Quantum Bell nonlocality as a form of entanglement

Eric Chitambar,[1,*] Gilad Gour,[2,3,†] Kuntal Sengupta,[2,‡] and Rana Zibakhsh[3,§]
[1]*Department of Electrical and Computer Engineering, Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*
[2]*Department of Mathematics and Statistics, Institute for Quantum Science and Technology, University of Calgary, AB, Canada T2N 1N4*
[3]*Department of Physics and Astronomy, University of Calgary, AB, Canada T2N 1N4*

(Received 24 February 2021; accepted 14 October 2021; published 15 November 2021)

Bell nonlocality describes a manifestation of quantum mechanics that cannot be explained by any local hidden variable model. Its origin lies in the nature of quantum entanglement, although understanding the precise relationship between nonlocality and entanglement has been a notorious open problem. In this paper, we develop a dynamical framework in which quantum Bell nonlocality emerges as a special form of entanglement and both are unified as resources under local operations and classical communication (LOCC). Our framework is built on the notion of classical and quantum processes, which are defined as channels that map elements between specific intervals in space and time. Entanglement is identified as a process that cannot be generated by LOCC while Bell nonlocality is the subset of these processes that have an *instantaneous* input-to-output delay time. LOCC preprocessing is a natural set of free operations in this theory, thereby providing previous nonlocality activation results a clear resource-theoretic foundation. We provide a systematic method to quantify the Bell nonlocality of a bipartite quantum channel. It is shown that both the relative entropy and the max relative entropy of nonlocality are nonadditive for a family of bipartite classical channels. This family includes the channel obtained when using the singlet state to maximally violate the CHSH inequality. We also find that the regularized relative entropy of Bell nonlocality provides an upper bound on the asymptotic rate of converting (i.e., simulating) many copies of one classical instantaneous resource to another.

DOI: 10.1103/PhysRevA.104.052208

I. INTRODUCTION

Entanglement and Bell nonlocality represent two of the most stunning nonclassical features of quantum mechanics. Both of them refer to a certain type of interdependence between two or more quantum systems that challenges the way we think about physical reality. More practically, entanglement and nonlocality have been recognized as potential resources for enhanced communication and information processing. Yet, from a fundamental perspective, the exact relationship between the two has remained unclear. This work reveals their unified nature and identifies the essential aspect of Bell nonlocality that makes it a resource for certain information tasks.

Entanglement is traditionally understood as the property of nonseparability [1,2]. More precisely, the bipartite state ρ^{AB} of systems A and B is entangled if it cannot be separated into a statistical mixture of product states,

$$\rho^{AB} \neq \sum_{\lambda} p(\lambda)\rho_{\lambda}^{A} \otimes \rho_{\lambda}^{B}. \tag{1}$$

On the other hand, quantum Bell nonlocality is the property that enables certain quantum states to implement a classical channel that does not admit a local hidden variable (LHV) model. These channels are called nonlocal, and they are characterized by transition probabilities $W(a,b|x,y)$ that cannot be separated into a statistical mixture of product channels,

$$W(a,b|x,y) \neq \sum_{\lambda} p(\lambda)W_{\lambda}(a|x)W_{\lambda}(b|y). \tag{2}$$

Here, (a,b) are channel outputs for Alice and Bob, respectively, while (x,y) are their inputs. Building on the pioneering insight of J.S. Bell [3], one can test whether a given channel is compatible with a LHV model by checking whether the probabilities $W(a,b|x,y)$ satisfy a finite family of inequalities, known as Bell inequalities [4]. The well-known Clauser-Horne-Shimony-Holt (CHSH) Inequality is one such inequality [5], and it completely characterizes the set of LHV channels for binary inputs and outputs [6,7]. Bell nonlocality is the underlying resource enabling many quantum information applications such as device-independent cryptography and randomness extraction [4].

The most direct way to generate a nonlocal channel from a quantum state ρ^{AB} is by performing local measurements on it. The channel generated by such a process has transition probabilities determined by Born's rule,

$$W(a,b|x,y) = \mathrm{Tr}\big[\rho^{AB}\big(M_a^x \otimes N_b^y\big)\big], \tag{3}$$

where $\{M_a^x\}_a$ represent measurement operators on Alice's system (A) for measurement choice x and outcome a, and similarly for the operators $\{N_b^y\}_b$ on Bob's system (B). For example, by choosing suitable pairs of local measurements,

*echitamb@illinois.edu
†giladgour@gmail.ca
‡kuntal.sengupta.in@gmail.com
§rana.zibakhshshabgah@ucalgary.ca

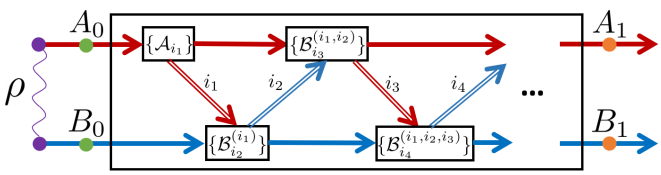

FIG. 1. Local operations and classical communications (LOCC) describes a class of distributed quantum information processing protocols in which only classical information is exchanged between spatially separated parties. Each "round" of the protocol consists one party, say Alice, performing a local quantum operation $\{\mathcal{A}_{i_1}\}$ and sending classical data i_1 to Bob who conditions his next local action on these data [22,23].

a bipartite maximally entangled state $|\phi^+\rangle = \sqrt{1/2}(|00\rangle + |11\rangle)$ can be used to generate a classical channel that violates the CHSH Inequality.

In reflecting on the similarity between Eqs. (1) and (2), prior work has shed some light on the physical relationship between bipartite entanglement and Bell nonlocality [1,2,8–19], although formulating a precise connection has remained a longstanding open problem. An early discovery was the existence of certain entangled states that cannot violate any Bell inequality when generating a classical channel according to Eq. (3) [1,11]. However, it was later shown that *every* bipartite entangled state $\rho^{A_0B_0}$ can violate the CHSH Inequality when combined with another state $\sigma^{A_0'B_0'}$ that itself cannot violate the CHSH Inequality, after the two are allowed to undergo processing by local operations and classical communication (LOCC) prior to receiving any classical inputs for the resultant channel (see Figs. 1 and 2) [14]. That is, a CHSH-violating classical channel can be generated by $\rho^{A_0B_0} \otimes \sigma^{A_0'B_0'}$ when Eq. (3) is modified to have the form

$$W(x_1, y_1|x_0, y_0) = \mathrm{Tr}\big[\mathcal{L}_{\mathrm{pre}}(\rho^{A_0B_0} \otimes \sigma^{A_0'B_0'})\big(M_{x_1}^{x_0} \otimes N_{y_1}^{y_0}\big)\big], \tag{4}$$

where $\mathcal{L}_{\mathrm{pre}}$ is a so-called preprocessing LOCC map that is applied to $\rho^{A_0B_0} \otimes \sigma^{A_0'B_0'}$ before the choice of measurements (x_0, y_0). A similar result also holds for multipartite entangled

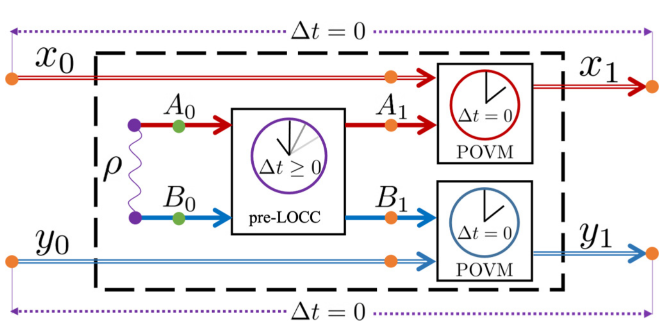

FIG. 2. A state ρ that cannot violate a Bell inequality can be transformed into a state that violates one by performing a pre-LOCC map, like that in Fig. 1. The key property we identify in this work is that *after* the LOCC map, what remains is a bipartite classical channel with essentially zero delay time between the inputs (x_0, y_0) and outputs (x_1, y_1). Dots with the same color correspond to the same time.

states [20], and other types of "nonlocality activation" have been discovered [10,15,21].

Despite the significance of this result, a remaining piece of the puzzle has remained: why is it physically reasonable to introduce pre-LOCC in the study of Bell nonlocality? On the one hand, in a Bell experiment it is crucial that Alice and Bob do not classically communicate after their input choices (x, y), or else they could exploit the so-called locality loophole [24,25] and classically violate any Bell inequality. On the other hand, the pre-LOCC map requires classical communication. Hence the resulting scenario has Alice and Bob freely communicating all the way up to the point of measurement choices, but then all of a sudden this freedom is removed. From a physical perspective, this sudden restriction might seem artificial, and one may wonder whether it is truly fundamental to the nature of nonlocality. Here, we introduce a new way to understand Bell nonlocality that removes any artificial restrictions in the operational model.

Every Bell experiment involves the simulation of a classical channel $W(a, b|x, y)$ that receives its input at some moment and time and yields its outputs at a later time. What is crucial here is not that Alice and Bob are able to generate a channel $W(a, b|x, y)$ capable of violating the CHSH Inequality, but rather that they are able to produce this channel using a local physical process having an input-to-output delay time shorter than the time it takes light to travel between the two laboratories. In contrast to other operational theories of nonlocality [26–32], we explicitly account for the significance of the input-to-output delay time by identifying it as *part of the nonlocality resource itself*.

To make this idea more rigorous, we invoke the machinery of dynamical resource theories (DRTs) [33,34], which have been extensively studied in recent years [35–46]. In fact, the whole field of quantum Shannon theory [47] can be seen as a DRT [48]. In a quantum resource theory, one identifies a restricted subset of quantum operations as being "free," and objects that cannot be generated by these free operations are deemed to possess a resource [33,34]. For example, in the resource theory of entanglement, LOCC represents the free class of operations and all processes that cannot be implemented by LOCC embody the resource of entanglement. Dynamical resource theories generalize static resource theories in that the former consider the resource-theoretic properties of channels whereas the latter study the resources in states. However note that every state can itself be interpreted as a channel with a one-dimensional input.

We then go one step forward and distinguish dynamical resources based on whether their input-to-output delay time is *instantaneous* versus *noninstantaneous*. Under this distinction, quantum Bell nonlocality is an instantaneous classical process that can be obtained by LOCC only when Alice and Bob are initially supplied with an entangled state, as in Fig. 2. One of the key features of our model is that pre-LOCC maps are naturally free operations; i.e., static-to-dynamic conversions having the form of Eq. (4) are allowed in this framework. At the same time, the use of LOCC maps *after* the classical inputs (x_0, y_0) in Eq. (4) is automatically prohibited as it leads to a noninstantaneous resource. Under our framework, Bell nonlocality and entanglement ultimately belong to the same species: resources under LOCC.

This paper carefully unpacks these high-level ideas and then uses them to motivate operationally-meaningful measures of Bell nonlocality. We begin in Sec. II by laying down the basic notation used throughout. In Sec. III, we introduce the abstraction of quantum processes and superprocesses, which provide the primary conceptual pieces to our approach. Section IV establishes our central claim that Bell nonlocality is indeed a special type of entanglement, in a precise resource-theoretic sense. In Sec. V, we identify measures of Bell nonlocality based on channel divergences and classical-to-quantum extensions of entanglement measures. Sections V A and V B show that both the relative entropy and the max relative entropies of Bell nonlocality are nonadditive. Section V C derives a bound on the rate of asymptotically converting instantaneous processes in terms of the regularized relative entropy of Bell nonlocality. Finally, concluding remarks are provided in Sec. VI.

