Resource Theory of Steering

Rodrigo Gallego[1] and Leandro Aolita[1,2]

[1]*Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany*
[2]*Instituto de Física, Universidade Federal do Rio de Janeiro, P.O. Box 68528,*
Rio de Janeiro, RJ 21941-972, Brazil

We present an operational framework for Einstein-Podolsky-Rosen steering as a physical resource. For arbitrary-dimensional bipartite systems composed of a quantum subsystem and a black-box device, we show that local operations assisted by one-way classical communication (1W-LOCCs) from the quantum part to the black box cannot create steering. Based on this, we build a resource theory of steering with 1W-LOCCs as the free operations. We introduce the notion of convex steering monotones as the fundamental axiomatic quantifiers of steering. As a convenient example thereof, we present the relative entropy of steering. In addition, we prove that two previously proposed quantifiers, the steerable weight and the robustness of steering, are also convex steering monotones. To end up with for minimal-dimensional systems, we establish, on the one hand, necessary and sufficient conditions for pure-state steering conversions under stochastic 1W-LOCCs and prove, on the other hand, the nonexistence of steering bits, i.e., measure-independent maximally steerable states from which all states can be obtained by means of the free operations. Our findings reveal unexpected aspects of steering and lay the foundations for further research, with potential implications in Bell nonlocality.

I. INTRODUCTION

Steering, as Schrödinger named it [1], is an exotic quantum effect by which ensembles of quantum states can be remotely prepared by performing local measurements at a distant lab. It allows [2,3] one to certify the presence of entanglement between a user with an untrusted measurement apparatus, Alice, and another user with a trusted quantum-measurement device, Bob. Thus, it constitutes a fundamental notion between quantum entanglement [4], whose certification requires quantum measurements on both sides, and Bell nonlocality [5], where both users possess untrusted black-box devices. Steering can be detected through simple tests analogous to Bell inequalities [6], and it has been verified in a variety of remarkable experiments [7], including steering without Bell nonlocality [8] and a loop-hole-free steering demonstration [9]. Apart from its fundamental relevance, steering has been identified as a resource for one-sided (1S) device-independent (DI) quantum key distribution (QKD), where only one of the parts has an untrusted apparatus while the other ones possess trusted devices [10,11]. In 1S-DI-QKD the experimental requirements for unconditionally secure keys are less stringent than in fully (both-sided) DI-QKD [12].

The formal treatment of a physical property as a resource is given by a resource theory. The basic component of this is a restricted class of operations, called the free operations. These are typically the operations at hand in physical scenarios where the property in question acts as a useful resource. They fulfill the essential requirement of mapping every free state, i.e., every one without the property, into a free state. Furthermore, they provide a formal recipe for the quantification of the resource: The fundamental necessary condition for a function to be a measure of the resource is that it is monotonous—nonincreasing—under the free operations. In other words, the operations that do not increase the resource on the free states do not increase it on all other states either.

Entanglement theory is the most popular and best understood [13,14] resource theory. In this theory, local operations assisted by classical communication (LOCCs) are usually the natural free operations [15]. However, other sets of operations that do not create entanglement either have also been considered [16,17]. On the other hand, resource theories have also been formulated for states out of thermal equilibrium [18], asymmetry [19], reference frames [20], and nonlocality [21,22], for instance.

In steering theory, systems are described by a collection of ensembles of quantum states on Bob's side and a conditional probability distribution of measurement outcomes (outputs) given measurement settings (inputs) on Alice's. Each input of Alice's is associated with an ensemble of Bob's, and each output is associated with a member state of such an ensemble. In other words, each pair of inputs and outputs of Alice's is correlated with a state in an ensemble of Bob's. Such systems are called assemblages [23–28]. The free operations for steering must

thus arise from natural constraints native to a physical scenario where steerable assemblages are useful for some task. Up to now, no attempt for an operational framework of steering as a resource has been reported.

In this work, we develop the resource theory of steering. First, we observe that one-way LOCCs (1W-LOCCs) from Bob to Alice are allowed operations—in the sense of not compromising the security—in 1S-DI-QKD protocols, the physical scenario where steering is a known resource [10,11]. Then, for arbitrarily many inputs and outputs for Alice's black box and an arbitrary Hilbert-space dimension for Bob's quantum system, we show that 1W-LOCCs from Bob to Alice do not create steering. These two facts give us physical motivations to take 1W-LOCCs as natural free operations for steering. We present the explicit parametrization of a generic 1W-LOCC acting on assemblages. With this, we provide a formal definition of steering monotones. As an example thereof, we present the relative entropy of steering, for which we also introduce, in turn, the notion of relative entropy between assemblages. In addition, we prove 1W-LOCC monotonicity for two other recently proposed steering measures, the steerable weight [24] and the robustness of steering [25], as well as convexity for all three measures. We prove two theorems on steering conversion under stochastic 1W-LOCCs for the lowest-dimensional case, i.e., qubits on Bob's side and 2 inputs × 2 outputs on Alice's. In the first one, we show that it is impossible to transform via 1W-LOCCs, even probabilistically, an assemblage composed of pairs of pure orthogonal states into another assemblage also composed of pairs of pure orthogonal states but with a different pair overlap, unless the latter is unsteerable. This no-go result implies that there exist yields infinitely many inequivalent classes of steering already for systems of lowest dimension. In the second theorem, we show that there exists no assemblage composed of pairs of pure states that can be transformed into any assemblage by stochastic 1W-LOCCs. This implies, in striking contrast to entanglement theory, that there exists no operationally well-defined, measure-independent, maximally steerable assemblage of minimal dimension.

The paper is organized as follows. In Sec. II, we formally define assemblages and present their basic properties. In Sec. III, we find an explicit parametrization of all stochastic 1W-LOCCs from assemblages into assemblages and show that the resulting maps are steering nonincreasing operations. In Sec. IV, we discuss the role of 1W-LOCCs as the natural operations available to Alice and Bob in 1S-DI-QKD. In Sec. V, we introduce the notion of convex steering monotones. In Sec. VI, we present the relative entropy of steering. In Sec. VII, we show convexity and 1W-LOCC monotonicity of the steerable weight and the robustness of steering. In Sec. VIII, we study, for minimal-dimensional systems, assemblage conversions under 1W-LOCCs and prove the nonexistence of pure-assemblage steering bits.

Finally, in Sec. IX, we present our conclusions and mention some future research directions that our results offer.

II. ASSEMBLAGES AND STEERING

We consider two distant parties, Alice and Bob, who each have half of a bipartite system. Alice holds a so-called black-box device, which, given a classical input $x \in [s]$, generates a classical output $a \in [r]$, where s and r are natural numbers and the notation $[n] := \{0, \ldots, n-1\}$, for $n \in \mathbb{N}$, is introduced. Bob holds a quantum system of dimension d (qudit), whose state can be perfectly characterized tomographically via trusted quantum measurements. The joint state of their system is thus fully specified by an assemblage

$$\rho_{A|X} := \{P_{A|X}(a, x), \varrho(a, x)\}_{a \in [r], x \in [s]}, \qquad (1)$$

of normalized quantum states $\varrho(a, x) \in \mathcal{L}(\mathcal{H}_B)$, with $\mathcal{L}(\mathcal{H}_B)$ the set of linear operators on Bob's subsystem's Hilbert space $\mathcal{H}_B$, each one associated with a conditional probability $P_{A|X}(a, x)$ of Alice getting an output a given an input x. We denote by $P_{A|X}$ the corresponding conditional probability distribution.

Equivalently, each pair $\{P_{A|X}(a, x), \varrho(a, x)\}$ can be univocally represented by the unnormalized quantum state

$$\varrho_{A|X}(a, x) := P_{A|X}(a, x) \times \varrho(a, x). \qquad (2)$$

In turn, an alternative representation of the assemblage $\rho_{A|X}$ is given by the set $\hat{\rho}_{A|X} := \{\hat{\rho}_{A|X}(x)\}_x$ of quantum states

$$\hat{\rho}_{A|X}(x) := \sum_a |a\rangle\langle a| \otimes \varrho_{A|X}(a, x) \in \mathcal{L}(\mathcal{H}_E \otimes \mathcal{H}_B), \qquad (3)$$

where $\{|a\rangle\}$ is an orthonormal basis of an auxiliary extension Hilbert space $\mathcal{H}_E$ of dimension r. The states $\{|a\rangle\}$ do not describe the system inside Alice's box; they are just abstract flag states to represent her outcomes with a convenient bra-ket notation. Expression (3) gives the counterpart for assemblages of the so-called extended Hilbert space representation used for ensembles of quantum states [29]. We refer to $\hat{\rho}_{A|X}$ for short as the quantum representation of $\rho_{A|X}$ and use either notation upon convenience.

We restrict ourselves throughout to no-signaling assemblages, i.e., those for which Bob's reduced state $\varrho_B \in \mathcal{L}(\mathcal{H}_B)$ does not depend on Alice's input choice x:

$$\varrho_B := \sum_a \varrho_{A|X}(a, x) = \sum_a \varrho_{A|X}(a, x') \quad \forall \; x, x'. \qquad (4)$$

In the bipartite case, the assemblages fulfilling the no-signaling condition (4) are the ones that possess a quantum realization. In other words, they can be obtained from local quantum measurements by Alice on a joint quantum state $\varrho_{AB} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$ shared with Bob, where $\mathcal{H}_A$ is the

Hilbert space of the system inside Alice's box. For any no-signaling assemblage $\rho_{A|X}$, we refer to the following x-independent quantity as the trace of the assemblage

$$\mathrm{Tr}[\rho_{A|X}] \coloneqq \mathrm{Tr}_{EB}[\hat{\rho}_{A|X}] = \mathrm{Tr}[\varrho_B] = \sum_a P_{A|X}(a,x), \quad (5)$$

and say that the assemblage is normalized if $\mathrm{Tr}[\rho_{A|X}] = 1$ and unnormalized if $\mathrm{Tr}[\rho_{A|X}] \leq 1$.

An assemblage $\sigma_{A|X} \coloneqq \{\varsigma_{A|X}(a,x)\}_{a\in[r],x\in[s]}$, with $\varsigma_{A|X}(a,x) \in \mathcal{L}(\mathcal{H}_B)$ being unnormalized states, is called unsteerable if there exist a probability distribution P_Λ, a conditional probability distribution $P_{A|X\Lambda}$, and normalized states $\xi(\lambda) \in \mathcal{L}(\mathcal{H}_B)$ such that

$$\varsigma_{A|X}(a,x) = \sum_\lambda P_\Lambda(\lambda) P_{A|X\Lambda}(a,x,\lambda) \xi(\lambda) \quad \forall \, x, a. \quad (6)$$

Such assemblages can be obtained by sending a classical random variable λ to Alice, correlated with the state $\xi(\lambda)$ sent to Bob, and letting Alice classically post-process her random variable according to $P_{A|X\Lambda}$, with $P_{X,\Lambda} = P_X \times P_\Lambda$ so that condition (4) holds. The variable λ is called a local-hidden variable, and the decomposition (6) is accordingly referred to as a local-hidden state (LHS) model. We refer to the set of all unsteerable assemblages as LHS. Any assemblage that does not admit a LHS model as in Eq. (6) is called steerable.

