

Clauser-Horne-Shimony-Holt game, Tsirelson's bound, and causal locality

Jacob A. Barandes*

Departments of Philosophy and Physics, [Harvard University](#), 17 Oxford Street, Cambridge, Massachusetts 02138, USA

Mahmudul Hasan[†] and David Kagan[‡]

Department of Physics, [University of Massachusetts Dartmouth](#), North Dartmouth, Massachusetts 02747, USA



(Received 2 January 2026; revised 29 April 2026; accepted 9 June 2026; published 17 August 2026)

Starting with heuristic assumptions about locality and causality, Bell established a numerical upper bound on statistical correlations between spatially separated systems and showed that quantum systems could violate it. This upper bound, also known as the Bell or Clauser-Horne-Shimony-Holt (CHSH) inequality, can be derived from an experimental setup known as the CHSH game. Tsirelson used the CHSH formulation to derive a higher, maximal bound on quantum correlations. Later, Popescu and Rohrlich (PR) studied correlations beyond quantum theory that were constrained only by the no-signaling principle, a condition that disallows superluminal communication between observers. They showed that such correlations could theoretically violate the Tsirelson bound. Is there some conceptually simple or transparent principle that explains the large gap between the Tsirelson bound and the PR bound? Or do we have to appeal to the technical axiomatic structure of quantum theory? In this paper, we reformulate the CHSH game in terms of indivisible stochastic processes. Using Barandes's stochastic-quantum correspondence and its associated definition of causal locality, we present a proof of the Tsirelson bound. In particular, we show that, unlike the no-signaling principle alone, the postulates defining causally local, indivisible stochastic processes are precisely strong enough to allow for violations of the Bell inequality up to, but not beyond, the Tsirelson bound.

DOI: [10.1103/xfpf-734d](#)

I. INTRODUCTION

Consider the Clauser-Horne-Shimony-Holt (CHSH) inequality interpreted as arising from a coordination game [1,2]: Before any given round, two players, Alice and Bob, agree on a strategy. They are unable to communicate directly once a round has begun. During each round of the game, Alice randomizes a bit and then sends the resulting value, along with a second bit value of her choosing, to the referee, Charlie. Bob does the same. Alice knows the values of the two bits she sends to Charlie, and Bob likewise knows the values of his two bits, but neither can be aware of the other's randomized bit. To win the round, Alice and Bob's generated bits must somehow add up (modulo 2) to match the product of their random bits.

Under the assumptions of various versions of Bell's theorem [3–5], Alice and Bob should not be able to win more than 3/4 of the rounds, even if they are allowed to share correlated bits as resources. However, if the players share a suitably prepared entangled pair of qubits, then they can exceed the

limitations of Bell's theorem and saturate Tsirelson's bound, winning $\cos^2(\pi/8) \approx 85.4\%$ of the time [6].

Known physical systems, whether classical or quantum, obey the no-signaling principle, which is roughly that probabilities for local events depend only on locally available data, thereby ruling out the possibility of superluminal signals [7,8]. Popescu and Rohrlich (PR) [9] studied whether imposing no-signaling was sufficient to explain the existence of the Tsirelson bound. Utilizing a construction by Rastall [10], Popescu and Rohrlich demonstrated the existence of no-signaling models that maximally exceed Tsirelson's bound. If such resources exist in nature, then they would allow Alice and Bob to win the CHSH game 100% of the time.

Why, then, does our world appear to realize the intermediate range, violating Bell's inequality while respecting Tsirelson's bound? This question constitutes the quantum-correlation problem [11]. In [12], the authors put forward their principle of information causality, which states that local systems can gain only as much information about a remote system as is communicated to them. According to [12,13], information causality can, indeed, be used to rule out no-signaling theories that exceed Tsirelson's bound, although arriving at the precise bound requires considering protocols that involve either a large number of bits or noisy channels.

Moreover, information causality belongs to an instrumentalist orientation toward quantum theory. While such approaches to reformulating or reconstructing quantum theory have been fruitful, they leave the agents who implement operations outside of the explanatory framework [11,14–17].

*Contact author: jacob_barandes@harvard.edu

[†]Contact author: mahmudul.hasan@umassd.edu

[‡]Contact author: dkagan@umassd.edu

Is there another way to understand why our world allows correlations stronger than Bell's but not the theoretical maximum consistent with no signaling? The stochastic-quantum correspondence [18–20] and the principle of causal locality formulated in [21] suggest an alternative. Consider the following postulates.

(1) *The existence of system configurations.* A physical system's configurations form a set, and the system is in one of its possible configurations at any given instant.¹ The configuration set of a composite system is the Cartesian product of the corresponding subsystem configuration sets.

(2) *A stochastic description of time-directed dynamics.* The dynamical laws governing a physical system's evolution are described via time-directed conditional probabilities relating probabilities of configurations at a given conditioning time to probabilities of configurations at target times. The dynamics are generically non-Markovian. We will often take the conditioning time to be an initial time preceding a later target time, but there need not be any fundamental breaking of time-reversal symmetry or fundamental arrow of time. Subsystems are governed by marginal distributions derived from the dynamical laws of their parent systems.

(3) *Causal locality.* Physical influences as described by a system's stochastic dynamics cannot propagate faster than the speed of light. That is, the stochastic dynamics of a localized system during a specific time interval cannot be conditionally dependent on any other systems beyond the range of lightlike signals during that time interval.

These postulates are meant to apply both to systems and to the agents who study and manipulate them. Barandes [18–20] showed that the first two postulates are sufficient to describe all known physical systems, including quantum systems. Here, we illustrate how stochastic systems that respect the third postulate can violate Bell inequalities up to Tsirelson's bound. Note the implication that any nonsignaling scenario involving Tsirelson-violating correlations, such as PR boxes, cannot be modeled using the indivisible stochastic systems of [18].