A Comment on LOSR theories

Before we describe our framework in more detail, let us comment on local operations and shared randomness (LOSR), along with the role that it plays in our theory. If one moves beyond the setting of Bell nonlocality and considers semiquantum nonlocality, then a tight connection between entanglement, nonlocality, and LOSR is already known. Semiquantum nonlocality involves replacing the fully classical channel of Eq. (3) with one that receives quantum inputs. Buscemi has elegantly shown that all entangled states provide an advantage in some semiquantum nonlocal game, and moreover, one bipartite state can be converted into another by LOSR if and only if the first scores no worse than the second in every semiquantum nonlocal game [49]. However, when considering Bell nonlocality as it is traditionally understood, such a relationship between nonlocal games and LOSR convertibility no longer holds.

Nevertheless, it often claimed that LOSR should still provide the operational foundation to any resource theory of Bell nonlocality [32] (although not always [26]). Here, we take a critical stance on this perspective. Unlike LOCC, LOSR lacks a clear operational interpretation when understood in the context of quantum information protocols. LOSR is typically justified using a common-cause model [30] in which a helper distributes shared randomness to spatially separated parties. This common randomness, however, is useless for carrying out some protocol without the parties first agreeing on some pre-established strategy, an agreement which inevitably will require some communication. For example, the only way that Alice, Bob, and perhaps some other party, can recognize which physical system encodes the shared randomness is by prior interactive communication. A second issue is that LOSR misses certain important aspects of Bell nonlocality such as it being "hidden" in some states; i.e., certain channels generated in Eq. (4) cannot be realized when the pre-LOCC map $\mathcal{L}_{\text{pre}}$ is restricted to LOSR. The framework presented here overcomes these problems since it starts with the physically motivated class of LOCC and then arrives at a restricted subset of operations, which includes pre-LOCC, by incorporating the practically-relevant property of input-to-output delay time. LOSR still plays an important role in this theory; however it emerges not by appealing to some common cause, but rather by considering bipartite LOCC channels that are implemented with zero input-to-output delay time.

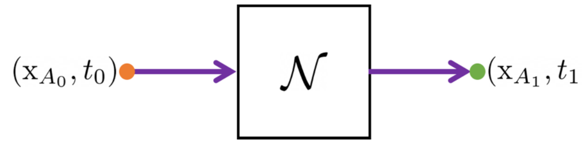

FIG. 3. A quantum process $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$ takes an input ρ at arbitrary space-time point $(\mathbf{x}_{A_0}, t_0)$ and generates an output $\mathcal{N}(\rho)$ at space-time point $(\mathbf{x}_{A_1} = \mathbf{x}_{A_0} + \Delta\mathbf{x}, t_1 = t_0 + \Delta t)$. The input-to-output delay time of $\mathcal{N}$ in this process is Δt.

II. NOTATION

In this paper, quantum physical systems and their corresponding Hilbert spaces will be denoted by A and B, while classical systems will be denoted by X and Y. We will make a distinction between *static* versus *dynamical* systems. Static systems will be denoted with subscripts such as A_0, A_1, B_0, B_1. Dynamical systems will be denoted without subscripts; for example, A refers to an input-output system (A_0, A_1), where the zero subscript will always refer to the input subsystem and the subscript one to the output subsystem. Similarly, $B = (B_0, B_1)$, $X = (X_0, X_1)$, etc. One exception of this notation is the auxiliary systems (e.g. environment system, reference systems) E and R which will always correspond to static systems. The notation $A_1B_1 \equiv A_1 \otimes B_1$ will indicate a bipartite (static) system. We write $|A| = |A_0||A_1|$ to denote the dimension of a dynamical system A, where $|A_0|$ and $|A_1|$ are the dimensions of its corresponding input and output subsystems.

The set of all density matrices (i.e., positive semidefinite hermitian matrices with trace one) acting on A_1 will be denoted by $\mathfrak{D}(A_1)$. As customary, we use ρ and σ to represent density matrices and ψ and ϕ to represent pure states. A maximally entangled state in $\mathfrak{D}(A_1B_1)$ is denoted by $\phi_+^{A_1B_1}$.

The set of all completely positive and trace preserving (CPTP) maps is denoted by $\text{CPTP}(A) := \text{CPTP}(A_0 \to A_1)$. Similarly, we will use the notation $\text{CPTP}(AB)$ in short for $\text{CPTP}(A_0B_0 \to A_1B_1)$. Quantum channels will be denoted with calligraphic letters $\mathcal{M}, \mathcal{N}, \mathcal{E}, \mathcal{F}$. We will use the superscripts $\mathcal{N}^A$ and $\mathcal{N}^{AB}$ to indicate $\mathcal{N} \in \text{CPTP}(A)$ and $\mathcal{N} \in \text{CPTP}(AB)$, respectively. The identity channel in $\text{CPTP}(A_0 \to A_0)$ is denoted by id^{A_0}.

III. QUANTUM PROCESSES

A quantum process, denoted by $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$, is a quantum channel $\mathcal{N} \in \text{CPTP}(A)$ that transforms any state $\rho \in \mathfrak{D}(A_0)$ at arbitrary space-time point $(\mathbf{x}_{A_0}, t_0)$ into state $\mathcal{N}(\rho) \in \mathfrak{D}(A_1)$ at space-time point $(\mathbf{x}_{A_1}, t_1)$, where $\mathbf{x}_{A_1} = \mathbf{x}_{A_0} + \Delta\mathbf{x}$ and $t_1 = t_0 + \Delta t$ (see Fig. 3). Here we assume for simplicity that systems A_0 and A_1 are at rest in the same inertial frame and the coordinates are measured with respect to this frame. The time interval $\Delta t \geqslant 0$ in a quantum process $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$ is called the *input-to-output delay time* of $\mathcal{N}$ (or simply the "delay

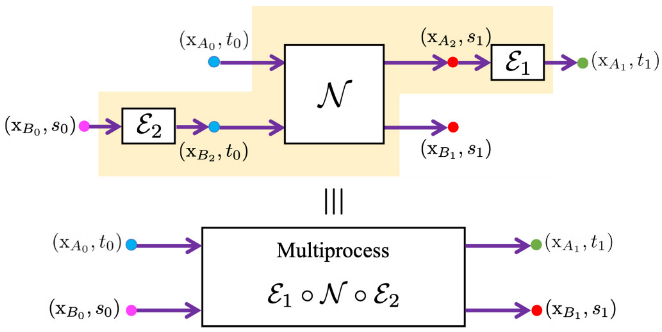

FIG. 4. The bipartite channel $\mathcal{E}_2 \circ \mathcal{N} \circ \mathcal{E}_1$ can have differing delay times for the different subsystems if it is built by composing multiple processes (top). The result is a multiprocess (bottom). Dots with the same color correspond to the same time.

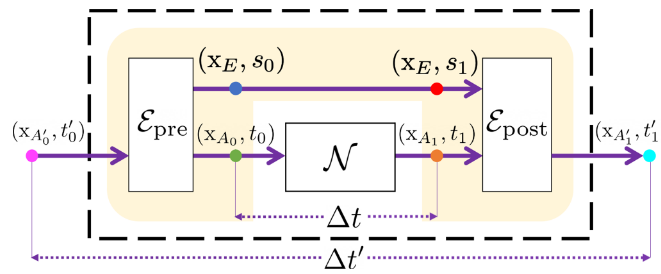

FIG. 5. A quantum superprocess $(\Theta, \Delta\mathbf{x} \to \Delta\mathbf{x}', \Delta t \to \Delta t')$ converts process $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$ to process $(\mathcal{N}', \Delta\mathbf{x}', \Delta t')$. The shaded yellow area represents the action of the superprocess. Since $\mathcal{E}_{\rm pre}$ and $\mathcal{E}_{\rm post}$ might belong to multiprocess, the times s_0 and s_1 are not necessarily equal to t_0 or t_1. For example, in Fig. 6, we will see an example in which $t_0 < t'_0 = s_0 = s_1 = t_1$.

time" of $\mathcal{N}$), and it quantifies how quickly the channel input propagates to the output.

Some quantum processes can be physically implemented and some cannot. For example, the no-signaling principle of special relativity prohibits $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$ from being physically realizable whenever $\mathcal{N}$ is able to transmit information and $\Delta t < \Delta\mathbf{x}/c$, where c is the speed of light. Quantum processes with zero input-to-output delay time are known as *instantaneous*, and they play an important role in this theory. Whenever $\Delta\mathbf{x} > 0$, an instantaneous process $(\mathcal{N}, \Delta\mathbf{x}, 0)$ is physically realizable if and only if $\mathcal{N}$ is a replacement channel, i.e., $\mathcal{N}$ has the form $\mathcal{N}_\rho(X) := \mathrm{Tr}[X]\rho$ for $\rho \in \mathfrak{D}(A_1)$. An instantaneous implementation of $\mathcal{N}_\rho$ is done as follows: knowing that Alice will receive the channel input at time t_0, she simply prepares the state ρ at t_0.

The abstraction of quantum processes also applies to bipartite channels $\mathcal{N} \in \mathrm{CPTP}(AB)$. In this case, however, the input and output of the channel are each distributed across two points in space. Hence, a bipartite process $(\mathcal{N}, \Delta\mathbf{x}_A, \Delta\mathbf{x}_B, \Delta t)$ is the channel $\mathcal{N}$ that transforms a bipartite state $\rho^{A_0B_0}$ held at $(\mathbf{x}_{A_0}, t_0)$ and $(\mathbf{x}_{B_0}, t_0)$ into the state $\mathcal{N}(\rho^{A_0B_0})$, held at $(\mathbf{x}_{A_1}, t_1)$ and $(\mathbf{x}_{B_1}, t_1)$, where $\mathbf{x}_{A_1} = \mathbf{x}_{A_0} + \Delta\mathbf{x}_A$, $\mathbf{x}_{B_1} = \mathbf{x}_{B_0} + \Delta\mathbf{x}_B$, and $t_1 = t_0 + \Delta t$. Note that the number of inputs and outputs can differ by treating one of the systems as trivial. Multipartite processes are defined in the same way: all inputs at spatial coordinates $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ evolve to outputs at spatial coordinates $(\mathbf{x}_1 + \Delta\mathbf{x}_1, \mathbf{x}_2 + \Delta\mathbf{x}_2, \cdots, \mathbf{x}_n + \Delta\mathbf{x}_n)$ in the *same* time interval Δt. By composing processes together, we obtain a multiprocess. The only difference between a process and a multiprocess is that the latter can have differing time delays for different subsystems due to the composition, whereas the former cannot. The general idea is depicted in Fig. 4.

We next go one step forward and define a quantum *superprocess* as a superchannel Θ that transforms one process $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$ into another $(\mathcal{N}', \Delta\mathbf{x}', \Delta t')$. We denote such objects by $(\Theta, \Delta\mathbf{x} \to \Delta\mathbf{x}', \Delta t \to \Delta t')$. Recall that the action of every superchannel $\Theta : \mathrm{CPTP}(A) \to \mathrm{CPTP}(A')$ on a channel $\mathcal{N} \in \mathrm{CPTP}(A)$ can be represented as

$$\Theta^{A\to A'}[\mathcal{N}^A] = \mathcal{E}_{\rm post}^{EA_1\to A'_1} \circ \mathcal{N}^{A_0\to A_1} \circ \mathcal{E}_{\rm pre}^{A'_0\to EA_0}, \quad (5)$$

where $\mathcal{E}_{\rm pre} \in \mathrm{CPTP}(A'_0 \to EA_0)$ and $\mathcal{E}_{\rm post} \in \mathrm{CPTP}(EA_1 \to A'_1)$ are fixed quantum channels corresponding to preprocessing and post-processing of the channel $\mathcal{N}^A$ [50,51]. A superprocess is constructed by invoking Eq. (5) and considering $\mathcal{E}_{\rm pre}$ and $\mathcal{E}_{\rm post}$ as processes, as they would be in any physical implementation of a superchannel. A general superprocess then has the form of Fig. 5. The wires entering and leaving each channel might be multiple systems bundled together.

