average, Alice and Bob will check the basis not used by Eve half the times, and half of these times, they would obtain the answer not encoded by Alice. On the remaining fraction, their measured outcome will be consistent with Alice's encoded value. Thus, the error observed by Alice and Bob is $e = \frac{f}{4}$, so that on average Bob receives $I(A:B) = \left(1 - h\left[\frac{f}{4}\right]\right)$ bits of information per transmitted gbit.

The protocol can be shown to be secure if $I(A:B) \geq I(A:E)$ [40], which in this case becomes

$$\left(1 - h\left[\frac{f}{4}\right]\right) \geq \frac{f}{2} \tag{3}$$

or $f \approx 0.68$, so that the tolerable error rate $e_{\max} = 0.48/4 = 17\%$. The probability that Eve is not detected on an

attacked particle is $\frac{3}{4}$. Therefore the probability that she escapes detection on all the m bits she attacks is $(3/4)^m$, which falls exponentially with security parameter m. This exponential drop characterizes *unconditional* security.

However, the security described above is not *device independent* (cf. [41], and references therein). Alice and Bob implicitly assume in the BB84 LKD protocol that the preparation and measurement devices are trustworthy. Now suppose that the device has been manufactured by Eve such that each actual particle is replaced by a clandestine random 2-bit preparation, such that when Alice measures X, the pre-existing bit is presented, and similarly for if she measures Z. If (subsequently) Bob measures the same observable, he would obtain the same bit as obtained by Alice, else an uncorrelated bit. This reproduces the BB84 statistics. It is entirely insecure once Eve learns about their respective bases during their public discussion. This is just the higher-dimension attack [42] adapted from BB84 to the present protocol. Thus, BB84 LKD is not secure in the device independent scenario.

4 Contextuality

A GPT having the property of incompatibility of observables, even though it possesses many nonclassical properties, in some sense, lacks the full essence of nonclassicality because of the possibility of the existence of a deterministic hidden variable model for such theories. For single systems, the stronger form of nonclassicality emerges through the impossibility of providing a non-contextual deterministic hidden variable model [43]. A theory which possesses the contextuality feature is referred to as a contextual theory.

Consider a theory with three observables A, B and C. If all of them are compatible, then clearly there exists a joint probability distribution for them, obtained by measuring the observables simultaneously. Thus, contextuality requires that there should be incompatibility among the observables, meaning that incompatibility is a necessary condition for contextuality. On the other hand, it is not sufficient. Consider a case where observables A and B are pairwise compatible, as well as B and C are pairwise compatible, but A and C are incompatible. In this case, one can always write a joint distribution given by

$$P(A,B,C) = \frac{P(A,B)P(B,C)}{P(B)}, \tag{4}$$

which would be sufficient to reproduce the experimentally observed probabilities. This shows that incompatibility is not sufficient to produce contextuality. Thus, in the axiomatic hierarchy, contextuality can be considered as a natural superstructure to include in theories with incompatibility, and naturally forms the next layer of nonclassicality.

Let $P(m|\omega, M)$ represent the outcome statistics from the state ω and the measurement M with outcome labelled as m. Suppose, λ is denoted as an underlying state in an ontological model that reproduces the given GPT, then

$$P(m|\omega, M) = \int \mu_\omega(\lambda)\xi(m|\lambda, M)d\lambda,$$

where $0 \leq \mu_\omega(\lambda) \leq 1$ is the probability density over an ontic variable λ, $\xi(m|\lambda, M)$ is an indicator function which represents the probability of getting an outcome m upon measuring M on the ontic state λ.

An ontological model is said to be preparation noncontextual if for any set of equivalent preparations i.,e given $P(m|\omega, M) = P(m|\omega', M), \forall M$, it holds true that this equivalence holds in the ontological model, i.e, $\mu_\omega(\lambda) = \mu_{\omega'}(\lambda)$. Similarly, an ontological model is said to be measurement noncontextual if for any equivalent set of measurements, i.e., given $P(m|\omega, M) = P(m|\omega, M'), \forall \omega$, it holds true that $\xi(m|\lambda, M) = \xi(m|\lambda, M')$.

Outcome determinism is the assumption that the indicator function outputs deterministically, i.e., $\xi(m|\lambda, M) \in \{0, 1\}$. Measurement noncontextuality along with outcome determinism is the assumption to derive noncontextuality inequality, violation of which determines the theory as contextual.

