Colloquium: Bell's theorem and locally mediated reformulations of quantum mechanics

K. B. Wharton*

Department of Physics and Astronomy, San José State University, San José, California 95192-0106, USA

N. Argaman†

Department of Physics, Nuclear Research Center—Negev, P.O. Box 9001, Be'er Sheva 84190, Israel

(published 18 May 2020)

Bell's theorem rules out many potential reformulations of quantum mechanics, but within a generalized framework it does not exclude all locally mediated models. Such models describe the correlations between entangled particles as mediated by intermediate parameters that track the particle worldlines and respect Lorentz covariance. These locally mediated models require the relaxation of an arrow-of-time assumption that is typically taken for granted. Specifically, some of the mediating parameters in these models must functionally depend on measurement settings in their future, i.e., on input parameters associated with later times. This option, often called retrocausal, has been repeatedly pointed out in the literature, but the exploration of explicit locally mediated toy models capable of describing specific entanglement phenomena has begun only in the past decade. A brief survey of such models is included here. These models provide a continuous and consistent description of events associated with spacetime locations, with aspects that are solved "all at once" rather than unfolding from the past to the future. The tension between quantum mechanics and relativity that is usually associated with Bell's theorem does not occur here. Unlike in conventional quantum models, the number of parameters needed to specify the state of a system does not grow exponentially with the number of entangled particles. The promise of generalizing such models to account for all quantum phenomena is identified as a grand challenge.

DOI: 10.1103/RevModPhys.92.021002

CONTENTS

I. INTRODUCTION

Bell's theorem places a strong restriction on reformulations of quantum mechanics (QM): any mathematical model that produces the same output predictions as QM, given the same inputs, must violate *local causality*[1] (Bell, 1964). In this sense, QM is nonlocal, but locality is not a simple yes or no question;

*kenneth.wharton@sjsu.edu
†argaman@mailaps.org

[1]Italic font is used for some prominent mathematical conditions that are explicitly defined in Sec. II.B.

QM is still local according to the operational definition, as it does not allow signaling at a distance, or outside the future light cone. Thus, QM may even be compatible with a generalized nonoperational definition of locality, not as strict as *local causality*, but in the spirit of Einstein's arguments against action at a distance. This Colloquium will examine a category of potential reformulations of QM that are as "local" as allowed by Bell's theorem.

To assess a model's locality, even at the operational level of inputs and outputs, the model must define its spacetime-based parameters [those that Bell (1976) called local beables]. Such parameters are mathematical variables that are clearly associated with a specific time and place, such as the local values of classical fields $Q(\boldsymbol{x}, t)$. This association allows concepts of locality to be meaningfully applied.

Conventional QM utilizes inputs corresponding to values that can be controlled by experimental physicists, including the settings of preparation and measurement devices, and it predicts the probability distribution of the values of measureable outputs. Reformulations of QM must be operationally equivalent, utilizing these same inputs and outputs and providing the same predictions. At minimum, this requires the use of spacetime-based parameters for the inputs and the outputs.

These input and output parameters do not continuously span the intermediate spacetime regions where preparations and measurements are not performed; nonoperational definitions of locality (such as *local causality*) concern parameters associated with these intermediate regions. If a model has no spacetime-based parameters associated with these regions, it will be nonlocal according to these definitions, i.e., it will have unmediated action at a distance.

Many physicists see Bell's theorem as a reason not to introduce such mediating parameters; see, e.g., Mermin (1985). It is difficult to map an entangled configuration-space wave function $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ onto spacetime-based parameters $Q(\boldsymbol{x}, t)$ (Norsen, 2010; Stoica, 2019), and even if there were such a mapping, Bell's theorem tells us that no reformulation could possibly conform to *local causality*. However, such a viewpoint presumes that there is no other type of locality worth saving, no subset of assumptions inside of *local causality* that might be beneficial for some future reformulation of QM.

In fact, there exists a category of quantum reformulations for which an essential aspect of locality can be retained; see, e.g., Costa de Beauregard (1953), Price (1997), and Argaman (2010). These models utilize spacetime-based parameters associated with intermediate regions between preparations and measurements, allowing these models to be "locally mediated," in the sense that correlations cannot be introduced or altered except via intermediate spacetime-based mediators. This condition will be explicitly defined in Sec. II, using the term *continuous action* to contrast with the phrase action at a distance. We are most interested in cases where this local mediation is always restricted to timelike or lightlike worldlines, allowing those models to also respect Lorentz covariance.

Such locally mediated reformulations of QM must violate a certain time-asymmetric assumption inherent to *local causality*. Specifically, the relevant assumption presumes that no model parameter associated with time t can be dependent upon model inputs associated with times greater than t. The most local reformulations of QM, those with *continuous action*, violate this assumption and are therefore future-input dependent. While some may view such "retrocausality" as unreasonable, we emphasize that only models with the same predictions as QM are of interest here, with no signaling into the past; see, e.g., Price (1997).

If one considers inputs to a model to also include boundary conditions, it is evident that future-input-dependent models are ubiquitous throughout physics. For example, models employing the stationary action principle fall into this category: mathematical inputs constrain both initial and final parameters, and the model determines the classical history at intermediate times. Any calculation of a closed-timelike curve in general relativity requires a similar all-at-once analysis. Quantum future-input-dependent models, such as the transactional interpretation (Cramer, 1980) and the two-state-vector formalism (Aharonov and Vaidman, 1991), have also been developed, motivated primarily by time symmetry rather than locality.

Attempts to develop a locally mediated account of quantum entanglement using future-input dependence have been promoted by a number of forward-thinking authors, beginning even before Bell's work (Costa de Beauregard, 1953, 1977, 1979; Pegg, 1982; Sutherland, 1983; Price, 1984). Still, mathematical future-input-dependent models reproducing the QM predictions for entangled particles, while explicitly maintaining local mediation, have been put forward mainly in the last decade. One purpose of this Colloquium is to survey the admittedly modest achievements of this recent and still-developing line of enquiry, and to indicate some intriguing directions for further exploration.

A formal development of these arguments will also result in a useful categorization of all of the ways in which a reformulation of QM can violate *local causality*; we say that such models are "Bell compatible". This is accomplished by presenting the assumptions of Bell's theorem in terms consistent with the recently developed framework of "causal models" (Pearl, 2009), which emphasizes the role of QM's controllable inputs. Bell himself spoke of the special importance of inputs, calling them "free external variables in addition to those internal to and conditioned by the theory" (Bell, 1977). Unfortunately, the mathematics of causal models was not well developed during his lifetime, and Bell (1976, 1981, 1990) adopted a neutral notation, e.g., $\{A|a, \lambda\}$ for the probability distribution of an output A given an input a and an internal parameter λ. Instead, we will denote this by $p_a(A|\lambda)$, emphasizing the input status of a and allowing a clear categorization of Bell-compatible reformulations of QM (while setting aside extraneous issues such as "superdeterminism").

Developing reformulations of existing theories has historically been very useful: think of the advances of Lagrangian and Hamiltonian classical mechanics. In quantum theory, the path integral has similarly led to new insights, and there is no indication that this strategy of seeking further reformulations has run its course (Feynman, 1965). In particular, an alternative quantum model with parameters restored to functions on spacetime, instead of a multidimensional configuration space (or a Hilbert space), would have significant advantages. Such a model would have a natural interpretation, with one allowed combination of the spacetime-based parameters corresponding to physical reality, and all other combinations being mere possibilities (as in classical statistical mechanics). As a result, the number of parameters describing an actual system would grow linearly (rather than exponentially) with the extent of that system. This would substantially lessen the disconnect between quantum theory and our linearly scaling classical theories of relativistic spacetime.

Note that a successful reformulation of QM in terms of spacetime-based parameters would certainly not imply that quantum theory was incorrect. Quantum states could still represent our best possible knowledge about measurable aspects of those parameters, given accessible information. In this case, quantum states could be viewed as states of knowledge, a popular perspective in the field of quantum information (Caves, Fuchs, and Schack, 2002; Spekkens, 2007; Leifer and Spekkens, 2013).

Section II carefully walks us through Bell's theorem, identifying all of the assumptions leading to the contradiction with quantum phenomena. Section III then categorizes Bell-compatible reformulations of QM. Several examples of locally mediated toy models are detailed in Sec. IV; those who want to look at a concrete mathematical model, rather than follow general reasoning, are referred to the model of Sec. IV.B, for which a detailed derivation is given in the Appendix. Section V discusses the approach and indicates avenues for further development. Alternative approaches are briefly discussed in Sec. VI. Section VII provides the conclusion, encouraging the future development of locally mediated reformulations of QM.

II. BELL'S THEOREM

Our first task is to prove Bell's theorem. Starting with a certain set of natural assumptions, we will give a mathematical proof of a Bell inequality: specifically, the Clauser-Horne-Shimony-Holt (CHSH) inequality (Clauser *et al.*, 1969). This inequality can be tested operationally (without reference to any underlying theory), and it is experimentally violated, just as predicted by QM. It follows that, for any model of these phenomena, at least one of the assumptions that lead to Bell's theorem must be violated. All such Bell-compatible models can then be usefully categorized in terms of which assumptions are relaxed.

The analysis is conducted as follows: Sec. II.A defines the framework rules for the models to be discussed, Sec. II.B lists the relevant reasonable-but-optional assumptions that could characterize such models, Sec. II.C provides some historical context, and Sec. II.D provides a derivation of the theorem.

A. Framework: Spacetime-based models

In all of physics, one uses mathematical models to generate falsifiable predictions that can be compared with empirical observations. The sort of models that accomplish this are essentially functions that take some parameters as inputs and generate other parameters as outputs.[2] We are therefore interested in models that come with well-defined input parameters (inputs for short), which will be denoted by the set I, and also well-defined output parameters (outputs), denoted by O. Models can have other parameters in addition to the inputs and outputs, and the set of these will be denoted by U. We will often discuss the set of all noninput parameters Q (the union of O and U). Parameters here are not limited to simple scalars: vectors or more complicated mathematical constructs such as functions may be utilized.

As discussed in the Introduction, we are interested in models of spacetime-based parameters, each associated with a particular location in ordinary spacetime. Examples of such parameters include the values of physical fields, such as $\boldsymbol{E}(\boldsymbol{x}, t)$ in classical electromagnetism and $g^{\mu\nu}(x^\gamma)$ in general relativity. Other examples include instrument settings and measurement results, which are associated with definite regions rather than points in spacetime. These parameters correspond to what Bell (1976) called "local beables" (pronounced "be-ables"). Unless otherwise noted, our use of the term parameters will be restricted to spacetime-based parameters, including the sets I and Q.

Of course, some models employ additional mathematical entities that are not spacetime based. For example, for $N > 1$, the N-particle configuration-space wave function in QM consists of values that do not correspond to particular locations in spacetime. For the purposes of Bell's analysis and the discussion here, non-spacetime-based parameters such as configuration-space wave functions are simply omitted from Q, even if they are mathematically utilized in a given model. It is also possible to construct nonlocalized parameters out of spacetime-based parameters, such as the total energy of an extended system, but such values are not to be included as elements of I or Q.

Deterministic models are those for which specification of all inputs I, including boundary conditions and external forces (if present) always exactly determines the noninput parameters Q. Stochastic models do not predict unique values for Q, but for any full set of inputs the model assigns a probability for every possible combination of noninput parameters. Thus, a fully specified mathematical model can always be written as $P_I(Q)$, a unique joint probability distribution function for the set of noninput parameters, given specific values for the inputs.[3] For deterministic models, these distributions are δ functions, but the analysis is not limited to such cases. This definition, which suffices for the present purposes, is minimal in the sense that it does not include algorithmic details (regarding how one parameter is deduced from another within the model) or the physical interpretation of a model.

According to the standard rules for probabilities, the full joint probability distribution $P_I(Q)$ of all noninput parameters of a model can be used to generate marginal distributions $P_I(Q_1)$ for any subset $Q_1 \subset Q$. It also generates conditional probabilities $P_I(Q_1|q_2)$, where q_2 are specific values of parameters in another subset Q_2. In some cases, a model may predict that Q_1 and Q_2 are statistically independent,

[2]We use the term "parameter" instead of "variable," as the latter sometimes implies a time-dependent quantity, while inputs and outputs are generally localized in time, as well as space.

[3]In many cases, "distribution functional" rather than "function" should be used here, as Q itself typically includes fields, functions of spacetime. Similarly, when Q is continuous, $P_I(Q)$ denotes probability densities rather than probabilities.

meaning that $P_I(Q_1, Q_2) = P_I(Q_1)P_I(Q_2)$. When statistical independence holds, knowledge of the values of parameters in Q_2 does not inform the marginal $P_I(Q_1)$, as represented by the condition $P_I(Q_1|Q_2) = P_I(Q_1)$.

Two models that use identical input and output sets I and O (when applied to a given system) and that also have the same marginal output probabilities $P_I(O)$ are said to be in input-output (IO) agreement: they always yield the same joint probability of the output parameters for a given set of inputs. Note that IO-agreement does not preclude different predictions at the level of $P_I(Q)$. Two models can even be in IO-agreement if they utilize different parameters U in Q. For example, in classical electromagnetism, one can change the gauge condition on parameters corresponding to the electromagnetic potentials without changing observable model predictions. The discussion to follow treats any such parameter-changing reformulations as different models because they might generally have different properties at the level of nonobservable parameters.[4] (As we shall see, just changing the associated spacetime location of a parameter can significantly change the model.)

In the following, we will focus on models that are in IO-agreement with QM, at least for a specific setup under consideration. Such models are guaranteed to share the empirical success of QM but are strictly constrained by Bell's theorem.

B. Physical assumptions

The following properties may or may not hold for any specific mathematical model, allowing for a categorization of models into classes and subclasses. To maintain an appropriate scope, we define here only key properties that play significant roles in the discussion to follow, with a few more given in Sec. III.B. For example, the need to formally define relativistic covariance of models does not arise here, although the light cones of Minkowski spacetime do play a role.

1. Continuous action (CA)

Instead of beginning with Bell's approach to defining locality, we first define the weaker condition CA that encodes the spirit of no action at a distance without requiring any light-cone structure from relativity, or even a distinction between past and future. Consider spacetime regions **1** and **2**, with **1** completely surrounded by a screening region **S**, as shown in Fig. 1(a). These are not merely spatial regions; **S** spans the past and future of **1** as well as its spatial extent. We will denote the set of all inputs in regions **1** and **2** by I_1 and I_2, respectively. If there are further inputs besides I_1 and I_2, their values are assumed to be fixed in the definitions that follow. The noninput parameters in each region are denoted by the corresponding Q_1, Q_2, and Q_S.