If the overall process transformation is $(\mathcal{N}, \Delta\mathbf{x}, \Delta t) \to (\mathcal{N}', \Delta\mathbf{x}', \Delta t')$, then there are four possibilities for how $\mathcal{E}_{\rm pre}^{A'_0\to EA_0}$ and $\mathcal{E}_{\rm post}^{EA_1\to A'_1}$ can facilitate this transformation.

(1) The input to $\mathcal{N}$ at the point $(\mathbf{x}_A, t_0)$ depends on the input to $\mathcal{E}_{\rm pre}$ at the point $(\mathbf{x}_{A'_0}, t'_0)$ (and hence $t'_0 \leqslant t_0$) and the output of $\mathcal{E}_{\rm post}$ at $(\mathbf{x}_{A'_1}, t'_1)$ depends on the output of $\mathcal{N}$ at $(\mathbf{x}_{A_1}, t_1)$ (and hence $t_1 \leqslant t'_1$). In this case, the input-to-output delay time is always nondecreasing, i.e., $\Delta t \leqslant \Delta t'$.

(2) The input to $\mathcal{N}$ at the point $(\mathbf{x}_A, t_0)$ does not depend on the input to $\mathcal{E}_{\rm pre}$ at the point $(\mathbf{x}_{A'_0}, t'_0)$ but the output of $\mathcal{E}_{\rm post}$ at $(\mathbf{x}_{A'_1}, t'_1)$ depends on the output of $\mathcal{N}$ at $(\mathbf{x}_{A_1}, t_1)$ (and hence $t_1 \leqslant t'_1$). This means that $\mathcal{E}_{\rm pre}^{A'_0\to A_0E}$ in Fig. 5 is a quantum channel whose output at A_0 is fixed and independent on the input of the channel. This is a special case of a semi-causal channel [52] and was proven in [53] to have for all $\omega \in \mathfrak{D}(A'_0)$ the form

$$\mathcal{E}_{\rm pre}^{A'_0\to EA_0}(\omega^{A'_0}) = \mathcal{L}^{A'_0A_2\to E}(\omega^{A'_0} \otimes \rho^{A_0A_2}), \quad (6)$$

where A_2 is some auxiliary system, $\rho \in \mathfrak{D}(A_0A_2)$ is a fixed quantum state, and $\mathcal{L} \in \mathrm{CPTP}(A'_0A_2 \to E)$ is a quantum channel. However, the form above implies that $\mathcal{L}^{A'_0A_2\to E}$ can be "absorbed" into $\mathcal{E}_{\rm post}$ (of Fig. 5) so that w.l.o.g. we can replace $\mathcal{E}_{\rm post}^{EA_1\to A'_1} \circ \mathcal{L}^{A'_0A_2\to E}$ with $\mathcal{E}_{\rm post}^{A'_0A_1A_2\to A'_1}$ (see Fig. 6). Hence when t_0 is sufficiently earlier than $t'_0 = t_1$, one can attain $\Delta t = t_1 - t_0 > \Delta t' = t'_1 - t'_0$. Note that the difference $\Delta t'$ is determined entirely by the delay time of $\mathcal{E}_{\rm post}$, whereas Δt is the delay time of the input process $(\mathcal{N}, \Delta\mathbf{x}, \Delta t)$.

(3) The input to $\mathcal{N}$ at the point $(\mathbf{x}_A, t_0)$ depends on the input to $\mathcal{E}_{\rm pre}$ at the point $(\mathbf{x}_{A'_0}, t'_0)$ but the output of $\mathcal{E}_{\rm post}$ at $(\mathbf{x}_{A'_1}, t'_1)$ does not depend on the output of $\mathcal{N}$ at $(\mathbf{x}_{A_1}, t_1)$. The superprocess in this case acts as *replacement* superchannel, always outputting the same channel

$$\Theta[\mathcal{N}] = \mathcal{E}_{\rm post}^{E\to A'_1} \circ \mathcal{E}_{\rm pre}^{A'_0\to E} \quad (7)$$

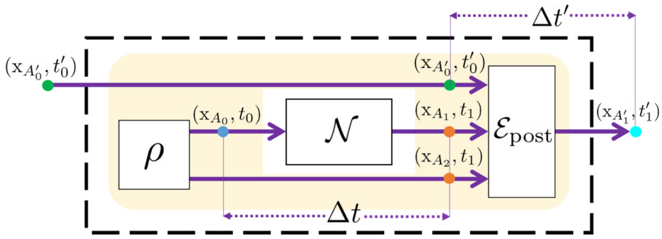

FIG. 6. When the input to $\mathcal{N}$ does not depend on the input to $\mathcal{E}_{\text{pre}}$ it is possible to construct a superprocess $(\Theta, \Delta\mathbf{x} \to \Delta\mathbf{x}', \Delta t \to \Delta t')$ that decrease the input-to-output delay time, $\Delta t' < \Delta t$. Observe in particular that we absorbed $\mathcal{L}^{A_0' \to E}$ that appears in the expression (6) of $\mathcal{E}_{\text{pre}}$ into $\mathcal{E}_{\text{post}}$. In this case the delay time $\Delta t'$ depends only on the implementation of $\mathcal{E}_{\text{post}}$ and in general can be much smaller than Δt. Dots with the same color correspond to the same time.

irrespective of the input channel $\mathcal{N}$. Here, $\mathcal{E}_{\text{pre}}^{A_0' \to E} := \text{Tr}_{A_0} \circ \mathcal{E}_{\text{pre}}^{A_0' \to EA_0}$ and for all $\rho \in \mathfrak{D}(E)$, $\mathcal{E}_{\text{post}}^{E \to A_1'}(\rho^E) := \mathcal{E}_{\text{post}}^{EA_1 \to A_1'}(\rho^E \otimes \omega^{A_1})$ for some arbitrary fixed $\omega \in \mathfrak{D}(A_1)$. Therefore, while this case might result with $\Delta t > \Delta t'$ [1], this is a somewhat trivial case and will not play an important role in our formalism.

(4) The input to $\mathcal{N}$ at the point $(\mathbf{x}_A, t_0)$ does not depend on the input to $\mathcal{E}_{\text{pre}}$ at the point $(\mathbf{x}_{A_0'}, t_0')$ and the output of $\mathcal{E}_{\text{post}}$ at $(\mathbf{x}_{A_1'}, t_1')$ does not depend on the output of $\mathcal{N}$ at $(\mathbf{x}_{A_1}, t_1)$. As in the previous case, also here the superprocess acts as the *replacement* superchannel that always output the channel given in (7).

The most crucial aspect of this analysis is that the superprocess depicted in Fig. 6 can produce processes with $\Delta t' = 0$. This happens when $\mathcal{E}_{\text{post}}$ is physically implementable with zero delay time, or more generally, if $\mathcal{E}_{\text{post}}$ has the form (see Fig. 7)

$$\mathcal{E}_{\text{post}}^{A_0' A_1 A_2 \to A_1'} = \mathcal{F}_2^{A_0' R \to A_1'} \circ \mathcal{F}_1^{A_1 A_2 \to R} \tag{8}$$

where $\mathcal{F}_1^{A_1 A_2 \to R}$ corresponds to a quantum process with arbitrary delay time, whereas $\mathcal{F}_2^{A_0' R \to A_1'}$ corresponds to an instantaneous quantum process. Since $\mathcal{F}_2$ is instantaneous,

[1]For example, one could remove $\mathcal{E}_{\text{post}}$ from Fig. 5, discard the output to $\mathcal{N}$ at $(\mathbf{x}_{A_1}, t_1)$, and extend $(\mathbf{x}_E, t_0)$ with a straight line to the output $(\mathbf{x}_{A_1'}, t_1')$ so that $t_0 = t_1'$.

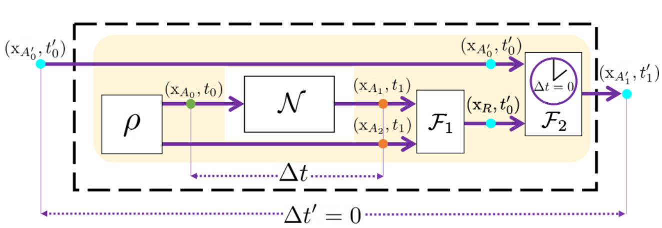

FIG. 7. The most general superprocess that can convert a quantum process $(\mathcal{N}^A, \Delta\mathbf{x}, \Delta t)$ with $\Delta t > 0$ into an instantaneous quantum process. Irrespective of the time delays of $\mathcal{N}$ and $\mathcal{F}_1$, since $\mathcal{F}_2$ is instantaneous we have $t_1' = t_0' \geqslant t_1 > t_0$. In particular, $\Delta t' = t_1' - t_0' = 0$.

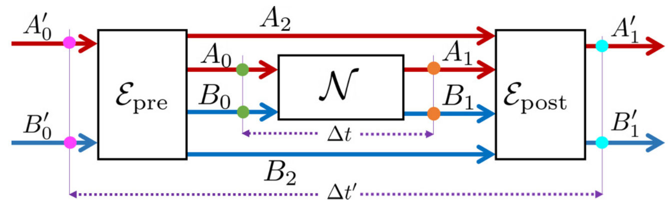

FIG. 8. A bipartite superprocess $(\Theta, \Delta t \to \Delta t')$ converts process $(\mathcal{N}, \Delta t)$ into process $(\Theta[\mathcal{N}], \Delta t)$, where Θ is a superchannel that transforms $\mathcal{N}$ using pre- and postprocessing maps $\mathcal{E}_{\text{pre}}$ and $\mathcal{E}_{\text{post}}$. The processes implementing $\mathcal{E}_{\text{pre}}$ and $\mathcal{E}_{\text{post}}$ must themselves have delay times consistent with the overall change $\Delta t \to \Delta t'$.

$\Delta t' = 0$ irrespective of the delay times associated with $\mathcal{N}$ and $\mathcal{F}_1$. In Fig. 7, we depict the most general superprocess that converts a noninstantaneous quantum process into an instantaneous one.

To isolate the essential features of this theory, going forward we will assume that the spatial intervals $\Delta\mathbf{x}$ and $\Delta\mathbf{x}'$ are given and fixed. Then, the only processes we consider are specified by the channel and its delay time, $(\mathcal{N}, \Delta t)$. The relevant superprocesses in this case are characterized by a superchannel and its induced change in delay times, $(\Theta, \Delta t \to \Delta t')$.

IV. ENTANGLEMENT THEORY WITH INSTANTANEOUS RESOURCES, ALIAS BELL NONLOCALITY

Having introduced the underlying concepts of this work, we now restrict attention to processes that can be implemented by LOCC (as in Fig. 1). We will make the simplifying assumption that all local operations are instantaneous, a reasonable assumption to make when Alice and Bob's laboratories are separated relatively far apart. Consequently, the delay time Δt of an LOCC process will always be proportional to the number of communication exchanges conducted in the particular implementation of the channel. Since there is always some nonzero spatial separation between Alice and Bob, every communication protocol will have $\Delta t > 0$.