A comment on steering as a property of assemblages, as opposed to quantum states, is in order. In the pioneering works [2,3], steering is defined as a property of quantum states. Namely, a state is said to be steerable if it can give rise, under local measurements, to correlations without a LHS model, i.e., to a steerable assemblage. The definition considered here, directly in terms of assemblages, and to which our resource theory applies, follows the treatment of Refs. [23–28], for instance, where one embeds a share of the bipartite quantum state in the untrusted measurement device (Alice's, in our case) and treats the entire embedding as a black box with unknown internal functioning. This definition is the scenario of 1S-DI-QKD, the very task for which steering is a known useful resource. In 1S-DI-QKD, quantum states alone are not useful since, without a trusted measurement device, correlations without a LHS model cannot, in general, be obtained from them. Working with assemblages, in contrast, is advantageous precisely because it removes the need of measurement specification in the untrusted part. Finally, it is important to note that post-quantum steering, i.e., steerable assemblages that, in spite of satisfying the no-signalling principle, do not admit a quantum realization, has been recently discovered [28].

III. OPERATIONAL FRAMEWORK

In this section, we show that stochastic 1W-LOCCs from Bob to Alice do not create steering and can, therefore, be

taken as free operations for steering. We consider the general scenario of stochastic 1W-LOCCs, i.e., 1W-LOCCs that do not necessarily occur with certainty, which map the initial assemblage $\rho_{A|X}$ into a final assemblage $\rho_{A_f|X_f}$ (see Fig. 1). Bob's generic quantum operation can be represented by a (possibly incomplete) generalized measurement. This measurement is described by a completely positive non-trace-preserving map $\mathcal{E}: \mathcal{L}(\mathcal{H}_B) \to \mathcal{L}(\mathcal{H}_{Bf})$ defined by

$$\mathcal{E}(\cdot) \coloneqq \sum_\omega \mathcal{E}_\omega(\cdot), \quad \text{with} \quad \mathcal{E}_\omega(\cdot) \coloneqq K_\omega \cdot K_\omega^\dagger, \quad (7a)$$

$$\text{such that} \sum_\omega K_\omega^\dagger K_\omega \leq \mathbb{1}, \quad (7b)$$

where $\mathcal{H}_{Bf}$ is the final Hilbert space, of dimension d_f, and $K_\omega: \mathcal{H}_B \to \mathcal{H}_{Bf}$ is the measurement operator corresponding to the ωth measurement outcome. For any normalized $\varrho_B \in \mathcal{L}(\mathcal{H}_B)$, the trace $\mathrm{Tr}[\mathcal{E}(\varrho_B)] \leq 1$ of the map's output $\mathcal{E}(\varrho_B)$ represents the probability that the physical transformation $\varrho_B \to \mathcal{E}(\varrho_B)/\mathrm{Tr}[\mathcal{E}(\varrho_B)]$ takes place. In turn, the map $\mathcal{E}_\omega(\cdot)$ describes the post-selection of the ωth outcome, which occurs with a probability
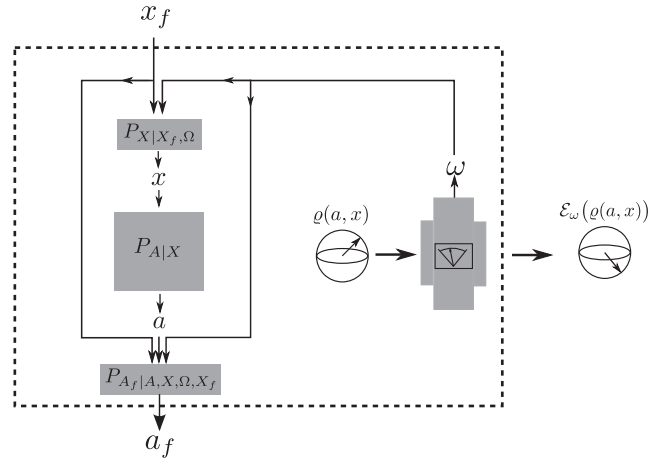

FIG. 1. Schematic representation of a 1W-LOCC map $\mathcal{M}$: The initial assemblage $\rho_{A|X}$ consists of a black box, with inputs x and outputs a, governed by the probability distribution $P_{A|X}$, in Alice's hand, and a quantum subsystem in one of the states $\{\varrho(a,x)\}_{a,x}$, in Bob's hands. The final assemblage $\rho_{A_f|X_f} = \mathcal{M}(\rho_{A|X})$ is given by a final black box, represented by the dashed-lined rectangle, of inputs x_f and outputs a_f, and a final subsystem, outside the dashed-lined rectangle, in the state $\varrho(a_f,x_f) = \mathcal{E}_\omega(\varrho(a,x))$. To implement $\mathcal{M}$, first Bob applies, with a probability $P_\Omega(\omega)$, a stochastic quantum operation $\mathcal{E}_\omega$ that leaves his subsystem in the state $\mathcal{E}_\omega(\varrho(a,x))$. He communicates ω to Alice. Then, Alice generates x by processing the classical bits ω and x_f with a local wiring described by a conditional distribution $P_{X|X_f,\Omega}$. She inputs x to her initial device, upon which the bit a is produced. Finally, Alice generates the output a_f of the final device by processing x_f, ω, x, and a, with a local wiring described by a distribution $P_{A_f|A,X,\Omega,X_f}$.

$$P_\Omega(\omega) := \text{Tr}[\mathcal{E}_\omega(\rho_B)] = \text{Tr}[K_\omega \varrho_B K_\omega^\dagger] \leq 1. \quad (8)$$

Since Alice can only process classical information, the allowed one-way communication from Bob to her must be classical, too. Thus, it can only consist of the outcome ω of his quantum operation. Classical bit processing is usually referred to as wiring [5]. Alice's wiring maps $a \in [r]$ and $x \in [s]$ into output and input bits $a_f \in [r_f]$ and $x_f \in [s_f]$, respectively, of the final assemblage, where s_f and r_f are natural numbers. The most general wirings respecting the above constraints are described by conditional probability distributions $P_{X|X_f,\Omega}$ and $P_{A_f|A,X,\Omega,X_f}$ of generating x from ω and x_f and a_f from x_f, ω, x, and a, respectively, as sketched in Fig. 1. Finally, Alice's wirings must be deterministic [30]. This means that, in particular, Alice cannot abort the experimental run. Without this condition the resulting maps would not be steering nonincreasing. The physical motivation for this condition will become clear in the next section. The condition implies that $P_{X|X_f,\Omega}$ and $P_{A_f|A,X,\Omega,X_f}$ must be well-normalized probability distributions.

The general form of the resulting maps is parametrized in the following definition (see Appendix A for details).

Definition 1 (Stochastic assemblage 1W-LOCCs): We define the class 1W-LOCC of *(stochastic) 1W-LOCCs* as the set of (stochastic) maps $\mathcal{M}$ that take an arbitrary assemblage $\hat{\rho}_{A|X}$ into a final assemblage $\hat{\rho}_{A_f|X_f} := \mathcal{M}(\hat{\rho}_{A|X})$, where

$$\mathcal{M}(\hat{\rho}_{A|X}) := \sum_\omega (\mathbb{1} \otimes K_\omega) \mathcal{W}_\omega(\hat{\rho}_{A|X})(\mathbb{1} \otimes K_\omega^\dagger), \quad (9)$$

with $\mathcal{W}_\omega$ being a *deterministic* wiring map given by

$$[\mathcal{W}_\omega(\hat{\rho}_{A|X})](x_f) := \sum_{a_f,a,x} P(x|x_f,\omega)P(a_f|a,x,\omega,x_f)$$
$$\times (|a_f\rangle\langle a| \otimes \mathbb{1})\hat{\rho}_{A|X}(x)(|a\rangle\langle a_f| \otimes \mathbb{1}), \quad (10)$$

where $P(x|x_f,\omega)$ and $P(a_f|a,x,\omega,x_f)$ are short-hand notations for the conditional probabilities $P_{X|X_f,\Omega}(x,x_f,\omega)$ and $P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)$, respectively.

Note that the final assemblage (9) is, in general, not normalized: Introducing

$$\mathcal{M}_\omega(\cdot) := (\mathbb{1} \otimes K_\omega)\mathcal{W}_\omega(\cdot)(\mathbb{1} \otimes K_\omega^\dagger), \quad (11)$$

such that $\mathcal{M}(\cdot) = \sum_\omega \mathcal{M}_\omega(\cdot)$, we obtain, using Eqs. (3), (4), (5), (8), (9), and (10),

$$\text{Tr}[\mathcal{M}(\hat{\rho}_{A|X})] = \sum_\omega \text{Tr}[\mathcal{M}_\omega(\hat{\rho}_{A|X})] = \sum_\omega P_\Omega(\omega) \leq 1. \quad (12)$$

As with quantum operations, the trace (12) of $\mathcal{M}(\hat{\rho}_{A|X})$ represents the probability that the physical transformation $\hat{\rho}_{A|X} \to \mathcal{M}(\hat{\rho}_{A|X})/\text{Tr}[\mathcal{M}(\hat{\rho}_{A|X})]$ takes place. Analogously,

the map $\mathcal{M}_\omega$ describes the assemblage transformation that takes place when Bob post-selects the ωth outcome, which occurs with probability $\text{Tr}[\mathcal{M}_\omega(\hat{\rho}_{A|X})] = P_\Omega(\omega)$. In the particular case where $\mathcal{M}$ is trace preserving, we refer to it as a deterministic 1W-LOCC.

Finally, in Appendix B, we prove the following theorem.

Theorem 1 (1W-LOCC invariance of LHS): Any map of the class 1W-LOCC takes every unsteerable assemblage into an unsteerable assemblage.

IV. PHYSICAL MOTIVATION FOR FREE OPERATIONS: 1W-LOCCS AS SAFE OPERATIONS IN 1S-DI-QKD

As shown in the previous section, stochastic 1W-LOCCs from Bob to Alice satisfy the basic requirement of mapping every unsteerable assemblage into an unsteerable assemblage. However, there may, in general, exist other sets of operations with this feature. In entanglement theory, for instance, apart from the LOCCs, the separable operations [16], the entanglement-assisted (catalytic) LOCCs [17], or simply the local operations, as well as any of these supplemented with particle swapping [4], are known not to create entanglement either. Each of these classes of operations leads, strictly speaking, to a valid resource theory of entanglement. The choice of a given class over others is based upon actual constraints from the physical scenario in question. In what follows, we discuss the role of 1W-LOCCs as the allowed operations in 1S-DI-QKD, in the sense of being those that do not compromise the security. This gives us a physical motivation to choose 1W-LOCCs as free operations for steering over other classes of operations that may also map LHS into itself. To this end, and for pedagogic reasons, we first discuss the allowed safe operations for QKD and fully DI QKD.

QKD consists of the extraction of a secret key from the correlations of local-measurement outcomes on a bipartite quantum state. The most fundamental constraint to which any generic QKD protocol is subject is, of course, the lack of a private, safe, classical-communication channel between distant labs. If such a channel were available, the whole enterprise of QKD would be pointless. This imposes restrictions on the operations allowed, so as not to break the security of the protocol. For instance, clearly, the local-measurement outcomes cannot be communicated, as they can be intercepted by potential eavesdroppers who could use them to extract the key. Of particular relevance for this work are the assumptions on the measurement devices. In non-DI QKD protocols, entanglement is the resource and security is proven under the assumption that the users have a specific quantum state and perfectly characterized measurement devices [31]. Knowledge of the state by an eavesdropper does not compromise the security. Therefore, prior to the measurements producing the key, the users are allowed to preprocess the state in any way and exchange information about it, for instance with LOCCs, or even to

discard the state aborting the protocol run. Preprocessing LOCCs or abortions can, at most, provide an eavesdropper with knowledge about the state, not about the key, and therefore do not affect the security.