How can our result depend on a third postulate, beyond the first two, if the first two postulates yield a formalism that is empirically equivalent to quantum theory? After all, Tsirelson's bound emerges as a result of textbook or orthodox quantum theory. The crucial point is that in the orthodox approach, external observers and measuring devices are treated instrumentally as part of the measurement axioms of quantum theory. Projection operators associated with local measurements are then given by a standard recipe, but that recipe for the form of these local projection operators requires further technical assumptions built out of the measurement axioms themselves. Within the framework of Barandes's indivisible stochastic theory, empirical equivalence with the predictions of quantum theory requires that agents and measurement de-

vices be included as parts of the overall system in our analysis. The independence of spatially separated measurement processes is derived from our third postulate, providing a simpler and conceptually more transparent explanation for the forms of measurement projectors.

Section II reviews the CHSH game and the no-signaling (NS) box framework that demonstrates the existence of probabilistic strategies that win the game all the way up to the theoretical maximum, assuming suitable nonsignaling resources shared by Alice and Bob. Section III outlines the stochastic-quantum correspondence and some of the key ideas for the definition of causal locality that we use in this work. In Sec. IV, we recast an arbitrary round of the CHSH game in stochastic terms. We then prove that adopting causal locality yields the Tsirelson bound. In Sec. V, we consider some questions about the nature of causality and locality in quantum theory versus the stochastic picture and suggest some directions for future research.

II. REVIEW OF THE CHSH GAME

A. Setup

In a given round of the CHSH game, two players, Alice and Bob, are spacelike separated. Alice randomizes a bit $x \in \{0, 1\}$, which she sends to Charlie, along with another bit $q \in \{0, 1\}$ of her choosing. Alice is aware of both bit values. Bob follows the same procedure, randomizing $y \in \{0, 1\}$ and choosing $r \in \{0, 1\}$, sending both to Charlie. Charlie tests whether the condition $q \oplus r = xy$ holds, where the operation $\oplus$ denotes addition modulo 2. If Alice and Bob win, they are awarded a point; if they lose, they are docked a point.

Alice and Bob share a bipartite resource, prepared by Charlie in advance according to instructions that Alice and Bob agreed to before the start of the game. Charlie sends Alice one part of the resource and Bob the other. Only one bit of information is extractable from each part of the resource.

Let $P(q, r, x, y)$ denote the probability that Alice and Bob return the bit values q, r, x , and y . Their expected score in any given round of the game is

$$\begin{aligned} \langle \text{score} \rangle &= \sum_{x,y,q,r} [P(q, r, x, y) \delta(q \oplus r = xy) \\ &\quad - P(q, r, x, y) \delta(q \oplus r \neq xy)] \\ &= \sum_{x,y,q,r} (-1)^{xy - q \oplus r} P(q, r | x, y) P(x, y) \\ &= \sum_{x,y} (-1)^{xy} [P(0, 0 | x, y) - P(0, 1 | x, y) \\ &\quad - P(1, 0 | x, y) + P(1, 1 | x, y)] P(x, y), \end{aligned} \quad (1)$$

where $P(x, y)$ is the marginal joint probability of the bit value pair (x, y) and $P(q, r | x, y) \equiv P(q, r, x, y) / P(x, y)$. The function $\delta(\text{condition}) = 1$ if the condition in its argument is true and is 0 otherwise. We can reexpress the score by substituting

¹We do not assume a richer mathematical structure, such as a Hilbert space. The case of systems with finite sets of configurations is sufficient for our purposes.

$P(1, 1|x, y) = 1 - P(0, 0|x, y) - P(0, 1|x, y) - P(1, 0|x, y)$ into the last expression, yielding

$$\begin{aligned}\langle \text{score} \rangle &= \sum_{x,y} (-1)^{xy} [1 - 2P(0, 1|x, y) - 2P(1, 0|x, y)] P(x, y) \\ &= \sum_{x,y} (-1)^{xy} P(x, y) - 2 \sum_{x,y} (-1)^{xy} [P(0, 1|x, y) + P(1, 0|x, y)] P(x, y) \\ &= 1 - 2P(x = 1, y = 1) - 2 \sum_{x,y} (-1)^{xy} [P(0, 1|x, y) + P(1, 0|x, y)] P(x, y),\end{aligned}\quad (2)$$

where the last equality in (2) follows from

$$\begin{aligned}\sum_{x,y} (-1)^{xy} P(x, y) &= P(0, 0) + P(0, 1) + P(1, 0) - P(1, 1) \\ &= 1 - 2P(1, 1).\end{aligned}\quad (3)$$

If Alice and Bob generate the values x and y according to a uniform distribution, then $P(x, y) = 1/4$. Thus,

$$\langle \text{score} \rangle = \frac{1}{2} \left\{ 1 - \sum_{x,y} (-1)^{xy} [P(0, 1|x, y) + P(1, 0|x, y)] \right\}.\quad (4)$$

Equivalently, from (1), we have

$$\begin{aligned}\langle \text{score} \rangle &= \frac{1}{4} \sum_{x,y} (-1)^{xy} [P(0, 0|x, y) \\ &\quad - P(0, 1|x, y) - P(1, 0|x, y) + P(1, 1|x, y)].\end{aligned}\quad (5)$$

Let us, for the moment, assume that the conditional probability distribution $P(q, r|x, y)$ is Bell local in the sense that²

$$P(q, r|x, y) = \sum_{\lambda} P(q|x, \lambda) P(r|y, \lambda) P(\lambda),\quad (6)$$

where λ denotes a set of local hidden variables that are established at the start of a round of the game. Substituting (6) into (5) yields

$$\begin{aligned}\langle \text{score} \rangle &= \frac{1}{4} \sum_{x,y} \sum_{\lambda} (-1)^{xy} [P(0|x, \lambda) P(0|y, \lambda) - P(0|x, \lambda) P(1|y, \lambda) \\ &\quad - P(1|x, \lambda) P(0|y, \lambda) + P(1|x, \lambda) P(1|y, \lambda)] P(\lambda) \\ &= \frac{1}{4} \sum_{x,y} \sum_{\lambda} (-1)^{xy} [P(0|x, \lambda) - P(1|x, \lambda)] [P(0|y, \lambda) - P(1|y, \lambda)] P(\lambda).\end{aligned}\quad (7)$$