LOCC superprocesses constitute the free operations in this resource theory, and these are superprocesses $(\Theta, \Delta t \to \Delta t')$ having the form of Fig. 8 with $\mathcal{E}_{\text{pre}}$ and $\mathcal{E}_{\text{post}}$ being LOCC maps. The fundamental question is whether one quantum process can be converted to another using an LOCC superprocess. When $(\mathcal{N}, \Delta t) \to (\mathcal{N}', \Delta t')$ is achievable by LOCC, we write

$$(\mathcal{N}, \Delta t) \xrightarrow{\text{LOCC}} (\mathcal{N}', \Delta t'). \tag{9}$$

The free objects in this resource theory are the quantum processes that can be generated by LOCC "from scratch" (i.e., from the trivial process),

$$(\mathsf{id}^1, 0) \xrightarrow{\text{LOCC}} (\mathcal{N}, \Delta t), \tag{10}$$

where id^1 represents here the trivial state/channel; i.e., the only element of $\mathfrak{D}(\mathbb{C}) \cong \text{CPTP}(\mathbb{C} \to \mathbb{C})$. Notice that every LOCC channel $\mathcal{L}$ belongs to some free process $(\mathcal{L}, \Delta t)$ with $\Delta t \in [0, +\infty]$. As a special case, a channel $\mathcal{L}$ belongs to an instantaneous LOCC processes $(\mathcal{L}, 0)$ if and only if it can be implemented by local operations and shared randomness

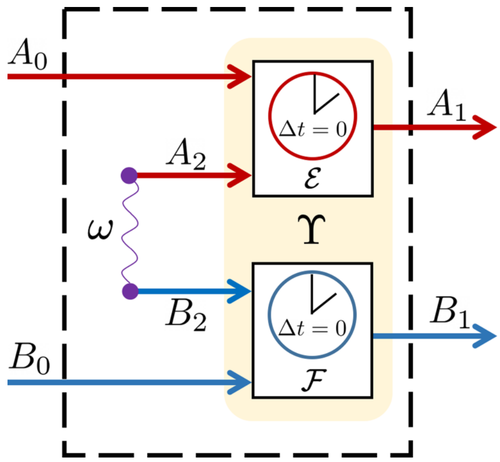

FIG. 9. Implementation of an LOSE quantum process with an entangled state $\omega \in \mathfrak{D}(A_2B_2)$ and local superprocess Υ that takes the quantum state $\omega^{A_2B_2}$ as its input and then outputs the bipartite quantum channel $\mathcal{N}^{AB} := \Upsilon[\omega^{A_2B_2}]$.

(LOSR). This means that

$$\mathcal{L} = \sum_{\lambda} p(\lambda)\mathcal{E}_{\lambda}^{A} \otimes \mathcal{F}_{\lambda}^{B} \tag{11}$$

with $p(\lambda)$ forming a probability distribution, $\mathcal{E}_{\lambda}^{A} \in \text{CPTP}(A)$ and $\mathcal{E}_{\lambda}^{B} \in \text{CPTP}(B)$, and the collection of such maps corresponds precisely to the family of instantaneous LOCC processes.

There are non-LOCC bipartite channels $\mathcal{N}$ whose instantaneous process $(\mathcal{N}, 0)$ can also be physically implemented. These are channels that belong to the class of local operations and shared entanglement (LOSE). As depicted in Fig. 9, we say that $\mathcal{N} \in \text{LOSE}(AB)$ if there exists some entangled state $\omega \in \mathfrak{D}(A_2B_2)$ such that

$$\mathcal{N}^{AB} = \Upsilon[\omega^{A_2B_2}], \tag{12}$$

where Υ is an LOSR superchannel comprising of the two local channels $\mathcal{E} \in \text{CPTP}(A_0A_2 \to A_1)$ and $\mathcal{F} \in \text{CPTP}(B_0B_2 \to B_1)$ such that for any $\rho \in \mathfrak{D}(A_0B_0)$

$$\Upsilon[\omega^{A_2B_2}](\rho^{A_0B_0}) := \mathcal{E}^{A_0A_2\to A_1} \otimes \mathcal{F}^{B_0B_2\to B_1}(\rho^{A_0B_0} \otimes \omega^{A_2B_2}). \tag{13}$$

Every bipartite quantum state $\rho \in \mathfrak{D}(A_1B_1)$ can be viewed as a bipartite quantum process $(\rho, 0)$ with a one-dimensional input and zero input-to-output delay time since if Alice and Bob hold $\rho^{A_1B_1}$ and they know they will receive an input at time t_0, they can immediately output $\rho^{A_1B_1}$ at time t_0. In this way, our framework captures all of state-based entanglement theory. We emphasize that even though states have an instantaneous input-to-output delay time, this does not mean that every bipartite state $\rho^{A_1B_1}$ can be created instantaneously. The LOCC preparation of a state is described by the superprocess

$$(\mathsf{id}^1, 0) \xrightarrow{\text{LOCC}} (\rho, 0), \tag{14}$$

which is possible if and only if ρ is separable. In fact, even if ρ is separable but not a product state, the LOCC superprocess carrying out Eq. (14) cannot be implemented instantaneously

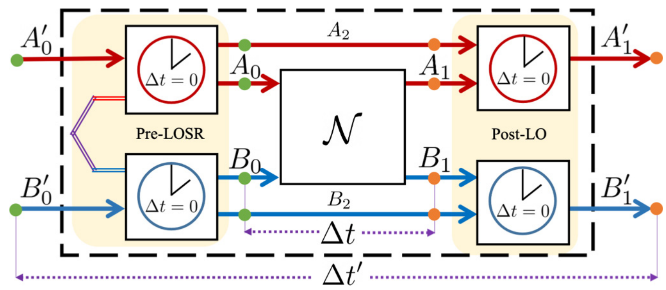

FIG. 10. An LOSR superprocess with $\Delta t = \Delta t'$.

since shared randomness must be distributed from one party to another, and this will require time. In general, bipartite processes with instantaneous input-to-output delay time will require nonzero time to build from scratch, if it is even possible to do so by LOCC.

The fundamental result of this approach is that Bell nonlocality can now be understood as a special type of entanglement. Every experimental test for Bell nonlocality involves the conversion of an entangled bipartite state $\rho^{A_1B_1}$ into a classical channel $\mathcal{N}^{XY}$, like in Eq. (3) or (4). This is sometimes described as transforming a static resource $(\rho^{A_1B_1})$ into a dynamical classical resource $(\mathcal{N}^{XY})$. However, this description does not make explicit the resource-theoretic aspect of nonlocality. While every classical channel can be implemented by LOCC, not every *instantaneous* classical process admits such an implementation. Hence, in the resource theory developed here, $(\mathcal{N}^{XY}, \Delta t)$ is a free object for $\Delta t > 0$, but $(\mathcal{N}^{XY}, 0)$ is a resource if $\mathcal{N}^{XY}$ is non-LOSR, i.e., not admitting a LHV model. Entanglement, whether it be static or dynamic, can be defined as any process $(\mathcal{N}^{AB}, \Delta t)$ that cannot be generated by LOCC,

$$\text{entanglement:} \quad (\text{id}^1, 0) \overset{\text{LOCC}}{\not\longrightarrow} (\mathcal{N}^{AB}, \Delta t), \tag{15}$$

and it is the resource in this theory. Bell nonlocality is simply the restriction of these superprocesses to those that output classical channels,

$$\text{Bell nonlocality:} \quad (\text{id}^1, 0) \overset{\text{LOCC}}{\not\longrightarrow} (\mathcal{N}^{XY}, 0). \tag{16}$$

When referring to *quantum* Bell nonlocality, we mean some Bell nonlocal process that can be obtained from a quantum process using LOCC,

$$(\mathcal{N}^{AB}, \Delta t) \xrightarrow{\text{LOCC}} (\mathcal{N}^{XY}, 0) \tag{17}$$

$$\text{but} \qquad (\text{id}^1, 0) \overset{\text{LOCC}}{\not\longrightarrow} (\mathcal{N}^{XY}, 0). \tag{18}$$

The following theorem characterizes the structure of such superprocesses.

Theorem 1. Let $(\Theta^{AB\to A'B'}, \Delta t \to \Delta t')$ be an LOCC superprocess.

(1) If $\Delta t = \Delta t' = 0$, then the superprocess has the form of either Figs. 10 or 11.

(2) If $\Delta t > \Delta t' = 0$, then the superprocess has the form of Fig. 11.

Conversely, if Θ has the form of either Figs. 10 or 11, then it is instantaneous-preserving; i.e., $\Delta t = 0 \Rightarrow \Delta t' = 0$.

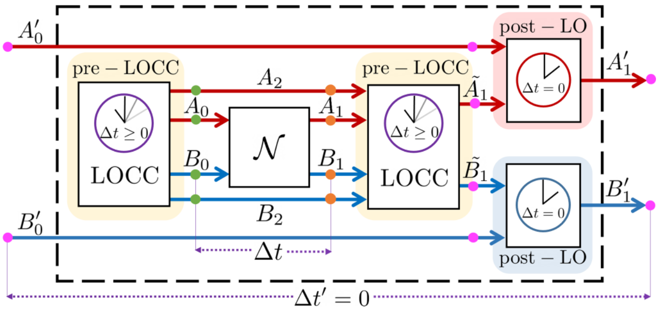

FIG. 11. An LOCC superprocess containing preprocessing LOCC simulating an instantaneous dynamical resource.

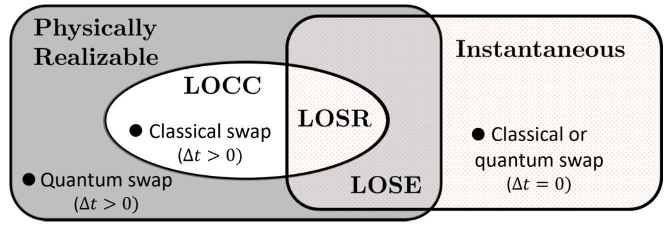

FIG. 12. A comparison between the different classes of processes relevant to the study of quantum entanglement and Bell nonlocality. All processes belonging to the physically-realizable region outside LOCC possess entanglement. The instantaneous classical LOSE channels lying in the intersecting region are the Bell nonlocal processes.

Proof. For the first part, assume that $\Delta t' = \Delta t = 0$. A general bipartite superprocess is depicted in Fig. 8. Recall the four cases that we considered in Sec. III on how $\mathcal{E}_{\rm pre}$ and $\mathcal{E}_{\rm post}$ can facilitate a superprocess. In the first of these cases, we have that $t_0' \leqslant t_0$ and $t_1 \leqslant t_1'$. When $\Delta t' = \Delta t = 0$ we must have that $t_0 = t_1 = t_0' = t_1'$ so that $\mathcal{E}_{\rm pre}$, and $\mathcal{E}_{\rm post}$ belongs to an instantaneous quantum process. Since we also have that Θ is LOCC, $\mathcal{E}_{\rm pre}$ and $\mathcal{E}_{\rm post}$ must be both LOCC *and* correspond to instantaneous quantum processes; i.e., both $\mathcal{E}_{\rm pre}$ and $\mathcal{E}_{\rm post}$ are LOSR which is precisely the form given in Fig. 10.

In the second case (out of the four cases) the superprocess has the form of Fig. 6 adjusted to the bipartite case. Now, since $\Delta t' = 0$ the superprocess has the form of Fig. 7 (again adjusted to the bipartite case). Note that since Θ is LOCC, ρ, $\mathcal{F}_1$, and $\mathcal{F}_2$ of Fig. 7 all must be LOCC since we want the superprocess to be both LOCC *and* instantaneous preserving. Hence, ρ is is precisely the LOCC process/state that appears on the left side of Fig. 11, $\mathcal{F}_1$ is the LOCC process that appears at the output of $\mathcal{N}^{AB}$ in Fig. 11, and since also $\mathcal{F}_2$ corresponds to an instantaneous quantum process it must be an LOSR channel corresponding to the two post-LO channels that appear on the right side on Fig. 11. Note that the shared randomness of $\mathcal{F}_2$ can be absorbed into the pre-LOCC channel preceding it. Since the third and fourth cases do not lead to new implementations of the superprocess, the proof is concluded for the case $\Delta t' = \Delta t = 0$.

Consider now the second part of the theorem in which $\Delta t > \Delta t' = 0$. The only implementation (out of the four) of the superprocess for this case is given in Fig. 7 adjusted to the bipartite case. As discussed in the first part of the proof, an LOCC superprocess of this form is given in Fig. 11.

