Operational Significance of the Quantum Resource Theory of Buscemi Nonlocality

Patryk Lipka-Bartosik[1,*] Andrés F. Ducuara[1,2,3,4,†] Tom Purves,[1] and Paul Skrzypczyk[1]

[1] *H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, United Kingdom*

[2] *Quantum Engineering Centre for Doctoral Training, University of Bristol, Bristol BS8 1FD, United Kingdom*

[3] *Quantum Engineering Technology Labs, University of Bristol, Bristol BS8 1FD, United Kingdom*

[4] *Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1FD, United Kingdom*

Although entanglement is necessary for observing nonlocality in a Bell experiment, there are entangled states that can never be used to demonstrate nonlocal correlations. In a seminal paper Buscemi [Phys. Rev. Lett. 108, 200401 (2012)] extended the standard Bell experiment by allowing Alice and Bob to be asked quantum, instead of classical, questions. This gives rise to a broader notion of nonlocality, one which can be observed for every entangled state. In this work we study a resource theory of this type of nonlocality referred to as *Buscemi nonlocality*. We propose a geometric quantifier measuring the ability of a given state and local measurements to produce Buscemi nonlocal correlations and prove the following results. First, we show that any distributed measurement that can demonstrate Buscemi nonlocal correlations provides strictly better performance than any distributed measurement that does not use entanglement in the task of distributed state discrimination, and that this advantage is quantified by the geometric quantifier we propose, thus establishing its operational significance. Second, we prove a quantitative relationship between Buscemi nonlocality, the ability to perform nonclassical teleportation, and entanglement. In particular, we show that the maximal amount of Buscemi nonlocality that can be generated using a given state is precisely equal to its entanglement content. Using this relationship, we propose new discrimination tasks for which nonclassical teleportation and entanglement lead to an advantage over their classical counterparts. Third, we interpret Buscemi nonlocality from the perspective of information theory and show that it is related to a single-shot capacity of a quantum-to-classical bipartite channel.

I. INTRODUCTION

Quantum entanglement is one of the most characteristic features of quantum theory [1]. During the early years of its development, however, it was recognized mainly as a bizarre property that distinguished it from classical physics. It was due to the discovery of Bell nonlocality [2] and subsequent development of Bell inequalities that allowed this distinction to be formulated quantitatively and to verify the predictions of quantum theory in an experimentally feasible setting.

*patryk.lipka.bartosik@gmail.com
†andres.ducuara@bristol.ac.uk

Bell nonlocality is today perceived as a phenomenon in its own right and can be defined and tested irrespectively of the underlying theory. In simple terms Bell nonlocality refers to the situation when correlations shared between spatially separated parties cannot be explained as arising from a shared classical resource. The concept of Bell nonlocality is perhaps best understood in terms of a Bell experiment, which is sometimes also called a "no-signaling game." In such a game, a referee distributes two physical systems to two spatially separated players, Alice (*A*) and Bob (*B*). Upon receiving their systems, each player is asked a question from a prearranged set of questions, labeled *x* for Alice and *y* for Bob. Depending on which of the questions was asked, Alice measures her system locally and obtains an outcome *a*. Similarly, based on his own question, Bob measures his share of the system and obtains *b*. The data produced from the experiment can be described using a conditional probability distribution $p(a, b|x, y)$, that is, the probability of producing outcomes

a and *b* given the choice of measurements labeled by *x* and *y*.

Importantly, not all entangled states can display Bell nonlocality [3–5]. Quantum states can actually demonstrate other forms of nonlocality that are not accessible in a Bell experiment but that may become apparent in different experimental settings. In a seminal work Buscemi [6] generalized Bell's original experiment by allowing the referee to ask "quantum questions." This amounts to replacing the original set of classical (and therefore mutually orthogonal) questions $\{|x\rangle\}$ with a set of quantum states $\{|\omega_x\rangle\}$ that need not be orthogonal. The correlation data $p(a, b|\omega_x, \omega_y)$ obtained in this modified experiment, dubbed *semiquantum nonsignaling games*, differs significantly from its archetypical counterpart. Perhaps the most striking consequence is that the new experiment is powerful enough to reveal the nonlocality of any entangled quantum state, even the nonlocality that would be hidden under a standard Bell test [6]. This semiquantum approach, also called measurement-device independent (MDI), has been a fruitful line of investigation during the last decade [7–19].

In this work we propose interpreting the correlation data obtained in a semiquantum nonsignaling game as an indicator of a this type of nonlocality that we refer to as *Buscemi nonlocality*. In order to formalize this notion, we utilize the framework of quantum resource theories (QRTs) [20,21]. This is a set of tools and techniques developed to systematically quantify different properties of quantum systems. QRTs can be classified in terms of *objects* and *resources* studied in a given theory. Classification of QRTs with respect to the object lead to the resource theories of states [21], measurements [22–27], channels [28–31], and boxes [32–35]. On the other hand, classifying QRTs with respect to the type of the studied resource leads to the resource theories of pure [36] and mixed-state entanglement [37], coherence [38], purity [39,40], athermality [41–46], nonlocality [47], asymmetry [48], measurement incompatibility [49], teleportation [50,51], magic [52], non-Markovianity [53–55] or non-Gaussianity [56], amongst many more. Its worth mentioning that although many QRTs use essentially the same mathematical formalism, their physical implications can be genuinely different. Hence, the wide applicability of the framework to otherwise unrelated problems is a truly surprising aspect of nature.

In this work, we focus on the quantum resource theory of Buscemi nonlocality, which is an instance of the resource theory from Refs. [34,35]. The natural object relevant for this theory is a positive operator-valued measure (POVM) performed by spatially separated parties that do not communicate (distributed measurement). We investigate a geometric measure that quantifies the amount of Buscemi nonlocality contained within a given distributed measurement termed robustness of Buscemi nonlocality

(RoBN). We then address Buscemi nonlocality as a property of states, by considering the maximal amount of Buscemi nonlocality that can be obtained using a given state by any local set of measurements on Alice's and Bob's side.

As our first and main result, we show that Buscemi nonlocality has operational significance, by finding an operational task for which Buscemi nonlocality is a natural resource. This can be seen as akin to several seminal results in the field of quantum information that showed the operational character of coherence [38], entanglement [56], steering [57], or Bell nonlocality [58] in terms of experimentally relevant information-processing tasks. Moreover, our task gives rise to a complete family of monotones for this resource theory, i.e., provides a sufficient and necessary characterization of Buscemi nonlocality contained in a distributed measurement. Consequently, the average probability of guessing in these family of tasks can be interpreted as a simple and complete set of "Buscemi inequalities" that characterize nonlocality of distributed measurements, in analogy with the celebrated Bell inequalities characterizing nonlocality of states.

Our second main result concerns how Buscemi nonlocality relates to other types of nonclassical phenomena studied in the literature: nonclassical teleportation and entanglement. We show that the maximal value of RoBN that can be achieved when Bob (Alice) is allowed to use any measurement is precisely the so-called robustness of teleportation (RoT) of a teleportation channel from Alice (Bob) to Bob (Alice). On the other hand, optimizing RoBN over all local measurements for both parties leads to the robustness of entanglement of the state shared by Alice and Bob. This result, despite its clarifying character being of independent interest, leads to new operational tasks for which both nonclassical teleportation and entanglement are natural resources. These quantitative relationships further expand the results presented in Refs. [47,56,59] by proposing new discrimination tasks for which both entanglement and nonclassical teleportation provide advantage over their classical (i.e., separable) counterparts.

As our third and final main result, we interpret Buscemi nonlocality from the perspective of single-shot quantum information theory. We show that Buscemi nonlocality, when viewed as a property of a communication channel between the sender (the referee) and receiver (Alice and Bob), quantifies the maximal amount of information that can be sent reliably when the channel is used only once (the so-called single-shot capacity of a quantum channel). This establishes an important link between Buscemi nonlocality and quantum communication.

The paper is organized as follows. In Sec. II we cover the relevant formalism, outline the idea of characterizing nonlocality in terms of nonsignaling games, and recall the robustness quantifier of Buscemi nonlocality. In Sec. II B we find its operational interpretation in terms of the

advantage in the task of distributed state discrimination (DSD). In Sec. III A we explore the relationship between Buscemi nonlocality and the concepts of nonclassical teleportation and entanglement. Finally, in Sec. II B we describe a tangential view of RoBN from the perspective of single-shot information theory. We conclude with Sec. IV where we summarize our findings and highlight several open questions.

II. FRAMEWORK

In what follows we denote a local bipartite measurement on Alice's side (system AA') with $\mathbb{M}^A = \{M_a^{AA'}\}$, where each $M_a^{AA'}$ is a positive semidefinite operator that adds up to the identity (POVM). Similarly, we use $\mathbb{M}^{AB}$ to indicate that the measurement is nonlocal, i.e., we treat systems labeled with different letters, e.g., A and B, as two spatially separated parties. We are interested in the most general type of measurement that can be performed in this bipartite scenario without the aid of classical or quantum communication. This can be realized by (i) allowing Alice and Bob to apply arbitrary bipartite measurements in their labs, respectively denoted $\mathbb{M}^A = \{M_a^{AA'}\}$ and $\mathbb{M}^B = \{M_b^{B'B}\}$, where $a \in \{1, \ldots, o_A\}$ and $b \in \{1, \ldots, o_B\}$ denote Alice's and Bob's outcomes, and (ii) allowing the two parties to share a quantum state $\rho^{A'B'}$. In this way Alice and Bob can store and share all types of classical information (e.g., classical memory or measurement strategy), as well as quantum information (i.e., shared entanglement). We denote such a measurement with $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$, where the corresponding POVM elements are of the general form

$$M_{ab}^{AB} = \mathrm{tr}_{A'B'}[(M_a^{AA'} \otimes M_b^{B'B})(\mathbb{1}^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B)]. \quad (1)$$

Since the sets of all quantum states and quantum measurements are both convex sets, it follows that the set of measurements of the form (1) is also a convex set. We refer to measurements of the form (1) as *distributed measurements* and denote the set of all such measurements by $\mathscr{R}_{\mathrm{BN}}$. These measurements are the main (*resourceful*) objects of the resource theory we consider here. Whenever the elements of measurement $\mathbb{M}^{AB}$ can be written as in Eq. (1) for some choice of shared state and local measurements, we write $\mathbb{M}^{AB} \in \mathscr{R}_{\mathrm{BN}}$. Later in Sec. II B we formally define the set of *free measurements* of this resource theory, which turn out to be distributed measurements with a separable shared state. In Fig. 1 we illustrate a distributed measurement and describe the relationship between different subsystems. These types of objects appear naturally in a wide range of contexts when studying nonlocal effects in a MDI setting [6,8,14,60].

We now specify the most general class of operations that the separated parties in A and B can perform, without communicating, to improve the properties of their distributed measurement $\mathbb{M}^{AB} = \{M_{ij}^{AB}\}$, where indices $i \in \{1, \ldots, o_A\}$
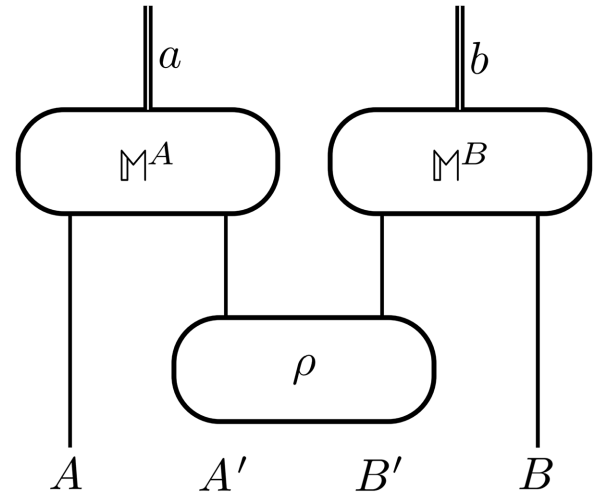

FIG. 1. A schematic diagram of a distributed measurement $\mathbb{M}^{AB}$ composed of local measurements for Alice $\mathbb{M}^A = \{M_a^{AA'}\}$, for Bob $\mathbb{M}^B = \{M_b^{B'B}\}$, and a state $\rho^{A'B'}$ shared between them. This is the most general type of measurement that Alice and Bob can perform in a distributed scenario that does not allow for communication.

and $j \in \{1, \ldots, o_B\}$ describe measurement outcomes. The free operations for the QRT of Buscemi nonlocality are the so-called local operations and shared randomness (LOSR) [32–35]. There, Alice and Bob are allowed to share any amount of classical memory described by a random variable λ. Formally, this is specified by providing a probability distribution $p(\lambda)$ that is available to both parties. Moreover, before measuring their systems both parties are allowed to locally perform any completely positive and trace-preserving map, potentially conditioned on the value of the shared memory, i.e., we allow for applying $\mathcal{E}_\lambda$ on Alice's and $\mathcal{N}_\lambda$ on Bob's side. Finally, the parties are allowed to postprocess their measurement outcomes using arbitrary classical channels $p(a|i, \lambda)$ and $p(b|j, \lambda)$ to produce their final guesses. This procedure leads to the most general type of LOSR operation that can be performed on a measurement of form (1) [34,35]. In what follows we refer to this as *quantum simulation*.

Definition 1 (Quantum simulation). *A quantum simulation of a bipartite measurement* $\mathbb{M} = \{M_{ij}\}$ *with a subroutine*

$$\mathscr{S} = \{p(\lambda), p(a|i, \lambda), p(b|j, \lambda), \mathcal{E}_\lambda, \mathcal{N}_\lambda\} \quad (2)$$

is a transformation that maps the POVM elements of $\mathbb{M}$ *into*

$$M_{ab}' = \sum_{i,j,\lambda} p(\lambda) p(a|i, \lambda) p(b|j, \lambda)(\mathcal{E}_\lambda^\dagger \otimes \mathcal{N}_\lambda^\dagger)[M_{ij}], \quad (3)$$

where $\mathcal{E}^\dagger$ *denotes the (unique) dual map to* $\mathcal{E}$.

In other words, any action that can be performed by Alice and Bob in their labs without access to communication can be described by some quantum simulation subroutine.

Quantum simulation induces a natural preorder on the set of all bipartite measurements. Formally, a preorder is an ordering relation that is reflexive ($a \succ a$) and transitive ($a \succ b$) and ($b \succ c$) implies ($a \succ c$). Here the preorder induced by quantum simulation will be denoted by "$\succ_q$," i.e., $\mathbb{M} \succ_q \mathbb{M}'$ if and only if there exists a subroutine $\mathscr{S}$ that allows $\mathbb{M}$ to simulate $\mathbb{M}'$, i.e., for the two measurements $\mathbb{M}$ and $\mathbb{M}'$, condition (3) in Definition 1 holds. The notion of simulation will turn out to be relevant for the operational tasks introduced later on.

A. Nonlocality from the perspective of no-signaling games

Bell nonlocality can be best understood from the perspective of no-signaling games, which also provide an intuitive understanding of Bell inequalities. Such games have been extensively studied in computer science for a long time, where they are a special instance of *interactive proof systems* [61].

The standard scenario of a no-signaling game involves two cooperating players (Alice and Bob) who play the game against a third party, the referee. The referee chooses a question $x \in \mathcal{X}$ for Alice and $y \in \mathcal{Y}$ for Bob according to some probability distribution $p(x,y) \colon \mathcal{X} \times \mathcal{Y} \to [0,1]$, where $\mathcal{X}$ and $\mathcal{Y}$ denote finite sets of questions. Without communicating, and, therefore, without knowing what question the other player was asked, Alice (Bob) returns an answer $a \in \mathcal{A}$ ($b \in \mathcal{B}$) from a finite set of possible answers $\mathcal{A}$ ($\mathcal{B}$). Based on the questions asked and the received answers, the referee determines whether the players win or lose the game, according to a prearranged set of rules. Such rules are typically expressed using a function $V \colon \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to [0,1]$, where $V(a,b,x,y) = 1$ if and only if Alice and Bob win the game by answering a and b for questions x and y.

Alice and Bob know the rules of the game, that is, they know the function V and the distribution of questions $p(x,y)$. Before the game starts they can agree on any strategy that provides them with the best chances of winning. However, once the game starts, they are not allowed to communicate any more. In the classical setting any strategy they can possibly devise can be encoded in a classical memory system, represented by a shared random variable λ and a probability distribution $p(\lambda)$. In the more general quantum case, any possible strategy can be described by a shared quantum state ρ and a choice of local measurements.