Let

$$\begin{aligned}\langle A_x(\lambda) \rangle &\equiv P(0|x, \lambda) - P(1|x, \lambda), \\ \langle B_y(\lambda) \rangle &\equiv P(0|y, \lambda) - P(1|y, \lambda).\end{aligned}\quad (8)$$

The quantities A_x and B_y are the standard CHSH observables, taking values of ± 1 . The bracketed quantities defined in (8) are the expectation values for A_x and B_y , conditioned on the hidden variables λ . We can now rewrite (7) in the usual CHSH form by expanding out the sums in x and y ,

$$\begin{aligned}\langle \text{score} \rangle &= \frac{1}{4} \sum_{x,y} \sum_{\lambda} (-1)^{xy} \langle A_x(\lambda) \rangle \langle B_y(\lambda) \rangle P(\lambda) \\ &= \frac{1}{4} \sum_{\lambda} \{ \langle A_0(\lambda) \rangle [\langle B_0(\lambda) \rangle + \langle B_1(\lambda) \rangle] \\ &\quad + \langle A_1(\lambda) \rangle [\langle B_0(\lambda) \rangle - \langle B_1(\lambda) \rangle] \} P(\lambda) \\ &= \frac{1}{4} \langle A_0[B_0 + B_1] + A_1[B_0 - B_1] \rangle.\end{aligned}\quad (9)$$

Thus, the typical expected value for the CHSH quantity is 4 times the expected score in our setup for the CHSH game.

B. No signaling and NS boxes

The conditional probability distribution $P(q, r|x, y)$ describes all possible correlations between Alice's and Bob's response bits, given the pair of input bits (x, y) they randomly generate and their use of the shared resource. Note that at this stage, $P(q, r|x, y)$ is merely descriptive and does not commit us to any particular underlying dynamical picture. The no-signaling principle is a physical criterion that is meant to constrain $P(q, r|x, y)$ and thus indirectly encode some physical information in the correlations the conditional probabilities describe. These no-signaling correlations satisfy

$$P(q|x, y) = P(q|x), \quad P(r|x, y) = P(r|y),\quad (10)$$

where the marginal probabilities $P(q|x, y)$ and $P(r|x, y)$ are defined by

$$P(q|x, y) \equiv \sum_r P(q, r|x, y), \quad P(r|x, y) \equiv \sum_q P(q, r|x, y).\quad (11)$$

The no-signaling constraints describe the intuition that outcomes seen only locally by either Alice or Bob should be

²Note that Bell locality is a distinct and stronger condition than the causal-locality condition that we define in Sec. III C.

affected by only local influences. In the CHSH context, Alice's outcomes for q should be correlated only with the bit value x , and similarly, Bob's outcomes for r should be correlated only with y .

Following [12,22], we can characterize nonsignaling resources using the family of distributions referred to as no-signaling boxes,

$$P_{\text{NS}}(q, r|x, y) = \frac{1}{4}[1 + (-1)^{xy-q\oplus r}E]. \quad (12)$$

Note that this distribution is locally unbiased in that Alice and Bob each have marginal probabilities of $1/2$ for any of the outcomes $q = 0, 1$ and $r = 0, 1$, respectively. In particular, it satisfies the no-signaling conditions (10). Substituting (12) into (5) shows that the parameter E is equivalent to $\langle \text{score} \rangle$. The case $E = 0$ corresponds to Alice and Bob ignoring their random bit values, x and y , respectively, and simply selecting q and r randomly with the equivalent of fair coin tosses. When $E = 1/2$, Alice and Bob have chosen a strategy that allows them to win $3/4$ of the time, saturating Bell's inequality. The quantum upper limit corresponds to $E = 1/\sqrt{2}$, while the "PR-box" construction of [9,10] corresponds to $E = 1$.

More generally, the set of all joint, nonsignaling conditional probability distributions $P(q, r|x, y)$ forms an eight-dimensional, convex NS polytope [23]. The Bell-local subpolytope corresponds to the subset of distributions satisfying the condition (6), which automatically impose the no-signaling constraints. Distributions constrained by the Bell-locality condition can achieve a maximum score of only $1/2$. By contrast, PR-box-type distributions correspond to the portions of the NS polytope that algebraically maximize the CHSH score.

These considerations, whether in terms of the unbiased NS-box construction (12) or the more generic construction of the NS polytope, show that the no-signaling principle alone places no restrictions on the achievable CHSH game score. No signaling thus lacks the explanatory resources to single out the Tsirelson bound.

III. THE STOCHASTIC-QUANTUM CORRESPONDENCE AND CAUSAL LOCALITY

A. Indivisible stochastic systems

Let $\mathcal{C}$ be the set of configurations of a system Q . For simplicity, we assume that $\mathcal{C}$ is finite, and let $q_t = 1, 2, \dots, N$ label the possible configuration of Q at time t . When we refer to a specific time t_n that carries an index of its own, we employ the notation $q_n \equiv q_{t_n}$. We assume at least one conditioning time t_0 at which Q has no correlations with any other systems.

Conditional probabilities $p(q_t|q_0)$ encode stochastic dynamics. Given an initial stand-alone probability distribution over configurations $p(q_0) \equiv p(q_0; t_0)$, we may define a joint probability distribution,

$$p(q_t, q_0) = p(q_t|q_0)p(q_0), \quad (13)$$

and marginalize over the configuration at the initial time to arrive at the stand-alone probability for configuration q at time t ,

$$p(q_t) = \sum_{q_0} p(q_t|q_0)p(q_0). \quad (14)$$

The formalism works for either $t < t_0$ or $t > t_0$, but for specificity, we will assume that t_0 is the earliest time under consideration at which our system is uncorrelated with anything else and that all other times in the process occur later.

Generically, systems evolve indivisibly, except for specific instants, including t_0 , called division events or division times. Indivisible evolution means that the system's configurations cannot be tracked, even in principle, which is reflected by the lack of any assignment of probabilities to system trajectories. By contrast, a divisible stochastic process during the time interval $t_0 \rightarrow t$ has division times at each instant within the given time interval, up to the given level of precision.