Loosely speaking, a mathematical model violates CA if it has unmediated action at a distance , i.e., if changes in **2** can be associated with changes in **1** without also being associated with changes within **S**. For example, a CA-respecting model of a light switch in **2** correlated with a lamp in **1** must include a description of the mediating parameters (e.g., the currents in the wires) in the intermediate screening region **S**. In such a model, knowledge of the values of all of the parameters in **S** makes additional information regarding **2** redundant for the purpose of predicting what happens in region **1**.

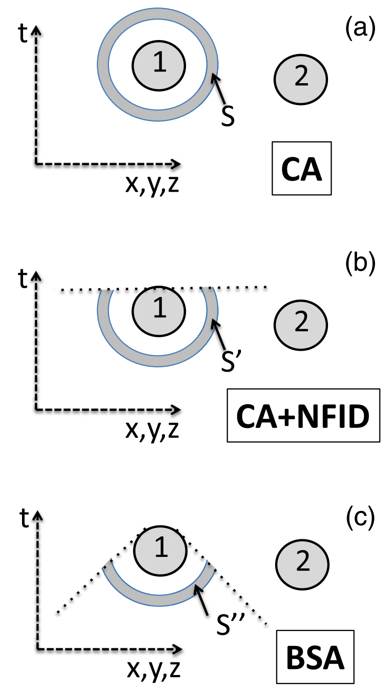

FIG. 1. The screening regions **S** used in different assumptions of locality. Given all modeled parameters in the screening region, a screened model will assign the same probabilities to parameters in region **1**, regardless of additional knowledge of parameters in region **2**. (a) shows the most general case of *continuous action* (CA); (b) breaks time symmetry by adding the *no future-input dependence* (NFID) assumption, and (c) references the light cones according to Bell's screening assumption (BSA).

Mathematically, CA corresponds to the condition

$$P_{I_1,I_2}(Q_1|Q_2, Q_S) = P_{I_1}(Q_1|Q_S) \tag{1}$$

for all combinations of the parameters in the regions depicted in Fig. 1(a). Equation (1) says that $P_{I_1}(Q_1|Q_S)$ is both statistically independent of Q_2 and functionally independent of I_2. When this occurs, we say that **S** "screens" **1** from **2**.[5] For CA models, this equality is required to hold for all simply connected, nonoverlapping regions **1**, **2**, and **S**, for which **S** completely separates **1** from **2** and is nowhere vanishingly thin. As there is no essential difference between regions **1** and **2**, a model with CA also must have **S** screen **2** from **1**.

Readers familiar with probabilistic modeling will notice that the role of the screening region **S** in CA is analogous to that of a "Markov blanket," a term coined by Pearl (1988). We avoid using this terminology not only because of required

[4]In the electromagnetism example, the use of Coulomb-gauge potentials as parameters will generally not respect the *local causality* condition defined in the following because those potentials can change instantaneously over all space.

[5]The word "shields" is often used in the literature, including Bell (1990), instead of "screens."

minor adjustments (discretizing the model in spacetime onto a set of nodes; properly representing the role of inputs) but also, primarily, because many physicists might be misled: Markov's name would likely be immediately associated with Markov processes, which propagate step by step from the past to the future, subject to a particularly strong arrow of time. This is exactly the opposite of our purpose here: generalizing "no action at a distance" to situations in which time symmetry is not broken at all, or is broken in a much weaker manner. Restricting attention to Markov processes would be an additional assumption, limiting attention to directed acyclic graphs with the directions of all edges determined by temporal order, one that will now be formally defined.

2. No future-input dependence (NFID)

There is a well-known tension between the time-symmetric equations characteristic of fundamental physical theories and the time-asymmetric manner in which models are utilized. For example, if one takes wave function "collapse" to be physically meaningful, this process defines a preferred direction of time, breaking the time symmetry evident in unitary evolution. More generally, a preferred direction of time is commonly chosen by limiting attention to models in which all events up to a time t' can be evaluated without regard to events in the future of t':

> NFID holds for a mathematical model $P_I(Q)$ if, for any time t' included in the relevant spacetime region, there exists a restricted model $P'_{I'}(Q')$, where I' is the set of all inputs belonging to times up to t' and Q' is the set of all noninput parameters up to t', such that
>
> $$P_I(Q') = P'_{I'}(Q') \tag{2}$$
>
> for all possible values of the parameters in I and Q'.

In other words, NFID means the marginal $P_I(Q')$ is functionally independent of future inputs.

When combined with CA, the assumption of NFID implies that there is no need to consider any parts of the screening region **S** that lie in the future of both regions **1** and **2**. As shown in Fig. 1(b), if CA holds for P and P' of a model respecting NFID, then the smaller region $\mathbf{S}'$ also screens **1** from **2**, $P_{I_1,I_2}(Q_1|Q_2,Q_{S'}) = P_{I_1}(Q_1|Q_{S'})$.

3. Bell's screening assumption (BSA)

If one accepts both of the previous assumptions (CA and NFID) and furthermore is interested in modeling only screening regions that remain applicable in all reference frames, it becomes appropriate to ignore any portion of **S** that is spacelike separated from both[6] regions **1** and **2**. This leads to the smaller region $\mathbf{S}''$ shown in Fig. 1(c). Bell (1990) proposed that this smaller region $\mathbf{S}''$ should screen region **1** from region **2**:

$$P_{I_1,I_2}(Q_1|Q_2,Q_{S''}) = P_{I_1}(Q_1|Q_{S''}). \tag{3}$$

It is important to note that this screening condition does not imply that parameters in **1** are independent of parameter values in **2**, merely that the latter values are redundant given the specification of all model parameters in $\mathbf{S}''$. We will call Eq. (3) Bell's screening assumption (BSA).

4. Local causality

Models that conform to both BSA and NFID are unable to describe certain quantum phenomena, as Bell's theorem establishes, and will be proved in Sec. II.D. We define this important combination of assumptions as *local causality*.[7] This definition may cause some initial confusion, because *local causality* is often identified with Eq. (3) in the literature, which is formally just BSA. However, in essentially all cases in which this is done, NFID is presupposed either explicitly or implicitly, and the addition of this assumption turns BSA into *local causality*. Bell (1990) himself introduced BSA after clearly assuming the past-to-future causal structure associated with NFID (see Fig. 6.3 there) and used the term *local causality* to convey this combination, often using the shorter "locality" as a synonym.

One should be cautioned about interpreting the phrase *local causality* as being the simple conjunction of locality and causality. There are many different meanings that could be ascribed to both of these words; we have already seen three different notions of locality in Fig. 1. All that is needed in the present analysis is for *local causality* to be understood as meaning the well-defined assumptions NFID and BSA.

An important condition that follows from NFID (or from *local causality*), but not BSA alone, can be derived by applying it to the S'' region from Fig. 1(c). Requiring the probabilities of parameters to be independent of future inputs and choosing S'' to lie entirely in the past of all of regions **1** and **2** (in some reference frame where NFID holds), one obtains the functional independence relation

$$P_{I_1,I_2}(Q_{S''}) = P(Q_{S''}). \tag{4}$$

A variant of this condition will play an important role in the following proof of Bell's theorem.

C. Historical interlude

At this point, it is appropriate to emphasize how natural it is to assume that all of the previous conditions, summarized by *local causality*, should hold in any detailed model describing

[6]Note that it is not sufficient to restrict the screening region to lie in the past light cone of region **1**; it must completely screen **1** from the overlap of the past light cones of **1** and **2**; see, e.g., note 7 in Bell (1986).

[7]The freedom in choosing the region **S** in the definition of CA is reflected in the definitions used for *local causality* (or "Einstein locality" or "local realism"). Figure 1 resembles Fig. 6.4 of Bell (1990), the "screening region" was effectively the entire past of **1** and **2** in Bell (1981), and Bell (1976) used the overlap of their past light cones. This affects the identification of λ in the separability condition of Eq. (7), but the subsequent derivation is unchanged.

real physics. It is convenient to do so by referencing Einstein and Bohr.

At the 1927 Solvay conference, Einstein noted that there were two possible "conceptions" of the single-particle quantum wave function $\psi(x,t)$ (Bacciagaluppi and Valentini, 2009). If viewed as a set of spacetime-based parameters $Q(x,t)$, the requirement that only a single particle is eventually measured implies some form of wave-function collapse that, for Einstein, "implies to my mind a contradiction with the postulate of relativity." Instead, he advocated a conception where "one does not describe the process solely by the Schrödinger wave," effectively pointing out the possibility that additional hidden parameters could indicate the particle's actual location.

Einstein, Podolsky, and Rosen (1935) (EPR) extended this analysis to a two-particle system (of a type to be analyzed in the following) and reached the logical conclusion that violations of *local causality* could be avoided only by adding new hidden parameters. If known, these new parameters would allow one to determine the outcomes in more detail than is possible within QM. EPR concluded that QM gave an incomplete description.

EPR did not use the formal mathematical language of Bell's analysis. Instead, they implied the existence of spacetime-based parameters Q that encoded "an element of physical reality," and deduced that hidden Q's must be present in a complete theory because in some cases it was possible to "predict with certainty ... the value of a physical quantity," such as position or momentum, "without in any way disturbing a system."

Bohr (1935) responded quickly to EPR, defending the completeness of QM on the basis of the notion of complementarity he had developed earlier (in connection with the quantum uncertainty principle). He advocated "a radical revision of our attitude towards the problem of physical reality," and argued that the phrase "without in any way disturbing a system" used by EPR "contains an ambiguity."

Bohr considered in detail a situation in which the properties of a particle can be discerned by first allowing it to pass through a slit in a diaphragm, and later making a "free choice" of measuring either the momentum or the position of the diaphragm. (It is remarkable, especially in the context of this work, that the guarantee for "without in any way disturbing" was spatial separation for EPR, but temporal order for Bohr.) "Of course there is … no question of a mechanical disturbance … during the last critical stage of the measuring procedure," he wrote. But as one can only measure either the position or the momentum of the diaphragm, "even at this stage" there still might be "an influence on … the possible types of predictions regarding the future behavior of the system." Bohr thus advocated[8] accepting some violations of *local causality* that are present in the formalism of QM, while at the same time excluding other violations: those corresponding to a "mechanical disturbance." (Similarly, his notion of "completeness" clearly differs from that of EPR.)

[8]We believe that it is appropriate to interpret Bohr in this manner but acknowledge that it is probably impossible to uncontroversially translate his writing into the formal language introduced later.

Most physicists simply adopted Bohr's complementarity, in either its original form or a variant thereof (Bell, 1992), and continued to develop and apply QM to a variety of physical systems (Mermin, 1985; Mann and Crease, 1988). But Einstein was not convinced. Summarizing the situation in 1948, he wrote the following (Born, 1971):

> [T]hose physicists who regard the descriptive methods of quantum mechanics as definitive in principle would … drop the requirement … for the independent existence of the physical reality present in different parts of space …. [W]hen I consider the physical phenomena known to me, and especially those which are being so successfully encompassed by quantum mechanics, I still cannot find any fact anywhere which would make it appear likely that [that] requirement will have to be abandoned. I am therefore inclined to believe that the description of quantum mechanics … has to be regarded as an incomplete and indirect description of reality, to be replaced at some later date by a more complete and direct one.

Here Einstein is essentially advocating for models to be built from spacetime-based parameters Q, while offering the opinion that other physicists had prematurely abandoned this possibility. But there was indeed a "fact" that he was not aware of, a theorem that would be proved by Bell in 1964 (after both Einstein and Bohr had passed away). We now turn to Bell's theorem, and the fact that all models in IO-agreement with QM must violate the package of assumptions that is *local causality*. Subsequently, we shall address the question of whether the hidden-parameter models advocated by Einstein should still be pursued, even given the necessary violation of *local causality*.

D. Statement and proof

Bell's theorem demonstrates the following:

> No model conforming with *local causality* can be in IO-agreement with QM.

We emphasize that the disagreement is with not only the predictions of QM but also the results of empirical observations: experiments that have been performed. The following proof is based on the CHSH inequality (Clauser *et al.*, 1969), which concerns a particular application of QM to the experimental scenario shown in Fig. 2. Specifically, a source emits a pair of particles, and these are later analyzed and detected in spacelike-separated regions **1** and **2**.

Mathematical models describing such situations will have an input parameter c specifying the particular settings and arrangement of the common source of the two particles. Additional input parameters a and b specify the settings and arrangement of the detectors in regions **1** and **2**, respectively. The results of the experiment are the output parameters A in region **1** and B in region **2**. The set of all of the model's spacetime-based parameters in region Λ is denoted by λ. The parameters a, b, and c are inputs, and A and B are outputs, just

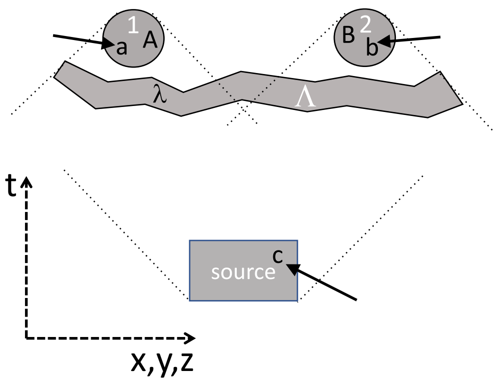

FIG. 2. The essential geometry of a Bell-type experiment. The parameters a, b, and c are inputs; the arrows indicate that their values come from outside the model. The parameters A and B are observable outputs. λ is the set of all localized model parameters in the region Λ, which screens regions **1** and **2** from the overlap of their backward light cones.

as in QM. The set λ can be quite general, with possibilities ranging from complex combinations of functions and operators to the simplest possibility: the empty set, for the case with no spacetime-based parameters associated with the region Λ.

The proof of the CHSH inequality, following Bell (1976, 1981, 1990) and Peres (1978), proceeds in Secs. II.D.1 and II.D.2 by assuming only *local causality* without making any reference to QM. Section II.D.3 then proves Bell's theorem by comparing this Bell inequality with quantum theory and experiments. (A disadvantage of Bell's original 1964 proof is discussed in Sec. VI.D.)

1. Bell's separability condition

Any mathematical model capable of producing predictions for the setup of Fig. 2 will provide a joint probability distribution $P_{a,b,c}(A,B,\lambda)$. The marginal distribution $P_{a,b,c}(A,B)$ can be compared with experiment and with QM. Models will generically also have other parameters, located between the designated regions, but these are not necessary for the main argument. Also not included in λ are nonspacetime-based entities, such as multiparticle wave functions, which may be utilized in some models.