The converse statement of theorem 1 follows by inspection. This completes the proof. ■

One of the most important consequences of theorem 1 is that it provides a physical justification for allowing pre-LOCC processes in experiments of Bell nonlocality. If the channel $\mathcal{N}$ is replaced by a quantum state $\rho^{A_1B_1}$, then Fig. 11 has precisely the same form as Fig. 2. Thus hidden Bell nonlocality and superactivation of nonlocality are both features that are operationally accessible in this resource theory. More precisely, channels having the form of Eq. (4) reflect the resource conversion

$$(\rho^{A_1B_1} \otimes \sigma^{A_1'B_1'}, 0) \xrightarrow{\rm LOCC} (\mathcal{N}^{XY}, 0), \qquad (19)$$

and so the result of [14] implies that every entangled state can activate some Bell nonlocal resource using the free operations of this theory.

Furthermore, all so-called "anomalies" of bipartite entanglement [13] vanish under this resource theory. These refer to certain nonlocal effects (such as how large Bell inequalities can be violated or how robust violations are to noise) that are more prominent in partially entangled states than maximally entangled ones. However, every partially entangled bipartite state can be obtained from a maximally entangled one under LOCC. Hence, in this resource theory whatever process can be freely generated by a weakly entangled state can also be generated by a maximally entangled one. This allows the partial order of states defined by any measure of Bell nonlocality to reflect the partial order defined by LOCC convertibility.

Let us further remark that theorem 1 unifies previous resource-theoretic formulations of Bell nonlocality. Figure 10 describes the transformation of channels by LOSR superchannels. This model has been studied in the abstract setting of nonlocal "boxes" that are governed by a common cause [30–32]. An alternative operational model has been proposed in Ref. [26] which is known as wiring and prior-to-input classical communication (WPICC) (see also [29]). In the context of quantum resources, WPICC amounts to applying a preprocessing map implementable by LOCC, as in Fig. 11. Our resource theory introduces the notion of input-to-output delay time as a way to physically motivate both of these types of channel transformations in a unified way.

We close this section with Fig. 12, which provides an overview of the different type of processes discussed here and their relationships. Everything within the LOCC oval is free, whereas processes outside LOCC are not free and possess the resource of entanglement. It is important to stress that this resource theory and the categories in Fig. 12 pertain to processes, not channels. To best illustrate this distinction, consider the swap channel, which interchanges the states in Alice and Bob's laboratory. The noninstantaneous classical swap channel can be implemented by LOCC and is therefore a free process in our model, whereas an instantaneous swap is not even physically realizable. On the other hand, the quantum swap channel can be physically implemented only through quantum communication between Alice and Bob's laboratories. Hence, the noninstantaneous quantum swap is a physical resource in this theory. Bell nonlocality refers to the classical processes lying in the LOSE region of Fig. 12.

V. QUANTIFICATION OF BELL NONLOCALITY

We next turn to the question of quantifying Bell nonlocality. As in any resource theory, a valid measure of Bell nonlocality must be monotonically decreasing under the free operations of the theory, and its value must vanish for all free objects. In Ref. [38], the notion of entanglement monotones for channels was introduced as any functional $\mathrm{E} : \cup_{A,B}\mathrm{CPTP}(AB) \to \mathbb{R}_+ \cup \{0\}$ satisfying $\mathrm{E}(\Theta[\mathcal{N}]) \leqslant \mathrm{E}(\mathcal{N})$ for any $\mathcal{N} \in \mathrm{CPTP}(AB)$ and for any $\Theta \in \mathrm{LOCC}(AB \to A'B')$. We can extend this to the level of processes by incorporating the input-to-output delay time associated with each quantum channel.

In the following definition we denote by $\mathfrak{Q}(AB)$ the set of all realizable bipartite quantum processes of the form $(\mathcal{N}^{AB}, \Delta t)$ over all $\Delta t \in [0, \infty]$, where $\mathcal{N} \in \mathrm{CPTP}(AB)$. Similarly, we denote by $\mathfrak{Q}(XY)$ the set of all such classical processes $(\mathcal{N}^{XY}, \Delta t)$.

Definition V..1. The function

$$E : \bigcup_{A,B} \mathfrak{Q}(AB) \to \mathbb{R},$$

where the union is over all finite dynamical systems A and B, is called a *measure of dynamical entanglement* if on the trivial process $(\mathsf{id}^1, \Delta t)$, we have $E(\mathsf{id}^1, \Delta t) = 0$, and for any two bipartite processes $(\mathcal{N}, \Delta t)$ and $(\mathcal{N}', \Delta t')$ such that

$$(\mathcal{N}, \Delta t) \xrightarrow{\text{LOCC}} (\mathcal{N}', \Delta t') \tag{20}$$

we have $E(\mathcal{N}, \Delta t) \geqslant E(\mathcal{N}', \Delta t')$.

The condition that $(\mathsf{id}^1, \Delta t) = 0$ in this definition assures that all free processes $(\mathcal{N}, \Delta t)$ satisfy $E(\mathcal{N}, \Delta t) = 0$. When $\mathcal{N}$ and $\mathcal{N}'$ are bipartite quantum states this definition reduces to a measure of (static) entanglement on bipartite states. When $\mathcal{N}$ and $\mathcal{N}'$ are classical bipartite channels and $\Delta t = \Delta t' = 0$ (i.e., instantaneous processes) the definition above reduces to a measure of Bell nonlocality. This means that any measure of dynamical entanglement, E, reduces to a measure of Bell nonlocality when the domain is restricted to instantaneous classical processes. We denote by $E_{\rm cl}$ the restriction of E to the classical domain, and we call it a classical entanglement measure. Note that if $\Delta t > 0$ (i.e., the process is noninstantaneous) and $\mathcal{N} \in \mathrm{CPTP}(XY)$, then the transformation $(\mathsf{id}^1, \Delta t) \xrightarrow{\text{LOCC}} (\mathcal{N}, \Delta t)$ is always physically realizable and so $E_{\rm cl}(\mathcal{N}^{XY}, \Delta t) = 0$. Hence, in what follows, the only classical processes we will consider are instantaneous because only these can possess Bell nonlocality, and we write $E_{\rm cl}(\mathcal{N}^{XY})$ to denote $E_{\rm cl}(\mathcal{N}^{XY}, 0)$. Similarly, we let $\mathrm{LOCC}_0(XY \to X'Y')$ indicate the set of all classical superprocesses having the form of either Fig. 10 or 11. The converse of theorem 1 assures that superprocesses from $\mathrm{LOCC}_0(XY \to X'Y')$ leave the collection of instantaneous classical processes invariant.

Monotones of Bell nonlocality for classical channels have been previously explored in which the domain of the functionals is restricted to nonsignaling bipartite classical channels [30]. As depicted in Fig. 12, nonsignaling processes can either be (i) noninstantaneous (and so free), (ii) instantaneous LOSR (and so free), (iii) instantaneous LOSE but not LOSR (and so physically realizable and not free), or (iv) instantaneous but not LOSE (and so not physically realizable). The measures we present below can quantify the resource content in processes of both type (iii) and (iv), similar to the results of Ref. [30]. In fact, since our resource measures are based on different channel "distances" from the set of LOSR, the measures can also be applied to instantaneous signaling channels, despite them being nonphysical like those of type (iv).

In the next three sections, we define and study the classical relative, max relative, and regularized relative entropies of nonlocality. While these are defined on classical systems, in Sec. V D, we extend them to the domain of quantum channels using the extension techniques developed in Ref. [54].

A. Relative entropy of Bell nonlocality

We begin by introducing a classical Bell nonlocality measure that is based on channel divergences [45,46,51,55–58]. Typically, in resource theories such monotones are constructed as the "distance" (as measured by the divergence) of the resource to the set of free objects. Since we are restricting to instantaneous processes in the classical case, the free objects can be identified with the set of LOSR channels.

Let D be the Kullback-Leibler divergence (relative entropy) defined on any two n-dimensional probability vectors $\mathbf{p}$ and $\mathbf{q}$ as

$$D(\mathbf{p}\|\mathbf{q}) := \sum_{x=1}^{n} p(x)[\log_2 p(x) - \log_2 q(x)]. \tag{21}$$

Its extension to classical channels is defined as [58]

$$D(\mathcal{N}\|\mathcal{M}) := \max_{x\in[|X_0|]} D(\mathcal{N}(|x\rangle\langle x|^{X_0})\|\mathcal{M}(|x\rangle\langle x|^{X_0})),$$

for all classical channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(X)$. The above function is called a channel divergence since it satisfies the data processing inequality; i.e., for all $\Theta \in \mathrm{SC}(X \to Y)$, we have $D(\Theta[\mathcal{N}]\|\Theta[\mathcal{M}]) \leqslant D(\mathcal{N}\|\mathcal{M})$. In [58] it was shown that the above classical channel relative entropy is the only channel relative entropy that reduces to the Kullback-Leibler divergence on classical states (i.e., when $|X_0| = 1$). We define the relative entropy of Bell nonlocality as

$$E_r(\mathcal{N}^{XY}) := \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} D(\mathcal{N}^{XY}\|\mathcal{M}^{XY}). \tag{22}$$

This function provides a valid measure of dynamical Bell nonlocality for instantaneous classical processes. In particular, since the channel divergence D satisfies the data-processing inequality, it follows that for any $\Theta \in \mathrm{LOCC}_0(XY \to X'Y')$

$$E_r(\Theta[\mathcal{N}^{XY}]) \leqslant E_r(\mathcal{N}^{XY}). \tag{23}$$

An appealing property of a resource measure is additivity. This says that the total amount of resource in independent states accumulates in an additive way:

$$E(\rho \otimes \sigma) = E(\rho) + E(\sigma). \tag{24}$$

In other words, the whole is equal to the sum of its parts. However, it turns out that this adage is generally not true for entanglement, and most known measures of static entanglement are nonadditive. Here we show that the same holds for instantaneous classical channels. For $|X_0| = |Y_0| = |X_1| = |Y_1| = 2$, consider the one-parameter family of channels $\mathcal{W}_\lambda^{X_0Y_0\to X_1Y_1}$ having the form

$$\mathcal{W}_\lambda \sim \frac{1}{4}\times \begin{array}{|c|c|c|}\hline x_0/y_0 & 0 & 1 \\ \hline 0 & \begin{matrix}1+\lambda & 1-\lambda \\ 1-\lambda & 1+\lambda\end{matrix} & \begin{matrix}1+\lambda & 1-\lambda \\ 1-\lambda & 1+\lambda\end{matrix} \\ \hline 1 & \begin{matrix}1+\lambda & 1-\lambda \\ 1-\lambda & 1+\lambda\end{matrix} & \begin{matrix}1-\lambda & 1+\lambda \\ 1+\lambda & 1-\lambda\end{matrix} \\ \hline \end{array}$$

with $0 \leqslant \lambda \leqslant 1$, in which each box encodes the channel probabilities

$$\begin{array}{|c|c|}\hline & y_0 \\ \hline x_0 & \begin{matrix} p(0,0|x_0,y_0) & p(0,1|x_0,y_0) \\ p(1,0|x_0,y_0) & p(1,1|x_0,y_0)\end{matrix} \\ \hline \end{array}$$

for $\mathcal{W}_\lambda$. Note that when $\lambda = \frac{1}{\sqrt{2}}$ the channel $\mathcal{W}_\lambda =: \mathcal{W}^{\text{CHSH}}$ generates the maximal quantum violation of the CHSH inequality, while when $\lambda = 1$ the channel $\mathcal{W}_\lambda =: \mathcal{W}^{\text{PR}}$ corresponds to a PR-box [59].

We will show that these channels have a nonadditive relative entropy under tensor product. One might already expect nonadditivity of the channel $\mathcal{W}^{\text{CHSH}}$ due to nonparallelization of the optimal classical strategy in XOR games [60]. However, the connection between an optimal strategy in nonlocal games and an optimal LOSR map in the definition of E_r is not direct. Nevertheless, the general intuition of nonadditivity remains correct. Our approach resembles that taken by Vollebrecht and Werner who showed that the relative entropy of static entanglement is nonadditive [61]. Namely, we will exploit the high degree of symmetry that the channels $\mathcal{W}_\lambda$ enjoy. Each $\mathcal{W}_\lambda$ is invariant under the following operations: (i) applying a pre- and post- swap gate between Alice and Bob's systems; (ii) flipping Alice and Bob's output bits; (iii) flipping Alice's output conditioned on her input being 1, while flipping Bob's input bit; and (iv) flipping Alice's output conditioned on her input being 0, while flipping Bob's output conditioned on his input being 1, and then flipping the input bits of both parties.