We define a stochastic matrix $\Gamma(t \leftarrow t_0)$ with components $\Gamma_{q_t, q_0}(t \leftarrow t_0) \equiv p(q_t|q_0)$. An indivisible stochastic process is called divisible at a time t_1 if the stochastic matrix factorizes at that time,

$$\Gamma(t \leftarrow t_0) = \Gamma(t \leftarrow t_1)\Gamma(t_1 \leftarrow t_0), \quad (15)$$

where $\Gamma(t \leftarrow t_1)$ is a well-defined stochastic matrix of its own. In that case, we axiomatically take the matrix elements of $\Gamma(t \leftarrow t_1)$ to be conditional probabilities, so that, given the stand-alone distribution at time $t_1 > t_0$,

$$p(q_1) = \sum_{q_0} p(q_1|q_0)p(q_0), \quad (16)$$

we may consistently write

$$p(q_t) = \sum_{q_1} p(q_t|q_1)p(q_1), \quad (17)$$

where $t > t_1$. As with t_0 , the time t_1 is called a division event or division time. Correlations between the system Q and other systems are effectively washed out during a division event. Division events are often only approximate or effective, such as when they arise due to interactions between a system and a larger environment. We adopt the pragmatic perspective that all dynamical probabilistic statements are defined up to some given level of precision, which demarcates the validity of approximating a process as a division event.

Composite systems

As stipulated in the first postulate in Sec. I, a configuration of a composite system QR can be expressed as an ordered pair (q_t, r_t) of configurations q_t and r_t of the individual systems Q and R , respectively. Thus, the configurations of QR form a Cartesian product, $\mathcal{C}_{QR} = \mathcal{C}_Q \times \mathcal{C}_R$, where $\mathcal{C}_Q$ and $\mathcal{C}_R$ are the configuration sets of the systems Q and R , respectively. Given an initial division event at t_0 , the joint stochastic dynamics for QR are defined by conditional probabilities of the form $p(q_t, r_t|q_0, r_0)$ that describe the evolution of the initial probability distribution to later times t ,

$$p(q_t, r_t) = \sum_{q_0, r_0} p(q_t, r_t|q_0, r_0)p(q_0, r_0). \quad (18)$$

Marginalizing yields mixed stochastic dynamics for Q and R ,

$$\begin{aligned} p(q_t|q_0, r_0) &= \sum_{r_t} p(q_t, r_t|q_0, r_0), \\ &= \sum_{q_t} p(q_t, r_t|q_0, r_0). \end{aligned} \quad (19)$$

If the conditional probabilities factorize,

$$p(q_t, r_t | q_0, r_0) = p(q_t | q_0) p(r_t | r_0), \quad (20)$$

then we say that the subsystems Q and R do not interact during the interval $t_0 \rightarrow t$. If Q and R interact at a later time, then their composite stochastic dynamics will cease to factorize until a subsequent division event. Note that interactions are an example of the broader notion of causal influence defined in [21] and described in more detail in Sec. III C.

B. The stochastic-quantum correspondence

1. Trace formula for conditional probabilities

A system's stochastic dynamics can be (nonuniquely) represented in terms of complex-valued elements of a matrix $\Theta(t \leftarrow t_0)$,

$$\Theta_{q_t q_0}(t \leftarrow t_0) = e^{i\theta_{q_t q_0}(t)} \sqrt{p(q_t | q_0)} \times p(q_t | q_0) = |\Theta_{q_t q_0}(t \leftarrow t_0)|^2, \quad (21)$$

where $\theta_{q_t q_0}(t)$ is an arbitrary phase. For the moment, let us consider the case where $\Theta(t \leftarrow t_0) = U(t \leftarrow t_0)$ is a unitary matrix. In that case, the corresponding matrix $\Gamma(t \leftarrow t_0)$ of conditional probabilities is called unistochastic [24]. The matrix $U(t \leftarrow t_0)$ acts linearly on a Hilbert space H_C of dimension N . The configurations of the system define a preferred basis, $\{|q\rangle\}_q$, for H_C . However, $U(t \leftarrow t_0)$ generically carries configuration-basis vectors to other vectors that are superpositions of configuration-basis vectors. They are not regarded as physical configurations of the system, but rather as a way to characterize the behavior of the system when it undergoes indivisible evolution.

Returning to the case of generic $\Theta(t \leftarrow t_0)$, we let $\mathcal{P}_q = |q\rangle\langle q|$ be a projection matrix onto the configuration-basis vector $|q\rangle$. The stochastic dynamic probabilities can now be expressed via the trace

$$p(q_t | q_0) = \text{Tr}[\mathcal{P}_{q_t} \Theta(t \leftarrow t_0) \mathcal{P}_{q_0} \Theta^\dagger(t \leftarrow t_0)]. \quad (22)$$

The relation (22) is the dictionary of the stochastic-quantum correspondence.

2. Unistochastic dynamics

Any indivisible stochastic system's quantum description can be dilated so that the enlarged system evolves according to a unitary time-evolution matrix $U(t \leftarrow t_0)$ with matrix elements $U_{q_t q_0}$, as demonstrated in [19]. Processes that can be represented unitarily are called unistochastic processes. Their conditional probabilities satisfy

$$\begin{aligned} p(q_t | q_0) &= \text{Tr}[\mathcal{P}_{q_t} U(t \leftarrow t_0) \mathcal{P}_{q_0} U^\dagger(t \leftarrow t_0)] \\ &= |\langle q_t | U(t \leftarrow t_0) | q_0 \rangle|^2. \end{aligned} \quad (23)$$

Unitarity permits us to define $U(t \leftarrow t') \equiv U(t \leftarrow t_0) U^\dagger(t' \leftarrow t_0)$ for any $t' < t$, which implies that $U(t \leftarrow t_0) = U(t \leftarrow t') U(t' \leftarrow t_0)$, and hence,

$$\begin{aligned} p(q_t | q_0) &= \sum_{q', q''} U_{q_t q'}^* U_{q' q_0}^* U_{q_t q''} U_{q'' q_0} \\ &= \sum_{q'} |U_{q_t q'}|^2 |U_{q' q_0}|^2 + \sum_{q' \neq q''} U_{q_t q'}^* U_{q' q_0}^* U_{q_t q''} U_{q'' q_0} \end{aligned}$$

$$\begin{aligned} &= \sum_{q'} \Gamma_{q_t, q'}'(t \leftarrow t') \Gamma_{q', q_0}(t' \leftarrow t_0) \\ &\quad + \text{qCor}_{q_t, q_0}(t \leftarrow t' \leftarrow t_0). \end{aligned} \quad (24)$$

Here, the index q' corresponds to the intermediate configuration label at time t' , and the matrix $\text{qCor}(t \leftarrow t' \leftarrow t_0)$ captures the discrepancy between actual indivisible unistochastic evolution and fictitious unistochastic evolution with a division event at t' . Such discrepancies are typically referred to as quantum interference effects. The stochastic-quantum correspondence thus situates quantum interference within the broader context of discrepancies between exact non-Markovian processes and any Markovian approximation.