The situation is different in DI-QKD [12]. In that case, the resource is given by Bell nonlocal correlations, and no assumption is made either on the quantum state or on the measurement devices. The users effectively hold black-box devices, whose inputs and outputs are all to which they have access. Since such inputs and outputs are precisely the bits with which the key is established, both classical communication and abortions are forbidden. Communication of outputs can directly reveal the key, as mentioned, whereas abortions and communication of inputs can, because of the locality and detection loop holes, respectively, be maliciously exploited by an eavesdropper to obtain information about the key, too. Hence, the security constraints of DI-QKD naturally yield local classical information processing assisted by shared randomness or prior-to-input classical communication as classes of allowed operations [21,22].

In 1S-DIQKD, in contrast, while no assumption is made on the bipartite quantum state or Alice's apparatus, Bob's measurement device is perfectly characterized. This is effectively described by assemblages of the form given in Eq. (1). The asymmetry in Alice's and Bob's devices leads to an asymmetry in the operations allowed for each of them. Alice is subject to the same restrictions as in both-sided DI QKD, while Bob is subject to those of non-DI QKD. Hence, Alice cannot abort or transmit any information, whereas, before measuring, Bob is allowed to implement arbitrary preprocessing quantum operations to his subsystem, including stochastic ones with possible abortions, and send any classical feedback about them to Alice. Altogether, this singles out a natural set of operations that do not compromise the security: all the assemblage transformations involving only deterministic classical maps on Alice's side and arbitrary—possibly stochastic—quantum operations on Bob's, assisted by one-way classical communication from Bob to Alice. These are, namely, the stochastic 1W-LOCCs from Bob to Alice (see Fig. 1). Note that shared randomness or prior-to-input classical communication from Alice to Bob [21,22], which do not introduce any security compromise, can always be recast as 1W classical communication from Bob to Alice and, therefore, do not need to be explicitly considered.

Finally, we emphasize that what is assumed to be untrusted is one of the measurement apparatuses, not the users. Both users are trusted and reliably operate on their systems, carrying out the allowed assemblage transformations.

V. STEERING MONOTONICITY

Once the free operations for steering are established, the natural next step is to introduce an axiomatic approach to define steering measures, i.e., a set of postulates that a bona fide quantifier of steering should fulfill.

Definition 2 (1W-LOCC monotonicity and convexity): A function $\mathcal{S}$, from the space of assemblages into $\mathbb{R}_{\geq 0}$, is a *steering monotone* if it fulfills the following two axioms:

(i) $\mathcal{S}(\hat{\rho}_{A|X}) = 0$ for all $\hat{\rho}_{A|X} \in LHS$.

(ii) $\mathcal{S}$ does not increase, on average, under deterministic 1W-LOCCs, i.e.,

$$\sum_\omega P_\Omega(\omega)\mathcal{S}\left(\frac{\mathcal{M}_\omega(\hat{\rho}_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\hat{\rho}_{A|X})]}\right) \leq \mathcal{S}(\hat{\rho}_{A|X}) \quad (13)$$

for all $\hat{\rho}_{A|X}$, with $P_\Omega(\omega) = \text{Tr}[\mathcal{M}_\omega(\hat{\rho}_{A|X})]$ and $\sum_\omega P_\Omega = 1$.

In addition, $\mathcal{S}$ is a *convex steering monotone* if it also satisfies the following property:

(iii) Given any real number $0 \leq \mu \leq 1$, and assemblages $\hat{\rho}_{A|X}$ and $\hat{\rho}'_{A|X}$, we have

$$\mathcal{S}(\mu\hat{\rho}_{A|X} + (1-\mu)\hat{\rho}'_{A|X}) \leq \mu\mathcal{S}(\hat{\rho}_{A|X})$$
$$+ (1-\mu)\mathcal{S}(\hat{\rho}'_{A|X}). \quad (14)$$

Condition (i) reflects the basic fact that unsteerable assemblages should have zero steering. Condition (ii) formalizes the intuition that, analogously to entanglement, steering should not increase—on average—under 1W-LOCCs, even if the flag information ω produced in the transformation is available. Finally, condition (iii) states the desired property that steering should not increase by probabilistically mixing assemblages. The first two conditions are taken as mandatory necessary conditions, the third one only as a convenient property. Importantly, there exists a less demanding definition of monotonicity. In this definition, the left-hand side of Eq. (13) is replaced by $\mathcal{S}(\mathcal{M}(\hat{\rho}_{A|X})/\text{Tr}[\mathcal{M}(\hat{\rho}_{A|X})])$. In other words, (ii′) only demands that the steering itself, instead of its average over ω, is nonincreasing under the free operations. The latter is actually the most fundamental necessary condition that a quantifier of a given resource must fulfill. However, monotonicity (ii) is in many cases (including the present work) easier to prove and, together with condition (iii), implies monotonicity (ii′). Hence, we focus throughout on monotonicity as defined by Eq. (13) and refer to it simply as 1W-LOCC monotonicity. All three known quantifiers of steering, the two introduced in Refs. [24,25] as well as the one we introduce in the next section, turn out to be convex steering monotones in the sense of Definition 2.

VI. RELATIVE ENTROPY OF STEERING

In this section, we introduce a convex steering monotone called the relative entropy of steering. To this end, we first define the notion of relative entropy between assemblages. For any two density operators ϱ and ϱ', let us recall the quantum von-Neumann relative entropy

$$S_Q(\varrho\|\varrho') := \mathrm{Tr}[\varrho(\log\varrho - \log\varrho')] \tag{15}$$

of ϱ with respect to ϱ' and, for any two probability distributions P_X and P_X', the classical relative entropy, or Kullback-Leibler divergence,

$$S_C(P_X\|P_X') := \sum_x P_X(x)[\log P_X(x) - \log P_X'(x)], \tag{16}$$

of P_X with respect to P_X'. The quantum and classical relative entropies (15) and (16) measure the distinguishability of states and distributions, respectively. To find an equivalent measure for assemblages, we note, for $\hat\rho_{A|X}(x)$ given by Eq. (3) and $\hat\rho'_{A|X}(x) := \sum_a P_{A|X}'(a,x)|a\rangle\langle a| \otimes \varrho'(a,x)$, that

$$S_Q(\hat\rho_{A|X}(x)\|\hat\rho'_{A|X}(x)) = S_C(P_{A|X}(\cdot,x)\|P_{A|X}'(\cdot,x))$$
$$+ \sum_a P_{A|X}(a,x)S_Q(\varrho(a,x)\|\varrho'(a,x)), \tag{17}$$

where $P_{A|X}(\cdot,x)$ and $P_{A|X}'(\cdot,x)$ are, respectively, the distributions over a obtained from the conditional distributions $P_{A|X}(\cdot,x)$ and $P_{A|X}'(\cdot,x)$ for a fixed x. In other words, the distinguishability between the states $\hat\rho_{A|X}(x)$ and $\hat\rho'_{A|X}(x) \in$

$\mathcal{L}(\mathcal{H}_E \otimes \mathcal{H}_B)$ equals the sum of the distinguishabilities between $P_{A|X}(\cdot,x)$ and $P_{A|X}'(\cdot,x)$ and between $\varrho(a,x)$ and $\varrho'(a,x) \in \mathcal{L}(\mathcal{H}_B)$, weighted by $P_{A|X}(a,x)$ and averaged over a.

The entropy (17), which depends on x, does not measure the distinguishability between the assemblages $\rho_{A|X}$ and $\rho'_{A|X}$. Since the latter are conditional objects, i.e., with inputs, a general strategy to distinguish them must allow for Alice choosing the input for which the assemblages' outputs are optimally distinguishable. Furthermore, Bob can first apply a generalized measurement on his subsystem and communicate the outcome γ to her, which she can then use for her input choice. This is the most general procedure within the allowed 1W-LOCCs. Hence, a generic distinguishing strategy under 1W-LOCCs involves probabilistically chosen inputs that depend on γ. Note, in addition, that distributions P_Γ or P_Γ', encode differences between $\rho_{A|X}$ and $\rho'_{A|X}$, too, and must therefore also be accounted for by a distinguishability measure. The following definition incorporates all of these considerations.

Definition 3 (Relative entropy between assemblages): Given any two assemblages $\rho_{A|X}$ and $\rho'_{A|X}$, we define the *assemblage relative entropy* of $\rho_{A|X}$ with respect to $\rho'_{A|X}$ as

$$S_A(\rho_{A|X}\|\rho'_{A|X}) := \max_{P_{X|\Gamma},\{E_\gamma\}} \left[S_C(P_\Gamma\|P_\Gamma') + \sum_{\gamma,x} P(x|\gamma)P_\Gamma(\gamma)S_Q\left(\frac{\mathbb{1} \otimes E_\gamma\hat\rho_{A|X}(x)\mathbb{1} \otimes E_\gamma^\dagger}{P_\Gamma(\gamma)} \middle\| \frac{\mathbb{1} \otimes E_\gamma\hat\rho'_{A|X}(x)\mathbb{1} \otimes E_\gamma^\dagger}{P_\Gamma'(\gamma)}\right) \right], \tag{18}$$

where $E_\gamma: \mathcal{H}_B \to \mathcal{H}_B$ are generalized-measurement operators such that $\sum_\gamma E_\gamma^\dagger E_\gamma = \mathbb{1}$, $P_{X|\Gamma}$ is a conditional probability distribution of x given γ, the short-hand notation $P(x|\gamma) := P_{X|\Gamma}(x,\gamma)$ has been used, and

$$P_\Gamma(\gamma) := \mathrm{Tr}[\mathbb{1} \otimes E_\gamma\hat\rho_{A|X}(x)\mathbb{1} \otimes E_\gamma^\dagger] = \mathrm{Tr}_B[E_\gamma\varrho_B E_\gamma^\dagger], \tag{19a}$$

$$P_\Gamma'(\gamma) := \mathrm{Tr}[\mathbb{1} \otimes E_\gamma\hat\rho'_{A|X}(x)\mathbb{1} \otimes E_\gamma^\dagger] = \mathrm{Tr}_B[E_\gamma\varrho_B' E_\gamma^\dagger], \tag{19b}$$

where ϱ_B' is Bob's reduced state for the assemblage $\rho'_{A|X}$.

In Appendix C, we show that S_A does not increase–on average–under deterministic 1W-LOCCs and, as its quantum counterpart S_Q, is jointly convex. Hence, S_A is a proper measure of distinguishability between assemblages under 1W-LOCCs [32]. The first term inside the maximization in Eq. (18) accounts for the distinguishability between the distributions of measurement outcomes γ, and the second one accounts for that between the distributions of Alice's outputs and Bob's states resulting from each γ, averaged over all inputs and measurement outcomes. In turn, the maximization over $\{E_\gamma\}$ and $P_{X|\Gamma}$ ensures that these output distributions and states are distinguished using the optimal 1W-LOCC-compatible strategy.

We are now in a good position to introduce a convex steering monotone. We do this with a theorem.

Theorem 2 (1W-LOCC monotonicity and convexity of $\mathcal{S}_R$): The *relative entropy of steering* $\mathcal{S}_R$, defined for an assemblage $\rho_{A|X}$ as

$$\mathcal{S}_R(\rho_{A|X}) := \min_{\sigma_{A|X}\in\mathrm{LHS}} S_A(\rho_{A|X}\|\sigma_{A|X}), \tag{20}$$

is a convex steering monotone.

The theorem is proven in Appendix C.

VII. OTHER CONVEX STEERING MONOTONES

Apart from $\mathcal{S}_R$, two other quantifiers of steering have recently been proposed: the steerable weight [24] and the robustness of steering [25]. In this section, we show that these are also convex steering monotones.