From the assumption of *local causality*, specifically from BSA, Eq. (3), it follows that

$$P_{a,b,c}(A|\lambda,B) = P_{a,c}(A|\lambda) \tag{5}$$

because $\mathbf{\Lambda}$ screens **1** from **2**, in the sense that the necessary S'' region can be chosen to be fully contained in $\mathbf{\Lambda}$. Similarly, $P_{a,b,c}(B|\lambda,A) = P_{b,c}(B|\lambda)$. It also follows from NFID that any model-generated probabilities of λ must be independent of the settings a and b because those settings lie in the future of λ. In equation form, following Eq. (4), this reads

$$P_{a,b,c}(\lambda) = P_c(\lambda). \tag{6}$$

This is often known as "measurement independence," a term that unfortunately obscures the input nature of the measurement settings. It is clearer to refer to it as the $\boldsymbol{\lambda}$-independence condition, as the equation specifies that λ is independent of the a and b inputs via a direct application of NFID.[9]

Basic probability theory provides a product expression for the joint conditional probability: $P_{a,b,c}(A,B|\lambda) = P_{a,b,c}(A|B,\lambda)P_{a,b,c}(B|\lambda)$. Since λ is hidden, the observable joint probability is found by summing or integrating this over all possible values of λ. Applying BSA and NFID by substituting in Eqs. (5) and (6) yields Bell's "separability condition":

$$P_{a,b,c}(A,B) = \int d\lambda P_c(\lambda)P_{a,c}(A|\lambda)P_{b,c}(B|\lambda), \tag{7}$$

where the integral is understood as a sum if λ is discrete, or as a functional integral if λ is a function. This must hold for every applicable model respecting *local causality*.

2. A Bell inequality

From Bell's separability condition, Eq. (7), one can derive the CHSH inequality (Clauser *et al.*, 1969), a generalized version[10] of Bell's original inequality (Bell, 1964). It applies to models for which the output parameters in regions **1** and **2**, i.e., the outcomes A and B, have two possible values.[11] Assigning ± 1 to the outcome values on each side, the product AB must then also be ± 1. Its expectation value for given inputs, i.e., the correlator of the outcomes, is denoted as

$$\langle AB\rangle_{a,b,c} \equiv \sum_{A,B} ABP_{a,b,c}(A,B). \tag{8}$$

The CHSH inequality restricts the values of a combination of correlators, which involves two of the possible settings of the input parameter a in region **1**, labeled a and a', and two possibilities for the input setting in region **2**, labeled b and b'. The source input setting c is held constant while the four possible combinations of inputs are manipulated, and it will be suppressed from here on; we will later consider only particular Bell states for which only one value of c is relevant. It is customary to transfer the primes to the A and B parameters so that $\langle A'B\rangle$ stands for $\langle AB\rangle_{a',b}$.

With this notation, the CHSH inequality concerns the combination $\langle AB\rangle + \langle A'B\rangle + \langle AB'\rangle - \langle A'B'\rangle$. It is easiest to evaluate this combination by sampling the probability distributions in Eq. (7) many times, in the style of a Monte Carlo simulation.[12] Denoting the nth value sampled from $P(\lambda)$ by λ_n,

[9]A different perspective results if one denies free-input-parameter status to the measurement settings, treating a and b as stochastic variables instead of inputs. This is the "superdeterministic" scenario discussed in Sec. VI.C, which allows a version of Eq. (6) to be considered a "no conspiracy," a "freedom of choice," or even a "free will" condition.

[10]See, e.g., Bell (1971) for details.

[11]The proof can be generalized to measurements with continuous results provided that their ranges are restricted, $|A|, |B| \leq 1$.

[12]The proof here follows Peres (1978). The discussion of the mathematical model rather than the modeled physical experiments avoids the need for any additional assumptions, such as "counterfactual definiteness" (the assumption that when A_n is measured, it is legitimate to discuss A'_n as well).

we have A_n sampled from $P_a(A|\lambda_n)$ and A'_n sampled from $P_{a'}(A|\lambda_n)$, and similarly with B_n, B'_n. The previous combination of correlators is then obtained by averaging over $(A_n + A'_n)B_n + (A_n - A'_n)B'_n$, and it follows from $A_n, A'_n = \pm 1$ that for each n one of the parentheses must vanish. The averaged combination therefore is of absolute magnitude 2 for each n, and the combination of correlators cannot be larger in magnitude:

$$|\langle AB\rangle + \langle A'B\rangle + \langle AB'\rangle - \langle A'B'\rangle| \leq 2. \tag{9}$$

This is the CHSH inequality.

3. Contradiction with QM and experiment

When the Bell inequalities were first derived, they were shown to be in conflict with the predictions of QM. Now they are known to be in direct conflict with actual experiments (Giustina *et al.*, 2015; Hensen *et al.*, 2015; Shalm *et al.*, 2015; Rosenfeld *et al.*, 2017), independent of the formalism of QM, demonstrating the failure of *local causality*.

It is simple to demonstrate that at least some QM predictions violate the CHSH inequality, Eq. (9). Consider two photons entangled in a spin-zero Bell state, as in some of the early experiments (Clauser and Shimony, 1978; Aspect, Grangier, and Roger, 1981). (Equivalently, two spin-1/2 particles can be analyzed.) Suppose that each photon encounters a polarizing beam splitter, with outputs directed onto two single-photon detectors. The two beam splitters are aligned at angles a and b in regions **1** and **2**, respectively (these are the measurement settings, defined modulo π). For the outcome parameters A and B, assign a value of $+1$ when the detectors imply a measured polarization aligned with the setting, and -1 for a measurement of the perpendicular polarization. The predictions of QM are then given by the probabilities

$$p_{a,b}(A, B) = \tfrac{1}{4}[1 + AB\cos(2a - 2b)]. \tag{10}$$

The expectation value of the product AB is therefore $\langle AB\rangle = \cos(2a - 2b)$.

For certain combinations of settings, this violates the CHSH inequality by a wide margin. The largest violation obtains for $a = 0$, $a' = \pi/4$, and $b = -b' = \pi/8$, for which the left-hand side of Eq. (9) is $2\sqrt{2}$ (each of the four terms contributes $+1/\sqrt{2}$).[13] These nonclassical correlations between the two photons served historically as an early and striking example of the much wider family of phenomena associated with quantum entanglement; see, e.g., Brunner *et al.* (2014) and Streltsov, Adesso, and Plenio (2017).

The observed violations of the inequalities are by impressive margins greatly exceeding the experimental accuracy. Indeed, as an empirical test of a mathematical model or a class of models, the confidence with which the CHSH inequality is rejected approaches the certainty of a mathematical proof. For example, the experimental results of Giustina *et al.* (2015) boast a value less than 3.7×10^{-31} for the probability that the results could be obtained under the assumption of *local causality*, according to the standard statistical analysis. Furthermore, this result belongs to the recent generation of "loophole-free" experiments (those cited previously), which are free from all of the simplifying assumptions that were necessary for Bell tests with earlier technology. Not only do the observations violate the CHSH inequality, the quantitative results follow the predictions of QM in fine detail. We now turn to models that can be consistent with these experiments, Bell's theorem notwithstanding.

[13] Cirel'son (1980) showed that this is the maximal value achievable in QM, and Popescu and Rohrlich (1994) devised a synthetic model that reaches even higher values, up to 4, the maximum possible.

III. IMPLICATIONS

The upshot of Bell's theorem is that there is no longer any hope of finding a reformulation of QM that respects *local causality*. But the use of spacetime-based parameters has not been ruled out altogether, and the motivations for using them remain intact. Furthermore, given such parameters, there are still live options for saving CA, the generalized form of locality defined in Sec. II.B.1. In this sense, Bell's theorem does not necessarily imply unmediated action at a distance.

The rest of this Colloquium is dedicated to an analysis of the possibility of reformulating QM in a locally mediated manner, consistent with both CA and Lorentz covariance. Recall that a "reformulation" here means a model in IO-agreement with QM, with the same inputs I, the same outputs O, and the same model-generated joint probabilities $P_I(O)$. In preparation for this, Sec. III.A proposes a categorization scheme for all models in IO-agreement with QM, and Sec. III.B clarifies relevant issues of causation and signaling.

A. Categories of Bell-compatible reformulations of QM

As stated, Bell's theorem dictates that no model in IO-agreement with QM can respect *local causality*, which is the conjunction of two assumptions: NFID and BSA. Reformulations of QM must thus violate at least one of these in some nontrivial manner such that the CHSH inequality can also be violated. Of these two assumptions, we argue that the primary one for categorization purposes should be NFID because it is often taken for granted, and because the motivation for BSA in Sec. II included NFID, as depicted in Fig. 1. A useful secondary categorization is the CA condition, indicating whether or not action at a distance is implied by a given model.

Bell's theorem thus requires all models in IO-agreement with QM to fall into one of the following categories:

- Type I: Respects NFID
  - Type IA: Respects CA (must violate BSA)
  - Type IB: Violates CA
- Type II: Violates NFID
  - Type IIA: Respects CA (may violate BSA)
  - Type IIB: Violates CA

Models that violate CA necessarily also violate BSA. For convenience, the different types of models are also identified in Table I.

From the definition of NFID, type I models allow for the calculation of all spacetime-based parameters in temporal order, using inputs that enter into the calculation in that same order. But because of the necessary

TABLE I. Categories of possible reformulations of QM (and the sections in which they are discussed). The columns identify whether or not a model conforms with the CA locality condition, and the rows refer to the NFID arrow-of-time condition. Bell's theorem rules out the subset of type IA models that also conform to the stricter BSA locality rule; see Fig. 1. In the following, we focus on locally mediated models, which are of type IIA.

	Continuous action	
No future-input dependence	CA	Not CA
NFID (type I)	Type IA (Sec. VI.A)	Type IB (Sec. VI.A)
FID (type II)	Type IIA (Secs. IV and V)	Type IIB (Sec. VI.B)

BSA violation, such models cannot adhere to the light-cone-constrained Cauchy problem typically found in classical physics.

Type IA models would have to avoid CA violation using faster-than-light mediators, bypassing the screening region $\mathbf{S}''$ of Fig. 1(c) but passing through the larger screening region $\mathbf{S}'$ of Fig. 1(b). Such models have not been formally developed but have been promoted by various authors, including Bell (1981) himself. For this not to violate NFID in a different reference frame, one might propose a special frame in which the model uniquely applies, at the expense of Lorentz covariance. Moreover, to maintain IO-agreement with QM in all cases, it is necessary that the mediating signal should always pass through $\mathbf{S}'$, even if this region is blocked, say, by a brick wall. For these reasons, we judge such models to be of less interest, and our use of the term "locally mediated" will exclude such faster-than-light unblockable mediators.

The type IB category includes the standard Schrödinger-picture QM itself, as well as less conventional approaches such as de Broglie–Bohm guiding waves (Bohm, 1952). Such models utilize mathematical intermediaries R to connect distant spacetime-based parameters. Recall that the model parameters I and Q are defined as being associated with particular places and times. Values in R might be associated with multiple spacetime locations in some nonseparable manner and do not generally have a form such as $R(\boldsymbol{x}, t)$.[14] The most prominent example of a parameter R is the many-body wave function, which for an entanglement setup is defined on configuration space. Such an account involves no parameters λ in the relevant spacetime regions, directly violating both BSA and CA; Eq. (7) becomes a simple product, with λ representing a constant, the empty set. The role of the wave function in producing the predictions of QM might be described as an abstract mathematical object connecting events in spacetime.

Type II models violate the NFID assumption, so they are not temporally sequential calculations. It is natural to call such models *future-input dependent*, (FID). With well-known examples such as the stationary action principle, it is clear that FID models cannot be trivially dismissed, yet they are rarely brought up in discussions of Bell's theorem.[15]

Type IIA models are particularly interesting because they do not involve action at a distance, in the sense that the screening condition of CA is respected. Furthermore, since NFID is already violated, Bell's theorem does not rule out the possibility of retaining even the stricter BSA locality condition. As we shall see, in a type IIA model, the mediation between spacelike regions can take place entirely on timelike worldlines. Instead of the unblockable faster-than-light mediators of type IA models, all relevant parameters in type IIA models can be associated with the actual particle histories, allowing Lorentz covariance to be preserved. This in turn means that it is possible to build models without any abstract mathematical structures R "mediating" events in conventional spacetime. Of course, one could still use such structures if desired, say, by retaining the conventional QM configuration-space wave function in a model. This would fall into the category of type IIB models that violate the intuitive NFID condition while restoring neither CA nor BSA.

Sections IV and V will be devoted, respectively, to a review of the specific achievements of the type IIA toy models that have already been developed, and to a discussion of the drawbacks and promise of this category of models. The other categories, as well as approaches that do not fall within the framework used here, will be discussed in Sec. VI. Before this review, some additional clarifications are necessary, which we turn to next.

B. Causality and locality

The failure of *local causality* implied by Bell's theorem leads naturally to the following question: In what sense, if at all, does *local causality* correspond to assumptions of locality and causality? Before continuing, it is necessary to clarify these issues.

1. Cause and effect

The definition of NFID in Sec. II.B uses the distinction between input and noninput parameters, rather than the words "cause" and "effect." Nevertheless, the NFID condition is closely related to a definition of causality that arises naturally within the modern account of "interventionist" causation, where causes are identified as interventions (Woodward, 2005; Pearl, 2009). If the input parameters in question are deemed to be controllable parameters, then it is appropriate to identify them as causes according to this account.

QM itself clearly adopts this connection between inputs and controllable parameters: the mathematical formalism of QM is a procedure for making operational predictions for observations, given the values of the controllable inputs. As our goal is

[14]The efforts of Norsen (2010) and Stoica (2019) aim to overcome this, which could lead to type IA models.

[15]One factor that surely contributed to this is that Bell himself did not mention the FID possibility in any of his publications (Bell, 2004); see Sec. III.B.1. This omission continued in some major reviews (Goldstein *et al.*, 2011; Shimony, 2017), although the latter was recently updated with a recognition of retrocausation (Myrvold, Genovese, and Shimony, 2019).

to discuss models in IO-agreement with QM, it is natural for us to adopt this approach. Such models limit the inputs I to the parameters that QM tells us can be externally controlled.