In order to relate the above game setting with Bell inequalities, note that the referee's questions x and y can be thought of as labels for different measurement settings.

Similarly, the answers correspond to the outcomes of local measurements. Any measurement strategy (be it classical or quantum) leads to a conditional probability $p(a,b|x,y)$ that describes when Alice and Bob give answers a and b for questions x and y, respectively. In the language of Bell inequalities, $p(a,b|x,y)$ determines the probability that Alice and Bob obtain measurement outcomes a and b when performing the measurements labeled by x and y. The average probability that Alice and Bob win, maximized over all possible strategies, can be written as

$$p_{\text{guess}}^{V}(\mathcal{G}, \mathbb{M}) = \sum_{a,b,x,y} p(x,y)p(a,b|x,y)V(a,b,x,y), \quad (4)$$

where $\mathcal{G} = \{p(x,y), V\}$ defines the game and the conditional probabilities $p(a,b|x,y)$ are related to the local measurements $\{M_{a|x}^{A}\}$ for Alice and $\{M_{b|y}^{B}\}$ for Bob, via the Born rule

$$p(a,b|x,y) = \text{tr}[(M_{a|x}^{A'} \otimes M_{b|y}^{B'})\rho^{A'B'}]. \quad (5)$$

With this in mind, Bell inequalities can be thought of as upper bounds on the average guessing probability $p_{\text{guess}}(\mathcal{G}, \mathbb{M})$ with which Alice and Bob can win a nonlocal game $\mathcal{G}$ using a classical strategy (i.e., when $\rho^{A'B'}$ is a separable state), optimized over all local measurements $\{M_{a|x}^{A'}\}$ and $\{M_{b|y}^{B'}\}$. A violation of a Bell inequality corresponds to the situation when there is a quantum strategy that uses an entangled shared state and outperforms the best classical strategy in a particular game $\mathcal{G}$.

Importantly, there are entangled states that can never violate any Bell inequality [3–5]. In the language of no-signaling games this means that there are states $\rho^{A'B'}$ that, although entangled, can never outperform the best classical strategy. However, Buscemi [6] showed that, when we modify the rules of the no-signaling game and allow the referee to ask *quantum* instead of classical questions, then all entangled states can outperform the best classical strategy in some nonlocal game, or equivalently, violate the corresponding Bell inequality.

Before going into the details, let us note that "asking classical questions" can also be mathematically modeled by sending states from a collection of orthogonal states from a fixed basis, e.g., $\{|x\rangle\}$ such that $\sum_{x} |x\rangle\langle x| = \mathbb{1}$ and $\langle x|x'\rangle = \delta_{x,x'}$ and similarly for $\{|y\rangle\}$. Such states are perfectly distinguishable and hence Alice and Bob, after receiving their questions, may choose their measurements unambiguously. This can be viewed as giving Alice and Bob the ability to perform controlled bipartite measurements $\mathbb{M}^{AA'} = \{M_{a}^{AA'}\}$ and $\mathbb{M}^{B'B} = \{M_{b}^{B'B}\}$ with the POVM elements

$$M_{a}^{AA'} = \sum_{x} |x\rangle\langle x|^{A} \otimes M_{a|x}^{A'}, \quad (6)$$

$$M_b^{B'B} = \sum_y M_{b|y}^{B'} \otimes |y\rangle\langle y|^B. \qquad (7)$$

If Alice and Bob share a quantum state $\rho^{A'B'}$ then effectively they have access to a distributed measurement $\mathbb{M}^{AB}$ of the form (1). This measurement is then applied to the "questions" they receive, which we denote here by $\omega_x^A = |x\rangle\langle x|^A$ for Alice and $\omega_y^B = |y\rangle\langle y|^B$ for Bob. Therefore, their behavior $p(a,b|\omega_x,\omega_y)$ can be written as

$$
\begin{aligned}
p(a,b|\omega_x,\omega_y) &:= \mathrm{tr}[M_{ab}^{AB}(\omega_x^A \otimes \omega_y^B)] \\
&= \mathrm{tr}[M_{ab}^{AB}(|x\rangle\langle x|^A \otimes |y\rangle\langle y|^B)] \\
&= \mathrm{tr}[(M_{a|x}^{A'} \otimes M_{b|y}^{B'})\rho^{A'B'}] \\
&= p(a,b|x,y). \qquad (8)
\end{aligned}
$$

With this in mind we can now formalize the process of asking "quantum questions." This happens precisely when the states sent by the referee are chosen from an arbitrary collection of states $\{\omega_x\}$. Crucially, these states need not be distinguishable and so each of them can be in a superposition of different orthogonal states.

Note, however, that using quantum states as inputs to the distributed measurement $\mathbb{M}^{AB}$ with local measurements of the forms (6) and (7) can only lead to a probabilistic version of the standard no-signaling game, i.e., Alice and Bob randomize their choices of measurements according to the respective overlaps $p(x'|x) = \langle x'|\omega_x|x'\rangle$ and $p(y'|y) = \langle y'|\omega_y|y'\rangle$. Thus, in order to use the power of asking genuinely quantum questions, one needs to allow for arbitrary bipartite local measurements on both sides. This leads to the general form of a distributed measurement (1) with the local POVM elements $\{M_a^{AA'}\}$ and $\{M_b^{BB'}\}$ being now fully general bipartite measurements, and therefore a Buscemi behavior is of the form

$$
\begin{aligned}
&p(a,b|\omega_x,\omega_y) \\
&= \mathrm{tr}[(M_a^{AA'} \otimes M_b^{B'B})(\omega_x^A \otimes \rho^{A'B'} \otimes \omega_y^B)]. \quad (9)
\end{aligned}
$$

The above extension of a no-signaling game leads to a novel type of nonlocality that was noticed for the first time in Ref. [6]. Here we refer to this type of nonclassical correlation as *Buscemi nonlocality*. In this language the main result of Ref. [6] states that all entangled states are Buscemi nonlocal.

In what follows we present a consistent way of quantifying Buscemi nonlocality. First we define a proxy quantity that quantifies how much Buscemi nonlocality can be evidenced using a fixed distributed measurement. This provides a natural quantifier for the resource theory of Buscemi nonlocality of distributed measurements, which is our main focus here. Optimizing this quantity over all choices of local measurements for Alice and Bob gives

rise to a quantity that measures the maximal degree of Buscemi nonlocality that can ever be obtained using a given quantum state.

B. Quantitative measure of Buscemi nonlocality

The fact that Alice and Bob may share entanglement in Eq. (1) and use it to perform a measurement means that the measurement is inherently nonlocal and can lead to interesting correlations, even when measured on completely independent systems. Our central question then is how to quantify this nonlocality present in a bipartite measurement. To build a valid reference point, we first consider the case when the measurement does not lead to any type of quantum correlation. This means that the behavior $p(a,b|\omega_x,\omega_y) = \mathrm{tr}[M_{ab}^{AB}(\omega_x^A \otimes \omega_y^B)]$ results from the measurement $\{M_{ab}^{AB}\}$ formed using a separable shared state $\rho^{A'B'} \in \mathsf{SEP}$, where SEP denotes the set of all separable operators. Any separable state can be written as

$$\rho^{A'B'} = \sum_\lambda p(\lambda)\rho_\lambda^{A'} \otimes \rho_\lambda^{B'}, \qquad (10)$$

where $p(\lambda)$ is a classical probability distribution corresponding to a shared random variable λ and $\{\rho_\lambda^{A'}\}$ and $\{\rho_\lambda^{B'}\}$ are collections of local quantum states. The associated distributed measurement from Eq. (1) takes the form

$$M_{ab}^{AB} = \sum_\lambda p(\lambda)M_{a|\lambda}^A \otimes M_{b|\lambda}^B, \qquad (11)$$

where we have denoted $M_{a|\lambda}^A := \mathrm{tr}_{A'}[M_a^{AA'}(\mathbb{1}^A \otimes \rho_\lambda^{A'})]$ for Alice and $M_{b|\lambda}^B := \mathrm{tr}_{B'}[M_b^{B'B}(\rho_\lambda^{B'} \otimes \mathbb{1}^B)]$ for Bob. This is the most general classical measurement scheme that can be realized if Alice and Bob have access only to classical randomness λ and the ability to locally prepare quantum states in their labs. The set of all measurements that can be written as in Eq. (11) will be denoted by $\mathscr{F}_{\mathrm{BN}}$. These measurements are the most natural candidates for free objects in the resource theory of Buscemi nonlocality. Note that measurements from this set have POVM elements that are all separable (SEP) and admit a quantum realization ($\mathscr{R}_{\mathrm{BN}}$), i.e., can be written as in Eq. (1) for some choice of local measurements and shared state. Such measurements can never demonstrate Buscemi nonlocality, regardless of the state being measured.

In order to better understand the difference between the sets $\mathscr{R}_{\mathrm{BN}}$ (all distributed measurements) and $\mathscr{F}_{\mathrm{BN}}$ (free-distributed measurements), let us consider the following simple example.

Example 1. *Let Alice and Bob share a two-qubit Werner state,*

$$\rho^{A'B'} = p\phi_+^{A'B'} + (1-p)\frac{\mathbb{1}^{A'B'}}{4}, \qquad (12)$$

where $p \in [0, 1]$, the state $\phi_+ = |\phi_+\rangle\langle\phi_+|$, and $|\phi_+\rangle :=$ $(1/\sqrt{2})(|0\rangle|0\rangle + |1\rangle|1\rangle)$ is a maximally entangled state. It is widely known that the Werner state (12) is separable for all $p \leq \frac{1}{3}$. Let $\{U_a\}$ for $a = 1, \ldots, 4$, be a set of Pauli operators. Consider a measurement $\mathbb{M}^A = \{M_a^{A'A}\}$ with elements

$$M_a^{A'A} = (U_a^{A'} \otimes \mathbb{1}^A)\phi_+^{A'A}(U_a^{A'} \otimes \mathbb{1}^A)^\dagger. \quad (13)$$

Defining an analogous measurement for Bob $\mathbb{M}^B = \{M_b^{BB'}\}$ and using definition (1) allows us to write the distributed measurement $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$ for Alice and Bob as

$$M_{ab}^{AB} = (U_a^{A'} \otimes U_b^{B'})\rho^{A'B'}(U_a^{A'} \otimes U_b^{B'})^\dagger \quad (14)$$

$$= p\phi_{ab}^{A'B'} + (1-p)\frac{\mathbb{1}^{A'B'}}{4}, \quad (15)$$

where $\phi_{ab}^{A'B'} := (U_a^{A'} \otimes U_b^{B'\dagger})\phi_+^{A'B'}(U_a^{A'} \otimes U_b^{B'\dagger})^\dagger$. Clearly, $\phi_{ab}^{A'B'}$ is again a maximally entangled state and therefore each POVM element of $\mathbb{M}^{AB}$ is a Werner state, up to local unitaries. Since entanglement is preserved under local unitary operations, all elements of the distributed measurement $\mathbb{M}^{AB}$ are entangled operators for $p > \frac{1}{3}$.

We can therefore conclude that, for $p \leq \frac{1}{3}$, the distributed measurement $\mathbb{M}^{AB}$ can be written as in Eq. (11), which, by definition, means that $\mathbb{M}^{AB} \in \mathscr{F}_{BN}$. Moreover, for $p > \frac{1}{3}$, we know that each $M_{ab}^{AB} \notin \mathsf{SEP}$ and therefore $\mathbb{M}^{AB} \notin \mathscr{F}_{BN}$. This implies that this distributed measurement is a resourceful measurement in the resource theory of Buscemi nonlocality.

A natural question at this point is: given an arbitrary bipartite measurement $\mathbb{M}^{AB} \in \mathscr{R}_{BN}$, how can its nonlocal properties be quantified, in particular its ability to generate Buscemi nonlocality? For this purpose, it is useful to define the following quantity.

Definition 2 (Robustness of Buscemi nonlocality [14])**.**
The robustness of Buscemi nonlocality of a distributed measurement $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$ is the solution to the optimization problem

$$\mathbf{R}_{BN}(\mathbb{M}^{AB}) = \min r \quad (16a)$$

$$such \ that \quad M_{ab}^{AB} + rN_{ab}^{AB} = (1+r)O_{ab}^{AB} \quad for \ all \ a, b, \quad (16b)$$

$$\{O_{ab}^{AB}\} \in \mathscr{F}_{BN}, \quad \{N_{ab}^{AB}\} \in \mathscr{R}_{BN}. \quad (16c)$$

Although this may not seem obvious at first sight, the above is a convex optimization problem and hence can be efficiently solved numerically [62–64] (see Appendix A for details). Moreover, due to the duality of convex

optimization problems, the dual formulation of the above has several nice properties that will be useful for our purposes. Robustness-based quantifiers were introduced in Refs. [37,65] as entanglement quantifiers and since then successfully applied in a wide range of QRTs. The above variant is closely related to the MDI-nonlocality robustness introduced in Ref. [14] at the level of probabilities (9). In particular, the two quantities are equivalent when the sets of input states $\{\omega_x\}$ and $\{\omega_y\}$ are tomographically complete. It is also worth mentioning that the quantity defined in Definition 2 is not a particular case of the robustness defined for general convex resource theories of measurements [26,66]. In particular, in Definition 2 the optimization is over all measurements $\{N_{ab}^{AB}\}$ and $\{O_{ab}^{AB}\}$ that have a quantum realization in the no-signaling scenario, whereas the quantifiers considered in [26] allow for arbitrary measurements (in particular, also those that require communication). In other words, the above general approach is valid only for measurements performed in a single location, whereas here we are explicitly interested in a distributed, multipartite scenario. Hence, our robustness measure is a genuinely different quantity than the generalized robustness of measurements studied in the above papers.

In Appendix A we derive the dual formulation of the RoBN, which will be used to study its operational characterization. Furthermore, we note that RoBN possesses the following three natural properties that one would expect from a reasonable measure of nonlocality.

(i) It is *faithful*, meaning that it vanishes if and only if the measurement is classical, i.e.,

$$\mathbf{R}_{BN}(\mathbb{M}^{AB}) = 0 \quad \Longleftrightarrow \quad \mathbb{M}^{AB} \in \mathscr{F}_{BN}. \quad (17)$$

(ii) It is *convex*, meaning that having access to two distributed measurements $\mathbb{M}_1^{AB}$ and $\mathbb{M}_2^{AB}$ one cannot obtain a better one by using them probabilistically, i.e., for $\mathbb{M}^{AB} = p\mathbb{M}_1 + (1-p)\mathbb{M}_2$ with $0 \leq p \leq 1$, we have

$$\mathbf{R}_{BN}(\mathbb{M}^{AB}) \leq p\mathbf{R}_{BN}(\mathbb{M}_1^{AB}) + (1-p)\mathbf{R}_{BN}(\mathbb{M}_2^{AB}). \quad (18)$$

(iii) It is *monotonic* (nonincreasing) under all quantum simulations. That is, if $\mathbb{N}^{AB}$ can be simulated by $\mathbb{M}^{AB}$ using some quantum simulation strategy (2) then

$$\mathbf{R}_{BN}(\mathbb{N}^{AB}) \leq \mathbf{R}_{BN}(\mathbb{M}^{AB}). \quad (19)$$

These properties were proven in Ref. [35] for a more general class of objects. For completeness, we give an independent proof in Appendix B.

Finally, we introduce a quantity that measures how much Buscemi nonlocality can be generated by using a

fixed shared state. In this way we define the *robustness of Buscemi nonlocality* of a state ρ_{AB} as

$$\mathbf{R}_{\mathrm{BN}}(\rho_{AB}) := \max_{\mathbb{M}^A, \mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}), \qquad (20)$$

where the optimization ranges over all local measurements on Alice's and Bob's side, $\mathbb{M}^{AB}$ is a distributed measurement of the form (1), and $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})$ is the robustness quantifier defined in Eq. (16). In this way the quantity from Eq. (20) is only a function of the *shared state*, rather than the whole *distributed measurement*. It quantifies the maximal "amount" of nonlocality of the corresponding behavior $\{p(a,b|\omega_x, \omega_y)\}$ that can be generated using a fixed $\rho^{A'B'}$, arbitrary local measurements $\mathbb{M}^A$, $\mathbb{M}^B$, and arbitrary input states $\{\omega_x\}, \{\omega_y\}$.