Operations (ii)–(iv) belong to LOSR, and while (i) does not, it still leaves the collection of local channels invariant. In fact, it is not difficult to verify that up to a relabeling of outputs, *any* channel in CPTP(XY) gets projected down into some channel $\mathcal{W}_\lambda$ when averaging over the action of the group generated by these four transformations. Note that a relabeling of, say, Alice's outputs simply cause the transformation $\lambda \to -\lambda$ in the channel probabilities of $\mathcal{W}_\lambda$.

Proposition 1. An arbitrary channel $\mathcal{N} \in$ CPTP(XY) undergoes the transformation $\mathcal{N} \to \mathcal{W}_\lambda$, with

$$\begin{aligned}\lambda = \tfrac{1}{4}(&p(0,0|0,0) + p(1,1|0,0) + p(0,0|0,1) \\ &+ p(1,1|0,1) + p(0,0|1,0) + p(1,1|1,0) + p(0,1|1,1) \\ &+ p(1,0|1,1)),\end{aligned}$$

when averaging over operations (i)—(iv) and all their compositions. This transformation leaves the set LOSR(XY) invariant.

We now compute the relative entropy of nonlocality for the channels $\mathcal{W}_\lambda$. By proposition 1 and convexity of the mapping

$$\mathcal{L}(x_0,y_0) \mapsto D(\mathcal{W}(x_0,y_0)\|\mathcal{L}(x_0,y_0))$$

for any choice of input (x_0, y_0), without loss of generality we can restrict attention to local channels $\mathcal{L}_\lambda$ that have the same symmetries as $\mathcal{W}_\lambda$. The following lemma characterizes the locality conditions.

Lemma 1. $\mathcal{L}_\lambda \in$ LOSR(XY) if and only if $\lambda \leqslant 1/2$.

Proof. It is well-known that membership of LOSR(XY) is decided entirely by whether or not the channel satisfies the CHSH inequality

$$\begin{aligned}-1 \leqslant\ & p(0,0|0,0) - p(0,1|0,1) \\ & - p(1,0|1,0) - p(0,0|1,1) \leqslant 0 \end{aligned} \tag{25}$$

and all equivalent inequalities obtained by input/output relabeling [7]. An exhaustive search finds that for $\mathcal{L}_\lambda$ this reduces to the constraint $\lambda \leqslant 1/2$. ■

Using lemma 1, we can then phrase the relative entropy as a simple optimization problem:

$$E_r(\mathcal{W}_\lambda) = \min_{|\mu|\leqslant 1/2} \frac{1+\lambda}{2}\log_2\frac{1+\lambda}{1+\mu} + \frac{1-\lambda}{2}\log_2\frac{1-\lambda}{1-\mu}.$$

This can be readily solved to yield the formula

$$E_r(\mathcal{W}_\lambda) = \frac{1+\lambda}{2}\log_2\frac{1+\lambda}{3/2} + \frac{1-\lambda}{2}\log_2\frac{1-\lambda}{1/2} \tag{26}$$

when $1/2 \leqslant \lambda \leqslant 1$, and $E_r(\mathcal{W}_\lambda) = 0$ when $0 \leqslant \lambda \leqslant 1/2$. In the former case, the optimal LOSR channel is given by

$$\mathcal{L}_{1/2} \sim \begin{array}{|c|c|c|}\hline x_0/y_0 & 0 & 1 \\ \hline 0 & \begin{matrix}3/8 & 1/8 \\ 1/8 & 3/8\end{matrix} & \begin{matrix}3/8 & 1/8 \\ 1/8 & 3/8\end{matrix} \\ \hline 1 & \begin{matrix}3/8 & 1/8 \\ 1/8 & 3/8\end{matrix} & \begin{matrix}1/8 & 3/8 \\ 3/8 & 1/8\end{matrix} \\ \hline \end{array},$$

and the specific channels of interest have values

$$E_r(\mathcal{W}^{\text{CHSH}}) \approx 0.046, \tag{27}$$

$$E_r(\mathcal{W}^{\text{PR}}) \approx 0.415. \tag{28}$$

Next, we consider two copies of the channel. Our main result is the following.

Theorem 2. The quantity E_r is nonadditive for all nonlocal channels $\mathcal{W}_\lambda$, i.e., whenever $\lambda > 1/2$.

Proof. Consider the local channel $\mathcal{L} \in$ LOSR$(X_0X_0'Y_0Y_0' \to X_1X_1'Y_1Y_1')$ defined by the functions

$$(x_1, x_1') = \begin{cases}(0,0) & \text{if } x_0x_0' = 0 \\ (1,0) & \text{if } x_0x_0' = 1\end{cases},$$

$$(y_1, y_1') = \begin{cases}(0,0) & \text{if } y_0y_0' = 0 \\ (0,1) & \text{if } y_0y_0' = 1\end{cases}. \tag{29}$$

This channel can also be depicted by the 16×16 grid

x_0/y_0	00	01	10	11
00	1 0 0 0	1 0 0 0	1 0 0 0	0 1 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
01	1 0 0 0	1 0 0 0	1 0 0 0	0 1 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
10	1 0 0 0	1 0 0 0	1 0 0 0	0 1 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
11	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0
	1 0 0 0	1 0 0 0	1 0 0 0	0 1 0 0
	0 0 0 0	0 0 0 0	0 0 0 0	0 0 0 0

in which each subblock encodes the transition probabilities $p(x_1, x_1'; y_1, y_1'|x_0, x_0'; y_0, y_0')$. We now symmetrize this channel by averaging over transformation (ii)—(iv) applied independently to each input/output group (x_0, y_0, x_1, y_1) and (x_0', y_0', x_1', y_1'), as well as an average over swapping of parties. This generates a total of $2^7 = 128$ transformations and the resulting channel is given by

$\overline{\mathcal{L}} \sim \frac{1}{32} \times$

x_0/y_0	00	01	10	11
00	5 1 1 1	5 1 1 1	5 1 1 1	5 1 1 1
	1 5 1 1	1 5 1 1	1 5 1 1	1 5 1 1
	1 1 5 1	1 1 5 1	1 1 5 1	1 1 5 1
	1 1 1 5	1 1 1 5	1 1 1 5	1 1 1 5
01	5 1 1 1	1 5 1 1	5 1 1 1	1 5 1 1
	1 5 1 1	5 1 1 1	1 5 1 1	5 1 1 1
	1 1 5 1	1 1 1 5	1 1 5 1	1 1 1 5
	1 1 1 5	1 1 5 1	1 1 1 5	1 1 5 1
10	5 1 1 1	5 1 1 1	1 1 5 1	1 1 5 1
	1 5 1 1	1 5 1 1	1 1 1 5	1 1 1 5
	1 1 5 1	1 1 5 1	5 1 1 1	5 1 1 1
	1 1 1 5	1 1 1 5	1 5 1 1	1 5 1 1
11	5 1 1 1	1 5 1 1	1 1 5 1	1 1 1 5
	1 5 1 1	5 1 1 1	1 1 1 5	1 1 5 1
	1 1 5 1	1 1 1 5	5 1 1 1	1 5 1 1
	1 1 1 5	1 1 5 1	1 5 1 1	5 1 1 1

Using this channel, we have

$$E_r(\mathcal{W}_\lambda^{\otimes 2}) \leqslant D(\mathcal{W}_\lambda^{\otimes 2} \| \overline{\mathcal{L}}) = \frac{(1+\lambda)^2}{4} \log_2 \frac{(1+\lambda)^2}{5} + \frac{1-\lambda^2}{2} \log_2 (1-\lambda^2) + \frac{(1-\lambda)^2}{4} \log_2 2(1-\lambda)^2 + 1. \quad (30)$$

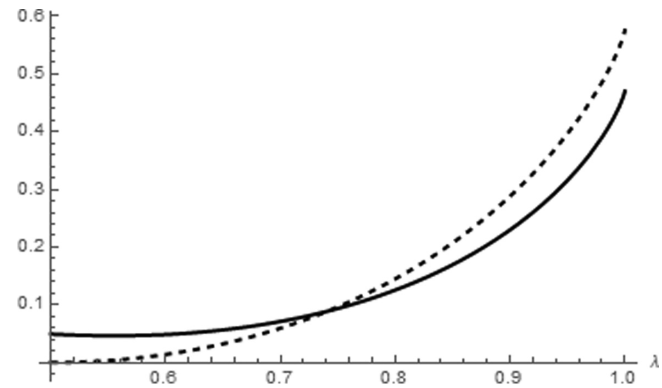

FIG. 13. The solid line is $2E_r(\mathcal{W}_\lambda)$ whereas the dashed line is $D(\mathcal{W}_\lambda^{\otimes 2} \| \overline{\mathcal{L}})$. We observe nonadditivity for $\lambda > 0.74$. The y axis is the relative entropy (units being defined as nonlocal bits) and the x axis is channel parameter λ.

In Fig. 13, we plot $D(\mathcal{W}_\lambda^{\otimes 2} \| \overline{\mathcal{L}})$ versus $2E_r(\mathcal{W}_\lambda)$. It can be seen that $D(\mathcal{W}_\lambda^{\otimes 2} \| \overline{\mathcal{L}})$ is strictly smaller whenever $\lambda > 0.74$. We next extend this construction by considering convex combinations of the two local channels $\mathcal{L}_{1/2}^{\otimes 2}$ and $\overline{L}$,

$$\mathcal{L}^{(z)} = z\mathcal{L}_{1/2}^{\otimes 2} + (1-z)\overline{L}. \quad (31)$$

Why might there be some advantage to mixing these channels? Intuitively, $\overline{\mathcal{L}}$ has a greater probability than $\mathcal{L}_{1/2}^{\otimes 2}$ of satisfying both the XOR conditions $x_1 \oplus y_1 = x_0 y_0$ and $x_1' \oplus y_1' = x_0' y_0'$ (5/8 versus 9/16, respectively). On the other hand, $\overline{\mathcal{L}}$ also has a greater probability than $\mathcal{L}_{1/2}^{\otimes 2}$ of not satisfying either of these XOR conditions (1/8 versus 1/16, respectively). Hence we might expect that a smaller channel divergence can be obtained by mixing the two strategies. In fact, this turns out to be the case. We consider

$$D(\mathcal{W}_\lambda^{\otimes 2} \| \mathcal{L}^{(z)}) = 2 + \frac{(1+\lambda)^2}{4} \log_2 \frac{(1+\lambda)^2}{9z + 10(1-z)} + \frac{1-\lambda^2}{2} \log_2 \frac{1-\lambda^2}{3z + 2(1-z)} + \frac{(1-\lambda)^2}{4} \log_2 \frac{(1-\lambda)^2}{z + 2(1-z)}, \quad (32)$$

and note that the choice $z = 1$ corresponds to the additive bound. With this in mind we compute

$$\frac{d}{dz} D(\mathcal{W}_\lambda^{\otimes 2} \| \mathcal{L}^{(z)}) \Big|_{z=1} = \frac{1}{9}(1-2\lambda)^2 \geqslant 0, \quad (33)$$

which implies that

$$D(\mathcal{W}_\lambda^{\otimes 2} \| \mathcal{L}^{(z)}) < 2E_r(\mathcal{W}_\lambda^{\otimes 2}) \quad (34)$$

in a neighborhood $z < 1$ for every $\lambda > 1/2$. As an example, for the CHSH channel, the minimization is obtained at the point $z_0 = 3 - \sqrt{2} - \sqrt{5(3 - 2\sqrt{2})}$, for which

$$D(\mathcal{W}^{\mathrm{CHSH}} \otimes \mathcal{W}^{\mathrm{CHSH}} \| \mathcal{L}^{(z_0)}) \approx 0.088 < 2E_r(\mathcal{W}^{\mathrm{CHSH}}) \approx 0.093. \quad (35)$$

■

Remark. Theorem 2 shows nonadditive of E_r for an entire class of channels. We conjecture that this result can extended to all channels $\mathcal{N} \in \mathrm{LOSE}(XY) \setminus \mathrm{LOSR}(XY)$. On the other hand, there are nonphysically realizable processes whose relative entropy of dynamical entanglement is additive. For instance, the instantaneous $d \times d$ classical swap has $E_r(\mathcal{N}_d^{\mathrm{swap}}, 0) = \log_2 d$, which is an additive quantity. The source of additivity here, appears to be because $\mathcal{N}_d^{\mathrm{swap}}$ is a deterministic channel. Since the only deterministic nonsignaling channels are also local, we suspect that all nonlocal yet physically realizable channels have nonadditive relative entropy.