Given this connection between controllable inputs and causes, one can identify different possible causal structures. In models that respect NFID, noninput parameters are typically functionally dependent on past inputs, but they are always functionally independent of future inputs. This "forward-causal" structure is clearly what Bell had in mind when he used the terms "causality" and "causal structure," with the controllable inputs called "free variables" or "free elements" (Bell, 1977, 1990).[16]

FID models, on the other hand, do not have a forward-causal structure. In other words, they cannot generally compute a given parameter $q(t')$ (or its probability distribution) without specifying certain inputs in the future of t'. In the framework of interventionist causality, if those future inputs are controllable, the FID models are retrocausal.[17]

Some FID models, such as classical action principles, are not retrocausal. In those cases, the final boundary constraints are required mathematical inputs, but not controllable inputs, and thus are not considered causes. Analysis of the causal structure of such a theory requires inverting the functional relation between some of the inputs and some of the outputs so that a different model is obtained, a model in which all inputs are controllable. Although it makes sense to refer to the action principle itself as a reformulation of Newton's equations, it is only after this inversion that one obtains a model fully in IO-agreement with the standard operational description of classical mechanics, which uses the controllable initial conditions as inputs.

At the time of Bell's work, the interventionist approach to causation had not yet been well developed. An older approach was taken for granted, dictating that if two parameters exhibit cause-effect correlations, it is appropriate to refer to the earlier one as a cause and the later one as an effect, i.e., to rely on their time ordering, regardless of which one can be externally controlled.

This is one topic where one's definition of causation directly impacts the types of mathematical models that one views as acceptable. Applied to the λ-independence condition, any violation of Eq. (6) would be viewed as retrocausal in the framework of interventionist causation, an instance of FID. But if one instead assumes that λ is the cause of the settings a and b because λ occurs before a and b were chosen, one would have to conclude that the settings were effects and could not be treated as free inputs (see footnote 9 and Sec. VI.C). The model would then not be in IO-agreement with QM.

2. Signals

Just as QM restricts the inputs I to be controllable, it also specifies that the outputs O are observable. If I is controllable and O is observable, $P_I(O)$ summarizes all possible signals. And as QM does not allow signals to be sent back in time, it follows that for models in IO-agreement with QM the outputs O cannot depend on future inputs. We call this requirement signal causality or, explicitly,

$$P_I(O') = P'_{I'}(O'), \tag{11}$$

where the primed sets of parameters are all those associated with times up to t', as in the similar Eq. (2).

A comparison with Eq. (2) indicates that any violation of NFID in a model in IO-agreement with QM must be at the level of unobservable (hidden) parameters U in Q. Such a FID model would be retrocausal (at a hidden level) but would not violate signal causality.[18]

Motivated by special relativity, it is natural to formulate a stronger restriction on signaling. This condition, called signal locality, limits signals to traveling no faster than light so that signals associated with a particular controlled input are limited to outputs in its future light cone. For outputs O_1 localized in region **1**, the relevant inputs I'' should thus lie in the past light cone of **1**, and the signal locality requirement corresponds to the existence of a restricted model P'' such that

$$P_I(O_1) = P''_{I''}(O_1). \tag{12}$$

This condition also holds in QM and must thus be maintained for any model in IO-agreement with QM.

As indicated in the Introduction, these signal-based definitions of locality and causality are operational, in the sense that they involve only controllable inputs and observable outputs. Bell's theorem states that models in IO-agreement with QM must violate either a distinct notion of locality (BSA) or a distinct notion of causality (NFID), which are not defined operationally, as they refer to hidden model parameters, not signals. Because of these different definitions, models can be local or causal in one sense, but not in another.[19]

IV. LOCALLY MEDIATED MODELS OF ENTANGLEMENT (TYPE IIA)

In this section, we discuss reformulations of QM that fall under type IIA, meaning that they are locally mediated; see Sec. III.A. These models conform to CA and are FID, allowing for compatibility with Bell's theorem without a necessary conflict with Lorentz covariance. As noted, such models are underrepresented in the literature on Bell's theorem, so this section and Sec. V provide a rather thorough discussion.

The essential strategy behind type IIA models of entanglement is to allow a violation of the λ-independence condition, Eq. (6), such that $P_{a,b}(\lambda)$ is not independent of the input

[16]There is an early exception: Bell (1964) used the word causality to imply complete causality, i.e., determinism.

[17]The word retrocausal conventionally implies there are some future causes of some past parameters, not a purely reverse-causal structure.

[18]If one demands not only that causes be identified with controllable inputs but also that effects be identified with observable outputs, one is led to take Eq. (11) as representing the causal arrow of time. However, the term retrocausal in the literature does not signify violations of signal causality. We use the more technical term NFID, which explicitly focuses on inputs, to minimize confusion.

[19]The literature on Bell's theorem involves quite a few additional locality conditions [see, e.g., Wiseman (2014)], but these are not needed here.

settings a and b. The relevant λ lies in the past light cones of a and b, so that such models are technically retrocausal, as defined in Sec. III.B.1. But as noted there, if IO-agreement with QM is to be maintained, any correlations with future settings must be sequestered in hidden variables, not observable outputs. By restricting attention to models in IO-agreement with QM, there is no possibility of signals being sent back in time, and thus no concern of generating paradoxes. These and other concerns with such models will be further discussed in Sec. V.

The promise of type IIA models is that, in any given case, there exist parameters λ that can act as local mediators of the actual correlations. It is always simple to find a distribution of shared parameters λ that will produce a given correlation for particular measurement settings; Bell showed that the problem was getting the same $P(\lambda)$ distribution to consistently work for all measurement settings; see Eq. (7). But for models $P_{a,b}(\lambda)$ where the distributions can be different for different settings, Bell's consistency problem disappears. This means that it is possible to retain BSA in some FID models, or at least the weaker locality condition CA.

At the current stage of development of type IIA models, there are none that are applicable to a wide range of quantum phenomena. Existing models aim at reproducing merely the known correlations for the Bell state, Eq. (10). Several are presented here, with schematic models in Sec. IV.A, and a model providing a more detailed description in Sec. IV.B.

A. Schematic models

Although the idea of using future-input dependence to explain entanglement has been around for a long time (Costa de Beauregard, 1953, 1977, 1979; Cramer, 1980; Pegg, 1982; Sutherland, 1983; Price, 1984, 1997), explicit type IIA mathematical models of entanglement have appeared in the literature mostly in the last decade. One notable exception is Pegg (1982), a description that could be simplified[20] and expressed in a manner quite similar to that of Argaman (2010), the model presented next.

1. A simplistic model

Consider again the correlations between the polarizations of a pair of entangled photons. Using the terminology of Sec. II.D.3, where a and b represent the angle settings of polarizers, the spin-zero Bell-state correlations can be obtained from the following toy model. First, take the two photons to both be initially polarized at an angle λ, distributed according to

$$P_{a,b}(\lambda) = \frac{1}{4}\left[\delta(a-\lambda) + \delta\left(a + \frac{\pi}{2} - \lambda\right) + \delta(b-\lambda) + \delta\left(b + \frac{\pi}{2} - \lambda\right)\right]. \tag{13}$$

Here $\lambda \in [0, \pi)$ and the δ functions are modulo π. In this model, the initial polarization λ is thus somehow constrained by the future settings to be a, $a+\pi/2$, b, or $b+\pi/2$ at equal probability; i.e., they are to be aligned with one of the detectors.

Next, apply Malus's law to obtain the results of the single-photon measurements A and B; i.e., $P_a(A=1|\lambda) = \cos^2(a-\lambda)$, etc.. Combining these using Eq. (7) reproduces the QM probabilities for the spin-zero Bell state, Eq. (10). While this model is schematic, it demonstrates that only mediation along the spacetime paths of the particles is required.

2. The Hall model

A number of additional type IIA schematic models follow a similar strategy. They consist of two components: (i) a specification of the sample space of the hidden variables and their distributions $P_{a,b}(\lambda)$, and (ii) models for the measurement outcomes $p_{a,[\lambda]}(A)$ and $p_{b,[\lambda]}(B)$ such that the combination of (i) and (ii) per Eq. (7) is in IO-agreement with QM for a specific setup of interest. (The notation $[\lambda]$ emphasizes that while λ is an input to the second component, it is not an external input.)

For lack of space, we provide the details of just one additional example, that of Hall (2010). The version adapted to photon polarizations (Argaman, 2018) has

$$P_{a,b}(\lambda) = \frac{1}{\pi}\frac{1 + \grave{A}\grave{B}\cos(2a-2b)}{1 + \grave{A}\grave{B}(1-z)}, \tag{14}$$

where $\grave{A} = \mathrm{sgn}[\cos(2a-2\lambda)]$, $\grave{B} = \mathrm{sgn}[\cos(2b-2\lambda)]$, and $z = (2/\pi)|2a-2b|$ are abbreviations. This model is "deterministic" in the sense that A is fully determined by a and λ through $p_{a,[\lambda]}(A) = \delta_{A,\grave{A}}$, with the same expression relating B to b and λ (all the randomness is in λ itself). It reproduces the results of QM for the Bell state, Eq. (10).

Here knowledge of λ provides only a rough idea of what a and b are. When properly quantified, the information about a and b that can be gleaned from the past parameter λ amounts to less than 0.07 bit per entangled pair (Hall, 2016). In this sense, one may view the toy model of Eq. (14) as a dramatic improvement over that of Eq. (13).

3. Additional toy models

A large number of additional schematic entanglement models exist in the literature, the majority of which are type IB models. The original Bell (1964) work contained such a model for "illustration" purposes, and many others were developed over the years, relying on different "resources": communication, shared randomness, and/or nonlocal boxes; see, e.g., Degorre, Laplante, and Roland (2005) and the references therein. Each of the models described so far proposes a novel distribution that may be denoted by $P_{a,b}(\lambda)$, and a way in which that λ generates the output statistics of QM, per the steps (i) and (ii). Simply associating λ with the worldlines of the entangled particles, rather than with the time of the measurements, can then change the type IB model into a type IIA model. An example was given by Barrett

[20]One complication is that, in its original form, the intermediate state appears to be output dependent rather than dependent on the future input setting.

and Gisin (2011), who modified the model of Toner and Bacon (2003) and Degorre, Laplante, and Roland (2005) by "moving" λ to the past.

Other changes in the spacetime location of λ can affect the assessment of the NFID and CA conditions, leading to a new model of a different type, even if the distribution $P_{a,b}(\lambda)$ is unchanged. For type IIA models, taking λ to be associated with the emission event at the source but not with the particle worldlines, as arguably done in the machine-learning-generated models of Weinstein (2017, 2018), formally results in a type IIB model, with no local mediators. But such models are easily transformed back into type IIA, simply by copying λ onto mediators along both worldlines. Alternatively, λ might be associated with the time of the measurements, rather than the emission or the worldlines, resulting in a model of type IB. Reinterpreting Eq. (13) in this manner leads to precisely the model of Di Lorenzo (2012). Yet another example is provided by the work of Sen (2019), who began with the model of Brans (1988) (itself obtained by associating the parameters of standard QM with the past) and explicitly transformed it into a Bohmian-style FID model.

The previously mentioned schematic type IIA models show both promise and limitations. On the positive side, they all serve as proof-of-principle examples, indicating that Bell inequalities can be violated without also introducing action at a distance, and they provide a variety of points of departure for future development. With the mediating parameters λ associated with the particle worldlines, other advantages quickly become evident. For example, a recent application (Sen, 2020) of a FID model to entanglement in accelerating reference frames indicates a nearly trivial reconciliation of quantum phenomena and general relativity, for a case that is quite difficult even for quantum field theory.

On the negative side, however, these models all simply assert the connection between the settings and λ without a proposed mechanism or explanation. One natural justification for such a connection would be an appeal to time symmetry: one could argue that the symmetry exhibited by microscale phenomena implies an equal role for both past and future. This would make the future settings a and b just as important as the initial state preparation c when modeling λ. But this justification seems inapplicable because these schematic models do not possess time symmetry in any sense. We now turn to a Bell-compatible FID reformulation that restores microscopic time symmetry, and does so in a manner that provides an account of both the $P_{a,b}(\lambda)$ distribution and the outcome probabilities.

B. The Schulman Lévy-flight model

Conventional QM is typically viewed as time symmetric, but its intermediate calculations are notably time asymmetric. For example, consider the polarization of a photon that is known to have passed through two consecutive polarizers set at angles θ_1 and θ_2. The conventional description associates the angle θ_1 with the polarization of the photon between the two polarizers, but time-reversal symmetry implies that θ_2 should be just as relevant to the intermediate description. Any time-symmetric account of the intermediate photon should therefore take both angles into account and would be a type IIA model.

Such a time-symmetric model was developed by Schulman (1997, 2012), using a time-varying polarization angle $q(t)$.[21] Schulman considered the possibility that $q(t)$ could be perturbed by microscopic rotations dq ("kicks") so that $q(t)$ evolves from θ_1 to θ_2 (or $\theta_2 + n\pi$) between the polarizers, without requiring a collapse at the last instant. If the magnitude of each microscopic kick is normally distributed (or has a finite second moment), one obtains diffusive behavior, which is inappropriate. However, if $q(t)$ describes a Lévy flight, e.g., if the magnitudes of the kicks are distributed according to the Cauchy (Lorentzian) distribution, $\propto d\gamma/[(dq)^2 + (d\gamma)^2]$ with a small width $d\gamma$, the net rotation Δq has a similar probability distribution:

$$P(\Delta q) = \frac{1}{\pi}\frac{\gamma}{(\Delta q)^2 + \gamma^2}, \tag{15}$$

where γ is the sum of the $d\gamma$ widths of all kicks along the path.

With $q(t)$ constrained to θ_1 at the time of the initial polarizer t_i and to θ_2 at t_f, the final time, $q(t)$ provides an appealing time-symmetric description of the dynamics constrained by initial and final boundaries. Moreover, and this is the main point of Schulman's derivation, the model correctly predicts the outcome probabilities for a single photon in the limit $\gamma \to 0$ if the measurement acts as a boundary constraint corresponding to discrete possibilities, requiring the photon polarization to be either aligned or perpendicular to the polarizer angle (either θ_2 or $\theta_2 + \pi/2$). Adding all of the equivalent contributions corresponding to $\theta_2 + n\pi$ per Eq. (15) gives a result $\propto 1/\sin^2(\theta_1 - \theta_2)$. Comparing this to the other possible outcome, summing over $\theta_2 + (n + 1/2)\pi$, reproduces Malus's law upon normalization: the probability for a photon of initial polarization θ_1 to align with a polarizer oriented at θ_2 is $\cos^2(\theta_1 - \theta_2)$. A detailed derivation can be found in the Appendix.