III. RESULTS

A. Operational characterization of RoBN

In the previous section we introduced a measure of Buscemi nonlocality quantifying how "close" a given measurement is to that which would arise from using only local measurements and shared randomness, i.e., a measurement of the form (11). In what follows we show that RoBN quantifies the advantage offered by a fixed distributed measurement over all classical measurements in a special type of state discrimination task relevant in the distributed scenario.

Let us now consider a task that is a special case of the no-signaling game described in Sec. II A. In this case we choose the function $V(a,b,x,y) = \delta_{ax}\delta_{by}$. This means that Alice and Bob win if they both manage to guess the values of $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ that were supplied to them by the referee. This is a variation of the standard state discrimination task in which a single player has to guess the realization of a single random variable x. Interestingly, due to the assumption that the players cannot communicate, distributed state discrimination cannot be reduced to the standard state discrimination task.

Task 1 (Distributed state discrimination). *The task consists of the following steps.*

(1) The referee chooses a bipartite state from the ensemble $\{p(x,y), \sigma_{xy}\}$ according to $p(x,y)$ and distributes it among parties by sending one part of it to Alice and the other part to Bob.

(2) After receiving their systems, Alice and Bob can preprocess them using arbitrary channels $\{\mathcal{E}_\lambda^A\}$ and $\{\mathcal{N}_\lambda^B\}$, potentially conditioned on a shared randomness λ.

(3) Alice and Bob apply fixed local measurements $\mathbb{M}^{AA'} = \{M_i^{AA'}\}$ and $\mathbb{M}^{B'B} = \{M_j^{AA'}\}$ to their shares of the state σ_{xy} and a part of the shared state $\rho^{A'B'}$.

They obtain outcomes i and j, respectively, which they can postprocess to produce their guesses a and b.

(4) Alice and Bob communicate their guesses a and b to the referee and win the game if they both correctly guess, i.e., when $a = x$ and $b = y$.

Note that the second and third steps can also be formulated as allowing Alice and Bob to apply any *quantum simulation* (3) to their distributed measurement $\mathbb{M}^{AB} \in \mathscr{R}_{\mathrm{BN}}$. Hence, the two players are effectively simulating a distributed measurement, denoted by $\mathbb{N}^{AB} \prec \mathbb{M}^{AB}$. The average probability of discriminating states in this discrimination game as specified by $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ can be expressed as

$$p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}, \mathbb{M}^{AB})$$
$$= \max_{\mathbb{N}^{AB} \prec_q \mathbb{M}^{AB}} \sum_{a,b,x,y} p(x,y)\mathrm{tr}[N_{ab}\sigma_{xy}]\delta_{xa}\delta_{yb}, \qquad (21)$$

where the optimization ranges over all measurements $\mathbb{N}^{AB} = \{N_{ab}\}$ that can be quantum simulated using $\mathbb{M}^{AB}$.

Let us now consider two different situations: (i) a classical scenario in which the distributed measurement performed by Alice and Bob is classical, i.e., $\mathbb{M}^{AB} \in \mathscr{F}_{\mathrm{BN}}$, and (ii) a quantum scenario in which the measurement performed by Alice and Bob is genuinely quantum, i.e., it cannot be written as in Eq. (11).

In the classical case (i) the optimal average probability of guessing which state from the ensemble $\{p(x,y), \sigma_{xy}\}$ was provided can be expressed as

$$p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}) = \max_{\mathbb{N}^{AB} \in \mathscr{F}_{\mathrm{BN}}} p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}, \mathbb{N}^{AB}). \qquad (22)$$

Note that the above optimization has to be performed over the convex set of measurements of the form (11), which is a subset of all separable measurements.

In the quantum case (ii) the above score can be further improved by exploiting Buscemi nonlocality contained in an entangled state that forms the distributed measurement $\mathbb{M}^{AB}$. The maximal amount by which quantum score outperforms classical score can be quantified by studying the ratio

$$\max_{\mathcal{G}} \frac{p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}, \mathbb{M}^{AB})}{p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G})}. \qquad (23)$$

In Appendix C we show that the maximal advantage that Alice and Bob can achieve when using $\mathbb{M}^{AB} \in \mathscr{R}_{\mathrm{BN}}$ over the best classical distributed measurement is precisely equal to the robustness of Buscemi nonlocality defined in Eq. (16). Formally, we have the following relation.

Result 1. *Let $\mathbb{M}^{AB} = \{M_{ab}^{AB}\}$ be a distributed measurement and let $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ be an ensemble of bipartite states. Then*

$$\max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^{AB})}{p_{\text{guess}}^{\text{DSD}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB}). \qquad (24)$$

This provides a direct operational meaning for Buscemi nonlocality. The proof of Result 1 consists of three parts. First, we use the primal formulation of problem (16) to show that the advantage from Eq. (24) is always upper bounded by the RoBN. Second, we identify a set of properties that characterize all distributed measurements and add them to the optimization problem (16) as superfluous constraints. Finally, using this characterization, we obtain a dual formulation of the problem that, after some simplifications, allows us to extract the optimal ensemble of states $\{p(x,y), \sigma_{xy}\}$ that achieves the optimum in Eq. (24). The full proof of this result is given in Appendix C.

The task of distributed state discrimination is a particular instance of a no-signaling game. In this respect we can further consider an advantage (23), with the average score $p_{\text{succ}}(\mathcal{G}, \mathbb{M}^{AB})$ given by Eq. (4), and optimize it over all ensembles $\mathcal{G}$ *and* scoring functions $V(a,b,x,y)$. This would allow us to find the largest possible advantage that can be achieved in any possible nonsignaling game. In this way Result 1 naturally leads to the following corollary.

Corollary 1. *Let $\mathbb{M}^{AB}$ and $\mathcal{G}$ be defined as above and let $V(a,b,x,y)\colon \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to [0,1]$. Then*

$$\max_{V,\mathcal{G}} \frac{p_{\text{guess}}^{V}(\mathcal{G}, \mathbb{M}^{AB})}{\max_{\sigma \in \text{SEP}} \max_{\mathbb{N}^{A}, \mathbb{N}^{B}} p_{\text{guess}}^{V}(\mathcal{G}, \mathbb{N}^{AB})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB}),$$
$$(25)$$

where the maximization in the denominator is over all distributed measurements $\mathbb{N}^{AB}$ that use a separable shared state $\sigma^{A'B'} \in \text{SEP}$. In this way we can also interpret RoBN as a quantifier of the Buscemi nonlocality contained within a given distributed measurement.

B. Connecting Buscemi nonlocality with other notions of nonclassicality

In this section we show that Buscemi nonlocality can be viewed as a type of nonlocality that is strictly stronger than two other well-known notions of nonlocal correlations: entanglement and nonclassical teleportation.

It is worth mentioning that the authors of Ref. [34] also studied the relationship between Buscemi nonlocality, nonclassical teleportation, and entanglement by studying a partial order between objects representing these resources: distributed measurements for Buscemi nonlocality. teleportation instruments for nonclassical teleportation, and

bipartite states for entanglement. Here we address an analogous problem using a more direct approach: we relate robustness quantifiers of these resource theories and find a direct and simple relationship between them.

Recall that a distributed measurement is composed of two local bipartite measurements and a shared state. This setting is very similar to the teleportation protocol in which Alice locally measures an input state provided by the referee and a part of an entangled state that she shares with Bob. Since the resource used in the teleportation task is effectively "contained" in the resource that is used in the task of distributed state discrimination, it is natural to ask if we can see some connection between these two tasks. In particular, how is the ability of performing nonclassical teleportation related to the ability of demonstrating Buscemi nonlocality? Furthermore, since teleportation is intrinsically related with entanglement [51], Buscemi nonlocality should also be quantitatively related to the entanglement content of a state. In the next section we show that in fact these three notions of nonclassical correlations are inherently connected and all describe different types of nonlocality.

1. Buscemi nonlocality and nonclassical teleportation

Quantum teleportation is one of the most important and thought-provoking discoveries in the whole quantum information theory. In the ideal version of the teleportation protocol proposed by Bennett *et al.* [67] two players, Alice and Bob, share a maximally entangled state. A third party, the referee, gives Alice an unknown quantum state. She then performs a Bell-state measurement on that system and her share of the entangled state and communicates her measurement result to Bob. With this new information, Bob applies an appropriate correcting unitary to his share of the entangled state, transforming it into the state that was initially given to Alice. This protocol can be naturally generalized to more realistic scenarios in which the shared entangled state and measurements performed by Alice are arbitrary.

A teleportation experiment can also be viewed as a way of testing nonlocality of a pair of objects: a state and measurement. In particular, the "teleportation resource" in that case is the teleportation channel or, more precisely, a collection of subchannels that form a *teleportation instrument* constructed using the shared state and Alice's measurement. Recall that an instrument $\mathbb{E} = \{\mathcal{E}_a\}$ for $\{a = 1, \dots, o_A\}$ is a collection of o_A completely positive and trace nonincreasing linear maps $\mathcal{E}_a$, so-called subchannels, such that $\sum_{a=1}^{o_A} \mathcal{E}_a$ is a channel. It was recently shown that the nonlocality present in a teleportation instrument can be exploited in several quantum-information theoretic tasks [59]. In order to relate nonclassical teleportation with Buscemi nonlocality, we first introduce the notion of a teleportation instrument.

Definition 3 (Teleportation instrument). *A teleportation instrument $\mathbb{A}^{A \to B'}$ from Alice to Bob is a collection of subchannels $\{\Lambda_a^{A \to B'}\}$ defined as*

$$\Lambda_a^{A \to B'}[\omega^A] = tr_{AA'}[(M_a^{AA'} \otimes \mathbb{1}^{B'})(\omega^A \otimes \rho^{A'B'})]. \quad (26)$$

The above notion fully captures the type of channel obtained during the generalized teleportation experiment. For some applications, it may be easier to work with states rather than subchannels. In that case for a collection of input states $\{\omega_x^A\}$ one can consider the so-called teleportation assemblages (teleportages) $\{\tau_{a|x}^{B'}\}$, where the elements of the assemblage are given by $\tau_{a|x}^{B'} := \Lambda_a^{A \to B'}[\omega_x^A]$.

Note that any teleportation instrument satisfies its own "no-signaling" constraint, which now reads $\sum_i \Lambda_i^{A \to B'}[\omega^A] = tr_{A'}[\rho^{A'B'}]$ for all input states ω^A. In fact, it can also be shown that teleportation instruments are the most general type of no-signaling instruments acting between two parties [59]. A teleportation instrument $\mathbb{A}^{A \to B'}$ is said to be classical (or free) if it describes a teleportation experiment performed using a separable shared state. We can find a general form of a classical teleportation instrument by taking $\rho^{A'B'} = \sum_\lambda p_\lambda \rho_\lambda^{A'} \otimes \rho_\lambda^{B'}$. The associated (classical) teleportation instrument reads

$$\Lambda_a^c(\omega_x) = \sum_\lambda p_\lambda tr_{AA'}[(M_a^{AA'} \otimes \mathbb{1}^B)(\omega_x^A \otimes \rho_\lambda^{A'} \otimes \rho_\lambda^{B'})]$$

$$= \sum_\lambda p_\lambda p(a|x, \lambda) \rho_\lambda^{B'}, \quad (27)$$

where $p(a|x, \lambda) = tr[M_a^{AA'}(\omega_x^A \otimes \rho_\lambda^{A'})]$. This is the most general classical teleportation scheme that can be realized if Alice and Bob have access only to classical randomness λ and the ability to locally prepare quantum states in their labs. In what follows we denote the set of all instruments that can be written as in Eq. (27) by $\mathscr{F}_T$. If a teleportation instrument cannot be written in this way, we refer to it as "nonclassical" and denote the set of all such instruments by $\mathscr{R}_T$. The quantity that quantitatively measures the amount of nonclassicality associated with a given teleportation instrument is called robustness of teleportation [51]. For a teleportation instrument $\mathbb{A}^{A \to B'} = \{\Lambda_a^{A \to B'}\}$, it is defined as

$$\mathbf{R}_T(\mathbb{A}^{A \to B'}) = \min_{r, \{\Gamma_a^{A \to B'}\}, \{\Omega_a^{A \to B'}\}} r \quad (28a)$$

such that $\quad \Lambda_a^{A \to B'} + r\Omega_a^{A \to B'} = (1+r)\Gamma_a^{A \to B'} \quad$ for all a, (28b)

$$\{\Gamma_a^{A \to B'}\} \in \mathscr{F}_T, \qquad \{\Omega_a^{A \to B'}\} \in \mathscr{R}_T. \quad (28c)$$

It turns out that the above is also a convex optimization problem that can be seen by formulating the constraints using the Choi-Jamiołkowski isomorphism (see

Appendix D for details). With the above notation we can now address our next result that relates Buscemi nonlocality with nonclassical teleportation.

Result 2. *Let $\mathbb{M}^{AB}$ be a distributed measurement composed of local bipartite measurements $\mathbb{M}^A$ and $\mathbb{M}^B$ and a shared state $\rho^{A'B'}$. Then*

$$\max_{\mathbb{M}^B} \mathbf{R}_{BN}(\mathbb{M}^{AB}) = \mathbf{R}_T(\mathbb{A}^{A \to B'}), \quad (29)$$

where the optimization is over all local measurements $\mathbb{M}^B = \{M_b^{B'\tilde{B}}\}$ for Bob. An analogous result holds for a teleportation instrument $\mathbb{A}^{B \to A'}$ if we instead optimize the left-hand side of Eq. (29) over all local measurements for Alice.

The proof of this result is given in Appendix D. Let us now use this result to show a new operational interpretation of the above teleportation quantifier.

Consider a task involving two players, Alice and Bob, who have access to a teleportation instrument $\mathbb{A}^{A \to B'}$ connecting their labs. Let the referee be in possession of an ensemble of bipartite quantum states $\mathcal{G} = \{p(x, y), \sigma_{xy}\}$. Just as before, the players may discuss their strategy before the game begins. This means that they may use a shared classical memory λ with a corresponding distribution $p(\lambda)$ and, conditioning on it, Alice may apply one of the channels $\{\mathcal{E}_\lambda^A\}$ to the input of the teleportation instrument and Bob may apply $\{\mathcal{N}_\lambda^{B'}\}$ to the output. The crucial difference here between the standard teleportation protocol is that Bob *does not* know Alice's measurement outcome and so his correction cannot depend on it. The task posed between Alice and Bob is the following.

Task 2 (Teleportation-assisted state discrimination [TSD]). *The task consists of the following steps.*

(1) The referee chooses a bipartite state from the ensemble $\mathcal{G} = \{p(x, y), \sigma_{xy}\}$ according to $p(x, y)$ and distributes it among parties by sending one part of it to Alice and the other part to Bob.

(2) Alice sends her part of the state to Bob using a teleportation instrument $\mathbb{A}^{A \to B'}$. She is also allowed to preprocess her part of the state conditioned on the classical randomness λ using a collection of channels $\{\mathcal{N}_\lambda^A\}$. Based on the outcome of the teleportation instrument i and potentially λ, she produces a guess a via $p(a|i, \lambda)$.

(3) Bob applies a correction $\{\mathcal{E}_\lambda^{B'}\}$ conditioned on the value of a shared random variable λ to the teleported state he received from Alice. He then measures both parts of the system using an arbitrary measurement $\mathbb{M}^B = \{M_b^{BB'}\}$ and produces a guess b.

(4) Alice and Bob win the game if they both simultaneously guess x and y.

The average probability of guessing in the above discrimination task can be expressed as

$$
\begin{aligned}
&p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{\Lambda}^{A \to B'}) \\
&= \max_{\mathbb{M}^B} \max_{\Phi \prec_q \mathbb{\Lambda}} \sum_{a,b,x,y} p(x,y) \text{tr}[M_b^{B'B}(\Phi_a^{A \to B'} \otimes \text{id}^B)\sigma_{xy}^{AB}]\delta_{xa}\delta_{yb},
\end{aligned}
$$
(30)

where the optimization ranges over all measurements $\mathbb{M}^B = \{M_b^{B'B}\}$ on Bob's side and all teleportation instruments $\Phi^{A \to B'} = \{\Phi_a^{A \to B'}\}$ that can be quantum simulated using the instrument $\mathbb{\Lambda}^{A \to B'} = \{\Lambda_i^{A \to B'}\}$. The elements of such a simulated instrument are of the form

$$
\Phi_a^{A \to B'}[\cdot] = \sum_{i,\lambda} p(\lambda)p(a|i,\lambda) \circ \mathcal{N}_\lambda^A \circ \Lambda_i^{A \to B'} \circ \mathcal{E}_\lambda^{B'}[\cdot]
$$
(31)

for some choice of local channels $\{\mathcal{E}_\lambda^{B'}\}$, $\{\mathcal{N}_\lambda^A\}$ and probabilities $p(a|i,\lambda)$, $p(\lambda)$.