Definition 4 (Steerable weight [24]): The steerable weight $\mathcal{S}_W(\rho_{A|X})$ of a normalized assemblage $\rho_{A|X}$ is the minimum $\nu \in \mathbb{R}_{\geq 0}$ such that

$$\rho_{A|X} = \nu\tilde\rho_{A|X} + (1-\nu)\sigma_{A|X}, \tag{21}$$

with $\tilde\rho_{A|X}$ an arbitrary a normalized assemblage $\sigma_{A|X}$ in LHS.

Definition 5 (Robustness of steering [25]): The robustness of steering $\mathcal{S}_{\text{rob}}(\rho_{A|X})$ of a normalized assemblage $\rho_{A|X}$ is the minimum $\nu \in \mathbb{R}_{\geq 0}$ such that the normalized assemblage

$$\sigma_{A|X} := \frac{1}{1+\nu}\rho_{A|X} + \frac{\nu}{1+\nu}\tilde{\rho}_{A|X} \qquad (22)$$

belongs to LHS, with $\tilde{\rho}_{A|X}$ an arbitrary normalized assemblage.

In Appendix D, we prove the following theorem.

Theorem 3 (1W-LOCC monotonicity and convexity of $\mathcal{S}_{\text{W}}$ and $\mathcal{S}_{\text{rob}}$): Both $\mathcal{S}_{\text{W}}$ and $\mathcal{S}_{\text{rob}}$ are convex steering monotones.

To conclude the section, we note that a steering measure for assemblages containing continuous-variable (CV) bosonic systems in Gaussian states has appeared very recently [33]. Our formalism can be straightforwardly extended to CV systems. However, such an extension is outside of the scope of the present paper.

VIII. ASSEMBLAGE CONVERSIONS AND NO STEERING BITS

We say that $\Psi_{A|X}$ and $\Psi'_{A|X}$ are pure assemblages if they are of the form

$$\Psi_{A|X} := \{P_{A|X}(a,x), |\psi(a,x)\rangle\langle\psi(a,x)|\}_{a,x}, \qquad (23a)$$

$$\Psi'_{A|X} := \{P'_{A|X}(a,x), |\psi'(a,x)\rangle\langle\psi'(a,x)|\}_{a,x}, \qquad (23b)$$

where $|\psi(a,x)\rangle$ and $|\psi'(a,x)\rangle \in \mathcal{H}_B$. Besides, we say that they are pure orthogonal assemblages if, in addition, $\langle\psi(a,x)|\psi(\tilde{a},x)\rangle = \delta_{a\tilde{a}} = \langle\psi'(a,x)|\psi'(\tilde{a},x)\rangle$ for all x. Note that pure orthogonal assemblages are the ones obtained when Alice and Bob share a pure maximally entangled state and Alice performs a von Neumann measurement on her share. We present two theorems about assemblage conversions under 1W-LOCCs.

The first one, proven in Appendix E, establishes necessary and sufficient conditions for stochastic-1W-LOCC conversions between pure orthogonal assemblages, therefore playing a similar role here to the one played in entanglement theory by Vidal's theorem [34] for stochastic-LOCC pure-state conversions.

Theorem 4 (Criterion for stochastic-1W-LOCC conversion): Let $\Psi_{A|X}$ and $\Psi'_{A|X}$ be any two pure orthogonal assemblages with $d = s = r = 2$. Then, $\Psi_{A|X}$ can be transformed into $\Psi'_{A|X}$ by a stochastic 1W-LOCC iff either $\Psi'_{A|X} \in LHS$ or $P'_{A|X} = P_{A|X}$ and

$$|\langle\psi'(a,0)|\psi'(a,1)\rangle| = |\langle\psi(a \oplus \alpha, 0)|\psi(a \oplus \alpha, 1)\rangle| \quad \forall a, \qquad (24)$$

for some $\alpha \in \{0, 1\}$.

In other words, no pure orthogonal assemblage of minimal dimension can be obtained via a 1W-LOCC, not even probabilistically, from a pure orthogonal assemblage of minimal dimension with a different state-basis overlap (except for trivial relabelings of a, given by α) unless the former is unsteerable. Hence, each state-basis overlap defines an inequivalent class of steering. It thus follows that there are infinitely many inequivalent classes of steering. In a way, this is reminiscent of the inequivalent classes of entanglement in multipartite [35] or infinite-dimensional bipartite [36] systems, but here the phenomenon takes place already for bipartite systems of minimal dimension.

The second theorem, proven in Appendix F, rules out the possibility of there being a (nonorthogonal) minimal-dimension pure assemblage from which all assemblages can be obtained.

Theorem 5 (Nonexistence of steering bits): There exists no pure assemblage with $d = s = r = 2$ that can be transformed into any assemblage by stochastic 1W-LOCCs.

Hence, among the minimal-dimension assemblages, there is no operationally well-defined unit of steering, or steering bit, i.e., an assemblage from which all assemblages can be obtained for free and can therefore be taken as a measure-independent maximally steerable assemblage. This is again in striking contrast to entanglement theory, where pure maximally entangled states can be defined without the need for entanglement quantifiers, and each one can be transformed into any state by deterministic LOCCs [34,37].

IX. DISCUSSION AND OUTLOOK

We have introduced the resource theory of Einstein-Podolsky-Rosen steering. The free operations of the theory are the 1W-LOCCs from the quantum part to the black box, i.e., all the assemblage transformations involving deterministic bit wirings on Alice's side and stochastic quantum operations on Bob's assisted by one-way classical communication from Bob to Alice. These operations satisfy the basic requirement of mapping all unsteerable assemblages into unsteerable assemblages and are also the allowed operations that naturally arise from the basic security constraints of one-sided device-independent QKD, where steering is a physical resource. With these operations, we introduced the notion of convex steering monotones, presented the relative entropy of steering as a convenient example thereof, and proved monotonicity and convexity of two other previously proposed steering measures. In addition, for minimal-dimensional systems, we established necessary and sufficient conditions for stochastic-1W-LOCC conversions between pure-state assemblages and proved the nonexistence of steering bits.

It is instructive to emphasize that the derived 1W-LOCCs are hybrids between the operations that map separable states into separable states, stochastic LOCCs, and those

that map Bell local correlations into Bell local correlations, local wirings assisted by prior-to-input classical communication [21,22,38]. In fact, our findings are also potentially useful for the quantification of Bell nonlocality. In addition, our work offers a number of challenges for future research. Namely, for example, the nonexistence of steering bits of minimal dimension can be seen as an impossibility of steering dilution of minimal-dimension assemblages in the single-copy regime. We leave as open questions what the rules for steering dilution and distillation are for higher-dimensional systems, mixed-state assemblages, or in asymptotic multicopy regimes, and what the steering classes are for mixed-state assemblages. Moreover, another fascinating question is whether one can formulate a notion of bound steering or an analogue to the positive-partial-transpose criterion for assemblages.

APPENDIX A: PARAMETRIZATION OF THE 1W-LOCC CLASS

In this appendix, we show that any generic assemblage map $\mathcal{M}$ involving stochastic local quantum operations on Bob's side, one-way classical communication from Bob to Alice, and deterministic (probability-preserving) local wirings on Alice's side is of the form given by Eqs. (9) and (10) and, therefore, belongs to the 1W-LOCC class of Definition 1.

Without loss of generality, such $\mathcal{M}$ can be decomposed into the following sequence (see Fig. 1) of operations:

(1) Bob applies an arbitrary stochastic generalized measurement, described by a completely positive non-trace-preserving map $\mathcal{E}$, defined by Eqs. (7), to his quantum subsystem before Alice introduces an input to her device. Note that, since the nonsignalling condition (4) is fulfilled, Bob has a well-defined reduced quantum state ρ_B, given by Eq. (4), independently of Alice still not having chosen her measurement input x. Therefore, Bob's measurement gives the outcome ω with the x-independent probability $P_\Omega(\omega)$ given by Eq. (8).

(2) Bob sends the outcome ω to Alice. Alice applies a local wiring, described by the normalized conditional probability distribution $P_{X|X_f,\Omega}$, to the input x_f of the final device and to ω, and uses the output of this wiring as the input x of her initial device. Given a value of a, x and ω, Alice's output a is generated with the conditional probability distribution $P_{A|X,\Omega,X_f}$. In that case, Bob's normalized state is given by the x_f-independent density operator

$$\varrho(a, x, \omega, x_f) := \frac{K_\omega \varrho(a,x) K_\omega^\dagger}{\mathrm{Tr}[K_\omega \varrho(a,x) K_\omega^\dagger]}. \qquad (A1)$$

(3) Alice applies a local wiring, described by the normalized conditional probability distribution $P_{A_f|A,X,\Omega,X_f}$, to all the previously generated classical bits, a, x, ω, and x_f, and uses the output of this wiring as the output a_f of her final device. This final processing of the bit a_f does not affect Bob's state. Thus, Bob's system ends up in the state $\varrho(a, x, \omega, x_f, a_f) := \varrho(a, x, \omega, x_f)$.

We denote by $P^i_{\Omega|A,X}$ the conditional distribution of Ω given A and X, for X chosen independently of Ω (in contrast to step 2 above), with elements $P^i_{\Omega|A,X}(\omega, a, x) := \mathrm{Tr}[K_\omega \rho(a,x) K_\omega^\dagger]$. With this notation the components $\varrho_{A_f|X_f}(a_f, x_f)$ of the final assemblage $\rho_{A_f|X_f}$ are explicitly given by

$$\varrho_{A_f|X_f}(a_f, x_f) := P_{A_f|X_f}(a_f, x_f) \times \varrho_f(a_f, x_f)$$

$$= \sum_{a,x,\omega} P_{A_f,A,X,\Omega|X_f}(a_f, a, x, \omega, x_f) \varrho(a_f, a, x, \omega, x_f) \qquad (A2)$$

$$= \sum_{a,x,\omega} P_{A_f,A,X,\Omega|X_f}(a_f, a, x, \omega, x_f) \frac{K_\omega \varrho(a,x) K_\omega^\dagger}{\mathrm{Tr}[K_\omega \varrho(a,x) K_\omega^\dagger]} \qquad (A3)$$

$$= \sum_{a,x,\omega} P_\Omega(\omega) P_{X|X_f,\Omega}(x, x_f, \omega) P_{A|X,\Omega}(a, x, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) \frac{K_\omega \varrho(a,x) K_\omega^\dagger}{P^i_{\Omega|A,X}(\omega, a, x)}. \qquad (A4)$$

Equation (A2) follows from basic properties of probability distributions and ensembles of states. Equation (A3) follows from the definition of $\rho(a, x, \omega, x_f, a_f)$. Equation (A4) follows from Bayes' theorem, together with the facts that $P_{A|X,\Omega,X_f} = P_{A|X,\Omega}$ (the output of Alice's initial device only depends on the input x and the measurement outcome ω) and

$P_{\Omega|X_f} = P_\Omega$ (the measurement outcome ω is independent of the input of Alice's final device), and from the definition of $P^i_{\Omega|A,X}$. Next, note that, since the statistics of A is fully determined by X and Ω regardless of whether X and Ω are independent or not, it holds that

$$
\begin{aligned}
P_{A|X,\Omega} &= P^i_{A|X,\Omega} \\
&= \frac{P^i_{A,X,\Omega}}{P^i_{X|\Omega}P^i_\Omega} \\
&= \frac{P^i_{A,X,\Omega}}{P^i_X P_\Omega} \\
&= \frac{P^i_{\Omega|A,X}P^i_{A,X}}{P^i_X P_\Omega} \\
&= \frac{P^i_{\Omega|A,X}P^i_{A|X}}{P_\Omega},
\end{aligned}
\tag{A5}
$$

where we have used Bayes' theorem, that $P^i_\Omega = P_\Omega$ and that, by definition, $P^i_{X|\Omega} = P^i_X$. Inserting Eq. (A5) into Eq. (A4), we obtain

$$
\begin{aligned}
\varrho_{A_f|X_f}(a_f, x_f) &= \sum_{a,x,\omega} P_{X|X_f,\Omega}(x, x_f, \omega) P^i_{A|X}(a,x) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) K_\omega \varrho(a,x) K_\omega^\dagger \\
&= \sum_{a,x,\omega} P_{X|X_f,\Omega}(x, x_f, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) K_\omega \varrho_{A|X}(a,x) K_\omega^\dagger, \quad \forall\ (a_f, x_f),
\end{aligned}
\tag{A6}
$$

where Eq. (A6) follows from the fact that $P^i_{A|X} = P_{A|X}$ and from the definition of $\varrho_{A|X}$. The right-hand side of Eq. (A6) gives the most general expression of the components $\varrho_{A_f|X_f}(a_f, x_f)$ of $\mathcal{M}(\rho_{A|X})$ explicitly as a function of the components $\varrho_{A|X}$ of $\rho_{A|X}$. The reader can straightforwardly verify that the quantum representation $\mathcal{M}(\hat\rho_{A|X})$ of the obtained final assemblage $\mathcal{M}(\rho_{A|X})$ is given by the right-hand side of Eq. (9).