Note that for small γ the path $q(t)$ is very close to being a constant, but the initial and final requirements enforce at least one significant kick, with a distribution $\propto d\gamma/(dq)^2$. In the $\gamma \to 0$ limit, paths with a single kick dominate. There is thus an event that corresponds to collapse in this description (unless $\theta_1 = \theta_2$ or $\theta_1 = \theta_2 + \pi/2$), but it happens at an arbitrary time between preparation and measurement, rather than at the time of the measurement, and thus respects time symmetry.

This model can be trivially extended to the case of two maximally entangled photons by combining two copies of the single-particle model $q_1(t)$ and $q_2(t)$, and constraining their unknown initial polarization angles to be identical $q_1(t_\mathrm{i}) = q_2(t_\mathrm{i})$ (Wharton, 2014; Almada *et al.*, 2016). Identifying this initial polarization as the hidden parameter λ reproduces precisely the probability distribution of the simplistic toy model of Eq. (13). This follows because the overwhelmingly

[21] Schulman's discussion of spin-1/2 particles is adapted here to photons.

most probable scenario is to have only one significant kick in the combination of the two paths, and this in turn requires λ to match one of the two future settings.

In this model, the screening region S'' from Fig. 1(c) contains the parameters $q_1(t)$. No inputs on the other arm of the experiment can affect the probability of the measured outcome $q_1(t_f)$ without also affecting the earlier values $q_1(t)$, conforming to BSA. (The earlier schematic models also respect BSA, for similar reasons.) The mechanism by which the correlations are enforced is NFID violating: the future settings (a, b) constrain the full histories $q_1(t)$ and $q_2(t)$, including the possible initial value of the hidden parameter λ. This explicitly violates NFID and Eq. (6), violating *local causality*. All locality conditions from Fig. 1 are thus preserved.

The Schulman type IIA model supplies a future boundary mechanism to explain the future-input dependence [an account missing from component (i) of the previous schematic models, as noted before], and the same mechanism provides the correct outcome probabilities [to explain component (ii)]. Indeed, this two-particle toy model is currently the most sophisticated example of how a model can yield the correct Bell-state correlations while retaining the BSA (or the CA) condition of locality. It demonstrates a spacetime-based mediation of the correlations involved in entanglement, via a mechanism that uses the entire history rather than instantaneous "states." By assigning probabilities to histories rather than states, this approach avoids the tension between entanglement and relativistic covariance. It demonstrates how type IIA models need not conflict with relativity (as noted in Sec. III.A) because all of the mediation is by parameters that reside on timelike or lightlike worldlines. If the relevant parameters λ reside on the classical worldlines of the entangled particles, this essentially looks similar in every reference frame, no matter which particle is measured first.

It is striking that the same set of rules is applicable to both one-photon and two-photon setups, and is also valid if additional measurements are considered. For a single photon, it provides the appropriate Malus-law probabilities for any number of sequential polarization measurements.[22] For the two-photon entanglement setup, as the hidden parameter λ is associated with the photon polarization, it is natural to ask whether an additional measurement of this polarization along the path of the photons could shed light on the mechanisms involved. QM itself describes how this would fail: after such a measurement, the two photons are no longer entangled. The Schulman model successfully describes this: the additional measurement would be associated with another boundary constraint, changing the entire history of the experiment, and requiring two "significant kicks" instead of one, reproducing again the often-perplexing results of standard QM.

The two-particle Schulman model can also be trivially generalized from the spin-zero state to any maximally entangled two-qubit state by performing polarization rotations on one of the two photons.[23] Further generalizations to scenarios with several particles (Bennett *et al.*, 1993; Pan *et al.*, 1998) might no longer respect BSA if some of the correlating parameters are localized on connected zigzagging worldlines (an entanglement-swapping setup), but they would continue to respect CA and would still be type IIA. The challenge of extending this type of model to partially entangled states remains an open problem.

V. DISCUSSION

In this section, we first address existing criticism of the type IIA approach and then discuss its potential and directions for future exploration.

A. Objections to type II models

Despite the availability of the simple models presented, much of the contemporary discussion of Bell's theorem fails to recognize such a possibility. For example, in a recent round of loophole-free experiments (Giustina *et al.*, 2015; Hensen *et al.*, 2015; Shalm *et al.*, 2015; Rosenfeld *et al.*, 2017), not one article mentioned the possibility of type II or FID models. In the rare case where experimental papers mention a retrocausal option, it is typically relegated to a footnote (Handsteiner *et al.*, 2017; Rauch *et al.*, 2018).

With this lack of attention, there are few published concerns about type II models in the recent literature, although a number of "intuitive" objections are likely to occur to most physicists upon first encountering these models. The most common such concerns will be addressed first, followed by a discussion of specific formal arguments that have appeared in the literature.

1. Intuitive objections

One common objection to FID models is that they violate some unwritten principle of causality. Formalizing this objection is difficult, but one evident concern is that such models might lead to logical difficulties with time-travel paradoxes. But time-travel paradoxes require communication with the past, with at least some level of observable signal, and this is forbidden in models in IO-agreement with QM, as they conform to signal causality, Eq. (11). For any FID model in IO-agreement with QM, the future-input dependence is always at the level of the hidden parameters λ, and as there is no protocol for observing the values of these parameters (without changing the whole setup), such models do not allow retrosignaling.

Another common concern is that FID models imply that future inputs must "exist" to constrain hidden parameters in

[22]It thus provides a "natural" mechanism or explanation for violations of Leggett-Garg inequalities (Leggett and Garg, 1985); these "beyond-Bell" inequalities facilitate experimental demonstrations of additional surprising quantum phenomena. Again, the type IIA approach describes a relationship between microscopic hidden variables and macroscopic observable results that appears to be quite perplexing from a "macroscopic realism" NFID-assuming point of view.

[23]The strategy of reducing a two-particle entanglement problem to two single-particle problems can be extended to all maximally entangled bipartite states (Wharton, Miller, and Price, 2011).

the past, and some find this block-universe view problematic (Sorkin, 2007; Kastner, 2017). But it appears to be ill advised to avoid developing a theory for such reasons: it would have been a pity, for example, if Newton were to avoid developing the law of universal gravitation because he perceived its nonlocality to be unacceptable. Furthermore, treating future events as valid model parameters and analyzing entire spacetime regions all at once is common in physics, e.g., in general relativity and with Wick rotations. And in any case, one can always wait until the whole relevant spacetime region is in the past and perform the model analysis retrospectively. We set aside this objection as an essentially anthropocentric restriction on mathematical models (Wharton, 2015).

As a related objection, some might take the view that because QM conforms to signal causality, as do all other established physical theories, there should never be any reason to consider FID reformulations of QM. However, as we have seen, the failure of *local causality* provides just such a reason. Bell's theorem does not formally tell us whether it is the locality (BSA) or causality (NFID) aspects of our models that require adjustment, so we should seriously consider both options rather than simply choosing the one we take to be more plausible.

And again, such a restriction is routinely ignored by physicists in practice. Histories approaches such as that of Griffiths (2001), and path integrals in general, encourage one to consider the past and future together as a single structure, violating the spirit of NFID. In the Heisenberg-picture QM, measurement operators are often evolved back in time to the previous measurement. And some analyses of "delayed-choice" experiments, such as that of Bohr (1935), briefly described in Sec. II.C, allow one to make incompatible inferences about past events for different future measurement choices. If those past events are parametrized, this also violates NFID.

2. Formal objections

An early technical argument against FID models is due to Maudlin (1994). Adapting it to the previously mentioned Bell-state setup, consider the case where one measurement is performed early enough that the result A can be sent ahead of the other particle (say, via a laser signal) to the other measurement device. This output parameter A could then be used to determine the other setting b via some algorithm $b = f(A)$. The challenge is one of self-consistency: if one uses a model that requires b as an input to generate λ and then uses λ to generate the outcomes A and B, the function $f(A)$ might be found to disagree with the value of b utilized in the calculation. This is of particular concern for the schematic models designed with one experiment in mind (such as those in Sec. IV.A) because this is essentially a different experimental configuration.

But it is unreasonable to expect precisely the same model, with the same inputs and outputs, to apply to this new configuration. In this version of Maudlin's challenge, the setting parameter b is no longer an input to the model (it cannot be freely set), so an analysis of this new experiment would require a type IIA model of the form $P_a(Q)$, rather than the original $P_{a,b}(Q)$. The Schulman model of Sec. IV.B is general enough to handle this new configuration because the boundary constraints imposed by the future measurements are still enforced in the global solution, no matter whether the settings are free inputs or calculated parameters. As long as the solution is calculated all at once, assigning probabilities to entire histories rather than states, every intermediate solution is self-consistent by definition (Berkovitz, 2008; Lewis, 2013; Wharton, 2014).[24]

A more recent objection that applies even to all-at-once accounts appeared in Wood and Spekkens (2015), although, notably, this stands as an objection to all accounts of entanglement phenomena, not specifically type IIA models. The essential point is that causal channels are typically accompanied by signal channels, absent some special "fine-tuning" of the underlying model. Such fine-tuning would require additional explanation. In any causal account of entanglement, such as the faster-than-light option of type IA models, signal locality (the inability to send a spacelike signal) must be the result of some perfect cancellation in the marginal probabilities. This is said to be fine-tuned because even a slight deviation would lead to spacelike signaling. For example, in quantum field theory, it is the perfect commutativity of spacelike-separated operators that guarantees the necessary fine-tuning.

The situation might appear to lead to an additional challenge to type II models, which have causal channels into the past, because another fine-tuning argument can be applied to signal causality (the inability to send signals into the past). But a more careful analysis reveals that the fine-tuning objection is not significantly worse for type II models than it is for type I models, because spacelike signaling violates signal causality in some reference frame. Further analysis of the Schulman model has revealed that the appearance of signal locality follows from a basic symmetry (Almada *et al.*, 2016), providing just the sort of explanation (from symmetry) that is most often used to explain fine-tunings in high-energy physics. A more comprehensive explanation of both signal locality and signal causality was recently proposed by Adlam (2018). Finding mathematical or physical principles underlying these signaling restrictions will be an important challenge for future reformulations of QM.

There is also a flip side to the Wood-Spekkens fine-tuning argument. If an underlying physics model indeed breaks time symmetry according to the NFID condition, it would take a finely balanced restriction to make microscopic physics look as time symmetric as it does. Leifer and Pusey (2017) weighed this argument against the Wood-Spekkens fine-tuning argument and proposed that the time-symmetry argument is stronger.

[24]In general, the IO-agreement-with-QM status of the original $P_{a,b}(Q)$ guarantees through signal causality that its operational version $P_{a,b}(O)$ can be restricted to times up to the first measurement, yielding $P'_a(A)$; subsequently, the full applicable model can be reconstructed:

$$P_a(Q) = \sum_A P'_a(A) P_{a,f(A)}(Q|A).$$

B. Potential of type IIA models

The examples in Sec. IV demonstrate that a number of type IIA models can successfully account for the Bell-state correlations. Thus, Bell's theorem cannot be said to stand in the way of a locally mediated reformulation of QM. In particular, the Schulman model achieves a description in IO-agreement with QM that conforms to CA, employing only spacetime-based parameters with local, time-symmetric interconnections, which pose no difficulties for Lorentz covariance.

A further advantage of such models relates to the exponential complexity of quantum states. By using only spacetime-based parameters Q, the model $P_I(Q)$ has an evident physical interpretation: it specifies the probability of each possible set of events in spacetime Q, while only one particular configuration actually occurs. This is analogous to the Liouville equation in classical mechanics, where the statistical distributions can be exponentially complex, but only one phase-space configuration is taken to represent an actual physical system (even when we do not know which it is). The complexity of this actual configuration scales linearly with the number of particles or the size of the modeled spacetime region.[25] The Schulman model provides a simple example of such linear scaling in that the parameters required for a two-particle experiment are merely two copies of the single-particle case.

The exponential growth of the conventional wave function $\psi(t)$ with particle number might lead one to think that achieving such linear scaling would be impossible, especially if one views the information contained in $\psi(t)$ as some physical entity that has to be translated into parameters $Q(t)$, the subset of Q pertaining to a time t. But note that $\psi(t)$ contains information about all possible measurement outcomes that might occur, for all possible future measurement settings. In a FID model, $Q(t)$ can be a function of those future settings and therefore needs only inform the outcomes for the actual future measurement, vastly reducing the required number of parameters; for further analysis, see Wharton (2014).

Beyond Bell-state correlations, there are plenty of other quantum phenomena that must be addressed to approach a full reformulation of QM. Single-particle interference appears challenging, but it may be resolved in a type IIA model by adopting a field-based rather than a particle-based viewpoint (Wharton, 2018).[26] Recent type IIA models have tackled other issues, including position measurements of entangled particles (Sen, 2019) and formal relativistic covariance (Wharton, 2010; Heaney, 2013; Sutherland, 2017). Presumably, more models will be developed in the near future, addressing additional issues such as three-particle and partial entanglement phenomena.

There are many avenues that could be pursued in searches for such models. Existing reformulations, such as stochastic mechanics (Nelson, 1966, 2012) and stochastic quantization (Damgaard and Hüffel, 1987), could perhaps provide starting points. There are also some recent efforts that first evaluate the probabilities for the outcomes $P_I(O)$ (using one of the standard methods of QM), then define additional mediating parameters so that, overall, the resulting model is type IIA [both Sutherland (2017) and Drummond (2019) can be read in this manner]. While such approaches may claim applicability to a wide range of quantum phenomena, in our view, additional development is necessary for these models to fulfill their promise, such that the mediating parameters explain the outputs rather than the other way around.

There are quite a number of additional results in the literature that should guide the development of type IIA locally mediated models. Many of these have been developed in the context of locality. For a review, see Brunner *et al.* (2014); a recent example is Carmi and Cohen (2019). A potentially important result that explicitly questions the arrow of time has recently been proven by Shrapnel and Costa (2018). By dropping the usual NFID assumption, their analysis indicates that such models must be "contextual," meaning that distinct hidden-parameter accounts would be required for situations not distinguished by standard QM. While it is not unreasonable to expect the details of intermediate hidden parameters to depend on the detailed intermediate context, this still raises the question of why standard QM cannot distinguish these differences. This might indicate the development of models with inherent hidden symmetries, where this contextuality could seem more natural.