The optimal average probability of guessing that can be achieved using only classical resources (i.e., a separable shared state, meaning that the teleportation instrument is classical) can be written as

$$
p_{\text{guess}}^{\text{TSD}}(\mathcal{G}) = \max_{\mathbb{F}^{A \to B'} \in \mathscr{F}_T} p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{F}^{A \to B'}),
$$
(32)

where $\mathbb{F}^{A \to B'}$ stands for a classical teleportation instrument from Alice to Bob. The maximal advantage that can be offered by any resourceful teleportation instrument $\mathbb{\Lambda}^{A \to B'}$ in the task of TSD is precisely equal to the quantifier of nonclassical teleportation defined in Eq. (28). This is captured by the following result.

Result 3. *Let $\mathbb{\Lambda}^{A \to B'} = \{\Lambda_a^{A \to B'}\}$ be a teleportation instrument from Alice to Bob and let $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ be an ensemble of bipartite states. Then the following statement holds:*

$$
\max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{\Lambda}^{A \to B})}{p_{\text{guess}}^{\text{TSD}}(\mathcal{G})} = 1 + \mathbf{R}_T(\mathbb{\Lambda}^{A \to B}).
$$
(33)

Proof. Consider maximizing both sides of Eq. (24) over all measurements $\mathbb{M}^B$ on Bob's side. Because of Result 2, the right-hand side of Eq. (24) is equal to $1 + \mathbf{R}_T(\mathbb{\Lambda}^{A \to B'})$. On the other hand, note that we can interchange maximization

over $\mathcal{G}$ with maximization over $\mathbb{M}^B$. Since $p_{\text{guess}}^{\text{DSD}}(\mathcal{G})$ does not depend on $\mathbb{M}^B$, the left-hand side of Eq. (24) becomes

$$
\begin{aligned}
\max_{\mathcal{G}} \frac{\max_{\mathbb{M}^B} p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^{AB})}{p_{\text{guess}}^{\text{DSD}}(\mathcal{G})} &= \max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{\Lambda}^{A \to B'})}{p_{\text{guess}}^{\text{DSD}}(\mathcal{G})} \\
&= \max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{\Lambda}^{A \to B'})}{p_{\text{guess}}^{\text{TSD}}(\mathcal{G})},
\end{aligned}
$$
(34)

where the last equality follows since

$$
\begin{aligned}
p_{\text{guess}}^{\text{DSD}}(\mathcal{G}) &= \max_{\mathbb{F}^{AB} \in \mathcal{F}_{\text{BN}}} p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{F}^{A \to B'}) \\
&= \max_{\mathbb{F}^{A \to B'} \in \mathcal{F}_T} \max_{\mathbb{M}^B} p_{\text{guess}}^{\text{TSD}}(\mathcal{G}, \mathbb{F}^{A \to B'}) \\
&= p_{\text{guess}}^{\text{TSD}}(\mathcal{G}).
\end{aligned}
$$
(35)

This completes the proof. ∎

2. Buscemi nonlocality and entanglement

Let us now explore the link between Buscemi nonlocality, which we defined as a property of a bipartite state and local measurements, and entanglement (a property of the state only). Among the large variety of known entanglement quantifiers [1,68–74], we are going to choose the one that most naturally relates to the RoBN—the so-called generalized robustness of entanglement (RoE), denoted here by $\mathbf{R}_E(\rho)$. This entanglement quantifier was considered for the first time in Ref. [37] and generalized in Ref. [65] and since then proved to be useful in several different contexts, e.g., in proving that all entangled states can demonstrate nonclassical teleportation [47], in exploring the connection between entanglement and permutation symmetry [75], or in studying the effects of local decoherence on multiparty entanglement [76]. This quantifier also has two interesting operational interpretations: it quantifies the maximal advantage that can be achieved in a bipartite subchannel discrimination task [77] and the maximal advantage in the task of local subchannel discrimination with a quantum memory [59]. It is defined in terms of the convex optimization problem

$$
\mathbf{R}_E(\rho^{AB}) = \min_{r, \eta^{AB}, \sigma^{AB}} r
$$
(36a)

such that $\quad \rho^{AB} + r\eta^{AB} = (1+r)\sigma^{AB}$, (36b)

$$
\eta^{AB} \geq 0, \qquad \text{tr}\eta^{AB} = 1
$$
(36c)

$$
\sigma^{AB} \in \mathsf{SEP}, \qquad \text{tr}\sigma^{AB} = 1.
$$
(36d)

Using this definition, we can now address our next result that relates Buscemi nonlocality with entanglement.

Result 4. *Let* $\mathbb{M}^{AB}$ *be a distributed measurement composed of local measurements* $\mathbb{M}^A$ *and* $\mathbb{M}^B$ *and a shared state* $\rho^{A'B'}$. *Then*

$$\max_{\mathbb{M}^A, \mathbb{M}^B} \mathbf{R}_{BN}(\mathbb{M}^{AB}) = \mathbf{R}_E(\rho^{A'B'}), \qquad (37)$$

where the optimization is over all local measurements for Alice $\mathbb{M}^A = \{M_a^{AA'}\}$ *and for Bob* $\mathbb{M}^B = \{M_b^{B'B}\}$.

The proof of this result is given in Appendix E. Note that the above relationship allows us to directly infer that the maximal amount of Buscemi nonlocality that can ever be generated using a given state, defined in Eq. (20), is precisely equal to its entanglement content. Therefore, we may write

$$\mathbf{R}_{BN}(\rho_{AB}) = \mathbf{R}_E(\rho^{AB}). \qquad (38)$$

Relationship (37) along with Result 1 also allows us to find a new operational interpretation of the RoE. Consider again the task of DSD with the relaxation that Alice and Bob may now apply arbitrary local measurements in their labs. The goal for Alice and Bob remains the same: to guess which state from the ensemble $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ was prepared, under the assumption that no communication is allowed. In this way the task posed between Alice and Bob is the following.

Task 3 (Entanglement-assisted state discrimination [ESD]). *The task consists of the following steps.*

(1) The referee chooses a bipartite state from the ensemble $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ *according to* $p(x,y)$ *and distributes it among parties by sending one part of it to Alice and the other part to Bob.*

(2) Alice and Bob apply arbitrary local measurements $\mathbb{M}^A$ *and* $\mathbb{M}^B$ *to the states they received and their part of the shared state* $\rho^{A'B'}$ *and receive outcomes a and b, respectively.*

(3) Alice and Bob win the game if they both guess which state was provided, i.e., guess both x and y.

The average probability of guessing in this task can be expressed as

$$p_{guess}^{ESD}(\mathcal{G}, \rho^{A'B'})$$
$$= \max_{\mathbb{M}^A, \mathbb{M}^B} \sum_{a,b,x,y} p(x,y) \mathrm{tr}[M_{ab}^{AB} \sigma_{xy}^{AB}] \delta_{xa} \delta_{yb}, \qquad (39)$$

where the optimization ranges over all measurements $\mathbb{M}^A = \{M_a^{AA'}\}$ on Alice's and $\mathbb{M}^B = \{M_b^{B'B}\}$ on Bob's side with measurement M_{ab}^{AB} of the form (1).

The best average probability of guessing in the classical scenario (i.e., when the shared state is separable) is given by

$$p_{guess}^{ESD}(\mathcal{G}) = \max_{\sigma^{A'B'} \in \mathsf{SEP}} p_{guess}^{ESD}(\mathcal{G}, \sigma^{A'B'})$$
$$= \max_{\mathbb{N}^{AB} \in \mathscr{F}_{BN}} p_{guess}^{ESD}(\mathcal{G}, \mathbb{N}^{AB})$$
$$= p_{guess}^{DSD}(\mathcal{G}). \qquad (40)$$

The maximal advantage that can be offered by an entangled state $\rho^{A'B'}$ in the ESD task can be quantified using the RoE. This is the content of our next result.

Result 5. *Let* $\rho^{A'B'}$ *be a bipartite state shared between Alice and Bob and let* $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ *be an ensemble of bipartite states. Then the following statement holds:*

$$\max_{\mathcal{G}} \frac{p_{guess}^{ESD}(\mathcal{G}, \rho^{A'B'})}{p_{guess}^{ESD}(\mathcal{G})} = 1 + \mathbf{R}_E(\rho^{A'B'}). \qquad (41)$$

Proof. The proof of Result 5 proceeds similarly to the case of nonclassical teleportation. Let us maximize both sides of Eq. (24) over all measurements on Alice's and Bob's side, i.e., over all $\mathbb{M}^A$ and $\mathbb{M}^B$. Because of Result 4, the right-hand side of Eq. (24) is equal to $1 + \mathbf{R}_E(\rho^{A'B'})$. On the other hand, due to Eq. (40) we can write the left-hand side of Eq. (24) as

$$\max_{\mathcal{G}} \frac{\max_{\mathbb{M}^A, \mathbb{M}^B} p_{guess}^{DSD}(\mathcal{G}, \mathbb{M}^{AB})}{p_{guess}^{DSD}(\mathcal{G})} = \max_{\mathcal{G}} \frac{p_{guess}^{ESD}(\mathcal{G}, \rho^{A'B'})}{p_{guess}^{ESD}(\mathcal{G})}. \qquad (42)$$

This completes the proof. ∎

Finally, let us note that entanglement-assisted state discrimination is a particular instance of a no-signaling game in which we fix $V(a,b,x,y) = \delta_{xa}\delta_{by}$ and allow for optimizing over local measurements. This exactly corresponds to the average score studied in Ref. [6]. Using this realization, we can now consider the maximal advantage in the task of entanglement-assisted state discrimination (41) and optimize it not only over ensembles $\mathcal{G}$, but also over all predicates $V(a,b,x,y)$, in a manner exactly similar as in the case of Corollary 1. This therefore yields the largest possible advantage that can be achieved in *any* no-signaling game. In this way Result 5 naturally leads to the following corollary.

Corollary 2. *Let* $\mathbb{M}^{AB}$ *and* $\mathcal{G}$ *be defined as before and let* $V(a,b,x,y)\colon \mathcal{A} \times \mathcal{B} \times \mathcal{X} \times \mathcal{Y} \to [0,1]$. *Then*

$$\max_{V, \mathcal{G}} \frac{\max_{\mathbb{M}^A, \mathbb{M}^B} p_{guess}^V(\mathcal{G}, \mathbb{M}^{AB})}{\max_{\sigma \in \mathsf{SEP}} \max_{\mathbb{N}^A, \mathbb{N}^B} p_{guess}^V(\mathcal{G}, \mathbb{N}^{AB})} = 1 + \mathbf{R}_E(\rho^{A'B'}), \qquad (43)$$

where $\mathbb{N}^{AB} = \{N_{ab}^{AB}\}$ *with the POVM elements defined as* $N_{ab} := tr_{A'B'}[(N_a^{AA'} \otimes N_b^{B'B})(\mathbb{1}^A \otimes \sigma^{A'B'} \otimes \mathbb{1}^B)].$

In this way we can now interpret the RoE as a quantifier of the Buscemi nonlocality contained within a given *state*. This not only rederives the main result of Ref. [6], but also makes it significantly stronger; the RoE can now be seen as the quantifier of the maximal advantage in any no-signaling game, therefore providing a completely new interpretation for this well-known entanglement quantifier.

3. Complete sets of monotones for quantum simulation

We finish this section by showing that the average guessing probability in the task of DSD completely describes the preorder induced by quantum simulation on distributed measurements $\mathbb{M}^{AB}$. Formally, this means that the average guessing probability $p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^{AB})$ when viewed as a function of $\mathcal{G}$ forms a complete set of monotones for quantum simulation of $\mathbb{M}^{AB}$. This is captured by the following result.

Result 6. *Any distributed measurement $\mathbb{M}^{AB}$ can quantum simulate another measurement $\mathbb{N}^{AB}$ if and only if, for all ensembles $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$, the following statement holds:*

$$p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^{AB}) \geq p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{N}^{AB}). \tag{44}$$

In other words, quantum simulation (or LOSR) can never improve the discrimination ability of any distributed measurement. The proof of this result is given in Appendix F.

C. RoBN as a quantifier in single-shot information theory

We now address another way of interpreting RoBN from the point of view of single-shot quantum information theory. In particular, in Appendix G we show that RoBN also quantifies the entanglement-assisted min-accessible information of a quantum-to-classical bipartite channel (i.e., a channel with quantum inputs and classical outputs). This connection parallels analogous results from the literature that correspond to single party quantum-to-classical channels [22,66].

We start by noting that any distributed measurement $\mathbb{M}^{AB}$ can be seen as an entanglement-assisted quantum-to-classical channel

$$\mathcal{N}^{AB \to XY}[\omega^A \otimes \omega^B]$$
$$= \sum_{a,b} p(a,b|\omega_x, \omega_y)|a\rangle\langle a|^X \otimes |b\rangle\langle b|^Y, \tag{45}$$

with $p(a,b|\omega_x, \omega_y)$ as in Eq. (9). In quantum information theory the standard quantifier of the maximal amount of classical information that can be reliably sent through a quantum channel is the accessible information that is defined for an arbitrary quantum channel $\mathcal{R}$ as

$$I^{\text{acc}}(\mathcal{R}) = \max_{\mathscr{E}, \mathscr{D}} I(X : G), \tag{46}$$

where $\mathscr{E} = \{p(x), \sigma_x\}$ is an ensemble of states that encode classical random variable X distributed according to $p(x)$, $\mathscr{D} = \{D_g\}$ is the decoding POVM that produces an outcome g with probability $p(g|x) := \text{tr}[D_g \mathcal{R}(\sigma_x)]$, and $I(X;G) = H(X) - H(X|G)$ is the mutual information of the distribution $p(x,g) := p(x)p(g|x)$. In the single-shot case a more relevant quantity is the *min-accessible information* $I_{\min}^{\text{acc}}(\mathcal{R})$ that is defined as [78]

$$I_{\min}^{\text{acc}}(\mathcal{R}) = \max_{\mathscr{E}, \mathscr{D}}[H_{\min}(X) - H_{\min}(X|G)], \tag{47}$$

where the optimization ranges over the same encodings and decodings as before and the single-shot entropies are given by [79]

$$H_{\min}(X) = -\log \max_x p(x), \tag{48}$$

$$H_{\min}(X|G) = -\log\left[\sum_g \max_x p(x,g)\right]. \tag{49}$$

Let us now consider an encoding of a bipartite random variable $X \times Y$, i.e., $\mathscr{E} = \{p(x,y), \sigma_{xy}\}$ and the associated decoding $\mathcal{D} = \{D_g\}$ for $g = 1, \ldots, |X| \cdot |Y|$. In Appendix G we show that, for this particular setting, RoBN quantifies the min-accessible information of the channel $\mathcal{N}^{AB \to XY}$. Formally, we have the following result.

Result 7. *Let $\mathcal{N}^{AB \to XY}$ be a quantum-to-classical channel of the form (45). Then the following statement holds:*

$$I_{min}^{acc}(\mathcal{N}^{AB \to XY}) = \log[1 + \mathbf{R}_{BN}(\mathbb{M}^{AB})]. \tag{50}$$

The proof of this result is given in Appendix G. The above result provides an alternative way of interpreting RoBN as the maximal amount of min-mutual information that can be obtained between the input and output of channel (45) when using it only once.

IV. CONCLUSIONS

In this work we have studied the notion of Buscemi nonlocality when it is formalized as a quantum resource theory

of distributed measurements. This formulation allows us to establish a direct operational interpretation of Buscemi nonlocality in terms of a practical information-theoretic task called distributed state discrimination (Result 1). We have shown that the average guessing probability in this task provides a complete set of monotones for the partial order of distributed measurements induced by quantum simulation (Result 6). This also gives rise to a simple and *complete* family of "Buscemi inequalities" that quantify nonlocal properties of distributed measurements.