3. Division events and decoherence

Stochastic evolution is generically indivisible. Given a system Q , let $\mathcal{T}_{\text{div}}$ denote the set of division times where

$$\begin{aligned} \Gamma(t \leftarrow t_0) &= \Gamma(t \leftarrow t_n) \Gamma(t_n \leftarrow t_{n-1}) \cdots \Gamma(t_1 \leftarrow t_0), \\ t_0, t_1, \dots, t_n &\in \mathcal{T}_{\text{div}}, \end{aligned} \quad (25)$$

and $t_0 < t_1 < \cdots < t_n < t$. The stochastic-quantum correspondence associates effective division events with environmental decoherence [18].³ Such decoherence can be modeled in terms of dephasing channels $\mathcal{D}_i$ that remove density-matrix coherence terms at the times $t_i \in \mathcal{T}_{\text{div}}$ as expressed in the configuration basis. We refer to these as division channels.

C. Causal influence and causal locality

Consider two systems Q and R . Following [21], we say that R does *not* causally influence Q during the time interval starting with the initial division time t_0 until the later target time t if

$$p(q_t | q_0, r_0) = p(q_t | q_0), \quad (26)$$

where $p(q_t | q_0, r_0)$ is the marginal conditional probability distribution derived from the joint conditional probability distribution $p(q_t, r_t | q_0, r_0)$. Otherwise, R does causally influence Q . Note that no external observers or measurements are invoked in this approach to causality. The laws formulated as conditional probabilities themselves provide the causal description in a noninterventionist variant of causal Bayesian network theory [25].

If Q and R are causally independent, then their marginal conditional probability distributions satisfy

$$p(q_t | q_0, r_0) = p(q_t | q_0), \quad (27)$$

$$p(r_t | q_0, r_0) = p(r_t | r_0). \quad (28)$$

The conditions (27) and (28) are weaker than the factorization condition (20) that characterizes noninteraction. Thus, interaction is a special case of causal influence.

Suppose that Q and R are spatially separated. They form a causally local system if they are causally independent when $c(t - t_0)$ is shorter than their separation distance, with c being

³Exact division events can also occur. See the example given in [18].

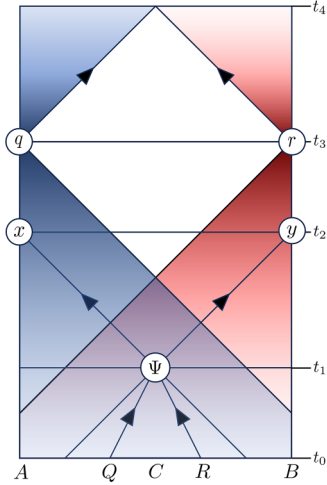


FIG. 1. A spacetime diagram depicting one round of the CHSH game.

the speed of light. Note how the stochastic picture, in postulating a level of physical reality beneath the Hilbert spaces of standard quantum theory, makes it possible to formulate this new criterion of causal locality, which will turn out to be stronger than the no-signaling principle and will make it possible to derive the Tsirelson bound.

IV. STOCHASTIC DESCRIPTION OF THE CHSH GAME

Consider the overarching system made up of Alice A , Bob B , Charlie C , particle Q , particle R , and their local environment E . Alice, Bob, and Charlie are monitored by the environment nearly continuously, so their local stochastic dynamics are effectively divisible. Particles Q and R may be shielded from the environment and may thus undergo indivisible unistochastic evolution.

Alice's, Bob's, and Charlie's configurations at any given time t are denoted by $a_t(\dots)$, $b_t(\dots)$, and $c_t(\dots)$, respectively, where " $\dots$ " represent the details of a particular configuration, potentially including memories of past configurations. Alice and Bob are each able to send Charlie two bits of information per round. Charlie can also transmit Q and R to Alice and Bob, respectively. Finally, Charlie, Alice, and Bob can perform local operations that are described by indivisible stochastic processes on systems Q and R .

Let t_0 denote a division event for all of the systems initiating one round of the game (Fig. 1). At this time, Charlie has particles Q and R in configurations q_0 and r_0 , respectively. During the interval $t_0 \rightarrow t_1$, Charlie implements the instructions given to him by Alice and Bob before the round began. At t_1 , Charlie's configuration will include information about the preparation of the composite system QR , which we denote $c_1(\Psi, \dots)$. The preparation may involve joining Q and R in an indivisible process, in which case the label Ψ in Charlie's memory is correlated with a process rather than with any particular configuration (q_1, r_1) of the joint system QR .

Charlie sends Q to Alice and R to Bob at time t_1 , with the particles arriving at time t_2 . On receiving their particles, Alice and Bob generate the random bits x and y , respectively. The bit values are reflected in Alice and Bob's configurations as

$a_2(x, \dots)$ and $b_2(y, \dots)$. Alice and Bob perform local operations on Q and R , respectively, during the interval $t_2 \rightarrow t_3$. Note that Alice's and Bob's choices of local operations may be influenced by the values of x and y , respectively, as part of their prearranged strategy. Causal locality plays an important role here, ensuring that Alice's choice of operation depends solely on x and Bob's depends solely on y .