Eventually, type IIA models must also provide a satisfactory treatment of quantum measurements, but at the present stage of development this goal is not yet in clear sight. Still, the type IIA Schulman model improves upon standard Schrödinger-picture-with-collapse QM in two ways. First, measurements do not correspond to any sudden collapse, so they look more like an ordinary interaction (the collapselike event occurs somewhere between preparation and measurement). Second, there is no confusion about whether the size of the relevant configuration space should expand (as in a QM interaction) or be reduced (as in a QM measurement) because nothing lives in configuration space; all parameters are associated with spacetime.

A future type IIA theory should provide an explanation for why the interaction between some large systems (measurement devices) and some smaller systems (such as the measured particles) can be described effectively by imposing boundary constraints on the smaller systems. It is worth noting that such behavior is evident near large conductors in electromagnetism and thermal reservoirs in classical thermodynamics. It is also well known that smaller systems exhibit an evident time symmetry in a way that larger, thermodynamic systems do not. Understanding this is particularly important if time symmetry is used as justification for introducing FID because this symmetry must somehow give way to the asymmetry of signal causality at larger, observable scales.

Taking the Schulman model as an illustrative example, the only time asymmetry enters via a subtle distinction between photon preparations and photon measurements. Both of these have controllable settings, but preparations have an additional point of control: the initial polarization is also treated as an

[25] Time plays a different role in the context of the Liouville equation, as within classical dynamics the configuration at one time determines the whole path.

[26] Particle-like phenomena could arise from the discreteness of measurement interactions (the detector "clicks"), enforced by boundary constraints, not by discreteness of the parameters.

input. (For the case of entanglement, the initial correlation between two polarizations is an input.) In contrast, the measurement does not allow this same level of control; one can input the final polarizer angle, but not the measured polarization (the latter is an output, not an input). This empirically based distinction between full control at preparations and mere setting control at measurements provides the symmetry-breaking mechanism that leads to the appearance of signal causality in the model. Everything else about the model respects time symmetry, most notably, the intermediate account between preparation and measurement.

It is possible to attribute this distinction between preparation and measurement to the involvement of macroscopic "agents" who have control of some quantities but not others (Price, 1997). Alternatively, one may attempt to include a description of the measurement process itself in the mathematical model. Because of the observation that quantum measurements must have irreversibly recorded results [see, e.g., Miller and Wheeler (1996)], one cannot expect a completely time-symmetric model to achieve this. Future research into this issue may look in detail at the effects of a thermal environment, which could be included in a type IIA description. In both classical and quantum cases, such treatments break time symmetry by fixing the initial states of the environment (averaging them over a known thermal distribution), while leaving its final states to be computed; see, e.g., Feynman and Vernon (1963). For an appropriate interaction between the system's degrees of freedom and the environment, the information regarding the values of some of the parameters pertaining to the "measured" system Q_M are effectively amplified and copied many times in the final state of the environment; see, e.g., Zurek (2018). An intriguing possibility, called lenient causality in Argaman (2018), is that the time symmetry breaking in models of this type could impose signal causality for parameters such as Q_M, without leading to NFID for the microscopic parameters.[27]

VI. ALTERNATIVES AND MISCONCEPTIONS

While in Secs. IV and V we discussed type IIA (locally mediated) models in detail, there are many other models in the literature that can reproduce the experimentally observed CHSH violations, including of course the existing formulations of QM. We briefly discuss each of the general possibilities in this section, giving references to some of the approaches not reviewed here. These either violate *local causality* in some other way, as categorized in Sec. III.A, or fall outside our framework, i.e., they are not in IO-agreement with QM.

Note that a specific approach can lead to a variety of models, and that a model must be fully specified to allow for a clear categorization. (For example, as we see in Sec. IV.A for the schematic models, a change in the spacetime location associated with λ is sufficient to change the type of the model.) We devote Secs. VI.A–VI.C to each of the possibilities, and Sec. VI.D to some misconceptions that might lead one to mistakenly believe that there are additional categories of models.

A. Type I models

Type I models have no parameters that are dependent upon future inputs. In Sec. III.A, such models are categorized as type IA ones that would have faster-than-light mediators, and type IB ones in which distant regions can directly influence each other via nonspacetime-based mathematical intermediaries, such as the configuration-space wave function of conventional Schrödinger-picture QM. The many-body wave function also enforces distant correlations in other type IB approaches, including Bohmian mechanics (Bohm, 1952) and also spontaneous-collapse models (Ghirardi, Rimini, and Weber, 1986), which achieve full IO-agreement with QM only in an appropriate limit. Development of such models continues, for example, with so-called flash models, which have parameters in spacetime (the flashes) but no intermediate screening parameters (Tumulka, 2006).

As noted in Sec. III, no representative type IA model has been formally developed [Norsen (2010) might be the closest]. Spekkens (2015) noted that one can convert standard QM into a corresponding type IA model by introducing "local copies" of the wave function $|\psi(t)\rangle$ at every point in space with time coordinate t; collapse due to a distant measurement would then instantaneously update all of these new spacetime-based parameters. Information is thus transferred from one region to another at an infinite speed, bypassing the S'' region of Fig. 1(c) while passing through the upper boundary of the region S' in Fig. 1(b).

Whichever type I technique one uses to enforce correlations across spacelike separations, such a connection makes it difficult to achieve Lorentz covariance, even when signal locality is satisfied. In such models, when entanglement correlations between regions **1** and **2** are described, some observers see **1** affecting **2**, while other observers see **2** affecting **1**. These descriptions do not properly transform into each other under Lorentz transformations,[28] motivating the possibility of omitting them altogether, resulting in a purely operational model, with just $P_I(O)$. Despite these difficulties, it is clear that type I models are overwhelmingly represented in the relevant discussions in the literature.

B. Type IIB models

While the previous sections focused on type IIA models with spacetime-based mediators, other future-input-dependent models can include non-spacetime-based entities, directly linking distant regions. Such type IIB models often use configuration-space wave functions, in addition to their spacetime-based parameters. Many of the prior concerns about type I models (failure of Lorentz covariance, nonlocal

[27] Better still, it could lead to a condition such as information causality (Pawłowski *et al.*, 2009), which is known to essentially guarantee compliance with the Tsirelson bound (Cirel'son, 1980).

[28] The requirement here is not only that individual parameters transform covariantly, but that the overall description of which events affect which is consistent among different frames; see, e.g., Gisin (2010).

influences, etc.) are therefore applicable to type IIB models as well.

One popular type IIB model is the two-state-vector formalism introduced by Aharonov and Vaidman (1991), which essentially doubles the state space of conventional QM. For single-particle cases, it adds to the ordinary wave function $\psi(\boldsymbol{x}, t)$ another wave function $\phi(\boldsymbol{x}, t)$, a solution of the Schrödinger equation that is determined by the setting and the outcome of the next strong measurement on the particle (essentially a future boundary constraint).[29] While these are naturally interpreted as spacetime-based parameters, for entanglement scenarios the relevant state vectors are conventional configuration-space wave functions $\psi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$ and $\phi(\boldsymbol{x}_1, \boldsymbol{x}_2, t)$, and these entangled two-particle wave functions cannot be easily mapped onto spacetime-based fields. These wave functions are not spacetime based but are at least time-based parameters, and in this generalized sense, they exhibit a violation of the essential ideas behind NFID. Having departed from spacetime, they no longer have any localized screening parameters and thus violate the CA locality condition; see also Vaidman (2013). It is therefore fair to categorize such a model as type IIB.

C. Models outside the framework

Various approaches in the literature raise more exotic possibilities, essentially claiming to not fall under any of the four model types listed in Sec. III.A. These approaches depart from our framework (Sec. II.A), either by violating the rules of probability theory or by dropping aspects of the requirement of IO-agreement with QM. The latter models risk losing the empirical content of QM, i.e., the comparison of $P_I(O)$ to experiment. To still claim some form of agreement with QM, the $P_I(O)$ predictions must be recovered, at least at an effective level. At that effective level, such models always fall within one of the Sec. III.A model types.

One example is the many worlds interpretation (Everett, 1957), sometimes claimed to be a way to avoid Bell's theorem because all possible measurement outcomes are represented in a never-collapsed wave function. In this approach, the measurement problem is avoided by removing the Born rule from the fundamental description, but the empirical success of QM, the $P_I(O)$, is then removed as well (Maudlin, 2010). Proponents of the many worlds interpretation would argue that, at an effective level, a version of the Born rule is still applicable, but the result is then a type IB effective model, in the same category as conventional QM.

The deviation from our model framework that appeared most frequently in the recent literature [perhaps because it was discussed repeatedly by Bell (1977, 1981, 1990)] is superdeterminism, which retains the implicit NFID assumption while considering violations of the $\boldsymbol{\lambda}$-independence assumption, Eq. (6). This cannot be done within our framework (Sec. II.A), which treats the measurement settings (a, b) as input parameters, corresponding to the mathematical concept of free variables. But if the settings are treated as statistical parameters, the $\boldsymbol{\lambda}$-independence condition becomes

$$P_c(\lambda|a, b) = P_c(\lambda), \tag{16}$$

where c encodes the free preparation setting, still treated as an input. This is a statistical-independence relation, and it permits a Bayesian inversion to an equation sometimes known as the no conspiracies assumption:

$$P_c(a, b|\lambda) = P_c(a, b). \tag{17}$$

Violations of this condition can then be pursued by expanding λ (or using additional variables) to include the systems that choose the measurement settings.

This approach has been seriously considered in the literature [see, e.g., 't Hooft (2016)], despite the fact that it is coherent only if it makes sense to talk about the probabilities of the settings a and b. But such probabilities cannot be defined without creating a conflict with standard QM, where a and b are free inputs.[30] Indeed, in the explicit superdeterministic toy models that have been proposed for the Bell-state correlations, the relevant hidden variables [λ_0 in Brans (1988) and μ in Hall (2016)] are simply copies of the measurement setting parameters a and b transferred to earlier times. The other elements of these models prescribe a specific form of $P_{a,b}(\lambda)$ and a role for λ in generating the outputs, as discussed in Sec. IV.A.2. In practice, therefore, explicit superdeterministic models that agree with QM are forced to treat the future settings a and b as free inputs. Once this is acknowledged, the model again falls within the framework, and its type can be identified.

There are additional well-established methods that can be more spacetime oriented, but that do not meet the probability rules of our framework. For example, path-integral accounts of QM utilize spacetime-localized paths. It might be tempting to think that each path might be represented by a set of parameters Q, but the path integral cannot be parsed into normalized probabilities $P_I(Q)$, where only one path Q can be taken to exist.[31] Similarly, quantum field theory can be viewed as assigning a complex amplitude to all possible field configurations in spacetime, but none of these configurations can be assigned a probability. A further example is given by the consistent histories approach [see, e.g., Griffiths (2011)], where the probability rules for the intermediate description $P_I(Q)$ are changed, while those for the outputs $P_I(O)$ are not. These approaches represent directions that are, in a sense, more radical than the search for type IIA models.

[29] Another similar example is the transactional interpretation (Cramer, 1980, 2016), where the individual "confirmation" waves correspond to ϕ.

[30] Note also that the original suggestion by Shimony, Horne, and Clauser (1976) aimed only to emphasize the importance of the free-variable assumption and argued that scientific exploration necessarily involves the assumption that "hidden conspiracies of this sort do not occur." The reply of Bell (1977) observed that, even if the settings were chosen by a mechanical pseudorandom generator that could be included in an enlarged model, they would still be "effectively free for the purpose at hand."

[31] Introducing a FID viewpoint, along with a different parsing of Q, might potentially resolve this problem (Wharton, 2016).

D. Misconceptions

It has often been claimed that Bell's theorem is based on additional assumptions not identified in Sec. II, including determinism and realism.[32] These erroneous claims have already been well addressed in the literature (Norsen, 2007, 2011, 2017; Maudlin, 2010, 2014), but some clarifications will be repeated here to alert the reader to some of the many controversies in the literature.

Bell did not originally present his proof as outlined in Sec. II.D; this unified approach came only later. The EPR paper (see Sec. II.C) had already demonstrated that certain perfect correlations between distant measurements violate *local causality*, unless one adds deterministic hidden parameters. Bell (1964) built upon this result and showed that even with deterministic hidden parameters *local causality* could not be saved, as other predictions of QM could not be obtained.

Unfortunately, the argumentation of EPR contained several additional elements that made it appear paradoxical even before Bell's work, and the notion that Bohr (1935) had refuted it was widespread; see, e.g., Clauser *et al.* (1969). As Bell (1964) did not go through the EPR part of the argument in any detail (Wiseman, 2014; Norsen, 2015), many have concluded that the implications could be avoided by not postulating hidden parameters in the first place, or by not requiring them to be "deterministic" (or "realistic" or "counterfactual definite," etc.). But such moves do not save *local causality*, for the reasons given in the EPR paper. Bell himself later wrote, "It is remarkably difficult to get this point across, that determinism is not a *presupposition* of the analysis" (emphasis in original) (Bell, 1981).[33] It is hoped that the explicit discussion of the framework and assumptions in this work will help alleviate such difficulties.

VII. CONCLUSIONS

We began this Colloquium by carefully framing the assumptions that lead to Bell's theorem in terms of input parameters I and noninput parameters Q, both associated with locations in space and time. By defining a model in terms of the probabilities $P_I(Q)$ that it generates, Bell's theorem indicates that any such model that is in agreement with QM must violate one of the original assumptions, one of the components of *local causality*. This allows a natural categorization of all possible reformulations of QM, as described in Sec. III.A.

To the extent that we require the parameters in our mathematical models to correspond to physical events, this *local causality* violation is quite significant. Einstein described the physical justification for *local causality* in 1948 as follows (Born, 1971):

> If one asks what, irrespective of quantum mechanics, is characteristic of the world of ideas of physics, one is first of all struck by the following: the concepts of physics relate to a real outside world….
> It is further characteristic of these physical objects that they are thought of as arranged in a space-time continuum. An essential aspect of this arrangement of things in physics is that they lay claim, at a certain time, to an existence independent of one another, provided these objects "are situated in different parts of space."
> The following idea characterizes the relative independence of objects far apart in space (A and B): external influence on A has no direct influence on B ….

But Bell showed that this line of thinking leads to limitations on distant correlations that are in direct conflict with QM. The outcomes of spatially separated experiments are correlated in a manner that cannot be explained only in terms of common past inputs. Still, it does not follow that our only option is to throw out the entirety of Einstein's analysis, giving up on "physical objects … arranged in a space-time continuum." At least one of the assumptions that make up *local causality* needs to go, but spacetime-associated parameters might still be retained. Indeed, if they are not retained to some extent, all concepts of locality lose their usual meaning.