This operational link is derived using a geometric quantity measuring the strength of nonlocal correlations generated using a given distributed measurement (RoBN). By connecting this quantifier with other measures of nonlocality we infer a quantitative relationship between distributed measurements, nonclassical teleportation, and quantum entanglement, a realization that we believe to be of independent interest. In particular, we have shown that the robustness of Buscemi nonlocality optimized over all local measurements for one party is equal to the robustness of nonclassical teleportation (Result 2). Similarly, optimizing RoBN over local measurements for both parties gives the robustness of entanglement (Result 4). This naturally leads to new operational interpretations for both of these quantifiers, in terms of appropriately tailored state discrimination tasks of teleportation-assisted state discrimination (Result 3) and entanglement-assisted state discrimination (Result 5).

We have also shown that the maximal amount of nonlocality that can ever be generated using a fixed bipartite state is directly proportional to its entanglement content. The entanglement content in this case is characterized by the robustness of entanglement, a widely known entanglement quantifier with direct operational significance. Importantly, this not only rederives the main result of Ref. [6], but also makes it significantly stronger; the generalized robustness of entanglement can now be seen as the quantifier of the maximal advantage in any no-signaling game (Corollaries 1 and 2).

As our last result, we have interpreted Buscemi nonlocality from the perspective of single-shot quantum information theory (Result 7). In particular, we have shown that Buscemi nonlocality, when viewed as a property of a communication channel between the sender (the referee) and receiver (Alice and Bob), quantifies the maximal amount of information that can be sent reliably when the channel is used only once (the so-called single-shot capacity of a quantum channel). We have shown that the RoBN can be viewed as the maximal single-shot capacity offered by a bipartite quantum-to-classical channel. This establishes an important link between Buscemi nonlocality and the single-shot theory of quantum communication.

Finally, we emphasize that while we focus exclusively on quantifying Buscemi nonlocality using a robustness-based measure, our results can be easily extended to address the so-called *weight-based* resource quantifiers [80,81]. These geometric measures find their operational meaning in the so-called exclusion tasks [24,82]. Consequently, the resource quantifiers of the weight of Buscemi nonlocality, the weight of nonclassical teleportation, and the weight of entanglement are quantifiers characterizing distributed state exclusion (DSE), teleportation-assisted state exclusion (TSE), and entanglement-assisted state exclusion (ESE), respectively.

We believe that the results presented in this work will shed new light on the complex structure of different types of nonclassical effects observed in nature, as well as on their practical relevance for physically motivated tasks.

This work also provides an example of a multiobject quantum resource theory that cannot be reduced to a theory of either measurements, states, channels, or state-measurement pairs [83]. This also means that the composite objects we study here constitute genuine multiobject quantum resources. It is an interesting open question to see if one can find additional examples of multiobject resource theories that address such irreducible resources. This is in sharp contrast to a recently introduced multiobject resource theory of state-measurement pairs, where the resources independently contribute to the benefit of the operational task of discrimination and exclusion of subchannels [83].

One of the standard questions addressed by quantum resource theories is determining when and at what rate a large number of copies of one resource can be converted into another. The fact that multiobject QRTs cannot be seen as resource theories of constituent objects leads to the natural question of whether this can be used to improve the existing asymptotic protocols. For example, in the resource theory of nonclassical teleportation one can ask whether n uses of the teleportation instrument can lead to a better teleportation than using n copies of the shared state. Similarly, we can ask whether access to n uses of a distributed measurement can be advantageous over using bipartite measurements and n copies of the shared state.

APPENDIX A: EQUIVALENT FORMULATION FOR THE ROBUSTNESS OF BUSCEMI NONLOCALITY

By definition, RoBN is a conic program. This means that we can use the tools of convex optimization theory to find its dual and from that obtain useful information about the primal problem. We assume knowledge of the tools of conic programming, and direct the interested reader to Ref. [62]. Let us start from the formulation given in the main text and substitute $\widetilde{N}_{ab}^{AB} = rN_{ab}^{AB}$ and $\widetilde{O}_{ab}^{AB} = (1+r)O_{ab}^{AB}$. After this substitution the primal problem can be written as

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \min r \tag{A1}$$

$$\text{such that} \quad M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} = \widetilde{O}_{ab}^{AB} \quad \text{for all } a, b, \tag{A2}$$

$$\{\widetilde{O}_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}, \qquad \{\widetilde{N}_{ab}^{AB}\} \in \mathscr{R}_{\mathrm{BN}}, \tag{A3}$$

where the optimization is performed over r, $\{\widetilde{N}_{ab}^{AB}\}$, and $\{\widetilde{O}_{ab}^{AB}\}$. Note that any collection of operators inside $\mathscr{R}_{\mathrm{BN}}$ or $\mathscr{F}_{\mathrm{BN}} = \mathscr{F}_{\mathrm{SEP}} \cap \mathscr{R}_{\mathrm{BN}}$ satisfies its own "no-signaling" constraint that can be easily deduced from the definition of the set $\mathscr{R}_{\mathrm{BN}}$. Moreover, any operator in $\mathscr{F}_{\mathrm{BN}}$ is separable. In this way, for any $\{X_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}$, we can write

$$\sum_a X_{ab}^{AB} = \mathbb{1}^A \otimes X_b^B \quad \text{for all } b, \qquad \sum_b X_{ab}^{AB} = X_a^A \otimes \mathbb{1}^B \quad \text{for all } a, \qquad X_{ab}^{AB} \in \mathsf{SEP}, \tag{A4}$$

$$\sum_b X_b^B = \mathbb{1}^B \quad \text{and} \quad \sum_a X_a^A = \mathbb{1}^A. \tag{A5}$$

Now we are going to add a family of such redundant constraints to our optimization problem. Note that we can always do that since adding constraints that are automatically satisfied by any operator in the feasible set does not change the optimal value of the program. Moreover, we can also relax constraint (A2) to an inequality $M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} \leq \widetilde{O}^{AB}$ without changing the optimal value of the conic program. To see why this is the case, suppose that we have solved the relaxed problem using variables r^{rel}, $\{\widetilde{N}_{ab}^{AB,\mathrm{rel}}\}$, $\{\widetilde{O}_{ab}^{AB,\mathrm{rel}}\}$, and $X_{ab}^{AB,\mathrm{rel}} \geq 0$ and such that, for all a and b, we have $M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,\mathrm{rel}} = \widetilde{O}_{ab}^{AB,\mathrm{rel}} - X_{ab}^{AB,\mathrm{rel}}$. Then the optimal value of the relaxed program becomes

$$\mathbf{R}_{\mathrm{BN}}{}^{\mathrm{rel}}(\mathbb{M}^{AB}) = -1 + \frac{1}{d^2} \sum_{ab} \mathrm{tr} \widetilde{O}_a^{AB,\mathrm{rel}} = -1 + \frac{1}{d^2} \sum_{ab} \mathrm{tr}[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,\mathrm{rel}} + X_{ab}^{AB,\mathrm{rel}}]$$

$$\geq -1 + \frac{1}{d^2} \sum_{ab} \mathrm{tr}[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB,\mathrm{rel}}]$$

$$\geq -1 + \frac{1}{d^2} \sum_{ab} \mathrm{tr}[M_{ab}^{AB} + \widetilde{N}_{ab}']$$

$$\geq -1 + \frac{1}{d^2} \sum_{ab} \mathrm{tr}[M_{ab}^{AB} + \widetilde{N}_{ab}^{AB}] = \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}), \tag{A6}$$

where $\{\widetilde{N}_{ab}'\}$ is a set of dual variables feasible for our initial problem (A1). In this way the conic program defining RoBN becomes

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{M}) = \min r \tag{A7}$$

$$\text{such that} \quad M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} \leq \widetilde{O}_{ab}^{AB} \quad \text{for all } a, b, \tag{A8}$$

$$\sum_a \widetilde{O}_{ab}^{AB} = \mathbb{1}^A \otimes \widetilde{O}_b^B \quad \text{for all } b, \qquad \sum_b \widetilde{O}_b^B = (1+r)\mathbb{1}^B, \tag{A9}$$

$$\sum_b \widetilde{O}_{ab}^{AB} = \widetilde{O}_a^A \otimes \mathbb{1}^B \quad \text{for all } a, \qquad \sum_a \widetilde{O}_a^A = (1+r)\mathbb{1}^A, \tag{A10}$$

$$\{\widetilde{O}_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}, \qquad O_{ab}^{AB} \in \mathsf{SEP}, \qquad \{\widetilde{N}_{ab}^{AB}\} \in \mathscr{R}_{\mathrm{BN}}, \quad \text{for all } a, b, \tag{A11}$$

where the minimization is performed over r, $\{\widetilde{O}_{ab}^{AB}\}$, $\{\widetilde{O}_a^A\}$, $\{\widetilde{O}_b^B\}$, and $\{\widetilde{N}_{ab}^{AB}\}$.

In what follows we denote a dual cone to $\mathscr{R}$ using $\mathscr{R}^*$, that is, $\mathscr{R}^* := \{X \,|\, \mathrm{tr} XQ \geq 0 \text{ for all } Q \in \mathscr{R}\}$. We now write the dual formulation of the above problem. To do so, we first write the associated Lagrangian using dual Hermitian variables associated with a corresponding set of constraints: $\{A_{ab}^{AB}\}$ such that $A_{ab}^{AB} \geq 0$ for all a, b, $\{B_b^{AB}\}$, $\{C_a^{AB}\}$, $D^A \geq 0$, $E^B \geq 0$, $\{F_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}^*$, meaning that $\sum_{ab} \mathrm{tr}[F_{ab}^{AB} X_{ab}^{AB}] \geq 0$ for all $\{X_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}$, $G_{ab}^{AB} \in \mathscr{F}_{\mathrm{SEP}}^*$ for all a, b, meaning that $\mathrm{tr}[G_{ab}^{AB} X^{AB}] \geq 0$ for all a, b and all separable operators $X^{AB} \in \mathscr{F}_{\mathrm{SEP}}$, and, finally, $\{H_{ab}^{AB}\} \in \mathscr{R}_{\mathrm{BN}}^*$. With this, the Lagrangian function of the conic program (A7—A11) becomes

$$
\begin{aligned}
\mathcal{L} = {} & r + \sum_{ab} \mathrm{tr} A_{ab}^{AB} [M_{ab}^{AB} + \widetilde{N}_{ab}^{AB} - \widetilde{O}_{ab}^{AB}] + \sum_b \mathrm{tr} B_b^{AB} \left[\sum_a \widetilde{O}_{ab}^{AB} - \mathbb{1}^A \otimes \widetilde{O}_b^B \right] \\
& + \sum_a \mathrm{tr} C_a^{AB} \left[\sum_b \widetilde{O}_{ab}^{AB} - \widetilde{O}_a^A \otimes \mathbb{1}^B \right] + \mathrm{tr} D^A \left[\sum_a \widetilde{O}_a^A - (1+r)\mathbb{1}^A \right] \\
& + \mathrm{tr} E^B \left[\sum_b \widetilde{O}_b^B - (1+r)\mathbb{1}^B \right] - \sum_{a,b} \mathrm{tr}[F_{ab}^{AB} \widetilde{O}_{ab}^{AB}] - \sum_{a,b} \mathrm{tr}[G_{ab}^{AB} \widetilde{O}_{ab}^{AB}] - \sum_{a,b} \mathrm{tr}[H_{ab}^{AB} \widetilde{N}_{ab}^{AB}] \\
= {} & r[1 - \mathrm{tr} D^A - \mathrm{tr} E^B] + \sum_{a,b} \mathrm{tr} \widetilde{N}_{ab}[A_{ab}^{AB} - H_{ab}^{AB}] + \sum_{a,b} \mathrm{tr} \widetilde{O}_{ab}[-A_{ab}^{AB} + B_b^{AB} + C_a^{AB} - F_{ab}^{AB} - G_{ab}^{AB}] + \sum_a \mathrm{tr} O_a^A [D^A - C_a^A] \\
& + \sum_b \mathrm{tr} O_b^B [E^B - B_b^B] + \sum_{ab} \mathrm{tr}[A_{ab}^{AB} M_{ab}^{AB}] - \mathrm{tr} D^A - \mathrm{tr} E^B .
\end{aligned}
\tag{A12}
$$

By demanding that the terms in the square brackets that appear along with the dual variables vanish we can ensure that $\mathcal{L} \leq r$. This leads to the (dual) conic program

$$
\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \max \sum_{ab} \mathrm{tr}[A_{ab}^{AB} M_{ab}^{AB}] - 1
\tag{A13}
$$

such that

$$
C_a^{AB} + B_b^{AB} = A_{ab}^{AB} + F_{ab}^{AB} + G_{ab}^{AB} \quad \text{for all } a, b,
\tag{A14}
$$

$$
A_{ab}^{AB} = H_{ab}^{AB} \quad \text{for all } a, b, \qquad C_a^A = D^A \quad \text{for all } a, \qquad B_b^B = E^B \quad \text{for all } b,
\tag{A15}
$$

$$
A_{ab}^{AB} \geq 0 \quad \text{for all } a, b, \qquad \{H_{ab}^{AB}\} \in \mathscr{R}_{\mathrm{BN}}^*, \qquad \{F_{ab}^{AB}\} \in \mathscr{F}_{\mathrm{BN}}^*, \qquad \mathrm{tr} D^A + \mathrm{tr} E^B = 1.
$$

Note now that the set $\mathscr{F}_{\mathrm{BN}} \in \mathscr{F}_{\mathrm{SEP}}$, which implies that the dual sets satisfy $\mathscr{F}_{\mathrm{SEP}}^* \in \mathscr{F}_{\mathrm{BN}}^*$. Hence, without loss of generality, we can assume that $G_{ab}^{AB} = 0$ for all a and b. In this way we can express the above program in the following way:

$$
1 + \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \max \sum_{ab} \mathrm{tr}[A_{ab}^{AB} M_{ab}^{AB}]
\tag{A16}
$$

$$
\text{such that} \quad C_a^{AB} + B_b^{AB} - A_{ab}^{AB} = F_{ab}^{AB} \in \mathscr{F}_{\mathrm{BN}}^* \quad \text{for all } a, b,
\tag{A17}
$$

$$
C_a^A = D^A \quad \text{for all } a, \qquad C_a^A, D^A \geq 0 \quad \text{for all } a,
\tag{A18}
$$

$$
A_{ab}^{AB} \geq 0 \quad \text{for all } a, b, \qquad \mathrm{tr} D^A + \mathrm{tr} E^B = 1.
\tag{A19}
$$

Using both primal (A1) and dual (A16) formulations, we can now describe some basic properties of the RoBN.

APPENDIX B: BASIC PROPERTIES OF THE RoBN

Here we prove the three basic properties of RoBN highlighted in the main text.

1. Faithfulness

If $\mathbb{M}^{AB} \in \mathscr{F}_{\mathrm{BN}}$ then we can always choose a feasible $r = 0$ in the primal form (A1). Since the solution is always nonnegative, $r = 0$ is also optimal.