At time t_3 , Alice and Bob measure the configurations of Q and R to be q and r , respectively. Alice and Bob will choose their discretionary bits to be their measurement-outcome bits q and r , respectively. They communicate all of their bit values x, y, q , and r to Charlie, who receives the information at time t_4 and whose configuration is now of the form $c_4(q, r, x, y, \Psi, \dots)$. Charlie is thus able to check the win condition for the round and keep a tally of the score in memory.

The overall process described above results in a hybrid stochastic distribution for q and r at times $t \geq t_3$ conditioned on Alice's knowledge $a_2(x)$ of x , Bob's knowledge $b_2(y)$ of y , and Charlie's preparation $c_1(\Psi)$,

$$p_{xy}(q, r|\Psi) \equiv p(q, r; t|a_2(x), b_2(y), c_1(\Psi)), \quad t \geq t_3, \quad (29)$$

where $a_2(x)$, $b_2(y)$, and $c_1(\Psi)$ lie along a single Cauchy hypersurface. The left-hand side of (29) introduces a notational shorthand, while the right-hand side makes explicit the conditional dependencies inherent in the stochastic process we are describing. We use lowercase p to distinguish the stochastic distribution (29) above from the generic distribution $P(q, r|x, y)$ introduced in Sec. II A.

A. The role of causal locality

In the CHSH context, causal locality implies further that

$$p_{xy}(q|\Psi) \equiv p_x(q|\Psi) = p(q; t|a_2(x), c_1(\Psi)), \quad (30)$$

$$p_{xy}(r|\Psi) \equiv p_y(r|\Psi) = p(r; t|b_2(y), c_1(\Psi)). \quad (31)$$

Causal locality thus defines what it means for Alice's and Bob's operations to be local: Configuration r is outside of the future light cone of random bit x and thus cannot be causally influenced by x (Fig. 2). Similarly, configuration q is outside of the future light cone of random bit y and thus cannot be causally influenced by y . However, the marginal probabilities for outcomes q and r both generically depend on the result of the preparation Ψ at time t_1 due to that indivisible process being in both of their causal pasts.

The no-signaling principle by itself does not constrain the nature of the processes underlying the descriptive joint conditional probability distribution $P(q, r|x, y)$ of Sec. II. By contrast, causal locality is more stringent, requiring a stochastic picture with time-directed conditional probabilities that describe the underlying laws governing the system.

At this stage, we claim the following.

Tsirelson's bound. The stochastic-quantum correspondence implies that the CHSH game score associated with a causally local unistochastic process given by $p_{xy}(q, r|\Psi)$ can be no greater than $1/\sqrt{2}$. Equivalently, the total win probability cannot exceed $(1 + \sqrt{2})/2\sqrt{2}$.

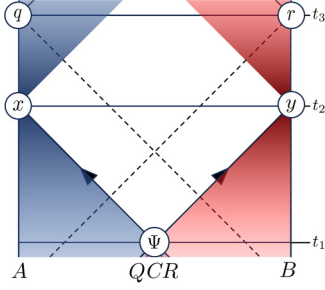


FIG. 2. At time t_3 , the configuration q of Q may depend on the random bit x as it lies in the blue shaded future light cone in the diagram. Similarly, the configuration r of R at time t_3 lies in the red shaded future light cone of random bit y and may therefore depend on y . Both configurations q and r will generically depend on the preparation Ψ because it lies in the overlap of their past light cones under the intersection of the dotted lines in the diagram. However, at t_3 , causal locality implies that r cannot depend on x and q cannot depend on y because they lie outside of the future light cones of x and y , respectively.

B. Proof

Suppose that QR evolves unistochastically up until Alice's and Bob's measurements at time t_3 . According to the stochastic-quantum dictionary established in [18], we have

$$\begin{aligned} p_{xy}(q, r|\Psi) &= \text{Tr}[(\mathcal{P}_q \otimes \mathcal{P}_r)(\mathcal{AB})_{xy} \mathcal{C}_\Psi |q_0, r_0\rangle \langle q_0, r_0| \\ &\quad \times \mathcal{C}_\Psi^\dagger (\mathcal{AB})_{xy}^\dagger] \\ &= \text{Tr}[(\mathcal{P}_q \otimes \mathcal{P}_r)(\mathcal{AB})_{xy} |\Psi\rangle \langle \Psi| (\mathcal{AB})_{xy}^\dagger], \end{aligned} \quad (32)$$

where $\mathcal{P}_q = |q\rangle\langle q|$, $\mathcal{P}_r = |r\rangle\langle r|$, $(\mathcal{AB})_{xy}$ is a unitary representation of Alice and Bob's joint operations on QR , and $\mathcal{C}_\Psi$ is a unitary representation of Charlie's preparation of QR .

Using properties of the trace, we may reexpress (32) as

$$p_{xy}(q, r|\Psi) = \langle \Psi | (\mathcal{AB})_{xy}^\dagger (\mathcal{P}_q \otimes \mathcal{P}_r) (\mathcal{AB})_{xy} | \Psi \rangle. \quad (33)$$

Causal locality requires Alice and Bob's joint operation to factorize: $(\mathcal{AB})_{xy} = \mathcal{A}_x \otimes \mathcal{B}_y$. Therefore,

$$p_{xy}(q, r|\Psi) = \langle \Psi | \mathcal{A}_x^\dagger \mathcal{P}_q \mathcal{A}_x \otimes \mathcal{B}_y^\dagger \mathcal{P}_r \mathcal{B}_y | \Psi \rangle. \quad (34)$$

We define the Hermitian operators

$$A_x^q \equiv \mathcal{A}_x^\dagger \mathcal{P}_q \mathcal{A}_x \otimes \mathbb{I}, \quad B_y^r \equiv \mathbb{I} \otimes \mathcal{B}_y^\dagger \mathcal{P}_r \mathcal{B}_y. \quad (35)$$