An example of such a contextual inequality

$$\langle \mathtt{VW} \rangle + \langle \mathtt{WX} \rangle + \langle \mathtt{XY} \rangle + \langle \mathtt{YZ} \rangle + \langle \mathtt{ZV} \rangle \geq -3, \tag{5}$$

where $\mathtt{V}, \mathtt{W}, \mathtt{X}, \mathtt{Y}, \mathtt{Z} = \pm 1$ [44]. One may check by direct substitution that no deterministic, non-contextual value assignments to these five observables can violate the inequality.

Now, we consider a nonclassical theory with nontrivial congruence structure, characterized by five observables $\mathtt{V}, \mathtt{W}, \mathtt{X}, \mathtt{Y}, \mathtt{Z}$ and the cyclic R_2-chain: $R_2(\mathtt{V}, \mathtt{W})$, $R_2(\mathtt{W}, \mathtt{X})$, $R_2(\mathtt{X}, \mathtt{Y})$, $R_2(\mathtt{Y}, \mathtt{Z})$, $R_2(\mathtt{Z}, \mathtt{V})$, and every other pair being incongruent. By the assumption of tomographic separability, an arbitrary state is assumed to be completely specified by the five fiducial probabilities $P_{\mathtt{V}}, P_{\mathtt{W}}, P_{\mathtt{X}}, P_{\mathtt{Y}}$ and $P_{\mathtt{Z}}$, which in turn determine $P_{\mathtt{VW}}, P_{\mathtt{WX}}$, $P_{\mathtt{XY}}, P_{\mathtt{YZ}}$ and $P_{\mathtt{ZV}}$, assumed to be consistent with contextual no-signaling. We shall refer to this theory (fragment) as $\mathcal{T}_{\mathrm{KCBS}}$, in view of the work where the contextuality of such correlations was studied [44].

Consider a state ρ in this contextual theory, where all the above pairs produce perfectly random but anticorrelated outcomes, i.e., 01 or 10. We can check directly that there is no way to assign values 0 and 1 to these five observables (indeed, any odd number of observables) in such a way as to satisfy this requirement, because there would be a clash of values on at least one observable. That is, if $\mathtt{A}$ has value $\mathtt{a}$, then $\mathtt{B}$ has $\overline{a}$, $\mathtt{C}$ has a, $\mathtt{D}$ has $\overline{a}$ so that $\mathtt{E}$ has a requiring $\mathtt{A}$ to have $\overline{a}$, contrary to assumption. Thus, ρ does not correspond to a state that has a JD over the five variables, where they take definite values. This is witnessed by the violation of the KCBS inequality, equation (5).

Now consider the following protocol implemented in $\mathcal{T}_{\mathrm{KCBS}}$: **KCBS LKD.**

1. Alice prepares n particles in state ρ, which she leaves at a pre-agreed location, after measuring each using one of the five observables $\mathtt{V}, \mathtt{W}, \mathtt{X}, \mathtt{Y}, \mathtt{Z}$.

2. Later Bob arrives at the location and measures the particles randomly in any one of these five bases.
3. Alice and Bob announce their measurement bases and throw away the (approximately 40%) data corresponding to instances where their bases aren't either identical or juxtaposed.
4. They publicly agree on certain coordinates of particles, and disclose their measurement outcomes for these. If their selected bases are identical (resp., juxtaposed), they verify that the outcomes are random and identical (resp., anti-correlated). If too many of them fail this criterion, then they abort the protocol.
5. Else Alice and Bob proceed to classically distil a secure, smaller key from the remaining bits.

In this case note that no pre-existing record can reproduce the KCBS perfect correlations. Thus, any *passive* cheat device of the type mentioned above will fail to pass the KCBS test, giving rise to a kind of device independence. Note that an *active* cheat device, meaning one that allows memory to be carried forward in time (which is a kind of local signaling), can defeat the protocol. For example, the device produces an arbitrary output when Alice measures. Her basis and outcome information are retained in the system's memory, so that if Bob measures in the basis as she did, then the device produces the identical (resp., anticorrelated) outcome if his basis matches (resp., is juxtaposed to) hers, and a random outcome otherwise. Note that the memory corresponds to a kind of signal formally, but which is not prohibited by special relativity, since the correlation is local. We shall refer to this scenario where Eve that is restricted from memory attacks in a local protocol, as *bounded device independence*.