B. The max relative entropy of Bell nonlocality

We next proceed to consider the max relative entropy of Bell nonlocality, which is defined as

$$E_{\max}(\mathcal{N}^{XY}) = \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} D_{\max}(\mathcal{N}^{XY}\|\mathcal{M}^{XY}), \tag{36}$$

where

$$D_{\max}(\mathcal{N}\|\mathcal{M}) = \log_2 \max\{\lambda \;:\; \lambda\mathcal{M} \geqslant \mathcal{N}\}. \tag{37}$$

Here the inequality $\lambda\mathcal{M} \geqslant \mathcal{N}$ is to be understood as nonnegativity over the cone of CP maps, i.e., $\lambda\mathcal{M} - \mathcal{N} \in \mathrm{CPTP}(XY)$. The max relative entropy has emerged as an important quantity in the study of different resource theories. For instance, it can be equivalently expressed as the log-robustness or resource [62], which measures how resilient a given state or channel is to losing all its resource under mixing. The max relative entropy (and its smooth variant) appears as a quantifier in certain channel discrimination problems [36,57]. It also captures the resource cost in the general task of catalytic resource erasing [42,63], as well as the one-shot cost for channel simulation in certain resource theories such as coherence theory [64] and general nonsignaling theories [65].

For classical channels $\mathcal{N}, \mathcal{M} \in \mathrm{CPTP}(XY)$, it is straightforward to see that

$$D_{\max}(\mathcal{N}\|\mathcal{M}) = \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} \max_{\substack{x_0,y_0\\x_1,y_1}} \log_2 \frac{p(x_0,y_0|x_1,y_1)}{q(x_0,y_0|x_1,y_1)},$$

where $p(x_0,y_0|x_1,y_1)$ are the channel probabilities of $\mathcal{N}$ and $q(x_0,y_0|x_1,y_1)$ the channel probabilities of $\mathcal{M}$. In fact, by introducing the notation $z_0 = (x_0,y_0)$, $z_1 = (x_1,y_1)$, and the normalized Choi matrices

$$\hat{J}_{\mathcal{N}} := \frac{1}{|X_0||Y_0|}\sum_{z_0,z_1} p(z_1|z_0)|z_0\rangle\langle z_0| \otimes |z_1\rangle\langle z_1|,$$

$$\hat{J}_{\mathcal{M}} := \frac{1}{|X_0||Y_0|}\sum_{z_0,z_1} q(z_1|z_0)|z_0\rangle\langle z_0| \otimes |z_1\rangle\langle z_1|,$$

one has

$$E_{\max}(\mathcal{N}^{XY}) = \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} D_{\max}(\hat{J}_{\mathcal{N}}\|\hat{J}_{\mathcal{M}}). \tag{38}$$

Interestingly, such a representation in terms of the Choi matrices does not hold for the standard relative entropy; i.e., $E_r(\mathcal{M}) \neq \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} D(\hat{J}_{\mathcal{N}}\|\hat{J}_{\mathcal{M}})$. Nevertheless, one can bound

$$E_r(\mathcal{M}) \geqslant \min_{\mathcal{M}\in\mathrm{LOSR}(XY)} D(\hat{J}_{\mathcal{N}}\|\hat{J}_{\mathcal{M}}). \tag{39}$$

Using the family of channels introduced in the previous section, nonadditivity of $E_{\max}$ can also be established. First, by employing the same symmetry argument as before, we have

$$\begin{aligned} E_{\max}(\mathcal{W}_\lambda) &= \min_{|\mu|\leqslant 1/2} \max\left\{\log_2\frac{1+\lambda}{1+\mu}, \log_2\frac{1-\lambda}{1-\mu}\right\} \\ &= \begin{cases} 1+\log_2\frac{1+\lambda}{3} & \text{if } \lambda \geqslant 1/2 \\ 0 \quad \text{if } \lambda < 1/2 \end{cases}. \end{aligned} \tag{40}$$

On the other hand,

$$\begin{aligned} &D_{\max}\big(\mathcal{W}_\lambda^{\otimes 2}\|\overline{\mathcal{L}}\big) \\ &\quad = 2 + \max\left\{\log_2\frac{(1+\lambda)^2}{10}, \log_2\frac{1-\lambda^2}{2}, \log_2\frac{(1-\lambda)^2}{2}\right\}. \end{aligned} \tag{41}$$

For the specific choice of $\lambda = \frac{1}{\sqrt{2}}$, we find

$$D_{\max}(\mathcal{W}^{\mathrm{CHSH}} \otimes \mathcal{W}^{\mathrm{CHSH}}\|\overline{\mathcal{L}}) \approx 0.22 \tag{42}$$

whereas

$$2E_{\max}(\mathcal{W}^{\mathrm{CHSH}}) \approx 0.37. \tag{43}$$

C. The regularized relative entropy of Bell nonlocality and asymptotic channel convertibility

Due to the nonadditivity of both the relative entropy and the max relative entropy, regularization is required when considering the rate of resource among multiple channels. For instance, the regularization of E_r is given by

$$E_r^{\infty}(\mathcal{N}) := \lim_{n\to\infty}\frac{1}{n}E_r((\mathcal{N}^{XY})^{\otimes n}). \tag{44}$$

Here we show that the relative entropy of Bell nonlocality provides an upper bound on the rate $\frac{n}{m}$ at which n copies of a bipartite instantaneous channel $\mathcal{N}^{XY}$ can be used to simulate m copies of $\mathcal{M}^{XY}$ under parallel use. More precisely, let $\mathcal{N} \in \mathrm{CPTP}(X_0Y_0 \to X_1Y_1)$ and $\mathcal{M} \in \mathrm{CPTP}(X_0'Y_0' \to X_1'Y_1')$ be two instantaneous classical resources. We define the asymptotic rate of converting $\mathcal{N}$ to $\mathcal{M}$ as

$$R(\mathcal{N}\to\mathcal{M}) := \lim_{\varepsilon\to 0^+}\sup\left\{\frac{m}{n} \;:\; d_{\mathrm{Bell}}(\mathcal{N}^{\otimes n}\to\mathcal{M}^{\otimes m}) \leqslant \varepsilon\right\}, \tag{45}$$

where d_{Bell} is the conversion distance

$$\begin{aligned} d_{\mathrm{Bell}}(\mathcal{N}\to\mathcal{M}) := \min_{\Theta\in\mathrm{LOSR}(XY\to X'Y')} \tfrac{1}{2}\|\mathcal{M}^{X'Y'} \\ - \Theta^{XY\to X'Y'}(\mathcal{N}^{XY})\|_{\diamond}. \end{aligned} \tag{46}$$

Note that a more general strategy for channel simulation involves adaptively using the n copies of $\mathcal{N}$, such that the output of some can be used as the input to others (see Ref. [36] and references within). We do not consider this problem here, and it remains an open question whether or not theorem 3 below still holds in the adaptive setting.

We say that the resource theory of Bell nonlocality is *reversible* if

$$R(\mathcal{N}\to\mathcal{M})R(\mathcal{M}\to\mathcal{N}) = 1. \tag{47}$$

It is unknown to the authors if the resource theory of Bell nonlocality is reversible. One might be tempted to try and apply the general reversibility criteria for state-based resource theories [66]. However, this approach does not hold in the channel domain. Moreover, Ref. [66] makes use of maximal set of free operations, whereas LOSR superchannels are not the maximal set of superchannels that leave the collection of LOSR channels invariant. To see this, not that a pre and post application of the swap channel will transform any LOSR channel into another LOSR channel; yet this superchannel is clearly not LOSR.

Despite the results of Ref. [66] being applicable here, we can still recover the same bound on the asymptotic convertibility rate for two channels.

Theorem 3. Using the same notations as above,

$$R(\mathcal{N} \to \mathcal{M}) \leqslant \frac{E_r^\infty(\mathcal{N}^{XY})}{E_r^\infty(\mathcal{M}^{X'Y'})} \tag{48}$$

and equality holds if the resource theory is reversible.

Proof. Let $\{\varepsilon_n\}_{n\in\mathbb{N}}$ be a sequence of positive numbers with zero limit, let $\{m_n\}_{n\in\mathbb{N}}$ be a sequence of integers, and $\{\Theta_n\}_{n\in\mathbb{N}}$ be a sequence of LOSR superchannels with $\Theta_n \in \mathrm{LOSR}(X^nY^n \to X'^{m_n}Y'^{m_n})$, such that

$$\left| R(\mathcal{N} \to \mathcal{M}) - \frac{m_n}{n} \right| \leqslant \varepsilon_n \tag{49}$$

and

$$\tfrac{1}{2}\|(\mathcal{M}^{X'Y'})^{\otimes m_n} - \Theta_n(\mathcal{N}^{XY})\|_1 \leqslant \varepsilon_n. \tag{50}$$

In Ref. [35], it was shown that the relative entropy of a dynamical resource is asymptotically continuous. Therefore, for all $n \in \mathbb{N}$,

$$\begin{aligned} E_r((\mathcal{N}^{XY})^{\otimes n}) &\geqslant E_r(\Theta_n((\mathcal{N}^{XY})^{\otimes n})) \\ &\geqslant E_r((\mathcal{M}^{X'Y'})^{\otimes m_n}) - \varepsilon_n\kappa_n \\ &\quad - (1+\varepsilon_n)h\left(\frac{\varepsilon_n}{1+\varepsilon_n}\right), \end{aligned}$$

where $\kappa_n := \max_{\mathcal{E}\in\mathrm{LOSR}(X^nY^n)} E_r(\mathcal{E})$ and h is the binary Shannon entropy. Dividing both sides by n and taking the limit $n \to \infty$ yields

$$\begin{aligned} E_r^\infty(\mathcal{N}) &\geqslant \lim_{n\to\infty} \frac{m_n}{n}\frac{1}{m_n} E_r((\mathcal{M}^{X'Y'})^{\otimes m_n}) \\ &= R(\mathcal{N} \to \mathcal{M}) E_r^\infty(\mathcal{M}^{X'Y'}), \end{aligned} \tag{51}$$

where we used the fact that

$$\lim_{n\to\infty} \frac{\kappa_n}{n} < \infty. \tag{52}$$

For the equality, observe that if the resource theory is reversible then we also have

$$R(\mathcal{N} \to \mathcal{M}) = \frac{1}{R(\mathcal{M} \to \mathcal{N})} \geqslant \frac{E_r^\infty(\mathcal{N}^{XY})}{E_r^\infty(\mathcal{M}^{X'Y'})} \tag{53}$$

since from the first part of the theorem we know that $R(\mathcal{M} \to \mathcal{N}) \leqslant \frac{E_r^\infty(\mathcal{M}^{XY})}{E_r^\infty(\mathcal{N}^{X'Y'})}$. Therefore, in this case, we must have equality in (48). ■

D. Extending classical measures to the quantum domain

We would like now to extend the above measure to all instantaneous bipartite quantum processes; i.e., to all channels $\mathcal{N} \in \mathrm{LOSE}(AB)$ as given in (12). We focus in this paper on the quantification of instantaneous resources since the quantification of noninstantaneous entanglement, both static and dynamic, are well understood [2,38,67]. Recall that LOSE also includes all bipartite quantum states, and so the extensions of E_r to all instantaneous bipartite quantum channels will also provide measures of Bell nonlocality in a quantum bipartite state. We follow the extension approach recently put forward in Ref. [54] for generalized resource theories.