We calculate the CHSH score for this setup by substituting $P(q, r|x, y) = p_{xy}(q, r|\Psi)$ into formula (5),

$$\begin{aligned} \langle \text{score} \rangle &= \frac{1}{4} \sum_{x,y} (-1)^{xy} [p_{xy}(0, 0|\Psi) - p_{xy}(0, 1|\Psi) \\ &\quad - p_{xy}(1, 0|\Psi) + p_{xy}(1, 1|\Psi)] \\ &= \frac{1}{4} \langle \Psi | \sum_{x,y} (-1)^{xy} (A_x^0 B_y^0 - A_x^0 B_y^1 - A_x^1 B_y^0 + A_x^1 B_y^1) | \Psi \rangle \\ &= \frac{1}{4} \langle \Psi | \sum_{x,y} (-1)^{xy} \mathcal{A}_x \mathcal{B}_y | \Psi \rangle \\ &= \frac{1}{4} \langle \Psi | \mathcal{A}_0 (\mathcal{B}_0 + \mathcal{B}_1) + \mathcal{A}_1 (\mathcal{B}_0 - \mathcal{B}_1) | \Psi \rangle, \end{aligned} \quad (37)$$

where

$$\mathcal{A}_x \equiv A_x^0 - A_x^1, \quad \mathcal{B}_y \equiv B_y^0 - B_y^1. \quad (38)$$

The quantity

$$\mathcal{C} \equiv \mathcal{A}_0 (\mathcal{B}_0 + \mathcal{B}_1) + \mathcal{A}_1 (\mathcal{B}_0 - \mathcal{B}_1) \quad (39)$$

is the usual CHSH operator, satisfying the Tsirelson bound

$$|\langle \mathcal{C} \rangle| \leq 2\sqrt{2}, \quad (40)$$

and thus, $\langle \text{score} \rangle \leq 1/\sqrt{2}$.

Recalling that the CHSH game awards +1 point for winning and -1 point for losing, the expected score per round is related to the probability of winning the round via

$$\langle \text{score} \rangle = P_{\text{win}} - P_{\text{loss}} = 2P_{\text{win}} - 1, \quad (41)$$

where P_{win} is the probability of winning the round and $P_{\text{loss}} = 1 - P_{\text{win}}$ is the probability of losing the round. Substituting Tsirelson's bound for the score and solving (41) for the win probability yields

$$P_{\text{win}} \leq \frac{1 + \sqrt{2}}{2\sqrt{2}}. \quad (42)$$

V. FUTURE DIRECTIONS AND CONCLUSION

Our work demonstrates that treating the CHSH game as a causally local unistochastic process reproduces the Tsirelson bound for the CHSH game. In particular, our intuitively natural condition of causal locality threads a fine needle between violating the CHSH inequality up to the Tsirelson bound while preserving the no-signaling principle.

Does causal locality imply that quantum systems whose dynamics satisfy the no-signaling condition are themselves local, despite some famous arguments to the contrary? The trouble with drawing this conclusion is that textbook quantum theory does not supply a way to define local causal influence at a microphysical level, as argued in [21]. The standard Copenhagen interpretation refrains from assigning any picture whatsoever to the microphysical goings-on of our systems. As such, it precludes any possibility of a causally local explanation underlying the behavior of the system from one measurement to another.

Causal locality relies on a different conceptual framework, one that, in contrast to standard quantum theory and the Copenhagen interpretation, does paint a picture of underlying laws governing the behavior of systems as they evolve from one configuration to another. All of the game's actions can be formulated in terms of generically indivisible stochastic processes that occur entirely within a light cone that encompasses the entangling of qubit systems Q and R , their transmission to Alice and Bob, Alice's and Bob's subsequent local manipulations and measurements of Q and R , and, ultimately, the tallying of the score by Charlie.

Our work implies that a violation of the Tsirelson bound either requires a violation of the principle of causal locality, and thus manifestly nonlocal dynamical laws, or requires laws that cannot be formulated in terms of unistochastic processes. Are there frameworks for dynamical laws that go beyond the indivisible stochastic systems formulated in [18]? We leave that question to future exploration.

Other approaches specialize from the very beginning to quantum systems that are explicitly formulated in terms of unitary time-evolution operators on Hilbert spaces [26]. According to these alternative approaches, one needs to impose a technical condition to obtain violations of Tsirelson's bound. This technical condition is that for a composite system consisting of localized subsystems at spacelike separation, the overall time-evolution operator should fail to tensor factorize into unitary factors that individually encode the time evolution of each localized subsystem. One can show that such arrangements would lead to faster-than-light signaling. Our work, by contrast, does not assume that we are limiting ourselves to the standard framework of textbook quantum theory based on a Hilbert-space formulation, thus allowing us to draw broader implications in the context of causally local unistochastic systems.

In developing the tools for our reexamination of the CHSH game, we explored the connection between division events and decoherence via the stochastic-quantum correspondence for systems that possess memory. Such systems may exhibit division events that produce a Markov-like washing away of information at times before the most recent division time, but marginalizing over their memory subsystem leaves the remaining subsystem with a non-Markov-like chain of conditional probabilities. More concretely, consider a system $A = SM$ with a subsystem S and a memory M that keeps track of the configurations $s_0, s_1, \dots$ at the division events of the overall system. The effective dynamics of S are non-Markovian in the sense that

$$p(s_t, s_n, \dots, s_1, s_0) = p(s_t | s_n, \dots, s_1, s_0) \times p(s_n | s_{n-1}, \dots, s_1, s_0) \cdots p(s_1 | s_0) p(s_0), \quad (43)$$

where $t > t_n > \dots > t_1 > t_0$ and the times $t_0, t_1, \dots, t_n \in \mathcal{T}_{\text{div}}^A$, where $\mathcal{T}_{\text{div}}^A$ is the set of division times for the system A . Note that we take time to be continuous, so the system evolves indivisibly between division times. Note, too, that t is not assumed to be a member of the set of division times.