2. Convexity

Let $\{N_{ab}^1, O_{ab}^1\}$ be optimal primal variables for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^1)$ and similarly let $\{N_{ab}^2, O_{ab}^2\}$ be primal optimal for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^2)$. Define $\mathbb{M}' = \{M_{ab}'\}$ as a convex combination of the two measurements, that is, $M_{ab}' = p\, M_{ab}^1 + (1-p)\, M_{ab}^2$ for each a and b. We can construct a set of feasible variables for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}')$ in the following way: $N_{ab}' = p\, N_{ab}^1 + (1-p)\, N_{ab}^2$ and $O_{ab}' = p\, O_{ab}^1 + (1-p)\, O_{ab}^2$. Substituting N_{ab}' and O_{ab}' into the constraints of the primal form for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M})$ shows that this choice is feasible. In this way we obtain an upper bound on $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}')$:

$$\mathbf{R}_{\mathrm{BN}}(\mathbb{M}') \leq \mathrm{tr}\sum_{a,b} N_{ab}' = p\,\mathrm{tr}\sum_{a,b} N_{ab}^1 + (1-p)\,\mathrm{tr}\sum_{a,b} N_{ab}^2 = p\,\mathcal{R}(\mathbb{M}^1) + (1-p)\,\mathcal{R}(\mathbb{M}^2). \tag{B1}$$

3. Monotonicity

Let us start with the assumption that there is a subroutine

$$\mathscr{S} = \{p(\lambda), p(a|i,\lambda), p(b|j,\lambda), \mathcal{E}_\lambda, \mathcal{N}_\lambda\},$$

which allows us to simulate $\mathbb{M}'$ using $\mathbb{M}$, i.e., $\mathbb{M} \succ_q \mathbb{M}'$. This means that the POVM elements $\{M_{ab}\}$ of $\mathbb{M}$ can be mapped into

$$M_{ab}' = \sum_{i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(\mathcal{E}_\lambda^\dagger \otimes \mathcal{N}_\lambda^\dagger)[M_{ij}].$$

Suppose now that we solved the dual problem for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}')$ using the optimal dual variables $\{A_{ab}'\}, \{B_b'\}, \{C_a'\}, D', E'$, and $\{F_{ab}'\}$. Using these, we construct an educated guess for $\mathbf{R}_{\mathrm{BN}}(\mathbb{M})$ in the following way:

$$A_{ij}^* = \sum_{a,b,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[A_{ab}'], \qquad B_j^* = \sum_{b,\lambda} p(\lambda)p(b|j,\lambda)(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[B_b'], \tag{B2}$$

$$C_i^* = \sum_{a,\lambda} p(\lambda)p(a|i,\lambda)(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[C_a'], \qquad D^* = \sum_\lambda p(\lambda)\mathcal{E}_\lambda[D'], \tag{B3}$$

$$E^* = \sum_\lambda p(\lambda)\,\mathcal{N}_\lambda[E'], \qquad F_{ij}^* = \sum_{a,b,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[F_{ab}']. \tag{B4}$$

It can be verified that the above choice of variables is feasible for dual problem (A16). In particular, note that by construction we have $C_i^* + B_j^* - A_{ij}^* = F_{ij}^*$ for all i,j since the primed dual variables satisfy the constraints of dual problem (A16). Furthermore, since $\mathrm{tr}_B(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[X^{AB}] = \mathcal{E}_\lambda[X^A]$, we can infer that $\mathrm{tr}_B C_i^* = D^*$ and $\mathrm{tr}_A B_j^* = E^*$. Moreover, as separable maps preserve both positivity and separability, we also have $A_{ij}^* \geq 0$ for all i,j and $\{F_{ij}^*\} \in \mathscr{F}_{\mathrm{BN}}^*$. Using the proposed set of dual variables, we find the lower bound

$$1 + \mathbf{R}_{\mathrm{BN}}(\mathbb{M}) \geq \sum_{i,j} \mathrm{tr}[M_{ij} A_{ij}^*]$$

$$= \sum_{a,b,i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)\mathrm{tr}\{M_{ij}\,(\mathcal{E}_\lambda \otimes \mathcal{N}_\lambda)[A_{ab}]\}$$

$$= \sum_{a,b,i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)\mathrm{tr}\{(\mathcal{E}_\lambda^\dagger \otimes \mathcal{N}_\lambda^\dagger)[M_{ij}]\,A_{ab}\}$$

$$= \sum_{a,b} \mathrm{tr}[M_{ab} A_{ab}]$$

$$= 1 + \mathbf{R}_{\mathrm{BN}}(\mathbb{M}'). \tag{B5}$$

This proves that RoBN is monotonic under quantum simulation.

APPENDIX C: PROOF OF RESULT 1

In this section we prove that RoBN can be seen as a quantifier of the advantage a given distributed measurement provides in the task of distributed state discrimination. To simplify the notation in this section, we omit subsystem labels whenever it is clear from the context. Let us recall that the average guessing probability in the task of distributed state discrimination using a distributed measurement $\mathbb{M}$ can be expressed as

$$p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}) = \max_{\mathbb{N} \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y)\text{tr}[N_{ab}\,\sigma_{xy}]\delta_{xa}\delta_{yb}, \tag{C1}$$

where the optimization ranges over all measurements $\mathbb{N} = \{N_{ab}\}$ that can be quantum simulated using $\mathbb{M} = \{M_{ij}\}$, where

$$M_{ij} = \text{tr}_{AB}[(M_i^{A'A} \otimes M_j^{BB'})(\mathbb{1}^{A'} \otimes \rho^{AB} \otimes \mathbb{1}^{B'})] \tag{C2}$$

is a distributed measurement and $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ is an ensemble of bipartite states. Suppose that we have solved the dual problem for RoBN (A16) using the set of dual variables $\{A_{ab}\}$, $\{C_a\}$, $\{B_b\}$, D, E, and $\{G_{ab}\}$. Note also that, due to the constraints in RoBN (A16), the matrix A_{ab} is positive semidefinite for all values of a and b. Let us now consider a particular game setting $\mathcal{G}^* = \{p^*(x,y), \sigma_{xy}^*\}$ defined as

$$C = \sum_{x,y} \text{tr}A_{xy}, \qquad p^*(x,y) = \frac{\text{tr}A_{xy}}{C}, \qquad \sigma_{xy}^* = \frac{A_{xy}}{\text{tr}A_{xy}}, \tag{C3}$$

where $x = 1, \ldots, o_A$, $y = 1, \ldots, o_B$, and o_A, o_B are the numbers of outcomes of local measurements performed by A and B. The best average guessing probability that can be achieved in the game $\mathcal{G}^*$ using a distributed measurement $\mathbb{M}$ is given by

$$\begin{aligned}
p_{\text{guess}}^{\text{DSD}}(\mathcal{G}^*, \mathbb{M}) &= \max_{\mathbb{N} \prec_q \mathbb{M}} \sum_{a,b,x,y} p^*(x,y)\text{tr}[N_{ab}\sigma_{xy}^*]\delta_{xa}\delta_{yb} \\
&\geq \sum_{x,y} \frac{\text{tr}A_{xy}}{C} \text{tr}\left[M_{xy}\frac{A_{xy}}{\text{tr}A_{xy}}\right] \\
&= \frac{1}{C}\sum_{x,y}\text{tr}[M_{xy}A_{xy}] \\
&= \frac{1}{C}[1 + \mathbf{R}_{\text{BN}}(\mathbb{M}^{AB})],
\end{aligned} \tag{C4}$$

where the inequality in the second line follows from choosing a particular subroutine $\mathscr{S}$ with $p(\lambda) = 1/|\lambda|$, $p(a|i,\lambda) = \delta_{ai}$, $p(b|j,\lambda) = \delta_{bj}$, and $\mathcal{E}_\lambda = \mathcal{N}_\lambda = \text{id}$. Let us now look at the corresponding classical (i.e., without access to entanglement)

probability of guessing:

$$
\begin{aligned}
p_{\text{guess}}^{\text{DSD}}(\mathcal{G}^*) &= \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} p_{\text{guess}}^{\text{DSD}}(\mathcal{G}^*, \mathbb{N}) = \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} \sum_{x,y} p^*(x,y) \text{tr}[N_{xy} \sigma_{xy}^*] \\
&= \frac{1}{C} \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} \sum_{x,y} \text{tr}[N_{xy} A_{xy}] \\
&= \frac{1}{C} \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} \sum_{x,y} \text{tr}[N_{xy}(C_x + B_y - F_{xy})] \\
&= \frac{1}{C} \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} \left(\sum_{x} \text{tr}[(N_x \otimes \mathbb{1}) C_x] + \sum_{y} \text{tr}[(\mathbb{1} \otimes N_y) B_y] - \sum_{x,y} \text{tr}[N_{xy} F_{xy}] \right) \\
&\leq \frac{1}{C} \max_{\mathbb{N} \in \mathscr{F}_{\text{BN}}} \left(\sum_{x} \text{tr}[N_x D] + \sum_{y} \text{tr}[N_y E] \right) \\
&= \frac{1}{C} (\text{tr} D + \text{tr} E) \\
&= \frac{1}{C},
\end{aligned}
\tag{C5}
$$

where the inequality follows since, for all $\mathbb{N} \in \mathscr{F}_{\text{BN}}$, we have $\sum_{xy} \text{tr}[N_{xy} F_{xy}] \geq 0$. Combining bounds (C4) and (C5) leads to

$$
\max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M})}{p_{\text{class}}^{\text{DSD}}(\mathcal{G})} \geq \frac{p_{\text{guess}}^{\text{DSD}}(\mathcal{G}^*, \mathbb{M})}{p_{\text{class}}^{\text{DSD}}(\mathcal{G}^*)} \geq 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}).
\tag{C6}
$$

In order to prove the upper bound, note that the first line of constraints in the primal formulation for RoBN (A1) implies that, for all a, b,

$$
M'_{ab} = \widetilde{O}'_{ab} - \widetilde{N}'_{ab},
\tag{C7}
$$

where $\widetilde{O}'_{ab} = [1 + \mathbf{R}_{\text{BN}}(\mathbb{M})] O'_{ab}$ for all a, b and $\{O'_{ab}\} \in \mathscr{F}_{\text{BN}}$. This allows us to write

$$
\begin{aligned}
p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}) &= \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y) \text{tr}[M'_{ab} \sigma_{xy}] \delta_{xa} \delta_{yb} \\
&= \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y) \text{tr}[(\widetilde{O}'_{ab} - \widetilde{N}'_{ab}) \sigma_{xy}] \delta_{xa} \delta_{yb} \\
&\leq \max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b,x,y} p(x,y) \text{tr}[\widetilde{O}'_{ab} \sigma_{xy}] \delta_{xa} \delta_{yb} \\
&= \max_{\mathbb{M}' \prec_q \mathbb{M}} [1 + \mathbf{R}_{\text{BN}}(\mathbb{M}')] \sum_{a,b,x,y} p(x,y) \text{tr}[O'_{ab} \sigma_{xy}] \delta_{xa} \delta_{yb} \\
&\leq \left\{ \max_{\mathbb{M}' \prec_q \mathbb{M}} [1 + \mathbf{R}_{\text{BN}}(\mathbb{M}')] \right\} \left\{ \max_{\{O_{ab}\} \in \mathscr{F}_{\text{BN}}} \sum_{a,b,x,y} p(x,y) \text{tr}[O_{ab} \sigma_{xy}] \delta_{xa} \delta_{yb} \right\} \\
&\leq [1 + \mathbf{R}_{\text{BN}}(\mathbb{M})] p_{\text{guess}}^{\text{DSD}}(\mathcal{G}),
\end{aligned}
\tag{C8}
$$

where the last inequality follows from the monotonicity of RoBN under quantum simulation. Combining bounds (C6) and (C8) yields

$$
\max_{\mathcal{G}} \frac{p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M})}{p_{\text{guess}}^{\text{DSD}}(\mathcal{G})} = 1 + \mathbf{R}_{\text{BN}}(\mathbb{M}).
\tag{C9}
$$

APPENDIX D: PROOF OF RESULT 2

Before proving the result we recall the primal and dual formulation of the RoT quantifier. Let $\mathbb{\Lambda} = \{\Lambda_a\}$ be a teleportation instrument whose elements are defined as

$$\Lambda_a^{A \to B'}[\omega] := \mathrm{tr}_{AA'}[(M_a^{AA'} \otimes \mathbb{1}^B)(\omega^A \otimes \rho^{A'B'})] \tag{D1}$$

for some measurement $M_a^{AA'}$ and a shared state $\rho^{A'B'}$. We denote the set of Choi-Jamiołkowski states corresponding to these subchannels by $\{J_a^{VB'}\}$, i.e., each $J_a^{VB'} := (\mathrm{id}^V \otimes \Lambda_a^{A \to B'})[\phi_+^{VA}]$ with system V isomorphic to A. With these definitions, the RoT for a teleportation instrument $\mathbb{\Lambda}^{A \to B'}$ can be written as

$$\mathbf{R}_T(\mathbb{\Lambda}^{A \to B'}) = \min \mathrm{tr}\tilde{\sigma}^{B'} \qquad \Longleftrightarrow \qquad \max \sum_a \mathrm{tr}[A_a^{VB'} J_a^{VB'}] - 1$$

$$\text{such that} \quad J_a^{VB'} \le F_a^{VB'} \quad \text{for all } a, \qquad\qquad B^{VB'} - A_a^{VB'} = W_a^{VB'} \in \mathscr{F}_{\mathrm{SEP}}^* \quad \text{for all } a,$$

$$\sum_a F_a^{VB'} = \frac{\mathbb{1}^V}{d} \otimes \tilde{\sigma}^{B'}, \qquad\qquad B^{B'} = \mathbb{1}^{B'}, \quad A_a^{VB'} \ge 0 \quad \text{for all } a,$$

$$F_a^{VB'} \in \mathscr{F}_{\mathrm{SEP}} \quad \text{for all } a, \quad \tilde{\sigma}^{B'} \ge 0.$$

Let us now proceed with the proof of Result 2.

Proof of Result 2. As before, the proof consists of two steps. First we show that $\mathbf{R}_T(\mathbb{\Lambda}^{A \to B'})$ lower bounds $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})$ for a particular choice of local measurement $\mathbb{M}^{B'B}$. Then we show that, for any choice of local measurements on Bob's side, $\mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})$ is never larger than the teleportation quantifier $\mathbf{R}_T(\mathbb{\Lambda}^{A \to B'})$.

Let $A_a^{VB'} \ge 0$, $W_a^{VB'} \in \mathscr{F}_{\mathrm{SEP}}^*$, and $B^{VB'}$ be optimal dual variables for $\mathbf{R}_T(\mathbb{\Lambda}^{A \to B'})$. Let $\{U_b^B\}$ for $b \in \{1, \ldots, d^2\}$ be a set of Pauli operators with respect to a basis $\{|i\rangle^B\}$. Consider the following measurement with $o_B = d^2$ outcomes:

$$M_b^{B'B} = (\mathrm{id}^{B'} \otimes \mathcal{U}_b^B)[\phi_+^{B'B}]. \tag{D2}$$

Here $\mathcal{U}_b[\cdot] := U_b(\cdot)U_b^\dagger$. We are interested in the lower bound for $\mathbf{R}_T(\mathbb{\Lambda}^{A \to B'})$. Let us choose a set of dual variables in Eq. (A16) inspired by the optimal dual variables for $\mathbf{R}_T(\mathbb{\Lambda}^{V \to B'})$:

$$A_{ab}^{AB} = [\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B][(A_a^{AB})^T], \quad F_{ab}^{AB} = [\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B][(W_a^{AB})^T], \quad B_b^{AB} = \frac{1}{d}[\mathrm{id}^A \otimes (\mathcal{U}_b^\dagger)^B][(B^{AB})^T],$$

$$C_a^{AB} = 0, \qquad D^B = \frac{1}{d}\mathbb{1}^B, \qquad E^A = 0.$$

It can be verified by direct substitution that the above choice is feasible. In particular, the above choice for $\{F_{ab}^{AB}\}$ is feasible as $\mathscr{F}_{\mathrm{SEP}}^{*} \in \mathscr{F}_{\mathrm{BN}}^{*}$ and both sets are invariant under local unitaries. This leads to the following chain of inequalities:

$$
\begin{aligned}
1 + \max_{\mathbb{M}^{B}} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) &\geq \sum_{ab} \mathrm{tr}[A_{ab}^{AB} M_{ab}^{AB}] \\
&= \sum_{ab} \mathrm{tr}\{[\mathrm{id}^{A} \otimes (\mathcal{U}_{b}^{\dagger})^{B}][(A_{a}^{AB})^{T}] \mathrm{tr}_{A'B'}[(M_{a}^{AA'} \otimes M_{b}^{B'B})(\mathbb{1}^{A} \otimes \rho^{A'B'} \otimes \mathbb{1}^{B})]\} \\
&= \sum_{ab} \mathrm{tr}\{[\mathrm{id}^{A} \otimes (\mathcal{U}_{b}^{\dagger})^{B}][(A_{a}^{AB})^{T}] \mathrm{tr}_{A'B'}[(M_{a}^{AA'} \otimes (\mathrm{id}^{B'} \otimes \mathcal{U}_{b}^{B})[\phi_{+}^{B'B}])(\mathbb{1}^{A} \otimes \rho^{A'B'} \otimes \mathbb{1}^{B})]\} \\
&= \sum_{ab} \mathrm{tr}\{\mathbb{1}^{A'B'} \otimes [\mathrm{id}^{A} \otimes (\mathcal{U}_{b}^{\dagger})^{B}][(A_{a}^{AB})^{T}][(M_{a}^{AA'} \otimes (\mathrm{id}^{B'} \otimes \mathcal{U}_{b}^{B})[\phi_{+}^{B'B}])(\mathbb{1}^{A} \otimes \rho^{A'B'} \otimes \mathbb{1}^{B})]\} \\
&= \sum_{ab} \mathrm{tr}\{[\mathbb{1}^{A'B'} \otimes (A_{a}^{AB})^{T}](M_{a}^{AA'} \otimes \phi_{+}^{B'B})(\mathbb{1}^{A} \otimes \rho^{A'B'} \otimes \mathbb{1}^{B})\} \\
&= \frac{1}{d^{2}} \sum_{ab} \mathrm{tr}\{A_{a}^{VB'} \, \mathrm{tr}_{AA'}[(\mathbb{1}^{V} \otimes M_{a}^{AA'} \otimes \mathbb{1}^{B'})(\phi_{+}^{VA} \otimes \rho^{A'B'})]\} \\
&= \sum_{a} \mathrm{tr}[A_{a}^{VB'} J_{a}^{VB'}] \\
&= 1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{V \to B'}).
\end{aligned}
\tag{D3}
$$