APPENDIX B: INVARIANCE OF LHS UNDER 1W-LOCC MAPS

We now show that if $\rho_{A|X} \in$ LHS then, for all $\mathcal{M} \in$ 1W-LOCC, $\mathcal{M}(\rho_{A|X}) \in$ LHS.
Proof of Theorem 1.— Replacing $\rho_{A|X}$ in Eq. (A6) by the right-hand side of Eq. (6), we write

$$
\begin{aligned}
\varrho_{A_f|X_f} &= \sum_{a,x,\omega,\lambda} P_\Lambda(\lambda) P_{A|X,\Lambda}(a,x,\lambda) P_{X|X_f,\Omega}(x, x_f, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) K_\omega \xi(\lambda) K_\omega^\dagger \\
&= \sum_{a,x,\omega,\lambda} P_\Lambda(\lambda) P_{\Omega|\Lambda}(\omega, \lambda) P_{A|X,\Lambda}(a,x,\lambda) P_{X|X_f,\Omega}(x, x_f, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) \xi(\lambda, \omega),
\end{aligned}
\tag{B1}
$$

where the conditional probability $P_{\Omega|\Lambda}(\omega, \lambda) := \mathrm{Tr}[K_\omega \xi(\lambda) K_\omega^\dagger]$ and the normalized state $\xi(\lambda, \omega) := [K_\omega \xi(\lambda) K_\omega^\dagger]/[P_{\Omega|\Lambda}(\omega, \lambda)]$ have been introduced. Using the fact that a_f does not explicitly depend on λ, we see that

$$
P_{A_f|A,X,\Omega,X_f} = P_{A_f|A,X,\Omega,\Lambda,X_f}.
\tag{B2}
$$

In turn, using the facts that x is independent of λ and a depends only on x and λ, and Bayes' theorem, we see that

$$
P_{A|X,\Lambda}P_{X|X_f,\Omega} = P_{A|X,\Omega,\Lambda,X_f}P_{X|\Omega,\Lambda,X_f} = P_{A,X|\Omega,\Lambda,X_f}.
\tag{B3}
$$

Substituting into Eq. (B1) yields

$$
\begin{aligned}
\rho_{A_f|X_f} &= \sum_{a,x,\omega,\lambda} P_\Lambda(\lambda) P_{\Omega|\Lambda}(\omega, \lambda) P_{A,X|\Omega,\Lambda,X_f}(a, x, \omega, \lambda, x_f) P_{A_f|A,X,\Omega,\Lambda,X_f}(a_f, a, x, \omega, \lambda, x_f) \xi(\lambda, \omega) \\
&= \sum_{\omega,\lambda} P_{\Omega,\Lambda}(\omega, \lambda) P_{A_f|\Omega,\Lambda,X_f}(a_f, \omega, \lambda, x_f) \xi(\lambda, \omega)
\end{aligned}
\tag{B4}
$$

$$
= \sum_{\tilde\lambda} P_{\tilde\Lambda}(\tilde\lambda) P_{A_f|X_f,\tilde\Lambda}(a_f, x_f, \tilde\lambda) \sigma(\tilde\lambda),
\tag{B5}
$$

where Eq. (B4) follows from Bayes' theorem and by summing over x and a, and Eq. (B5) follows from defining the hidden variable $\tilde{\lambda} := (\omega, \lambda)$ governed by the normalized probability distribution $P_{\tilde{\Lambda}} := P_{\Omega,\Lambda}$. Equation (B5) manifestly shows that $\rho_{A_f|X_f} \in \text{LHS}$. ∎

APPENDIX C: RELATIVE ENTROPY OF STEERING

In this appendix, we prove Theorem 2. The proof strategy is similar to that of the proof that the relative entropy of entanglement for quantum states is a convex entanglement monotone [39]. It relies on two Lemmas, which we state next but whose proofs we leave for Appendix G.

Lemma 1. The assemblage relative entropy S_A, defined by Eq. (18), does not increase, on average, under deterministic 1W-LOCCs. In other words, for any map $\mathcal{M}$ of the form given by Eqs. (9) and (10) but with $\sum_\omega K_\omega^\dagger K_\omega = \mathbb{1}$ and any two assemblages $\rho_{A|X}$ and $\rho'_{A|X}$, S_A satisfies the inequality

$$\sum_\omega P_\Omega(\omega) S_A\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \frac{\mathcal{M}_\omega(\rho'_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]}\right)$$
$$\leq S_A(\rho_{A|X}|\rho'_{A|X}), \quad (C1)$$

where $\mathcal{M}_\omega$ is the stochastic map defined in Eq. (11), $P_\Omega = \text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]$, and $P'_\Omega = \text{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]$, with $\sum_\omega P_\omega = 1 = \sum_\omega P'_\omega$.

Lemma 2. The assemblage relative entropy S_A, defined by Eq. (18), is jointly convex. In other words, given two sets $\{\rho_{A|X}^{(j)}\}_{j=1,\ldots,n}$ and $\{\rho'^{(j)}_{A|X}\}_{j=1,\ldots,n}$ of n arbitrary assemblages each and n positive real numbers $\{\mu^{(j)}\}_{j=1,\ldots,n}$ such that $\sum_j \mu^{(j)} = 1$, with $n \in \mathbb{N}$, S_A satisfies the inequality

$$S_A\left(\sum_i \mu^{(j)} \rho_{A|X}^{(j)} \middle\| \sum_j \mu^{(j)} \rho'^{(j)}_{A|X}\right)$$
$$\leq \sum_j \mu^{(j)} S_A(\rho_{A|X}^{(j)} \| \rho'^{(j)}_{A|X}). \quad (C2)$$

We are now in a good position to prove the theorem.

Proof of Theorem 2.— That the relative entropy of steering $\mathcal{S}_R$, defined in Eq. (20), satisfies condition (i) follows immediately from its definition and the positivity of the von Neumann relative entropy for quantum states. Conditions (ii) and (iii), 1W-LOCC monotonicity and convexity of $\mathcal{S}_R$, can be proven in analogous fashion to LOCC monotonicity and convexity of the relative entropy of entanglement, respectively. We include their proofs for completeness.

To prove condition (ii), we denote by σ^* an unsteerable assemblage for which the minimization in Eq. (20) is attained, i.e., such that

$$S_A(\rho_{A|X}\|\sigma^*) := \mathcal{S}_R(\rho_{A|X}), \quad (C3)$$

and by $\hat{\sigma}_\mu^*$ an unsteerable assemblage such that

$$S_A\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \hat{\sigma}_\mu^*\right) := \mathcal{S}_R\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}\right). \quad (C4)$$

Then, we write

$$\sum_\omega P_\Omega(\omega) \mathcal{S}_R\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}\right)$$
$$= \sum_\omega P_\Omega(\omega) S_A\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \sigma_\mu^*\right)$$
$$\leq \sum_\omega P_\Omega(\omega) S_A\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \frac{\mathcal{M}_\omega(\sigma^*)}{\text{Tr}[\mathcal{M}_\omega(\sigma^*)]}\right) \quad (C5)$$

$$\leq S_A(\rho_{A|X}\|\sigma^*) \quad (C6)$$

$$= \mathcal{S}_R(\rho_{A|X}), \quad (C7)$$

where Eq. (C5) follows because σ_μ^* minimizes the assemblage relative entropy in each ωth term in the sum and $\mathcal{M}_\omega(\sigma^*)/\text{Tr}[\mathcal{M}_\omega(\sigma^*)] \in \text{LHS}$, Eq. (C6) follows from Lemma 1, and Eq. (C7) follows from the definition of σ^*.

To prove condition (iii), we further introduce unsteerable assemblages σ'^* and σ_{mix}^* such that

$$S_A(\rho'_{A|X}\|\sigma'^*) = \mathcal{S}_R(\rho'_{A|X}) \quad (C8)$$

and

$$S_A(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}\|\sigma_{\text{mix}}^*) = \mathcal{S}_R(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}). \quad (C9)$$

Then, we write

$$\mu\mathcal{S}_R(\rho_{A|X}) + (1-\mu)\mathcal{S}_R(\rho'_{A|X})$$
$$= \mu S_A(\rho_{A|X}\|\sigma^*) + (1-\mu)S_A(\rho'_{A|X}\|\sigma'^*)$$
$$\geq S_A(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}\|\mu\sigma^* + (1-\mu)\sigma'^*) \quad (C10)$$

$$\geq S_A(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}\|\sigma_{\text{mix}}^*) \quad (C11)$$

$$=: \mathcal{S}_R(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}), \quad (C12)$$

where Eq. (C10) holds because of Lemma 2, Eq. (C11) because σ_{min}^* minimizes the corresponding assemblage relative entropy and $\mu\sigma^* + (1-\mu)\sigma'^* \in \text{LHS}$, and Eq. (C12) by the definition of σ_{min}^*. ∎

APPENDIX D: PROOF OF THEOREM 3

Proof of Theorem 3.— Let us first prove the theorem's statement concerning the steerable weight. That $\mathcal{S}_{\mathrm{W}}$ satisfies condition (i) of Definition 2 follows immediately from its definition. To prove that it fulfills condition (ii), first, we apply the map $\mathcal{M}_\omega$ to both sides of Eq. (21) and renormalize. This yields

$$
\begin{aligned}
\frac{\mathcal{M}_\omega(\rho_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} &= \nu \frac{\mathcal{M}_\omega(\tilde{\rho}_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \\
&\quad + (1-\nu)\frac{\mathcal{M}_\omega(\sigma_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \\
&= \nu \frac{\mathcal{M}_\omega(\tilde{\rho}_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})]} \frac{\mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})]}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \\
&\quad + (1-\nu)\frac{\mathcal{M}_\omega(\sigma_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\sigma_{A|X})]} \frac{\mathrm{Tr}[\mathcal{M}_\omega(\sigma_{A|X})]}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}.
\end{aligned}
$$
(D1)

Denoting by ν_ω^* the minimum $\nu \in \mathbb{R}_{\geq 0}$ such that a decomposition of the form of Eq. (D1) is possible, it is clear that

$$
\nu_\omega^* \leq \mathcal{S}_{\mathrm{W}}(\rho_{A|X}),
$$
(D2)

as any $\nu \in \mathbb{R}_{\geq 0}$ that allows for a decomposition as in Eq. (21) also allows for one as in Eq. (D1). Furthermore, taking into account that, since $\sigma_{A|X} \in \mathrm{LHS}$, it holds that $\mathcal{M}_\omega(\sigma_{A|X})/\mathrm{Tr}[\mathcal{M}_\omega(\sigma_{A|X})] \in \mathrm{LHS}$, it is also clear that

$$
\mathcal{S}_{\mathrm{W}}\!\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}\right) \leq \nu_\omega^* \frac{\mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})]}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}.
$$
(D3)

Hence, we obtain

$$
\begin{aligned}
\sum_\omega P_\Omega(\omega) &\mathcal{S}_{\mathrm{W}}\!\left(\frac{\mathcal{M}_\omega(\rho_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}\right) \\
&\leq \sum_\omega P_\Omega(\omega)\nu_\omega^* \frac{\mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})]}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \\
&\leq \sum_\omega \mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})]\mathcal{S}_{\mathrm{W}}(\rho_{A|X}) \\
&= \mathcal{S}_{\mathrm{W}}(\rho_{A|X}),
\end{aligned}
$$
(D4)

where the second inequality is due to the fact that $P_\omega = \mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]$ and Eq. (D2), and the last equality is due to the fact that, since $\tilde{\rho}_{A|X}$ is a normalized assemblage and $\mathcal{M}$ is a deterministic map, $\sum_\omega \mathrm{Tr}[\mathcal{M}_\omega(\tilde{\rho}_{A|X})] = 1$.