A particular concept of locality, *continuous action*, was defined here in a time-neutral manner that prevents unmediated action at a distance. Even given Bell's theorem, this definition of locality can be retained in two different styles of quantum models, categorized as types IA and IIA.[34] The former would require faster-than-light mediating parameters, so only the latter is compatible with Lorentz covariance. The price for retaining Lorentz covariance while forbidding action at a distance is the violation of an assumption arguably unrelated to locality: the premise that a model's parameters should not functionally depend on inputs associated with the future of those parameters, *no future-input dependence*. Without this assumption (or its corollary, the λ-independence condition), Bell's theorem cannot be derived.

This analysis therefore motivates type IIA models with future-input dependence and *continuous action* as the most "local" models compatible with QM. In rough terms, these models may be described as violating *local causality* by violating our intuition of causality rather than our intuition of locality.[35] Einstein saw no reason to relax either one of these,

[32]Once the discussion is cast purely in terms of mathematical models as done here, assumptions of "realism" can play no role [see Norsen (2007) for a discussion in a wider context]. Note that when realism is taken to imply that systems have properties prior to measurements, the NFID assumption is again being taken for granted, assuming not only that the systems have "objective" properties, but also that these properties are independent of the settings of future measurements.

[33]The original derivation of the CHSH inequality (Clauser *et al.*, 1969) simply assumed deterministic hidden parameters without using the EPR argument. It was rapidly understood that the same inequality also holds for indeterministic local hidden-variable models [see footnote 10 of Bell (1971) and Clauser and Horne (1974), or the unified type of proof found in Sec. II.D.2], but this is often ignored.

[34]The different types are identified in Table I in Sec. III.A.

[35]See Sec. III.B for clarification of these issues.

and Bohr effectively relaxed both, taking an operational view that keeps only the signal causality and signal locality conditions. Bell and his followers took the causality condition for granted, without realizing that an alternative exists,[36] and as a result studied type IA and type IB models. Others took an operational approach that drops both the causality and the locality requirements for the internal (hidden) variables, resulting in the development of type IIB models.

As analyzed in Sec. IV, type IIA models of quantum entanglement have effective connections associated only with the particle worldlines, either within the light cones or on the light cones for photons, i.e., there are no direct spacelike connections. Dropping the *no future-input dependence* assumption allows these to be effective two-way connections. Using this strategy, Einstein's "independence of objects far apart in space" can be softened without requiring connections that violate the spirit of relativity. In particular, this view accommodates entanglement scenarios by allowing an external influence on A to have an indirect influence on B, via mediating events in the intersection of their past light cones, without raising any difficulties with Lorentz covariance. As discussed in Sec. V.A, this need not lead to logical inconsistencies or deviations from conventional QM predictions.

Physics models with explicit future-input dependence have already been developed in the context of classical electrodynamics (Wheeler and Feynman, 1945, 1949), and their relevance to Bell-like scenarios was pointed out even before Bell's theorem emerged (Costa de Beauregard, 1953), and then repeatedly since; see, e.g., Pegg (1982) and Price (1997). Despite this, the development of explicit type IIA models of entanglement phenomena has only recently begun in earnest and is currently limited to a few applications, most notably the Bell-state correlations that typically serve to demonstrate the issue of Bell's theorem. The detailed discussion of the proof-of-principle examples of such models in Sec. IV is intended to introduce these possibilities to a wider audience, and Sec. V.B indicates several possible avenues for future development. These would include describing more complicated entanglement scenarios and developing a treatment of quantum measurements as interactions between small and large systems.

We emphasize that while future-input-dependent (or retrocausal) models of QM can have an underlying structure that is as time symmetric as classical physics, all such models must have a mechanism to recover the time-asymmetric condition of signal causality. Two possibilities for such a mechanism were suggested here. The first emphasizes the role of time-asymmetric agents employing the theory: they select which parameters of a theory to use as inputs of a specific model and which to use as outputs.[37] The second considers the possibility of a time-symmetry-breaking physical principle (perhaps due to the low entropy of the big bang), with possibly relatively minor effects on the mathematical model, e.g., a specification of initial conditions. As a result of this mild symmetry breaking, irreversibility could appear in the thermodynamic limit, and with it signal causality.[38]

A successful type IIA reformulation of QM would employ only spacetime-based parameters and would associate conventional probabilities with each fully specified configuration. An appropriate interpretation would take only one of these possibilities to actually occur in nature. In other words, the number of parameters describing a system would grow only linearly with its extent. This stands as a dramatic advantage over existing approaches, where the number of necessary parameters scales exponentially with the number of particles in the system. Combined with Lorentz covariance, this could greatly alleviate the disconnect between quantum theory and general relativity.

Such a reformulation would also shed light on an unresolved issue in quantum foundations: how to interpret the conventional wave function ψ and the collapse postulate. Although ψ is not included in the underlying model, it could still represent available knowledge about the actual parameters, a viewpoint that has become known as "ψ epistemic" (Spekkens, 2007). Such states of incomplete knowledge naturally reside in configuration space (as in classical statistical mechanics), as they have to represent a large number of possible correlations. Unitary evolution of these states would then correspond to time evolving the available information, in a manner analogous with Liouville dynamics. Learning additional information about future settings and future outcomes would then lead to a Bayesian updating of ψ corresponding to a nonphysical collapse. This is essentially the style of model advocated for by Einstein, where the actual state of the system was not ψ but rather something more fundamental (Harrigan and Spekkens, 2010).

While this work focuses on Bell's theorem, additional lines of research are also converging on the promise of future-input-dependent models. As already mentioned, Leifer and Pusey (2017) motivate such models via time symmetry. Another argument is motivated by the much-discussed Pusey-Barrett-Rudolph theorem (Pusey, Barrett, and Rudolph, 2012), recently reviewed by Leifer (2014), whereas yet another relies on arguments concerning the complexity achievable with quantum computation (Argaman, 2020). One of Leifer's conclusions matches ours exactly, promoting the development of "retrocausal … models that posit a deeper reality underlying quantum theory that does not include the quantum state." The spacetime-associated parameters Q in future-input-dependent models would mathematically represent this "deeper reality." Fully realizing this goal remains an open challenge.

[36] When prompted to consider the failure of the λ-independence condition, Eq. (6), which they called measurement independence, they always considered the conspiratorial superdeterministic option discussed in Sec. VI.C.

[37] See Price (1997) for further discussion.

[38] It is interesting to note that Bell (1990) already asked: "Could it be that causal structure emerges only in something like a 'thermodynamic' approximation?" But his tentative answer was negative, possibly due to his taking the causality condition of NFID for granted.

ACKNOWLEDGMENTS

The authors thank A. Briggs, P. Drummond, J. Finkelstein, S. Friederich, R. Sutherland, and several anonymous referees for the useful discussions and comments on the manuscript. This work is supported in part by the Fetzer Franklin Fund (Fetzer Pioneers Fund—P#2017-5) of the John E. Fetzer Memorial Trust.

APPENDIX: DERIVATION OF THE SCHULMAN MODEL

Schulman's original single-particle model applies to a single spin-$1/2$ particle; here we convert it to a photon polarization problem. The photon's classical trajectory is known, and it has a real hidden polarization direction $q(t)$ everywhere on its trajectory. The photon is prepared and measured by passing through two polarization cubes, with the first set at an angle θ_1 and the second set at θ_2. The initial polarization is constrained $q(t_1) = \theta_1$, as is usual for initial boundary conditions. Schulman enforced a similar final boundary condition at measurement, where the final polarization was constrained to be either $q(t_2) = \theta_2$ or $q(t_2) = \theta_2 + \pi/2$.

This final constraint is controllable (modulo $\pi/2$) and the model is FID. The time asymmetry (modulo $\pi/2$ at the output, but modulo π at the input) is external: an experimenter can choose to block a photon with an unwanted input polarization but does not know the output polarization until it is too late to interfere. Otherwise, everything in this model is fully time symmetric.

Such two-time-boundary problems can be solved only all at once, with probabilities assigned to entire histories $q(t)$, not instantaneous states. (One can extract the latter probabilities from the former.) Defining a net rotation

$$\Delta q \equiv \int_{t_1}^{t_2} \frac{dq(t)}{dt} dt \tag{A1}$$

(which is permitted to be larger than 2π for multiple rotations), the convolution of Schulman's proposed Cauchy kicks implies the following probability assignment of Eq. (15):

$$P(\Delta q) \propto \frac{1}{(\Delta q)^2 + \gamma^2}. \tag{A2}$$

This distribution recovers Malus's law as $\gamma \to 0$. Seeing this requires adding the probabilities for all rotations that end at the same polarization angle (modulo π) and normalization.

The evaluation requires summing over all possibilities of getting from θ_1 to $\theta_2 (\mathrm{mod}\, \pi)$, allowing for rotations through angles larger than π in both directions. The sum

$$\sum_{n=-\infty}^{\infty} \frac{1}{(\Delta\theta + n\pi)^2}, \tag{A3}$$

with $\Delta\theta = \theta_1 - \theta_2$, can be calculated as in Euler's solution of the Basel problem (finding the sum of N^{-2} for all integer N) by equating two different families of polynomial approximations to the same function, in this case,

$$f(x) = \sin(\Delta\theta + x)\sin(\Delta\theta - x). \tag{A4}$$

$(1/2)[\cos(2x) - \cos(2\Delta\theta)]$, the coefficient of x^2 is -1, yielding $f(x) = \sin^2(\Delta\theta) - x^2 + O(x^4)$. The other polynomial approximation scheme is obtained by multiplying the value of the function at $x = 0$ by a factor of $1 - x/z_k$ for each of the zeros z_k (the roots) of the original function (a specific approximation is obtained by including roots up to a certain absolute magnitude only). Treating the roots in pairs $z_n = -z_n' = \Delta\theta + n\pi$ gives

$$f(x) = \sin^2(\Delta\theta) \prod_{n=-\infty}^{\infty} \left(1 - \frac{x^2}{(\Delta\theta + n\pi)^2}\right), \tag{A5}$$

and expanding only up to terms quadratic in x gives the necessary sum:

$$\sum_{n=-\infty}^{\infty} \frac{1}{(\Delta\theta + n\pi)^2} = \frac{1}{\sin^2(\Delta\theta)}. \tag{A6}$$

Normalizing the probabilities for either $q(t_2) = \theta_2$ or $q(t_2) = \theta_2 + \pi/2$ is achieved by simply multiplying by the product of the corresponding denominators on the right-hand side of Eq. (A6), yielding Malus's law, $p = \cos^2(\Delta\theta)$, as required in Sec. IV.B.

Schulman also used this idea to prove the Born rule, in the sense of showing that probabilities $\propto |\psi|^x$ are compatible with this idea of multiple kicks only for $x = 2$, whether or not the Cauchy-Lorentz distribution is used for each kick.

REFERENCES

Adlam, Emily Christine, 2018, "Quantum mechanics and global determinism," Quanta **7**, 40–53.

Aharonov, Yakir, and Lev Vaidman, 1991, "Complete description of a quantum system at a given time," J. Phys. A **24**, 2315.

Almada, D., K. Ch'ng, S. Kintner, B. Morrison, and K. B. Wharton, 2016, "Are retrocausal accounts of entanglement unnaturally fine-tuned?," Int. J. Quantum Found. **2**, 1–14, https://www.ijqf.org/archives/3218.

Argaman, Nathan, 2010, "Bell's theorem and the causal arrow of time," Am. J. Phys. **78**, 1007–1013.

Argaman, Nathan, 2018, "A lenient causal arrow of time?," Entropy **20**, 294.

Argaman, Nathan, 2020, "Quantum computation and arrows of time," arXiv:2001.10517.

Aspect, Alain, Philippe Grangier, and Gérard Roger, 1981, "Experimental Tests of Realistic Local Theories via Bell's Theorem," Phys. Rev. Lett. **47**, 460–463.

Bacciagaluppi, Guido, and Antony Valentini, 2009, *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference* (Cambridge University Press, Cambridge, England).

Barrett, Jonathan, and Nicolas Gisin, 2011, "How Much Measurement Independence Is Needed to Demonstrate Nonlocality?," Phys. Rev. Lett. **106**, 100406.

Bell, J. S., 1964, "On the Einstein-Podolsky-Rosen paradox," Physics **1**, 195–200.

Bell, J. S., 1971, "Introduction to the hidden variable question," in *Proceedings of the International School of Physics 'Enrico Fermi,'* *Course IL*, edited by, B. d'Espagnat (Academic, New York), pp. 171–181.

Bell, J. S., 1976, "The theory of local beables," Epistemological Lett. **9**, 11–24, https://curate.nd.edu/show/c247dr29w4x.

Bell, J. S., 1977, "Free variables and local causality," Epistemological Lett. **15**, 79–84, https://curate.nd.edu/show/d791sf2976t.

Bell, J. S., 1981, "Bertlmann's socks and the nature of reality," J. Phys. (Paris), Colloq. **42**, C2-41.

Bell, J. S., 1986, "EPR correlations and EPW distributions," Ann. N.Y. Acad. Sci. **480**, 263–266.

Bell, J. S., 1990, "La nouvelle cuisine," in *Between Science and Technology*, edited by A. Sarlemijn and P. Kroes (Elsevier, New York), pp. 97–115.

Bell, J. S., 1992, "Six possible worlds of quantum mechanics," Found. Phys. **22**, 1201–1215.

Bell, J. S., 2004, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy*, 2nd ed. (Cambridge University Press, Cambridge, England).

Bennett, Charles H., Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters, 1993, "Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels," Phys. Rev. Lett. **70**, 1895–1899.

Berkovitz, Joseph, 2008, "On predictions in retro-causal interpretations of quantum mechanics," Stud. Hist. Philos. Sci. B **39**, 709–735.

Bohm, David, 1952, "A suggested interpretation of the quantum theory in terms of 'hidden' variables. I," Phys. Rev. **85**, 166–180.

Bohr, N., 1935, "Can quantum-mechanical description of physical reality be considered complete?," Phys. Rev. **48**, 696–702.

Born, Max, 1971, "*The Born-Einstein Letters*," (Walker & Co., New York).

Brans, Carl H., 1988, "Bell's theorem does not eliminate fully causal hidden variables," Int. J. Theor. Phys. **27**, 219–226.

Brunner, Nicolas, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, 2014, "Bell nonlocality," Rev. Mod. Phys. **86**, 419–478.

Carmi, Avishy, and Eliahu Cohen, 2019, "Relativistic independence bounds nonlocality," Sci. Adv. **5**, eaav8370.

Caves, Carlton M., Christopher A. Fuchs, and Rüdiger Schack, 2002, "Quantum probabilities as Bayesian probabilities," Phys. Rev. A **65**, 022305.