We now prove the upper bound. Note that, for any distributed measurement $\mathbb{M}^{AB}$, we can construct $\mathbb{M}^{VB} := \{M_{ab}^{VB}\}$ such that $M_{ab}^{VB} := d \mathrm{tr}_{A}[(\mathbb{1}^{V} \otimes M_{ab}^{AB})(\phi_{+}^{VA} \otimes \mathbb{1}^{B})]$. This in turn can be written as

$$
M_{ab}^{VB} := d \mathrm{tr}_{AA'B'}[(\mathbb{1}^{A} \otimes M_{a}^{AA'} \otimes M_{b}^{B'B})(\phi_{+}^{VA} \otimes \rho^{A'B'} \otimes \mathbb{1}^{B})] = d \mathrm{tr}_{B'}[(\mathbb{1}^{V} \otimes M_{b}^{B'B})(J_{a}^{VB'} \otimes \mathbb{1}^{B})].
\tag{D4}
$$

Note that we can always write $J_{a}^{VB'} \leq [1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})] F_{a}^{VB'}$, where $\{F_{a}^{VB'}\}$ are Choi-Jamiołkowski operators of some classical teleportation instrument. This allows us to further rewrite Eq. (D4) as

$$
M_{ab}^{VB} \leq d[1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})] \mathrm{tr}_{A}[(\mathbb{1}^{V} \otimes M_{ab}^{AB})(\phi_{+}^{VA} \otimes \mathbb{1}^{B})] = [1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})] O_{ab}^{VB},
\tag{D5}
$$

where $\{O_{ab}^{VB}\}$ is a free-distributed measurement. Hence, $M_{ab}^{AB} \leq [1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})] O_{ab}^{AB}$ for some free-distributed measurement $\{O_{ab}^{AB}\}$ also. This finally allows us to write

$$
\max_{\mathbb{M}^{B}} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) \leq [1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})] \max_{\mathbb{M}^{B}} \sum_{ab} \mathrm{tr}[A_{ab}^{AB} O_{ab}^{AB}] \leq [1 + \mathbf{R}_{T}(\mathbb{\Lambda}^{A \to B'})].
\tag{D6}
$$

This proves the lemma.   ∎

APPENDIX E: PROOF OF RESULT 4

Let us recall that the conic program formulation of RoE is given by

$$
\mathbf{R}_{E}(\rho^{A'B'}) = \min \mathrm{tr} \widetilde{\sigma}^{A'B'} \qquad \Longleftrightarrow \qquad \max \sum_{a} \mathrm{tr}[A^{A'B'} \rho^{A'B'}] - 1
\tag{E1}
$$

$$
\text{such that} \quad \rho^{A'B'} \leq \widetilde{\sigma}^{A'B'}, \qquad\qquad\qquad \mathbb{1}^{A'B'} - A^{A'B'} = W^{A'B'} \in \mathscr{F}_{\mathrm{SEP}}^{*},
$$

$$
\widetilde{\sigma}^{A'B'} \in \mathscr{F}_{\mathrm{SEP}}, \qquad\qquad\qquad\qquad A^{A'B'} \geq 0.
$$

The proof is based on three parts. First we use Result 2 to connect RoBN to the RoT. Then we essentially parallel the steps taken in the proof of Result 2 to link the RoT to the RoE. It is worth mentioning that the link between the RoT and RoE was obtained some time ago in Ref. [51]. Here, for convenience, we state an independent proof.

Proof. Let us begin by noting that Result 2 implies that

$$\max_{\mathbb{M}^A,\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB}) = \max_{\mathbb{M}^A}\left[\max_{\mathbb{M}^B} \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})\right] = \max_{\mathbb{M}^A} \mathbf{R}_T(\mathbb{A}^{A\to B'}). \tag{E2}$$

Let $A^{A'B'} \geq 0$, $W^{A'B'} \in \mathscr{F}^*_{\mathrm{SEP}}$ be optimal dual variables for $\mathbf{R}_E(\rho^{A'B'})$. Let $\{U_a^{A'}\}$ for $a \in \{1,\dots,d^2\}$ be a set of Pauli operators with respect to a basis $\{|i\rangle^{A'}\}$. Consider the following measurement with $o_A = d^2$ outcomes:

$$M_a^{AA'} = (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}]. \tag{E3}$$

We are interested in the lower bound for $\mathbf{R}_E(\rho^{A'B'})$. Let us construct a set of (potentially suboptimal) dual variables in maximization (E1) using the optimal set of dual variables for $\mathbf{R}_T(\mathbb{A}^{V\to B'})$, i.e.,

$$A_a^{VB'} = [(\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'}][A^{VB'}], \quad W_a^{VB'} = [(\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'}][W^{VB'}], \qquad B^{VB'} = \frac{1}{d}\mathbb{1}^{VB'}. \tag{E4}$$

It can be verified by direct substitution that the above choice is feasible. This leads to the following chain of inequalities:

$$\begin{aligned}
1 + \max_{\mathbb{M}^A} \mathbf{R}_T(\mathbb{A}^{A\to B'}) &\geq \sum_a \mathrm{tr}[A_a^{VB'} J_a^{VB'}] \\
&= \sum_a \mathrm{tr}\{[(\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'}][A_a^{VB'}]\,\mathrm{tr}_{AA'}[(\mathbb{1}^V \otimes M_a^{AA'} \otimes \mathbb{1}^{B'})(\phi_+^{VA} \otimes \rho^{A'B'})]\} \\
&= \sum_a \mathrm{tr}\{[(\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'}][A_a^{VB'}]\,\mathrm{tr}_{AA'}[\{\mathbb{1}^V \otimes (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}] \otimes \mathbb{1}^{B'}\}(\phi_+^{VA} \otimes \rho^{A'B'})]\} \\
&= \sum_a \mathrm{tr}\{\mathbb{1}^{AA'} \otimes [(\mathcal{U}_a^\dagger)^V \otimes \mathrm{id}^{B'}][A_a^{VB'}]\{\mathbb{1}^V \otimes (\mathrm{id}^A \otimes \mathcal{U}_a^{A'})[\phi_+^{AA'}] \otimes \mathbb{1}^{B'}\}(\phi_+^{VA} \otimes \rho^{A'B'})\} \\
&= \frac{1}{d^2} \sum_a \mathrm{tr}\{[(\mathcal{U}_a^\dagger)^{A'} \otimes \mathrm{id}^{B'}]A^{A'B'}\,(\mathcal{U}_a^{A'} \otimes \mathrm{id}^{B'})\rho^{A'B'}\} \\
&= \mathrm{tr}[A^{A'B'}\rho^{A'B'}] \\
&= 1 + \mathbf{R}_E(\rho^{A'B'}).
\end{aligned} \tag{E5}$$

We now prove the upper bound. Note that any teleportation instrument $\mathbb{A}^{A\to B'}$ expressed using Choi-Jamiołkowski operators $\{J_a^{VB'}\}$ satisfies

$$\begin{aligned}
J_a^{VB'} &:= \mathrm{tr}_{VA}[(M_a^{VA} \otimes \mathbb{1}^{B'})(\phi_+^A \otimes \rho^{A'B'} \otimes \mathbb{1}^B)] \\
&\leq [1 + \mathbf{R}_E(\rho^{A'B'})]\mathrm{tr}_{VA}[(M_a^{VA} \otimes \mathbb{1}^{B'})(\phi_+^A \otimes \sigma^{A'B'} \otimes \mathbb{1}^B)] \\
&= [1 + \mathbf{R}_E(\rho^{A'B'})]O_a^{VB'}
\end{aligned} \tag{E6}$$

for some state $\sigma^{A'B'} \in \mathscr{F}_{\mathrm{SEP}}$ and corresponding (classical) teleportation operators $\{O_a^{VB'}\}$. In this way we can write

$$\begin{aligned}
\max_{\mathbb{M}^A}[1 + \mathbf{R}_T(\mathbb{A}^{A\to B'})] &= \max_{\mathbb{M}^A} \max_{\{A_a^{VB'}\}} \sum_a \mathrm{tr}[A_a^{VB'} J_a^{VB'}] \\
&\leq [1 + \mathbf{R}_E(\rho^{A'B'})] \sum_a \mathrm{tr}[A_a^{VB'} O_a^{VB'}] \\
&\leq [1 + \mathbf{R}_E(\rho^{A'B'})].
\end{aligned} \tag{E7}$$

This completes the proof of the lemma.   ∎

APPENDIX F: PROOF OF RESULT 6

In this section, unless explicitly specified, all bipartite operators act on subsystems A and B. We begin by assuming that a distributed measurement $\mathbb{M}$ can be used to simulate $\mathbb{M}^*$, that is, $\mathbb{M} \succ_q \mathbb{M}^*$. We have

$$
\begin{aligned}
p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}) &= \max_{\mathbb{M} \succ_q \mathbb{M}'} \sum_{a,b} p(a,b)\,\text{tr}[M'_{ab}\sigma_{ab}] \\
&\geq \max_{\mathbb{M}^* \succ_q \mathbb{M}'} \sum_{a,b} p(a,b)\text{tr}[M'_{ab}\sigma_{ab}] \\
&= p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^*),
\end{aligned}
\tag{F1}
$$

since the set $\{\mathbb{M}'|\mathbb{M}^* \succ_q \mathbb{M}'\}$ is a subset of $\{\mathbb{M}'|\mathbb{M} \succ_q \mathbb{M}'\}$. Now we are going to assume that $p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}) \geq p_{\text{guess}}^{\text{DSD}}(\mathcal{G}, \mathbb{M}^*)$ holds for all games $\mathcal{G} = \{p(x,y), \sigma_{xy}\}$ and show that there always exist a subroutine $\mathscr{S}$ that allows us to simulate $\mathbb{M}^*$ using $\mathbb{M}$. We thus have, for all $\mathcal{G}$,

$$
\max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b} p(a,b)\,\text{tr}[M'_{ab}\sigma_{ab}] - \max_{\mathbb{M}'' \prec_q \mathbb{M}^*} \sum_{a,b} p(a,b)\,\text{tr}[M''_{ab}\sigma_{ab}] \geq 0.
\tag{F2}
$$

Let us now choose a particular subroutine in the second maximization, i.e., $\mathscr{S}^* = \{p(\lambda) = \delta_{\lambda 0}, \ p(a|i,\lambda) = \delta_{ai}, \ p(b|j,\lambda) = \delta_{bj}, \ U_\lambda = V_\lambda = \mathbb{1}\}$. In this way Eq. (F2) implies that, for all $\mathcal{G}$,

$$
\max_{\mathbb{M}' \prec_q \mathbb{M}} \sum_{a,b} p(a,b)\text{tr}[(M'_{ab} - M^*_{ab})\,\sigma_{ab}] \geq 0.
\tag{F3}
$$

Let $\Delta_{ab} := M'_{ab} - M^*_{ab}$. Since both M'_{ab} and M^*_{ab} are measurements, we have $\sum_{a,b} \Delta_{ab} = 0$. This also means that only one of the two situations can hold: either (i) $\Delta_{ab} = 0$ for all a, b or (ii) there exists at least one Δ_{ab} with at least one negative eigenvalue.

We now show by contradiction that (ii) cannot be true. Let us assume that (ii) holds and label the negative eigenvalue $\lambda_{a^*b^*}$ and the associated eigenvector $|\lambda_{a^*b^*}\rangle$. Then, since Eq. (F3) holds for all games $\mathcal{G}$, it also holds for a particular game $\mathcal{G}^* = \{p(a,b) = \delta_{aa^*}\delta_{bb^*}, \sigma_{ab} = |\lambda_{a^*b^*}\rangle\langle\lambda_{a^*b^*}|\}$. Hence, Eq. (F3) implies that

$$
\langle \lambda_{a^*b^*}| \, \Delta_{a^*b^*} \, |\lambda_{a^*b^*}\rangle = \lambda_{a^*b^*} \; < 0,
\tag{F4}
$$

which is a contradiction. Hence, we infer that (ii) cannot be true and the only possibility is that each operator Δ_{ab} is identically zero. This means that

$$
M^*_{ab} = M'_{ab} := \sum_{i,j,\lambda} p(\lambda)p(a|i,\lambda)p(b|j,\lambda)(U_\lambda^\dagger \otimes V_\lambda^\dagger)M_{ij}(U_\lambda \otimes V_\lambda),
\tag{F5}
$$

i.e., $\mathbb{M}^*$ can be simulated using $\mathbb{M}$.

APPENDIX G: PROOF OF RESULT 7

The accessible min-information $I_{\min}^{\text{acc}}(\mathcal{N})$ of a channel $\mathcal{N}$ is defined as [31]

$$
I_{\min}^{\text{acc}}(\mathcal{N}) = \max_{\mathscr{E},\mathscr{D}}[H_{\min}(X) - H_{\min}(X|G)],
\tag{G1}
$$

where the optimization is over all encodings $\mathscr{E} = \{p(x), \sigma_x\}$ and decodings $\mathscr{D} = \{D_g\}$ and the min-entropies are defined as

$$
H_{\min}(X) = -\log \max_x p(x),
\tag{G2}
$$

$$
H_{\min}(X|G) = -\log\left[\sum_g \max_x p(x,g)\right],
\tag{G3}
$$

and $p(x, g)$ is the probability distribution induced by channel $\mathcal{N}$, i.e.,

$$p(x, g) = p(x)p(g|x) = p(x)\text{tr}[\mathcal{N}[\sigma_x]D_g]. \tag{G4}$$

Consider now encoding a bipartite random variable $X \times Y$ in an ensemble of bipartite quantum states, i.e., $\mathscr{E} = \{p(x, y), \sigma_{xy}^{AB}\}$ and $\mathscr{D} = \{D_g^{A'B'}\}$ for $g = 1, \ldots, o_A \cdot o_B$. Moreover, consider the channel $\mathcal{N} = \mathcal{N}^{AB \to A'B'}$ to be a quantum-to-classical measurement channel, which can be written as

$$\mathcal{N}^{AB \to A'B'}(\rho^{AB}) = \sum_{a,b} \text{tr}[M_{ab}^{AB}\rho^{AB}]|a\rangle\langle a|_{A'} \otimes |b\rangle\langle b|_{B'}, \tag{G5}$$

where $\mathbb{M} = \{M_{ab}^{AB}\}$ is a distributed measurement. We have

$$
\begin{aligned}
I_{\min}^{\text{acc}}(\mathcal{N}^{AB \to A'B'}) &= \max_{\mathscr{E},\mathscr{D}} \log\left\{\sum_g \max_{x,y} p(x,y)\text{tr}[\mathcal{N}^{AB \to A'B'}(\sigma_{xy}^{AB})D_g^{A'B'}]\right\} - \log\max_{a,b} p(a,b)\\
&= \max_{\mathscr{E},\mathscr{D}} \log\left\{\sum_g \sum_{a,b} \max_{x,y} p(x,y)\text{tr}(M_{ab}^{AB}\sigma_{xy}^{AB})\text{tr}(D_g^{A'B'}|a\rangle\langle a|_{A'} \otimes |b\rangle\langle b|_{B'})\right\}\\
&\quad - \log\max_{a,b} p(a,b)\\
&= \log\left\{\sum_{a,b} \max_{\mathscr{E}} \max_{x,y} p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}]\right\} - \log\max_{a,b} p(a,b). \tag{G6}
\end{aligned}
$$