To prove the validity of condition (iii) for $\mathcal{S}_{\mathrm{W}}$, we first write

$$
\begin{aligned}
\mu\rho_{A|X} &+ (1-\mu)\rho'_{A|X} \\
&= \mu[\mathcal{S}_{\mathrm{W}}(\rho_{A|X})\tilde{\rho}_{A|X} + (1-\mathcal{S}_{\mathrm{W}}(\rho_{A|X}))\sigma_{A|X}] \\
&\quad + (1-\mu)[\mathcal{S}_{\mathrm{W}}(\rho'_{A|X})\tilde{\rho}'_{A|X} + (1-\mathcal{S}_{\mathrm{W}}(\rho_{A|X}))\sigma'_{A|X}]
\end{aligned}
$$
(D5)

$$
= \nu^{(\mu)}\tilde{\rho}^{(\mu)}_{A|X} + (1-\nu^{(\mu)})\sigma^{(\mu)}_{A|X},
$$
(D6)

where Eq. (D5) holds because of the definition of $\mathcal{S}_{\mathrm{W}}$ and, in Eq. (D6), we have introduced the positive real

$$
\nu^{(\mu)} := \mu\mathcal{S}_{\mathrm{W}}(\rho_{A|X}) + (1-\mu)\mathcal{S}_{\mathrm{W}}(\rho'_{A|X}),
$$
(D7)

the normalized assemblage

$$
\tilde{\rho}^{(\mu)}_{A|X} := \frac{1}{\nu^{(\mu)}}[\mu\mathcal{S}_{\mathrm{W}}(\rho_{A|X})\tilde{\rho}_{A|X} + (1-\mu)\mathcal{S}_{\mathrm{W}}(\rho'_{A|X})\tilde{\rho}'_{A|X}],
$$
(D8)

and the normalized unsteerable assemblage

$$
\begin{aligned}
\sigma^{(\mu)}_{A|X} := \frac{1}{1-\nu^{(\mu)}}[&\mu(1-\mathcal{S}_{\mathrm{W}}(\rho_{A|X}))\sigma_{A|X} \\
&+ (1-\mu)(1-\mathcal{S}_{\mathrm{W}}(\rho_{A|X}))\sigma'_{A|X}].
\end{aligned}
$$
(D9)

Thus, the expression (D6) gives a decomposition of the mixture $\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}$ of the form of Eq. (21). However, it is not necessarily the optimal one. Hence, we get

$$
\mathcal{S}_{\mathrm{W}}(\mu\rho_{A|X} + (1-\mu)\rho'_{A|X}) \leq \nu^{(\mu)},
$$
(D10)

which, together with Eq. (D7), finishes the proof of convexity of $\mathcal{S}_{\mathrm{W}}$.

Similar arguments can be employed to prove the theorem's statement concerning the robustness of steering. That $\mathcal{S}_{\mathrm{rob}}$ satisfies condition (i) of Definition 2 also follows immediately by definition. Condition (ii) can be proven with a strategy similar to that for $\mathcal{S}_{\mathrm{W}}$. Condition (iii) can be proven by noting that Definition 5 implies that

$$
\rho_{A|X} = [1 + \mathcal{S}_{\mathrm{rob}}(\rho_{A|X})]\sigma_{A|X} - \mathcal{S}_{\mathrm{rob}}(\rho_{A|X})\tilde{\rho}_{A|X}
$$
(D11a)

$$
\rho'_{A|X} = [1 + \mathcal{S}_{\mathrm{rob}}(\rho'_{A|X})]\sigma'_{A|X} - \mathcal{S}_{\mathrm{rob}}(\rho'_{A|X})\tilde{\rho}'_{A|X},
$$
(D11b)

where the unsteerable assemblage $\sigma'_{A|X}$ and the arbitrary assemblage $\tilde{\rho}'_{A|X}$ play, respectively, the same roles for $\rho'_{A|X}$ as the ones played by $\sigma_{A|X}$ and $\tilde{\rho}_{A|X}$ for $\rho_{A|X}$ in Definition 5. Then, one can introduce the positive real

$$
\nu^{(\mu)} = \mu\mathcal{S}_{\mathrm{rob}}(\rho_{A|X}) + (1-\mu)\mathcal{S}_{\mathrm{rob}}(\rho'_{A|X}),
$$

the normalized assemblage

$$
\tilde{\rho}^{(\mu)}_{A|X} = \frac{\mu\mathcal{S}_{\mathrm{rob}}(\rho_{A|X})}{\nu^{(\mu)}}\tilde{\rho}_{A|X} + (1-\mu)\mathcal{S}_{\mathrm{rob}}(\rho'_{A|X})\tilde{\rho}'_{A|X},
$$

and the normalized unsteerable assemblage

$$\sigma_{A|X}^{(\mu)} = \frac{\mu(1 + \mathcal{S}_{\rm rob}(\rho_{A|X}))}{1 + \nu^{(\mu)}} \sigma_{A|X}$$
$$+ (1 - \mu)(1 + \mathcal{S}_{\rm rob}(\rho'_{A|X}))\sigma'_{A|X},$$

such that

$$\mu\rho_{A|X} + (1 - \mu)\rho'_{A|X} = (1 + \nu^{(\mu)})\sigma_{A|X}^{(\mu)} - \nu^{(\mu)}\tilde{\rho}_{A|X}^{(\mu)},$$

and proceed with $\mathcal{S}_{\rm rob}$ analogously as with $\mathcal{S}_{\rm W}$ in Eq. (D10) above. ∎

APPENDIX E: PROOF OF THEOREM 4

Proof of Theorem 4.— One of the implications is trivial to prove. If $P'_{A|X} = P_{A|X}$ and Eq. (24) holds, there exists a unitary operator U such that $|\psi'(a,x)\rangle = U|\psi(a,x)\rangle$ for all a and x. Then, $\Psi_{A|X}$ can be transformed into $\Psi'_{A|X}$ by means of a deterministic 1W-LOCC: namely, the one consisting of Bob applying U to his subsystem and Alice doing nothing. Likewise, if $\Psi'_{A|X} \in$ LHS, then $\Psi_{A|X}$ can trivially be transformed into $\Psi'_{A|X}$ by 1W-LOCCs, as any unsteerable assemblage can be created by stochastic 1W-LOCCs by definition [see discussion after Eq. (6)].

Let us then prove the converse implication. In other words, assuming that $\Psi_{A|X}$ and $\Psi'_{A|X}$ are pure orthogonal assemblages and that the latter can be obtained from the former by a stochastic 1W-LOCC, we prove that either $\Psi'_{A|X} \in$ LHS or $P'_{A|X} = P_{A|X}$ and Eq. (24) is true. To this end, we first note that the no-signaling condition (4) restricts minimal-dimension pure-orthogonal assemblages to a rather specific form. Namely, the fact that $\Psi_{A|X}$ is no-signaling implies that

   (i) either $P_{A|X}(\cdot, x)$ is a deterministic distribution for all x,

   (ii) or $P_{A|X}(\cdot, x)$ is the uniform distribution for all x.

If case (i) holds, $\Psi_{A|X} \in$ LHS. Then, since, by assumption, $\Psi'_{A|X}$ can be obtained via a stochastic 1W-LOCC from $\Psi_{A|X}$, one automatically obtains that $\Psi'_{A|X} \in$ LHS.

To analyze case (ii), we use the fact that $\Psi'_{A|X}$ is also subject to the no-signaling condition (4):

   (i′) either $P'_{A|X}(\cdot, x)$ is a deterministic distribution for all x,

   (ii′) or $P'_{A|X}(\cdot, x)$ is the uniform distribution for all x.

The fact that case (i′) is possible if case (ii) holds is clear, as (i′) corresponds to $\Psi'_{A|X} \in$ LHS. So, it only remains to show that if cases (ii) and (ii′) hold, then either $\Psi'_{A|X} \in$ LHS or Eq. (24) holds. We show this result in what follows.

Assuming that (ii) and (ii′) hold and that there is a stochastic 1W-LOCC $\mathcal{M}$ that maps $\Psi_{A|X}$ into $\Psi'_{A|X}$, i.e., such that $\mathcal{M}(\Psi_{A|X}) \propto \Psi'_{A|X}$, where "$\propto$" stands for "is proportional to," we use Eq. (A6) to obtain

$$\sum_{a,x,\omega} P_{X|X_f,\Omega}(x, x_f, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) K_\omega |\psi(a,x)\rangle\langle\psi(a,x)| K_\omega^\dagger \propto |\psi'(a_f, x_f)\rangle\langle\psi'(a_f, x_f)| \quad \forall (a_f, x_f). \quad \text{(E1)}$$

Since the right-hand side of Eq. (E1) is composed of a rank-one projector onto a pure state, each term of the sum in the left-hand side must be either zero or proportional to $|\psi'(a_f, x_f)\rangle\langle\psi'(a_f, x_f)|$. In particular, this must also hold for each ωth term. In other word, for all ω, it must hold that

$$\sum_{a,x} P_{X|X_f,\Omega}(x, x_f, \omega) P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f) K_\omega |\psi(a,x)\rangle\langle\psi(a,x)| K_\omega^\dagger \sim |\psi'(a_f, x_f)\rangle\langle\psi'(a_f, x_f)| \quad \forall (a_f, x_f), \quad \text{(E2)}$$

where the symbol "$\sim$" is used to signify "is either equal to zero or proportional to." Indeed, using the facts that $K_\omega \neq 0$ and that $P_{X|X_f,\Omega}$ and $P_{A_f|A,X,\Omega,X_f}$ are normalized distributions, one can see by case analysis that there are always at least two different pairs (a_f, x_f) for which the left-hand side of Eq. (E2) is not zero and, therefore, proportional to $|\psi'(a_f, x_f)\rangle\langle\psi'(a_f, x_f)|$.