Thus, contextuality can provide security to LKD in the restricted device-independent scenario where memorylessness (with regard to Alice's input) is assumed, but a nonclassical theory without contextuality lacks security even in this memoryless device-independent scenario. Note that in the case where device independence is based on nonlocality, the memory of Alice's actions (and outcomes) cannot be transmitted in light-travel time to Bob's spacelike separated measurement event, thereby preventing Eve from launching the above kind of attack. In fact, without assumptions about device trustworthiness, that would be the only way to prevent Eve's device attack. Thus, full device independence requires nonlocality, and contextuality will be insufficient.

5 Entanglement

All axioms covered so only concern single system nonclassicalities. To move beyond single systems to multipartite systems, the most basic issue concerns how two or more systems can be combined. The natural question that arises here is: does the theory only allow the trivial combination of states of different systems by forming a direct product, or does the theory allow more generality, i.e., the nonclassical feature is the non-seperability or entanglement in the state space Ω of the multipartite system?

In the GPT framework, considering two systems, the state space Ω of a composite system must lie between two extremes: the *maximal tensor product* of state spaces of its constituents, Ω_1 and Ω_2, represented as $\Omega_1 \otimes_{\max} \Omega_2$ and the minimal tensor product $\Omega_1 \otimes_{\min} \Omega_2$. The latter is the convex hull of set of product states, whilst the former is the set of composite states that yield a valid probability distribution corresponding to product effects. The state of a composite system lying outside the minimal tensor product is called entangled.

Quantum entanglement, an important element which deviates our classical world view, acts as a resource for many information theoretic as well as many computational advantages, providing a dominent non-classical feature [6]. Even though, as few results claim, the existence of entanglement in classical optics, it lacks any information theoretic advantages. An important task which reveals entanglement and also of necessary is teleportation. It is been proved that entanglement is sufficient for teleportation in any GPTs. The important relation revealed by Spekkens toy model [23] as well as the toy model proposed by Hardy shows the difference between entanglement and non-locality [45]. They show that both entanglement and teleportation is possible in local theory, thus seperating nonlocality from entanglement.

6 Nonlocality

In a scheme for QKD, suppose at the end of the quantum part, Alice's and Bob's joint probability $P(ab|xy)$ is described by:

$$P(ab|xy) = \sum_{\mu} p(\mu), P(a|x,\mu)P(b|y,\mu), \tag{6}$$

with $p(\mu)$ being a probability distribution over parameter μ. Since this is potentially preparation information with Eve, the condition of security in the device-independent sense is that $P(ab|xy)$ shouldn't have the above form, i.e., should be a nonlocal correlation [10]. Thus, including the axiom of nonlocal correlations being allowed in the theory enables DI security in the usual paradigm of cryptography.

The most general state that lies within the minimal tensor product $\bigotimes_{\min}$ of two systems is given by

$$\rho = \sum_{j,k} p_{j,k} \sigma_j^{(I)} \otimes \sigma_k^{(II)}, \tag{7}$$

where the superscripts indicate the system. Suppose any product measurement is applied to the state in equation (7). In this case, the single system outcome probabilities are defined by local measurement, and it immediately follows that the resulting joint probability distribution will have the local form equation (6). It follows that entanglement is necessary for nonlocality. On the other hand, Werner states [46] provide simple example to show that entanglement is not sufficient for the generated correlations to be nonlocal, i.e. that the correlations lack a local hidden variable description.

Thus, nonlocality is a stronger nonclassical feature than entanglement, and forms a natural higher stage in our hierarchy.

This particular step in our hierarchy is of course well recognized from various considerations: for example, [47] notes how maximally and non-maximally entangled states place different requirements on a "nonlocal machine". In [48], an instance of states that are genuinely tripartite and yet local is pointed out. That a logarithmical higher degree of depolarizing noise is required to make two entangled qudits separable than to make them local, is shown in reference [49]. This inequivalence between nonlocality and entanglement is known to hold in going beyond bipartite correlations to an arbitrary number of parties [50], and also in going beyond projective measurements to positive operator valued measures [51].

Moreover, in reference [51] steering is shown to be inequivalent to entanglement and Bell nonlocal with respect to such general measurements, and to be intermediate between these two. In this light, one may place steering hierarchically between entanglement (axiom Q3) and nonlocality (axiom Q4). Such a placement is indeed supported by GPT considerations (see e.g., [52]), but for simplicity, we have not considered it in the present version of our scheme.

Indeed, in any non-signaling theory, nonlocality can be the basis for distilling shared secret randomness [53].

7 Indistinguishability of identical particles

Our final axiom is indistinguishability, relatively less studied aspect of quantum nonclassicality both in quantum information processing and the GPT framework. A task that separates quantum mechanics from classical mechanics, that wouldn't be possible even with our above axioms (i.e., quantum mechanics based purely on distinguishable particles), is boson sampling, a task that becomes easy when quantum indistinguishability is included.