We conjecture that the quantum description of a non-Markovian conditional probability distribution is given by a case of the quantum comb or equivalent process-tensor constructions of [27–29],

$$p(s_t | s_n, s_{n-1}, \dots, s_0) = \text{Tr}[\mathcal{P}_{s_t} \otimes \mathcal{P}_{s_n} \otimes \cdots \otimes \mathcal{P}_{s_0} \mathcal{E}_{t \leftarrow t_n}^{(n)} \times \{\mathcal{K}_n\{\mathcal{P}_{s_n} \otimes \cdots \otimes \mathcal{P}_{s_0}\}\}], \quad (44)$$

where

$$\mathcal{K}_n\{\rho^{(n)}\} = \sum_{s'_0, \dots, s'_n} \mathcal{P}_{s'_n} \otimes \cdots \otimes \mathcal{P}_{s'_0} \rho^{(n)} \mathcal{P}_{s'_n} \otimes \cdots \otimes \mathcal{P}_{s'_0} \quad (45)$$

acts as an $(n+1)$ -time dephasing channel on an $(n+1)$ -time density matrix $\rho^{(n)}$ and the time evolution of S from $t_n \rightarrow t$ is represented by the evolution channel

$$\mathcal{E}_{t \leftarrow t_n}^{(n)}\{\mathcal{P}_{s_n} \otimes \cdots \otimes \mathcal{P}_{s_0}\} = \mathcal{R}_{s_n, \dots, s_0}(t) \otimes \mathcal{P}_{s_n} \otimes \cdots \otimes \mathcal{P}_{s_0}, \quad (46)$$

with

$$\mathcal{P}_s \mathcal{R}_{s_n, \dots, s_0}(t) \mathcal{P}_s = p(s | s_n, \dots, s_0) \mathcal{P}_s. \quad (47)$$

Note that the division times of the whole system A do not reduce to division times for the subsystem S because S alone lacks the Markov-like property at these times. Nevertheless, joint and conditional probability distributions for S can depend on the division times of A in a non-Markovian manner. We thus refer to the division times of the whole system as effective conditioning times when describing subsystem S alone.

Finally, it would be interesting to explore the structure of interference terms such as $\text{qCor}(t \leftarrow t' \leftarrow t_0)$ in (24) that arise from the discrepancy between unistochastic processes and their Markovian approximations. We believe that the non-trivial structure of such matrices may provide a route toward a novel proof of Tsirelson's bound and similar results, bypassing the need for explicit use of the quantum side of the stochastic-quantum correspondence.

ACKNOWLEDGMENTS

M.H. was supported by the UMass Dartmouth Marine and Undersea Technology (MUST) Research Program sponsored by the Office of Naval Research (ONR) under Grant No. N00014-22-1-2012. D.K. was supported by the UMass Dartmouth MUST Research Program sponsored by the ONR under Grant No. N00014-25-1-2315. We are also grateful for support from the American Institute of Mathematics.

DATA AVAILABILITY

There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

-
- [1] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* **23**, 880 (1969).
 - [2] B. Tsirelson, Quantum information processing lecture notes, 2. Quanta in networks (unpublished), <https://www.webcitation.org/5f12WZOMI>.
 - [3] J. S. Bell, On the Einstein Podolsky Rosen paradox, *Phys. Phys. Fiz.* **1**, 195 (1964).
 - [4] J. S. Bell, The theory of local beables, CERN Technical Report No. 2053-CERN, 1975, <https://cds.cern.ch/record/980036>.
 - [5] J. S. Bell, *La Nouvelle Cuisine* (Cambridge University Press, Cambridge, 2004).
 - [6] B. S. Cirel'son, Quantum generalizations of Bell's inequality, *Lett. Math. Phys.* **4**, 93 (1980).
 - [7] G. C. Ghirardi, A. Rimini, and T. Weber, A general argument against superluminal transmission through the quantum mechanical measurement process, *Lett. Nuovo Cimento* **27**, 293 (1980).
 - [8] T. F. Jordan, Quantum correlations do not transmit signals, *Phys. Lett. A* **94**, 264 (1983).
 - [9] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, *Found. Phys.* **24**, 379 (1994).
 - [10] P. Rastall, Locality, Bell's theorem, and quantum mechanics, *Found. Phys.* **15**, 963 (1985).

- [11] A. Cabello, Simple explanation of the quantum violation of a fundamental inequality, *Phys. Rev. Lett.* **110**, 060402 (2013).
- [12] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, *Nature (London)* **461**, 1101 (2009).
- [13] N. Miklin and M. Pawłowski, Information causality without concatenation, *Phys. Rev. Lett.* **126**, 220403 (2021).
- [14] L. Hardy, Quantum theory from five reasonable axioms, [arXiv:quant-ph/0101012](#).
- [15] J. Barrett, Information processing in generalized probabilistic theories, *Phys. Rev. A* **75**, 032304 (2007).
- [16] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Probabilistic theories with purification, *Phys. Rev. A* **81**, 062348 (2010).
- [17] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Informational derivation of quantum theory, *Phys. Rev. A* **84**, 012311 (2011).
- [18] J. A. Barandes, The stochastic-quantum correspondence, *Philos. Phys.* **3**, 6 (2025).
- [19] J. A. Barandes, The stochastic-quantum theorem, [arXiv:2309.03085](#).
- [20] J. A. Barandes, Quantum systems as indivisible stochastic processes, [arXiv:2507.21192](#).
- [21] J. A. Barandes, New prospects for a causally local formulation of quantum theory, [arXiv:2402.16935](#).
- [22] Ll. Masanes, A. Acín, and N. Gisin, General properties of nonsignaling theories, *Phys. Rev. A* **73**, 012112 (2006).
- [23] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Nonlocal correlations as an information-theoretic resource, *Phys. Rev. A* **71**, 022101 (2005).
- [24] A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, *Am. J. Math.* **76**, 620 (1954).
- [25] J. Pearl, *Causality* (Cambridge University Press, Cambridge, 2009).
- [26] H. Buhrman and S. Massar, Causality and Tsirelson's bounds, *Phys. Rev. A*, **72**, 052103 (2005).
- [27] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum circuit architecture, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [28] S. Milz, D. Egloff, P. Taranto, T. Theurer, M. B. Plenio, A. Smirne, and S. F. Huelga, When is a non-Markovian quantum process classical? *Phys. Rev. X* **10**, 041049 (2020).
- [29] S. Milz and K. Modi, Quantum stochastic processes and quantum non-Markovian phenomena, *PRX Quantum* **2**, 030201 (2021).