Let us first consider the case

$$K_\omega |\psi(a,x)\rangle \neq 0 \quad \forall (a,x). \quad \text{(E3)}$$

The other case will be considered at the end. The first step is to note that Eqs. (E2) and (E3) imply that, unless $\Psi'_{A|X} \in$ LHS,

$$P_{X|X_f,\omega}(x|x_f, \omega) = \delta_{x_f x \oplus f(\omega)}, \quad \text{(E4a)}$$

$$P_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x \oplus f(\omega)) \in \{0, 1\} \quad \forall (a_f, a, x), \quad \text{(E4b)}$$

where $f(\omega) \in \{0, 1\}$. In other words, for any ω for which Eq. (E3) holds, unless $\Psi'_{A|X} \in$ LHS, the variables X and X_f must be either fully correlated or fully anticorrelated and $P_{A_f|A,X,\Omega,X_f}(\cdot, a_f, x, \omega, x \oplus f(\omega))$ must be a deterministic distribution for all (a_f, x).

To prove Eq. (E4a), suppose that it does not hold. Then, there must exist $\tilde{x}$ such that $P_{X|X_f,\omega}(\tilde{x}|x_f, \omega) \neq 0$ for all x_f. Because of Eq. (E2), this result implies that

$$\sum_a P_{A_f|A,X,\Omega,X_f}(a_f,a,\tilde{x},\omega,0)K_\omega|\psi(a,\tilde{x})\rangle\langle\psi(a,\tilde{x})|K_\omega^\dagger$$

$$\sim |\psi'(a_f,0)\rangle\langle\psi'(a_f,0)|, \tag{E5a}$$

$$\sum_a P_{A_f|A,X,\Omega,X_f}(a_f,a,\tilde{x},\omega,1)K_\omega|\psi(a,\tilde{x})\rangle\langle\psi(a,\tilde{x})|K_\omega^\dagger$$

$$\sim |\psi'(a_f,1)\rangle\langle\psi'(a_f,1)|. \tag{E5b}$$

In turn, choosing $\tilde{a}_f$ and $\bar{a}_f$ such that $P_{A_f|A,X,\Omega,X_f}\times(\tilde{a}_f,a,\tilde{x},\omega,0) > 0$ and $P_{A_f|A,X,\Omega,X_f}(\bar{a}_f,a,\tilde{x},\omega,1) > 0$, which is always possible because $P_{A_f|A,X,\Omega,X_f}$ is a normalized distribution and does not require any extra assumption, Eqs. (E3) and (E5) imply that

$$K_\omega|\psi(a,\tilde{x})\rangle \propto |\psi'(\tilde{a}_f,0)\rangle, \tag{E6a}$$

$$K_\omega|\psi(a,\tilde{x})\rangle \propto |\psi'(\bar{a}_f,1)\rangle. \tag{E6b}$$

This result finally leads to $|\psi'(\tilde{a}_f,0)\rangle = |\psi'(\bar{a}_f,1)\rangle$, which is true only if $\Psi'_{A|X} \in$ LHS.

To prove Eq. (E4b), we use a similar argument. If one assumes that Eq. (E4b) is false, then there must exist a pair $(\tilde{a},\tilde{x})$ such that $P_{A_f|A,X,\Omega,X_f}(a_f,\tilde{a},\tilde{x},\omega,\tilde{x}\oplus f(\omega)) > 0$ for all a_f. Using this fact and Eqs. (E2), (E3), and (E4a), one arrives at

$$K_\omega|\psi(\tilde{a},\tilde{x})\rangle \propto |\psi'(0,\tilde{x}\oplus f(\omega))\rangle, \tag{E7a}$$

$$K_\omega|\psi(\tilde{a},\tilde{x})\rangle \propto |\psi'(1,\tilde{x}\oplus f(\omega))\rangle, \tag{E7b}$$

which, since $|\psi'(0,\tilde{x}\oplus f(\omega))\rangle$ and $|\psi'(1,\tilde{x}\oplus f(\omega))\rangle$ are orthogonal, yields a contradiction.

The second step is to note that Eqs. (E3) and (E4) impose restrictions on which a's and x's can contribute to each a_f and x_f in Eq. (E2). More precisely, one can see by case analysis that, up to relabelings of a_f or x_f, only three different types of assignments are possible:

(a)	(b)	(c)
$K_\omega\|\psi(0,0)\rangle \propto \|\psi'(0,0)\rangle$	$K_\omega\|\psi(0,0)\rangle \propto \|\psi'(0,0)\rangle$	$K_\omega\|\psi(0,0)\rangle \propto \|\psi'(0,0)\rangle$
$K_\omega\|\psi(1,0)\rangle \propto \|\psi'(1,0)\rangle$	$K_\omega\|\psi(1,0)\rangle \propto \|\psi'(1,0)\rangle$	$K_\omega\|\psi(1,0)\rangle \propto \|\psi'(0,0)\rangle$
$K_\omega\|\psi(0,1)\rangle \propto \|\psi'(0,1)\rangle$	$K_\omega\|\psi(0,1)\rangle \propto \|\psi'(0,1)\rangle$	$K_\omega\|\psi(0,1)\rangle \propto \|\psi'(0,1)\rangle$
$K_\omega\|\psi(1,1)\rangle \propto \|\psi'(1,1)\rangle$	$K_\omega\|\psi(1,1)\rangle \propto \|\psi'(0,1)\rangle$	$K_\omega\|\psi(1,1)\rangle \propto \|\psi'(0,1)\rangle$

The third step is to show that all three cases, (a)–(c), are possible only if either $\Psi'_{A|X} \in$ LHS or Eq. (24) holds. Note that it is enough to show this for the case where all eight vectors $\{|\psi(a,x)\rangle, |\psi'(a_f,x_f)\rangle\}_{a,x,a_f,x_f}$ lie on the same plane of the Bloch sphere. This is due to the fact that, since $\Psi_{A|X}$ and $\Psi'_{A|X}$ are both pure no-signaling assemblages of minimal dimension, $\{|\psi(a,x)\rangle\}_{a_f,x_f}$ and $\{|\psi'(a_f,x_f)\rangle\}_{a_f,x_f}$ are both already contained in two planes of the Bloch sphere, as one can straightforwardly see using Eq. (4). These two planes can always be rotated so as to coincide by a unitary operation, which can in turn be absorbed in the definition of the Kraus operator K_ω. Hence, without loss of generality, we take

$$|\psi(0,0)\rangle = |0\rangle,$$

$$|\psi(1,0)\rangle = |1\rangle,$$

$$|\psi(0,1)\rangle = \cos(\varphi)|0\rangle + \sin(\varphi)|1\rangle,$$

$$|\psi(1,1)\rangle = -\sin(\varphi)|0\rangle + \cos(\varphi)|1\rangle,$$

$$|\psi'(0,0)\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle,$$

$$|\psi'(1,0)\rangle = -\sin(\theta)|0\rangle + \cos(\theta)|1\rangle,$$

$$|\psi'(1,1)\rangle = \cos(\phi)|0\rangle + \sin(\phi)|1\rangle,$$

$$|\psi'(1,1)\rangle = -\sin(\phi)|0\rangle + \cos(\phi)|1\rangle, \tag{E8}$$

for arbitrary φ, θ, and $\phi \in [0,\pi/2]$, where $|0\rangle$ and $|1\rangle$ represent the computational-basis states. First, we analyze case (a). Dividing both vector components (in the computational basis) of the first equation of this case, one obtains $[K_\omega]_{00}/[K_\omega]_{10} = \cos(\theta)/\sin(\theta)$, where $[K_\omega]_{ij} := \langle i|K_\omega|j\rangle$. Analogously, dividing both vector components of the second equation yields $[K_\omega]_{01}/[K_\omega]_{11} = -\sin(\theta)/\cos(\theta)$. Hence, introducing proportionality constants $\kappa_1 > 0$ and $\kappa_2 > 0$, the Kraus operator can be matrix represented in the computational basis as

$$K_\omega = \begin{pmatrix} \kappa_1\cos(\theta) & -\kappa_2\sin(\theta) \\ \kappa_1\sin(\theta) & \kappa_2\cos(\theta) \end{pmatrix}. \tag{E9}$$

Using Eq. (E9), the third equation of case (a) implies that

$$\kappa_1\cos(\varphi)\sin(\theta-\phi) = \kappa_2\sin(\varphi)\cos(\theta-\phi). \tag{E10}$$

Finally, the fourth equation leads to

$$\kappa_2\cos(\varphi)\sin(\theta-\phi) = \kappa_1\sin(\varphi)\cos(\theta-\phi). \tag{E11}$$

Equations (E10) and (E11) can be simultaneously satisfied only if $\theta - \phi = \varphi$ or $(\theta-\phi)\times\varphi = 0$. The former option yields Eq. (24). The latter one implies that $\Psi'_{A|X} \in$ LHS. In a similar fashion, for case (b), the first three equations lead to Eq. (E10) and the fourth one to

$$-\kappa_1 \sin(\varphi) \sin(\theta - \phi) = \kappa_2 \cos(\phi) \cos(\theta - \phi). \quad \text{(E12)}$$

This equation cannot be satisfied unless $(\theta - \phi) \times \varphi = 0$, which means that $\Psi'_{A|X} \in$ LHS. With a similar argument, the reader can straightforwardly verify that the same thing happens for case (c). This finishes the proof of the theorem for the ω's for which Eq. (E3) holds.

As the fourth and final step, it remains to treat the case where, for a certain ω, there exists a pair $(\tilde{a}, \tilde{x})$ for which $K_\omega |\psi(\tilde{a}, \tilde{x})\rangle = 0$. Since $K_\omega \neq 0$, the latter is true only if the support of K_ω is given by the span of $|\psi(\tilde{a} \oplus 1, \tilde{x})\rangle$. Using this and the fact that there are always at least two different pairs (a_f, x_f) for which the left-hand side of Eq. (E2) is not zero, one obtains $K_\omega \propto |\psi'(a_f, x_f)\rangle\langle\psi(\tilde{a} \oplus 1, \tilde{x})|$ for two different pairs (a_f, x_f), which, unless $\Psi'_{A|X} \in$ LHS, is a contradiction. ∎

APPENDIX F: NONEXISTENCE OF MINIMAL-DIMENSION STEERING BITS

In this appendix, we prove Theorem 5. This section bears many similarities to Appendix E.

Proof of Theorem 5.— We proceed by *reductio ad absurdum*. In other words, we show that if one supposes that there exists a pure normalized assemblage $\Psi_{A|X} := \{P_{A|X}(a, x), |\psi(a, x)\rangle\}_{a,x}$, with $d = s = r = 2$, from which all assemblages can be obtained via stochastic 1W-LOCCs, one obtains a contradiction.