Definition V.2. Let

$$E : \bigcup_{X,Y} \mathfrak{Q}(XY) \to \mathbb{R}$$

be a classical entanglement measure. Then, for any bipartite quantum channel $\mathcal{N} \in \mathrm{LOSE}(AB)$, the minimal and maximal extensions of E_{cl} to all instantaneous quantum processes, denoted by $\overline{E_{\mathrm{cl}}}$ and $\underline{E_{\mathrm{cl}}}$, respectively, are given by

$$\begin{aligned} \underline{E_{\mathrm{cl}}}(\mathcal{N}) &:= \sup E_{\mathrm{cl}}\big(\Theta[\mathcal{N}]\big) \quad \text{and} \\ \overline{E_{\mathrm{cl}}}(\mathcal{N}) &:= \inf E_{\mathrm{cl}}(\mathcal{C}), \end{aligned} \tag{54}$$

where the supremum and infimum are taken over all LOCC superchannels $\Theta \in \mathrm{LOCC}(AB \to XY)$ and all bipartite classical channels $\mathcal{C} \in \cup_{X,Y}\mathrm{CPTP}(XY)$ that satisfy $\mathcal{N} = \Upsilon[\mathcal{C}]$ for some superchannel $\Upsilon \in \mathrm{LOSR}(XY \to AB)$.

Remark. Since we are only considering here instantaneous resources, the superchannels Θ and Υ above are themselves instantaneous LOCC operations. In particular, Θ has either the form of Fig. 10 (i.e., LOSR) or the form of Fig. 11, while Υ can only have the LOSR form of Fig. 10 since its domain is classical (hence the state on systems $\tilde{A}_1$ and $\tilde{B}_1$ will always be separable which means it could be replaced with classical shared randomness with the quantum preparation being absorbed into the post-LO instantaneous channels; in this case the superchannel depicted in Fig. 11 becomes a special case of the LOSR superchannel of Fig. 10).

The following theorem follows directly from the general formalism introduced in Ref. [54] for the extension of resource measures from one domain to a larger one.

Theorem 4. Let E_{cl} be a classical entanglement measure, and $\underline{E_{\mathrm{cl}}}$ and $\overline{E_{\mathrm{cl}}}$ be as above. Then, $\underline{E_{\mathrm{cl}}}$ and $\overline{E_{\mathrm{cl}}}$ are entanglement measures for instantaneous bipartite quantum processes. Moreover, any other such measure E'_{cl} that reduces to E_{cl} on classical instantaneous processes satisfies

$$\underline{E_{\mathrm{cl}}}(\mathcal{N}) \leqslant E'_{\mathrm{cl}}(\mathcal{N}) \leqslant \overline{E_{\mathrm{cl}}}(\mathcal{N}) \quad \forall\, \mathcal{N} \in \mathrm{LOSE}(AB).$$

Note that from its definition, $\underline{E_{\mathrm{cl}}}$ cannot be increased even under pre-LOCC operations that results in an instantaneous classical channel, as depicted in Fig. 11.

VI. CONCLUSION

In this work, we have demonstrated that Bell nonlocality is a property of bipartite quantum systems that can be studied as a special form of entanglement. This is accomplished by constructing a resource theory based on the abstract notion of a quantum processes and input-to-output time delay. Resources in quantum information science can thus be diversely classified as (1) classical or quantum, (2) static or dynamic, (3) noisy or noiseless, (4) private or public, and as we add in this paper, (5) instantaneous or noninstantaneous. This last piece is the key to unifying entanglement and nonlocality as quantum resources.

One of the key features of our model is that pre-LOCC maps are naturally free operations; i.e., static-to-dynamic conversions having the form of Eq. (4) are allowed in this framework. At the same time, the use of LOCC maps *after* the classical inputs (x_0, y_0) in Eq. (4) is automatically prohibited as it leads to a noninstantaneous resource. By allowing pre-LOCC, a well-defined partial order among static and dynamic resource convertibility can be established. In particular, maximally entangled states will be "maximally resourceful" for generating all forms of quantum correlations, including those present in Hardy-like Bell tests [8]. This reflects our overall conclusion that Bell nonlocality belongs to the same resource theory as quantum entanglement.

There are a number of interesting open questions related to the quantification of Bell nonlocality. We have shown that a large class of bipartite classical channels is nonadditive with respect to the relative entropy of resource. We conjecture that this is a generic property in that all nonlocal, nonsignaling channels are nonadditive. A related question is how non-additive the relative entropy of Bell nonlocality can be. In principle, it is possible that the measure scales sub-linearly in the number of copies so that $E_r^{\infty}(\mathcal{N}) = 0$ for some non-LOSR channel $\mathcal{N}$. It is known that E_r^{∞} is a faithful measure for states (i.e., vanishing iff a given state is entangled) [68,69]; whether or not the same is true for channels remains an open problem.

ACKNOWLEDGMENTS

E.C. is funded by the National Science Foundation Award No. 1914440. G.G. is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

[1] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[2] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[3] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).
[4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[5] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[6] M. Froissart, Il Nuovo Cimento B **64**, 241 (1981).
[7] A. Fine, Phys. Rev. Lett. **48**, 291 (1982).
[8] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).
[9] P. H. Eberhard, Phys. Rev. A **47**, R747 (1993).
[10] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[11] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[12] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[13] A. A. Methot and V. Scarani, Quantum Inf. Comput. **7**, 157 (2007).
[14] L. Masanes, Y.-C. Liang, and A. C. Doherty, Phys. Rev. Lett. **100**, 090403 (2008).
[15] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
[16] M. T. Quintino, T. Vértesi, D. Cavalcanti, R. Augusiak, M. Demianowicz, A. Acín, and N. Brunner, Phys. Rev. A **92**, 032107 (2015).
[17] R. Augusiak, M. Demianowicz, J. Tura, and A. Acín, Phys. Rev. Lett. **115**, 030404 (2015).
[18] F. Hirsch, M. T. Quintino, J. Bowles, T. Vértesi, and N. Brunner, New J. Phys. **18**, 113019 (2016).
[19] V. Scarani, *Bell Nonlocality*, Oxford Graduate Texts (Oxford University Press, Oxford, 2019).
[20] Y.-C. Liang, L. Masanes, and D. Rosset, Phys. Rev. A **86**, 052115 (2012).
[21] C. Palazuelos, Phys. Rev. Lett. **109**, 190401 (2012).
[22] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[23] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter, Commun. Math. Phys. **328**, 303 (2014).
[24] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).
[25] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics: Collected Papers on Quantum Philosophy* (Cambridge University Press, Cambridge, 2004).
[26] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Phys. Rev. Lett. **109**, 070401 (2012).
[27] J. I. de Vicente, J. Phys. A: Math. Theor. **47**, 424017 (2014).
[28] K. Horodecki, A. Grudka, P. Joshi, W. Kłobus, and J. Łodyga, Phys. Rev. A **92**, 032104 (2015).
[29] R. Gallego and L. Aolita, Phys. Rev. A **95**, 032118 (2017).
[30] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Quantum **4**, 280 (2020).
[31] D. Schmid, D. Rosset, and F. Buscemi, Quantum **4**, 262 (2020).
[32] D. Schmid, T. C. Fraser, R. Kunjwal, A. B. Sainz, E. Wolfe, and R. W. Spekkens, arXiv:2004.09194.
[33] B. Coecke, T. Fritz, and R. W. Spekkens, Information and Computation **250**, 59 (2016), quantum Physics and Logic.
[34] E. Chitambar and G. Gour, Rev. Mod. Phys. **91**, 025001 (2019).
[35] G. Gour and A. Winter, Phys. Rev. Lett. **123**, 150401 (2019).
[36] X. Wang and M. M. Wilde, Phys. Rev. Research **1**, 033169 (2019).
[37] X. Wang, M. M. Wilde, and Y. Su, New J. Phys. **21**, 103002 (2019).
[38] G. Gour and C. M. Scandolo, Phys. Rev. Lett. **125**, 180505 (2020).
[39] S. Bäuml, S. Das, X. Wang, and M. M. Wilde, arXiv:1907.04181.
[40] K. Fang, O. Fawzi, R. Renner, and D. Sutter, Phys. Rev. Lett. **124**, 100501 (2020).
[41] Y. Liu and X. Yuan, Phys. Rev. Research **2**, 012035(R) (2020).
[42] Z.-W. Liu and A. Winter, arXiv:1904.04201.
[43] G. Gour and M. M. Wilde, Phys. Rev. Research **3**, 023096 (2021).
[44] E. Kaur and M. M. Wilde, J. Phys. A: Math. Theor. **51**, 035303 (2017).
[45] V. Katariya and M. M. Wilde, Quantum Inf. Process **20**, 78 (2021).

[46] K. Fang and H. Fawzi, Commun. Math. Phys. **384**, 1615 (2021).
[47] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).
[48] I. Devetak, A. W. Harrow, and A. J. Winter, IEEE Trans. Inf. Theory **54**, 4587 (2008).
[49] F. Buscemi, Phys. Rev. Lett. **108**, 200401 (2012).
[50] G. Chiribella, G. M. D'Ariano, and P. Perinotti, EPL (Europhysics Letters) **83**, 30004 (2008).
[51] G. Gour, IEEE Trans. Inf. Theory **65**, 5880 (2019).
[52] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, Phys. Rev. A **64**, 052309 (2001).
[53] T. Eggeling, D. Schlingemann, and R. F. Werner, Europhysics Letters (EPL) **57**, 782 (2002).
[54] G. Gour and M. Tomamichel, Phys. Rev. A **102**, 062401 (2020).
[55] T. Cooney, M. Mosonyi, and M. M. Wilde, Commun. Math. Phys. **344**, 797 (2016).
[56] F. Leditzky, E. Kaur, N. Datta, and M. M. Wilde, Phys. Rev. A **97**, 012332 (2018).
[57] M. Berta, C. Hirche, E. Kaur, and M. M. Wilde, Lett. Math. Phys. **110**, 2277 (2020).
[58] G. Gour, Phys. Rev. X Quantum **2**, 010313 (2021).
[59] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[60] R. Cleve, W. Slofstra, F. Unger, and S. Upadhyay, Computational Complexity **17**, 282 (2008).
[61] K. G. H. Vollbrecht and R. F. Werner, Phys. Rev. A **64**, 062307 (2001).
[62] N. Datta, Int. J. Quantum. Inform. **07**, 475 (2009).
[63] A. Anshu, M.-H. Hsieh, and R. Jain, Phys. Rev. Lett. **121**, 190504 (2018).
[64] M. G. Díaz, K. Fang, X. Wang, M. Rosati, M. Skotiniotis, J. Calsamiglia, and A. Winter, Quantum **2**, 100 (2018).
[65] R. Takagi, K. Wang, and M. Hayashi, Phys. Rev. Lett. **124**, 120502 (2020).
[66] F. G. S. L. Brandão and G. Gour, Phys. Rev. Lett. **115**, 070503 (2015).
[67] M. B. Plenio and S. Virmani, Quant. Inf. Comput. **7**, 1 (2007).
[68] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke, Phys. Rev. Lett. **95**, 190501 (2005).
[69] M. Piani, Phys. Rev. Lett. **103**, 160504 (2009).