Cirel'son, B. S., 1980, "Quantum generalizations of Bell's inequality," Lett. Math. Phys. **4**, 93–100.

Clauser, J. F., and A. Shimony, 1978, "Bell's theorem. Experimental tests and implications," Rep. Prog. Phys. **41**, 1881–1927.

Clauser, John F., and Michael A. Horne, 1974, "Experimental consequences of objective local theories," Phys. Rev. D **10**, 526.

Clauser, John F., Michael A. Horne, Abner Shimony, and Richard A. Holt, 1969, "Proposed Experiment to Test Local Hidden-Variable Theories," Phys. Rev. Lett. **23**, 880.

Costa de Beauregard, O., 1953, "A response to the argument directed by Einstein, Poldosky and Rosen against the Bohrian interpretation of quantum phenomena," C.R. Acad. Sci. **236**, 1632–1634.

Costa de Beauregard, O., 1977, "Time symmetry and the Einstein paradox," Nuovo Cimento Soc. Ital. Fis. **42B**, 41–64.

Costa de Beauregard, O., 1979, "Time symmetry and the Einstein paradox. II," Nuovo Cimento Soc. Ital. Fis. **51B**, 267–279.

Cramer, John G., 1980, "Generalized absorber theory and the Einstein-Podolsky-Rosen paradox," Phys. Rev. D **22**, 362.

Cramer, John G., 2016, "*The Quantum Handshake*," (Springer, New York).

Damgaard, Poul H., and Helmuth Hüffel, 1987, "Stochastic quantization," Phys. Rep. **152**, 227–398.

Degorre, Julien, Sophie Laplante, and Jérémie Roland, 2005, "Simulating quantum correlations as a distributed sampling problem," Phys. Rev. A **72**, 062314.

Di Lorenzo, Antonio, 2012, "A simple model for the spin singlet: Mathematical equivalence of non-locality, slave will and conspiracy," J. Phys. A **45**, 265302; **45**, 279501(E) (2012).

Drummond, Peter D., 2019, "Time evolution with symmetric stochastic action," arXiv:1910.00001.

Einstein, A., B. Podolsky, and N. Rosen, 1935, "Can quantum-mechanical description of physical reality be considered complete?," Phys. Rev. **47**, 777–780.

Everett, Hugh, 1957, ""Relative state" formulation of quantum mechanics," Rev. Mod. Phys. **29**, 454–462.

Feynman, R. P., and F. L. Vernon, 1963, "The theory of a general quantum system interacting with a linear dissipative system," Ann. Phys. (N.Y.) **24**, 118–173.

Feynman, Richard P., 1965, *The Character of Physical Law* (MIT Press, Cambridge, MA).

Ghirardi, Gian Carlo, Alberto Rimini, and Tullio Weber, 1986, "Unified dynamics for microscopic and macroscopic systems," Phys. Rev. D **34**, 470.

Gisin, Nicolas, 2010, "The free will theorem, stochastic quantum dynamics and true becoming in relativistic quantum physics," arXiv:1002.1392.

Giustina, Marissa *et al.*, 2015, "Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons," Phys. Rev. Lett. **115**, 250401.

Goldstein, S., T. Norsen, D. Victor Tausk, and N. Zanghi, 2011, "Bell's theorem," Scholarpedia **6**, 8378.

Griffiths, Robert B., 2001, *Consistent Quantum Theory* (Cambridge University Press, Cambridge, England).

Griffiths, Robert B., 2011, "EPR, Bell, and quantum locality," Am. J. Phys. **79**, 954–965.

Hall, Michael J. W., 2010, "Local Deterministic Model of Singlet State Correlations Based on Relaxing Measurement Independence," Phys. Rev. Lett. **105**, 250404.

Hall, Michael J. W., 2016, "The significance of measurement independence for Bell inequalities and locality," in *At the Frontier of Spacetime: Scalar-Tensor Theory, Bell's Inequality, Machs Principle, Exotic Smoothness*, edited by Torsten Asselmeyer-Maluga, Fundamental Theories of Physics Vol. 183 (Springer International Publishing, Cham, Switzerland), pp. 189–204.

Handsteiner, Johannes, *et al.*, 2017, "Cosmic Bell Test: Measurement Settings from Milky Way Stars," Phys. Rev. Lett. **118**, 060401.

Harrigan, Nicholas, and Robert W. Spekkens, 2010, "Einstein, incompleteness, and the epistemic view of quantum states," Found. Phys. **40**, 125–157.

Heaney, Michael B., 2013, "A symmetrical interpretation of the Klein-Gordon equation," Found. Phys. **43**, 733–746.

Hensen, Bas *et al.*, 2015, "Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres," Nature (London) **526**, 682–686.

't Hooft, Gerard, 2016, *The Cellular Automaton Interpretation of Quantum Mechanics* (Springer, New York).

Kastner, Ruth E., 2017, "Is there really "retrocausation" in time-symmetric approaches to quantum mechanics?," AIP Conf. Proc. **1841**, 020002.

Leggett, A. J., and Anupam Garg, 1985, "Quantum Mechanics versus Macroscopic Realism: Is the Flux There When Nobody Looks?," Phys. Rev. Lett. **54**, 857–860.

Leifer, Matthew S., and Matthew F. Pusey, 2017, "Is a time symmetric interpretation of quantum theory possible without retrocausality?," Proc. R. Soc. A **473**, 20160607.

Leifer, Matthew S., and Robert W. Spekkens, 2013, "Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference," Phys. Rev. A **88**, 052130.

Leifer, Matthew Saul, 2014, "Is the quantum state real? An extended review of ψ-ontology theorems," Quanta **3**, 67–155.

Lewis, Peter J., 2013, "Retrocausal quantum mechanics: Maudlin's challenge revisited," Stud. Hist. Philos. Sci. B **44**, 442–449.

Mann, C., and R. Crease, 1988, "Interview: John Bell," Omni **10**, No. 5, 84–92.

Maudlin, Tim, 1994, *Quantum Non-Locality and Relativity*, 1st ed. (Blackwell Publishing, Oxford).

Maudlin, Tim, 2010, "What Bell proved: A reply to Blaylock," Am. J. Phys. **78**, 121–125.

Maudlin, Tim, 2014, "What Bell did," J. Phys. A **47**, 424010.

Mermin, N. David, 1985, "Is the moon there when nobody looks? Reality and the quantum theory," Phys. Today **38**, No. 4, 38.

Miller, Warner A., and John A. Wheeler, 1996, "Delayed-choice experiments and Bohr's elementary quantum phenomenon," in *Foundations of Quantum Mechanics in the Light of New Technology*, edited by S. Nakajima, Y. Murayama, and A. Tonomura (World Scientific, Singapore), pp. 72–84.

Myrvold, Wayne, Marco Genovese, and Abner Shimony, 2019, "Bell's theorem," in *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, spring 2019 ed. (The Metaphysics Research Lab, Stanford), https://plato.stanford.edu/archives/spr2019/entries/bell-theorem/.

Nelson, Edward, 1966, "Derivation of the Schrödinger equation from Newtonian mechanics," Phys. Rev. **150**, 1079–1085.

Nelson, Edward, 2012, "Review of stochastic mechanics," J. Phys. Conf. Ser. **361**, 012011.

Norsen, Travis, 2007, "Against 'realism,'" Found. Phys. **37**, 311–340.

Norsen, Travis, 2010, "The theory of (exclusively) local beables," Found. Phys. **40**, 1858–1884.

Norsen, Travis, 2011, "John S. Bell's concept of local causality," Am. J. Phys. **79**, 1261.

Norsen, Travis, 2015, "Are there really two different Bell's theorems?," Int. J. Quantum Found. **1**, 65, https://www.ijqf.org/archives/1646.

Norsen, Travis, 2017, *Foundations of Quantum Mechanics* (Springer, New York).

Pan, Jian-Wei, Dik Bouwmeester, Harald Weinfurter, and Anton Zeilinger, 1998, "Experimental Entanglement Swapping: Entangling Photons That Never Interacted," Phys. Rev. Lett. **80**, 3891–3894.

Pawłowski, Marcin, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski, 2009, "Information causality as a physical principle," Nature (London) **461**, 1101.

Pearl, Judea, 1988, *Probabilistic Reasoning in Intelligent Systems* (Morgan Kaufmann, Burlington, MA).

Pearl, Judea, 2009, *Causality* (Cambridge University Press, Cambridge, England).

Pegg, D. T., 1982, "Time-symmetric electrodynamics and the Kocher-Commins experiment," Eur. J. Phys. **3**, 44.

Peres, Asher, 1978, "Unperformed experiments have no results," Am. J. Phys. **46**, 745–747.

Popescu, Sandu, and Daniel Rohrlich, 1994, "Quantum nonlocality as an axiom," Found. Phys. **24**, 379–385.

Price, Huw, 1984, "The philosophy and physics of affecting the past," Synthese **61**, 299–323.

Price, Huw, 1997, *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time* (Oxford University Press, New York).

Pusey, Matthew F., Jonathan Barrett, and Terry Rudolph, 2012, "On the reality of the quantum state," Nat. Phys. **8**, 475.

Rauch, Dominik, *et al.*, 2018, "Cosmic Bell Test Using Random Measurement Settings from High-Redshift Quasars," Phys. Rev. Lett. **121**, 080403.

Rosenfeld, Wenjamin, Daniel Burchardt, Robert Garthoff, Kai Redeker, Norbert Ortegel, Markus Rau, and Harald Weinfurter, 2017, "Event-Ready Bell Test Using Entangled Atoms Simultaneously Closing Detection and Locality Loopholes," Phys. Rev. Lett. **119**, 010402.

Schulman, Lawrence S., 1997, *Time's arrows and quantum measurement* (Cambridge University Press, Cambridge, England).

Schulman, Lawrence S., 2012, "Experimental test of the 'special state' theory of quantum measurement," Entropy **14**, 665–686.

Sen, Indrajit, 2019, "A local ψ-epistemic retrocausal hidden-variable model of Bell correlations with wavefunctions in physical space," Found. Phys. **49**, 83–95.

Sen, Indrajit, 2020, "The effect of time-dilation on Bell experiments in the retrocausal Brans model," Proc. R. Soc. A **476**, 20190546.

Shalm, Lynden K. *et al.*, 2015, "Strong Loophole-Free Test of Local Realism," Phys. Rev. Lett. **115**, 250402.

Shimony, Abner, 2017, "Bell's theorem," in *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, fall 2017 ed. (The Metaphysics Research Lab, Stanford), https://plato.stanford.edu/archives/fall2017/entries/bell-theorem/.

Shimony, Abner, Michael A. Horne, and John F. Clauser, 1976, "Comment on 'The theory of local beables,'" Epistemological Lett. **13**, 1, https://curate.nd.edu/show/cv43nv96424.

Shrapnel, Sally, and Fabio Costa, 2018, "Causation does not explain contextuality," Quantum **2**, 63.

Sorkin, Rafael D., 2007, "Relativity theory does not imply that the future already exists: A counterexample," in *Relativity and the Dimensionality of the World* (Springer, New York), pp. 153–161.

Spekkens, Robert W., 2007, "Evidence for the epistemic view of quantum states: A toy theory," Phys. Rev. A **75**, 032110.

Spekkens, Robert W., 2015, "The paradigm of kinematics and dynamics must yield to causal structure," in *Questioning the Foundations of Physics*, edited by Anthony Aguirre, Brendan Foster, and Zeeya Merali (Springer, New York), pp. 5–16.

Stoica, Ovidiu Cristinel, 2019, "Representation of the wave function on the three-dimensional space," Phys. Rev. A **100**, 042115.

Streltsov, Alexander, Gerardo Adesso, and Martin B. Plenio, 2017, "Colloquium: Quantum coherence as a resource," Rev. Mod. Phys. **89**, 041003.

Sutherland, Roderick I., 1983, "Bell's theorem and backwards-in-time causality," Int. J. Theor. Phys. **22**, 377–384.

Sutherland, Roderick I., 2017, "Lagrangian description for particle interpretations of quantum mechanics: Entangled many-particle case," Found. Phys. **47**, 174–207.

Toner, B. F., and D. Bacon, 2003, "Communication Cost of Simulating Bell Correlations," Phys. Rev. Lett. **91**, 187904.

Tumulka, Roderich, 2006, "A relativistic version of the Ghirardi-Rimini-Weber model," J. Stat. Phys. **125**, 821–840.

Vaidman, L., 2013, "Past of a quantum particle," Phys. Rev. A **87**, 052104.

Weinstein, Steven, 2017, "Learning the Einstein-Podolsky-Rosen correlations on a restricted Boltzmann machine," arXiv:1707.03114.

Weinstein, Steven, 2018, "Neural networks as 'hidden' variable models for quantum systems," arXiv:1807.03910.

Wharton, K. B., 2010, "A novel interpretation of the Klein-Gordon equation," Found. Phys. **40**, 313–332.

Wharton, Ken, 2014, "Quantum states as ordinary information," Information **5**, 190–208.

Wharton, Ken, 2015, "The Universe is not a computer," in *Questioning the Foundations of Physics*, edited by Anthony Aguirre, Brendan Foster, and Zeeya Merali (Springer, New York), pp. 177–189.

Wharton, Ken, 2016, "Towards a realistic parsing of the Feynman Path Integral," Quanta **5**, 1–11.

Wharton, Ken, 2018, "A new class of retrocausal models," Entropy **20**, 410.

Wharton, Ken B., David J. Miller, and Huw Price, 2011, "Action duality: A constructive principle for quantum foundations," Symmetry **3**, 524–540.

Wheeler, John Archibald, and Richard Phillips Feynman, 1945, "Interaction with the absorber as the mechanism of radiation," Rev. Mod. Phys. **17**, 157.

Wheeler, John Archibald, and Richard Phillips Feynman, 1949, "Classical electrodynamics in terms of direct interparticle action," Rev. Mod. Phys. **21**, 425.

Wiseman, Howard M., 2014, "The two Bell's theorems of John Bell," J. Phys. A **47**, 424001.

Wood, Christopher J., and Robert W. Spekkens, 2015, "The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning," New J. Phys. **17**, 033002.

Woodward, James, 2005, *Making Things Happen: A Theory of Causal Explanation* (Oxford University Press, New York).

Zurek, Wojciech Hubert, 2018, "Quantum theory of the classical: Quantum jumps, Born's rule and objective classical reality via quantum Darwinism," Phil. Trans. R. Soc. A **376**, 20180107.