Boson sampling, introduced in reference [54] and experimentally realized in references [55,56] and references therein, is the task of exactly or approximately sampling from the probability distribution of identical bosons scattered by a linear interferometer. It's widely believed that this task is intractable in the classical world (i.e, intractable for classical computers), but can be solved efficiently in the quantum world. In other words, inclusion of (Q6) in a nonclassical theory provides us the ability of solving boson sampling problem. The appearance of the permanent in the outcome statistics of single-photon measurements makes the task computationally hard, whereas the corresponding linear optics uses only polynomial resources to implement it practically. Here, we note that the permanent of a square matrix in linear algebra is a determinant-like function of the matrix, and a special case of *immanant*, a more general matrix function.

The reason this is interesting from a computational perspective is that the probability distribution that boson sampling device is required to sample from, which as noted above is connected to the permanent of a complex matrix. Computing the permanent, as well as approximating it within multiplicative error, are known in the general case to be in the #P-hard computational complexity class. Therefore, just as quantum nonlocality suggests a "behind-the-scenes" super-classical communication, so does bosonic sampling suggest a "behind-the-scenes" super-classical computing, that in some ways is more spectacular than the quantum speedup witnessed in Shor's prime factorization algorithm.

8 Conclusion

We have proposed a hierarchy of axioms for quantum mechanics, meant to bring out the increasing structure in the theory as a departure from classical mechanics, rather than to derive quantum mechanics per se.

These axioms, inspired by considerations from GPTs or convex operational theories, are in their proper order given by: (Q1) incompatibility and uncertainty ; (Q2) contextuality; (Q3) entanglement; (Q4) nonlocality and (Q5) indistinguishability of identical particles.

The hierarchy is ordered in such a way that the preceeding axioms are necessary for the succeeding axioms, but are not sufficient to imply them. Exceptions here are the step from contextuality (Q2) to entanglement (Q3), which is simply a movement from a nonclassical theory of single systems to that of composite systems; and, similarly from nonlocality (Q4) to indistinguishability (Q5). Where possible, we have employed the security of a quantum information processing task or of a GPT task as a manifestation that exemplifies the hierarchy, such that the security could not be guaranteed without the given level of nonclassicality in the hierarchy.

We hope that this work would shed light on the question of what makes quantum mechanics special to be singled out by Nature.

AP thanks Defense Research and Development Organization (DRDO), India for the support provided through the project number ERIPR/ER/1403163/M/01/1603. RS thanks Department of Science & Technology - Science and Engineering Research Board (DST-SERB), India, for financial support provided through the project EMR/ 2016/004019.

Author contribution statement

All the authors have contributed equally, and all of them were involved in the preparation of the manuscript. All the authors have read and approved the final manuscript.

References

1. S. Banerjee, A. Pathak, R. Srikanth, Physically inspired axioms for quantum mechanics (in press)
2. L. Mandel, E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995)
3. R. Loudon, *The Quantum Theory of Light* (Oxford University Press, 2000)
4. U. Leonhardt, *Measuring the Quantum State of Light*, Cambridge Studies in Modern Optics (Cambridge University Press, 1997)
5. M. Genovese, Phys. Rep. **413**, 319 (2005)