Note now that we can always express the optimization over (x, y) as

$$
\begin{aligned}
\max_{x,y} p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}] &= \max_{p(x|a)} \max_{p(y|b)} \sum_{x,y} p(x|a)p(y|b)p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}]\\
&= \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a,\lambda)p(y|b,\lambda)p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}]. \tag{G7}
\end{aligned}
$$

Note further that if we carry out the optimization of the above expression over $\mathscr{E}$ we can additionally write

$$
\begin{aligned}
&\max_{\mathscr{E}} \max_{x,y} p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}]\\
&= \max_{\mathscr{E}} \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a,\lambda)p(y|b,\lambda)p(x,y)\text{tr}[M_{ab}^{AB}\sigma_{xy}^{AB}]\\
&= \max_{\mathscr{E}} \max_{\{\mathcal{E}_\lambda\},\{\mathcal{F}_\lambda\}} \max_{p(\lambda)} \max_{p(x|a,\lambda)} \max_{p(y|b,\lambda)} \sum_{x,y,\lambda} p(x|a,\lambda)p(y|b,\lambda)p(x,y)\text{tr}[M_{ab}^{AB}(\mathcal{E}_\lambda^A \otimes \mathcal{F}_\lambda^B)(\sigma_{xy}^{AB})]\\
&= \max_{\mathscr{E}} \max_{\mathbb{N} \prec \mathbb{M}} \sum_x p(x,y)\text{tr}[N_{ab}^{AB}\sigma_{xy}^{AB}]. \tag{G8}
\end{aligned}
$$

Hence, we can further continue from Eq, (G6) and write

$$
I_{\min}^{\mathrm{acc}}(\mathcal{N}^{AB \to A'B'}) = \log\left[\sum_{a,b} \max_{\mathcal{E}} \max_{\mathbb{N} \prec \mathbb{M}} p(a,b) \mathrm{tr}(N_{ab}^{AB} \sigma_{ab}^{AB}) \right] - \log \max_{a,b} p(a,b)
$$

$$
= \max_{\mathcal{E}} \log\left[\max_{\mathbb{N} \prec \mathbb{M}} \sum_{a,b} p(a,b) \mathrm{tr}(M_{ab}^{AB} \sigma_{ab}^{AB}) \right] - \max_{a,b} p(a,b)
$$

$$
= \max_{\mathcal{E}} \log[p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}, \mathbb{M}^{AB})] - \log[p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G})]
$$

$$
= \log\left[\max_{\mathcal{E}} \frac{p_{\mathrm{guess}}^{\mathrm{DSD}}(\mathcal{G}, \mathbb{M}^{AB})}{p_{\mathrm{guess}}(\mathcal{G})} \right]
$$

$$
= \log[1 + \mathbf{R}_{\mathrm{BN}}(\mathbb{M}^{AB})]. \tag{G9}
$$

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Quantum entanglement, Rev. Mod. Phys. **81**, 865 (2009).

[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[3] R. F. Werner, Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model, Phys. Rev. A **40**, 4277 (1989).

[4] J. Barrett, Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality, Phys. Rev. A **65**, 042302 (2002).

[5] R. Augusiak, M. Demianowicz, and A. Acín, Local hidden-variable models for entangled quantum states, J. Phys. A: Math. Theor. **47**, 424002 (2014).

[6] F. Buscemi, All Entangled Quantum States Are Nonlocal, Phys. Rev. Lett. **108**, 200401 (2012).

[7] C. Branciard, D. Rosset, Y.-C. Liang, and N. Gisin, Measurement-Device-Independent Entanglement Witnesses for All Entangled Quantum States, Phys. Rev. Lett. **110**, 060405 (2013).

[8] E. G. Cavalcanti, M. J. W. Hall, and H. M. Wiseman, Entanglement verification and steering when Alice and Bob cannot be trusted, Phys. Rev. A **87**, 032306 (2013).

[9] P. Xu, X. Yuan, L.-K. Chen, H. Lu, X.-C. Yao, X. Ma, Y.-A. Chen, and J.-W. Pan, Implementation of a Measurement-Device-Independent Entanglement Witness, Phys. Rev. Lett. **112**, 140506 (2014).

[10] E. Verbanis, A. Martin, D. Rosset, C. C. W. Lim, R. T. Thew, and H. Zbinden, Resource-Efficient Measurement-Device-Independent Entanglement Witness, Phys. Rev. Lett. **116**, 190501 (2016).

[11] Q. Zhao, X. Yuan, and X. Ma, Efficient measurement-device-independent detection of multipartite entanglement structure, Phys. Rev. A **94**, 012343 (2016).

[12] C. C. W. Lim, Optimality of semiquantum nonlocality in the presence of high inconclusive rates, Phys. Rev. A **93**, 020101 (2016).

[13] F. Shahandeh, M. J. W. Hall, and T. C. Ralph, Measurement-Device-Independent Approach to Entanglement Measures, Phys. Rev. Lett. **118**, 150505 (2017).

[14] I. Šupić, P. Skrzypczyk, and D. Cavalcanti, Measurement-device-independent entanglement and randomness estimation in quantum networks, Phys. Rev. A **95**, 042340 (2017).

[15] F. Bischof, H. Kampermann, and D. Bruß, Measurement-device-independent randomness generation with arbitrary quantum states, Phys. Rev. A **95**, 062305 (2017).

[16] D. Rosset, F. Buscemi, and Y.-C. Liang, Resource Theory of Quantum Memories and Their Faithful Verification with Minimal Assumptions, Phys. Rev. X **8**, 021033 (2018).

[17] Z.-D. Li, Q. Zhao, R. Zhang, L.-Z. Liu, X.-F. Yin, X. Zhang, Y.-Y. Fei, K. Chen, N.-L. Liu, F. Xu, Y.-A. Chen, L. Li, and J.-W. Pan, Measurement-Device-Independent Entanglement Witness of Tripartite Entangled States and Its Applications, Phys. Rev. Lett. **124**, 160503 (2020).

[18] X. Zhang and Q. Zhao, Simultaneous certification of entangled states and measurements in bounded dimensional semiquantum games, Phys. Rev. Res. **2**, 033400 (2020).

[19] I. Šupić, M. J. Hoban, L. D. Colomer, and A. Acín, Self-testing and certification using trusted quantum inputs, New J. Phys. **22**, 073006 (2020).

[20] M. Horodecki and J. Oppenheim, (Quantumness in the context of) resource theories, Int. J. Mod. Phys. B **27**, 1345019 (2012).

[21] E. Chitambar and G. Gour, Quantum resource theories, Rev. Mod. Phys. **91**, 025001 (2019).

[22] P. Skrzypczyk and N. Linden, Robustness of Measurement, Discrimination Games, and Accessible Information, Phys. Rev. Lett. **122**, 140403 (2019).

[23] S. Designolle, M. Farkas, and J. Kaniewski, Incompatibility robustness of quantum measurements: A unified framework, New J. Phys. **21**, 113053 (2019).

[24] A. F. Ducuara and P. Skrzypczyk, Operational Interpretation of Weight-Based Resource Quantifiers in Convex Quantum Resource Theories, Phys. Rev. Lett. **125**, 110401 (2020).

[25] T. Guff, N. A. McMahon, Y. R. Sanders, and A. Gilchrist, A resource theory of quantum measurements, arXiv:1902.08490 (2019).

[26] M. Oszmaniec and T. Biswas, Operational relevance of resource theories of quantum measurements, Quantum **3**, 133 (2019).

[27] M. Oszmaniec, L. Guerini, P. Wittek, and A. Acín, Simulating Positive-Operator-Valued Measures with Projective Measurements, Phys. Rev. Lett. **119**, 190501 (2017).

[28] T. Theurer, D. Egloff, L. Zhang, and M. B. Plenio, Quantifying Operations with an Application to Coherence, Phys. Rev. Lett. **122**, 190405 (2019).

[29] Z.-W. Liu, K. Bu, and R. Takagi, One-Shot Operational Quantum Resource Theory, Phys. Rev. Lett. **123**, 020401 (2019).

[30] Y. Liu and X. Yuan, Operational Resource Theory of Quantum Channels, Preprint arXiv:1904.02680 (2019).

[31] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, Cambridge, 2013).

[32] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Quantifying Bell: the resource theory of nonclassicality of common-cause boxes, Quantum **4**, 280 (2020).

[33] D. Schmid, T. C. Fraser, R. Kunjwal, A. B. Sainz, E. Wolfe, and R. W. Spekkens, Why standard entanglement theory is inappropriate for the study of bell scenarios, arXiv:2004.09194 (2020).

[34] D. Schmid, D. Rosset, and F. Buscemi, The type-independent resource theory of local operations and shared randomness, Quantum **4**, 262 (2020).

[35] D. Rosset, D. Schmid, and F. Buscemi, Characterizing nonclassicality of arbitrary distributed devices, arXiv:1911.12462 (2019).

[36] M. A. Nielsen, Conditions for a Class of Entanglement Transformations, Phys. Rev. Lett. **83**, 436 (1999).

[37] G. Vidal and R. Tarrach, Robustness of entanglement, Phy. Rev. A **59**, 141 (1999).

[38] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, Robustness of Coherence: An Operational and Observable Measure of Quantum Coherence, Phys. Rev. Lett. **116**, 150502 (2016).

[39] M. Horodecki, P. W. Shor, and M. B. Ruskai, Entanglement Breaking Channels, Rev. Math. Phys. **15**, 629 (2003).

[40] A. Streltsov, H. Kampermann, S. Wölk, M. Gessner, and D. Bruß, Maximal coherence and the resource theory of purity, New J. Phys. **20**, 053058 (2018).

[41] D. Janzing, P. Wocjan, R. Zeier, R. Geiss, and T. Beth, Thermodynamic cost of reliability and low temperatures: Tightening Landauer's principle and the second law, Int. J. Theor. Phys. **39**, 2717 (2000).

[42] F. G. S. L. Brandão, M. Horodecki, J. Oppenheim, J. M. Renes, and R. W. Spekkens, Resource Theory of Quantum States Out of Thermal Equilibrium, Phys. Rev. Lett. **111**, 250404 (2013).

[43] F. G. S.L. Brandão and G. Gour, Reversible Framework for Quantum Resource Theories, Phys. Rev. Lett. **115**, 070503 (2015).

[44] M. Horodecki and J. Oppenheim, Fundamental limitations for quantum and nanoscale thermodynamics, Nat. Commun. **4**, 2059 (2013).

[45] N. H. Y. Ng and M. P. Woods, Resource theory of quantum thermodynamics: Thermal operations and second laws, Thermodyn. Quantum Regime, 625 (2018).

[46] K. Horodecki, A. Grudka, P. Joshi, W. Kłobus, and J. Łodyga, Axiomatic approach to contextuality and nonlocality, Phys. Rev. A **92**, 032104 (2015).

[47] D. Cavalcanti and P. Skrzypczyk, Quantitative relations between measurement incompatibility, quantum steering, and nonlocality, Phys. Rev. A **93**, 052112 (2016).

[48] M. Piani, M. Cianciaruso, T. R. Bromley, C. Napoli, N. Johnston, and G. Adesso, Robustness of asymmetry and coherence of quantum states, Phys. Rev. A **93**, 042107 (2016).

[49] F. Buscemi, E. Chitambar, and W. Zhou, Complete Resource Theory of Quantum Incompatibility as Quantum Programmability, Phys. Rev. Lett. **124**, 120401 (2020).

[50] I. Šupić, P. Skrzypczyk, and D. Cavalcanti, Methods to estimate entanglement in teleportation experiments, Phys. Rev. A **99**, 032334 (2019).

[51] D. Cavalcanti, P. Skrzypczyk, and I. Šupić, All Entangled States can Demonstrate Nonclassical Teleportation, Phys. Rev. Lett. **119**, 110501 (2017).

[52] M. Howard and E. Campbell, Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing, Phys. Rev. Lett. **118**, 090501 (2017).

[53] S. Bhattacharya, B. Bhattacharya, and A. S. Majumdar, Convex resource theory of non-Markovianity, arXiv:1803.06881 (2018).

[54] E. Wakakuwa, Operational resource theory of non-Markovianity, arXiv:1709.07248 (2017).

[55] N. Anand and T. A. Brun, Quantifying non-Markovianity: a quantum resource-theoretic approach, arXiv:1903.03880 (2019).

[56] R. Takagi and Q. Zhuang, Convex resource theory of non-Gaussianity, Phys. Rev. A **97**, 062337 (2018).

[57] M. Piani and J. Watrous, Necessary and Sufficient Quantum Information Characterization of Einstein-Podolsky-Rosen Steering, Phys. Rev. Lett. **114**, 060404 (2015).

[58] A. Acín, N. Gisin, and L. Masanes, From Bell's Theorem to Secure Quantum Key Distribution, Phys. Rev. Lett. **97**, 120405 (2006).

[59] P. Lipka-Bartosik and P. Skrzypczyk, Operational advantages provided by nonclassical teleportation, Phys. Rev. Res. **2**, 023029 (2020).

[60] D. Rosset, A. Martin, E. Verbanis, C. C. W. Lim, and R. Thew, Practical measurement-device-independent entanglement quantification, Phys. Rev. A **98**, 052332 (2018).

[61] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, in *Proceedings. 19th IEEE Annual Conference on Computational Complexity* (Amherst, MA, USA, 2004), p. 236.

[62] S. Boyd and L. Vandenberghe, *Convex Optimization* (Cambridge University Press, New York, NY, USA, 2004).

[63] J. Watrous, *The Theory of Quantum Information* (Cambridge University Press, Cambridge, 2018).

[64] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, http://cvxr.com/cvx (2014).

[65] M. Steiner, Generalized robustness of entanglement, Phys. Rev. A **67**, 054305 (2003).

[66] R. Takagi and B. Regula, General Resource Theories in Quantum Mechanics and Beyond: Operational Characterization via Discrimination Tasks, Phys. Rev. X **9**, 031053 (2019).

[67] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels, Phys. Rev. Lett. **70**, 1895 (1993).

[68] M. B. Plenio and S. Virmani, An introduction to entanglement measures, arXiv:quant-ph/0504163 (2005).

[69] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Mixed-state entanglement and quantum error correction, Phys. Rev. A **54**, 3824 (1996).

[70] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[71] V. Vedral, M. B. Plenio, K. Jacobs, and P. L. Knight, Statistical inference, distinguishability of quantum states, and quantum entanglement, Phys. Rev. A **56**, 4452 (1997).

[72] E. M. Rains, Rigorous treatment of distillable entanglement, Phys. Rev. A **60**, 173 (1999).

[73] B. Schumacher and M. D. Westmoreland, Relative entropy in quantum information theory, arXiv:quant-ph/0004045 (2000).

[74] P. M. Hayden, M. Horodecki, and B. M. Terhal, The asymptotic entanglement cost of preparing a quantum state, J. Phys. A: Math. Gen. **34**, 6891 (2001).

[75] A. K. Rajagopal and R. W. Rendell, Robust and fragile entanglement of three qubits: Relation to permutation symmetry, Phys. Rev. A **65**, 032328 (2002).

[76] C. Simon and J. Kempe, Robustness of multiparty entanglement, Phys. Rev. A **65**, 052327 (2002).

[77] R. Takagi, B. Regula, K. Bu, Z.-W. Liu, and G. Adesso, Operational Advantage of Quantum Resources in Sub-channel Discrimination, Phys. Rev. Lett. **122**, 140402 (2019).

[78] N. Ciganović, N. J. Beaudry, and R. Renner, Smooth max-information as one-shot generalization for mutual information, Trans. Inf. Theory **60**, 1573 (2014).

[79] R. Renner, Security of quantum key distribution, arXiv:quant-ph/0512258 (2005).

[80] A. C. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Phys. Lett. A **162**, 25 (1992).

[81] M. Lewenstein and A. Sanpera, Separability and Entanglement of Composite Quantum Systems, Phys. Rev. Lett. **80**, 2261 (1998).

[82] R. Uola, T. Bullock, T. Kraft, J.-P. Pellonpää, and N. Brunner, All Quantum Resources Provide an Advantage in Exclusion Tasks, Phys. Rev. Lett. **125**, 110402 (2020).

[83] A. F. Ducuara, P. Lipka-Bartosik, and P. Skrzypczyk, Multiobject operational tasks for convex quantum resource theories of state-measurement pairs, Phys. Rev. Res. **2**, 033374 (2020).