6. R. Horodecki, P. Horodecki, M. Horodecki, K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009)
7. F. Verstraete, J. Dehaene, B. de Moor, J. Mod. Opt. **49**, 1277 (2002)
8. H. Ollivier, W.H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001)
9. J. Bell, Physics **1**, 195 (1964)
10. J.F. Clauser, M.A. Horne, A Shimony, R.A. Holt, Phys. Rev. Lett. **23**, 880 (1969)
11. C. Brukner, S. Taylor, S. Cheung, V. Vedral, arXiv:quant-ph/0402127v1
12. J. Barrett, Phys. Rev. A **75**, 032304 (2007)
13. J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, D. Roberts, Phys. Rev. A **71**, 022101 (2005)
14. G. Brassard, H. Buhrman, N. Linden, A.A. Méthot, A. Tapp, F. Unger, Phys. Rev. Lett. **96**, 250401 (2006)
15. A. Broadbent, A.A. Methot, Theor. Comput. Sci. **3**, 358 (2006)
16. H. Barnum, J. Barrett, M. Leifer, A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007)
17. H. Barnum, Stud. Hist. Philos. Sci. B **34**, 343 (2003)
18. G. Chiribella, G.M. D'Ariano, P. Perinotti, Phys. Rev. A **81**, 062348 (2010)
19. P. Janotta, H. Hinrichsen, J. Phys. A: Math. Theor. **47**, 323001 (2014)
20. J. Oppenheim, S. Wehner, Science **330**, 1072 (2010)
21. M. Banik, M.R. Gazi, S. Ghosh, G. Kar, Phys. Rev. A **87**, 052125 (2013)
22. R.W. Spekkens, Phys. Rev. A **71**, 052108 (2005)
23. R.W. Spekkens, Phys. Rev. A **75**, 032110 (2007)
24. C. Pfister, S. Wehner, Nat. Commun. **4**, 1851 (2013)
25. L. Hardy, Quantum theory from five reasonable axioms, arXiv:quant-ph/0101012 (2001)
26. L. Hardy, Stud. Hist. Philos. Sci. B **34**, 381 (2003)
27. P. Mana, arXiv:quant-ph/0305117 (2003)
28. M. Navascués, S. Pironio, A. Acín, New J. Phys. **10**, 073013 (2008)
29. R. Clifton, J. Bub, H. Halvorson, Found. Phys. **33**, 1561 (2003)
30. A. Grinbaum, Br. J. Philos. Sci. **58**, 387 (2007)
31. S. Popescu, D. Rohrlich, Found. Phys. **24**, 379 (1994)
32. B. Toner, Proc. R. Soc. A **465**, 59 (2009)
33. H.K. Lo, H.F. Chau, Science, **283**, 2050 (1999)
34. W.K. Wootters, W.H. Zurek, Nature **299**, 802 (1982)
35. N. Gisin, Phys. Lett. A **242**, 1 (1998)
36. P. Busch, T. Heinosaari, J. Schultz, N. Stevens, EPL **103**, 10002 (2013)
37. S. Aravinda, R. Srikanth, A. Pathak, J. Phys. A: Math. Theor. **50**, 465303 (2017)
38. C.H. Bennett, G. Brassard, Quantum cryptography: Public key distribution and coin tossing, in *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore* (1984), p. 175
39. S. Aravinda, A. Banerjee, A. Pathak, R. Srikanth, Int. J. Quantum Inf. **12**, 1560020 (2014)
40. I. Csiszar, J. Korner, IEEE Trans. Inf. Theory **24**, 339 (1978)
41. U. Vazirani, T. Vidick, Phys. Rev. Lett. **113**, 140501 (2014)
42. J. Barrett, L. Hardy, A. Kent, Phys. Rev. Lett. **95**, 010503 (2005)
43. S. Kochen, E.P. Specker, J. Math. Mech. **17**, 59 (1967)
44. A.A. Klyachko, M. Ali Can, S. Binicioğlu, A.S. Shumovsky, Phys. Rev. Lett. **101**, 020403 (2008)
45. L. Hardy, arXiv:quant-ph/9906123 (1999)
46. R.F. Werner, Phys. Rev. A **40**, 4277 (1989)
47. N. Brunner, N. Gisin, V. Scarani, New J. Phys. **7**, 88 (2005)
48. G. Tóth, A. Acín, Phys. Rev. A **74**, 030306 (2006)
49. M.L. Almeida, S. Pironio, J. Barrett, G. Tóth, A. Acín, Phys. Rev. Lett. **99**, 040403 (2007)
50. R. Augusiak, M. Demianowicz, J. Tura, A. Acín, Phys. Rev. Lett. **115**, 030404 (2015)
51. M.T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, N. Brunner, Phys. Rev. A **92**, 032107 (2015)
52. R Srikanth, arXiv:1811.12409 (2018)
53. Ll. Masanes, A. Acin, N. Gisin, Phys. Rev. A **73**, 012112 (2006)
54. S. Aaronson, A. Arkhipov, The computational complexity of linear optics, in *Proceedings of the forty-third annual ACM symposium on Theory of computing* (ACM, 2011), pp. 333–342
55. M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, P. Walther, Nat. Photonics **7**, 540 (2013)
56. H. Wang, W. Li, X. Jiang, Y.-M. He, Y.-H. Li, X. Ding, M.-C. Chen, J. Qin, C.-Z. Peng, C. Schneider, M. Kamp, W.-J. Zhang, H. Li, L.-X. You, Z. Wang, J.P. Dowling, S. Hofling, C.-Y. Lu, J.-W. Pan, Phys. Rev. Lett. **120**, 230502 (2018)