Without loss of generality, we can choose the computational basis $\{|0\rangle, |1\rangle\}$ so that its first element coincides with $|\psi(0, 0)\rangle$ and the element $|\psi(1, 0)\rangle$ is in the plane that contains the vectors $|0\rangle$ and $(1/\sqrt{2})(|0\rangle + |1\rangle)$. What is more, clearly, $\Psi_{A|X}$ cannot have a LHS model; otherwise, $\Psi_{A|X}$ could not be mapped into all assemblages by stochastic 1W-LOCCs. Thus, we can safely assume that

$$\Psi_{A|X} \notin \text{LHS.} \quad \text{(F1)}$$

Hence, we take

$$|\psi(0, 0)\rangle = |0\rangle,$$
$$|\psi(1, 0)\rangle = \cos(\varphi_{10})|0\rangle + \sin(\varphi_{10})|1\rangle,$$
$$|\psi(0, 1)\rangle = \cos(\varphi_{01})|0\rangle + e^{i\alpha_{01}} \sin(\varphi_{01})|1\rangle,$$
$$|\psi(1, 1)\rangle = \cos(\varphi_{11})|0\rangle + e^{i\alpha_{11}} \sin(\varphi_{11})|1\rangle, \quad \text{(F2)}$$

with

$$\varphi_{10} \in]0, \pi[, \quad \text{(F3a)}$$

$$\alpha_{ax} \in [0, 2\pi], \quad \forall (a, x) \notin \{(0, 0), (1, 0)\}, \quad \text{(F3b)}$$

and

$$(\varphi_{a1}, \alpha_{a,1}) \neq (\varphi_{a'1}, \alpha_{a'1}) \quad \forall a \neq a'. \quad \text{(F4)}$$

Equations (F3) and (F4) hold due to the fact that $\Psi_{A|X} \notin$ LHS and because of the no-signaling condition (4). More precisely, if $\varphi_{10} = \{0, \pi\}$, $|\psi(1, 0)\rangle = |0\rangle$, which implies that Bob's reduced state is $\varrho_B = |0\rangle\langle0|$. Then, the no-signalling condition (4) implies that $|\psi(0, 1)\rangle = |0\rangle = |\psi(1, 1)\rangle$. Such assemblage clearly has a LHS model, which contradicts the assumption (F1). The same argument implies Eq. (F4). Furthermore, $\Psi_{A|X} \notin$ LHS and the no-signaling principle also imply that $P_{A|X}(a, x) \neq 0$ for all (a, x). To see the latter, suppose that there is a pair (a, x) for which $P_{A|X}(a, x) = 0$. Then, clearly, $P_{A|X}(a \oplus 1, x) = 1$. This result, together with Eq. (4), implies that there is an $\tilde{a}$ for which $P_{A|X}(\tilde{a}, x \oplus 1) = 1$, which in turn leads to $\Psi_{A|X} \in$ LHS.

Let us now consider pure orthogonal assemblages $\{\Psi^\theta_{A|X}\}_\theta$ with $d = s = r = 2$ of the form $\Psi^\theta_{A|X} := \{\frac{1}{2}, |\psi^\theta(a, x)\rangle\}_{a,x}$, where

$$|\psi^\theta(0, 0)\rangle = |0\rangle,$$
$$|\psi^\theta(1, 0)\rangle = |1\rangle, \quad \text{(F5a)}$$

and

$$|\psi^\theta(0, 1)\rangle = \cos(\theta)|0\rangle + \sin(\theta)|1\rangle,$$
$$|\psi^\theta(1, 1)\rangle = -\sin(\theta)|0\rangle + \cos(\theta)|1\rangle. \quad \text{(F5b)}$$

We restrict ourselves to $0 < \theta < \pi/2$ to ensure that $\Psi^\theta_{A|X} \notin$ LHS. If all assemblages can be obtained via stochastic 1W-LOCCs from $\Psi_{A|X}$, there must be a stochastic 1W-LOCC $\mathcal{M}^\theta$ such that $\mathcal{M}^\theta(\Psi_{A|X}) \propto \Psi^\theta_{A|X}$, where "$\propto$" stands for "is proportional to." Then, as in Appendix E, Eq. (A6) implies that, for all ω, the following must hold:

$$\sum_{a,x} P^\theta_{X|X_f,\Omega}(x, x_f, \omega) P^\theta_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x_f)$$
$$\times P_{A|X}(a, x) K^\theta_\omega |\psi(a, x)\rangle\langle\psi(a, x)| K^{\theta\dagger}_\omega$$
$$\sim |\psi^\theta(a_f, x_f)\rangle\langle\psi^\theta(a_f, x_f)| \quad \forall (a_f, x_f), \quad \text{(F6)}$$

where the symbol "$\sim$" is used to signify "is either equal to zero or proportional to." However, we note again that, since $K_\omega \neq 0$ and $P_{X|X_f,\Omega}$ and $P_{A_f|A,X,\Omega,X_f}$ are normalized distributions, there are always at least two different pairs (a_f, x_f) for which the left-hand side of Eq. (F6) is not zero and, therefore, proportional to $|\psi^\theta(a_f, x_f)\rangle\langle\psi^\theta(a_f, x_f)|$, as can be seen by direct case analysis.

Let us then consider the case

$$K^\theta_\omega |\psi(a, x)\rangle \neq 0 \quad \forall (a, x). \quad \text{(F7)}$$

The other case will be considered later. The first step is to note that Eqs. (F6) and (F7), together with the fact that $\Psi_{A|X} \notin$ LHS, imply that

$$P^\theta_{X|X_f,\omega}(x|x_f,\omega) = \delta_{x_f x \oplus f^\theta(\omega)}, \qquad \text{(F8a)}$$

$$P^\theta_{A_f|A,X,\Omega,X_f}(a_f, a, x, \omega, x \oplus f^\theta(\omega)) \in \{0,1\} \quad \forall \ (a_f, a, x),$$
$$\text{(F8b)}$$

where $f^\theta(\omega) \in \{0,1\}$. In other words, for any ω for which Eq. (F7) holds, X and X_f must be either fully correlated or fully anticorrelated, and $P^\theta_{A_f|A,X,\Omega,X_f}(\cdot, a_f, x, \omega, x \oplus f(\omega))$ must be a deterministic distribution for all (a_f, x). The proofs of Eqs. (F8) are almost identical to the proofs of Eqs. (E4) in Appendix E, with the only difference that, here, $\Psi_{A|X} \notin \mathrm{LHS}$ and $\Psi^\theta_{A|X} \notin \mathrm{LHS}$ are true by assumption. We therefore do not repeat the argument.

The second step is to note that Eqs. (F7) and (F8) impose restrictions on which a's and x's can contribute to each a_f and x_f in Eq. (F6). More precisely, one can see by case analyses that, up to relabelings of a_f or x_f, only one type of assignment is possible:

$$K^\theta_\omega|\psi(0,0)\rangle \propto |\psi^\theta(0,0)\rangle, \qquad \text{(F9a)}$$

$$K^\theta_\omega|\psi(1,0)\rangle \propto |\psi^\theta(1,0)\rangle, \qquad \text{(F9b)}$$

$$K^\theta_\omega|\psi(0,1)\rangle \propto |\psi^\theta(0,1)\rangle, \qquad \text{(F9c)}$$

$$K^\theta_\omega|\psi(1,1)\rangle \propto |\psi^\theta(1,1)\rangle. \qquad \text{(F9d)}$$

The third step is to show that Eqs. (F9) lead to a contradiction. To this end, together with Eqs. (F2) and (F5a), Eqs. (F9a) and (F9b), respectively, imply that $[K^\theta_\omega]_{10} = 0$ and $[K^\theta_\omega]_{00}/[K^\theta_\omega]_{01} = -\tan(\varphi_{10})$, where $[K^\theta_\omega]_{ij} := \langle i|K^\theta_\omega|j\rangle$. In turn, dividing both vector components in both Eqs. (F9c) and (F9d), one obtains, using Eqs. (F2) and (F5b), that

$$\frac{[K^\theta_\omega]_{11}}{[K^\theta_\omega]_{01}} = \tan(\theta)\left(-\frac{\tan(\varphi_{10})}{\tan(\varphi_{01})e^{i\alpha_{01}}} + 1\right), \qquad \text{(F10a)}$$

$$\frac{[K^\theta_\omega]_{11}}{[K^\theta_\omega]_{01}} = \frac{-1}{\tan(\theta)}\left(-\frac{\tan(\varphi_{10})}{\tan(\varphi_{11})e^{i\alpha_{11}}} + 1\right). \qquad \text{(F10b)}$$

Equating the right-hand sides of Eqs. (F10a) and (F10b) gives, after straightforward algebraic manipulation,

$$\frac{1}{\tan(\varphi_{10})} = \frac{\sin^2(\theta)}{\tan(\varphi_{11})e^{i\alpha_{11}}} + \frac{\cos^2\theta}{\tan(\varphi_{01})e^{i\alpha_{01}}}. \qquad \text{(F11)}$$

Since the last condition is independent of K^θ_ω and since $\Psi_{A|X}$ should be transformed by stochastic 1W-LOCCs into any member of the family $\{\Psi^\theta_{A|X}\}_\theta$, the same condition should be fulfilled for any $0 < \theta < \pi/2$. It actually suffices to choose just two assemblages $\Psi^{\theta_1}_{A|X}$ and $\Psi^{\theta_2}_{A|X}$, for any

$0 < \theta_1, \theta_2 < \pi/2$ with $\theta_1 \neq \theta_2$, to arrive at a contradiction. Indeed, since the angles φ_{10}, φ_{01}, φ_{11}, α_{01}, and α_{11} are fixed, the only way to satisfy Eq. (F11) for both θ_1 and θ_2 is if

$$\tan(\varphi_{10}) = \tan(\varphi_{01})e^{i\alpha_{01}} = \tan(\varphi_{11})e^{i\alpha_{11}}. \qquad \text{(F12)}$$

This, in turn, can happen only if $\alpha_{01} = 0 = \alpha_{11}$ and $\varphi_{10} = \varphi_{01} = \varphi_{11}$, which is clearly incompatible with Eq. (F4).

It remains to treat the case where, for a certain ω, there exists a pair $(\tilde{a}, \tilde{x})$ for which Eq. (F7) does not hold. By relabeling a or x, we can always choose $(\tilde{a}, \tilde{x}) = (0,0)$. Hence, we consider

$$K^\theta_\omega|\psi(0,0)\rangle = K^\theta_\omega|0\rangle = 0. \qquad \text{(F13)}$$

Since $K^\theta_\omega \neq 0$, the latter is true only if the support of K^θ_ω is given by the span of $|1\rangle$. Using this and the fact that there are always at least two different pairs (a_f, x_f) for which the left-hand side of Eq. (F6) is not zero, one arrives at a contradiction of the type $K^\theta_\omega \propto |\psi^\theta(a_f, x_f)\rangle\langle 1|$ for two different pairs (a_f, x_f). This finishes the proof for pure assemblages. ∎

We finish the appendix with a remark on the difficulty to generalize Theorem 5 to the case of mixed-state assemblages, i.e., to also rule out the existence of steering bits among mixed-state assemblages. Since any mixed-state assemblage can be decomposed as a convex combination of pure assemblages and since $\mathcal{M}$ is a linear transformation, one would be tempted to trivially extend the proof above to mixed-state assemblages by using similar reasonings to those presented just above, with each pure assemblage in the convex combination together with linearity arguments. However, such straightforward extension unfortunately fails. The reason for this is that each pure assemblage in the pure-assemblage decomposition of a mixed-state assemblage is, as far as we can see, not necessarily a no-signalling type. We emphasize that our formalism deals only with no-signalling objects. Hence, while we strongly believe that minimal-dimension steering bits do not exist, in general, i.e., even among the mixed-state assemblages, we leave the proof of this statement as an open question.

APPENDIX G: PROOFS OF LEMMAS 1 AND 2

Before we proceed, we recall some known mathematical facts necessary for the proofs.

First, the von Neumman relative entropy S_Q, defined by Eq. (15), fulfills the following properties [40].

(i) Given two sets $\{\varrho^{(j)}\}_{j=1,\ldots,n}$ and $\{\varrho'^{(j)}\}_{j=1,\ldots,n}$ of n arbitrary positive-semidefinite (not necessarily normalized) operators each and n positive real numbers $\{\mu^{(j)}\}_{j=1,\ldots,n}$ such that $\sum_j \mu^{(j)} = 1$, with $n \in \mathbb{N}$, S_Q satisfies the joint convexity property

$$S_Q\left(\sum_i \mu^{(j)}\varrho^{(j)} \middle\| \sum_j \mu^{(j)}\varrho'^{(j)}\right) \leq \sum_j \mu^{(j)} S_Q(\varrho^{(j)}\|\varrho'^{(j)}).$$

$$\text{(G1)}$$

(ii) Given any completely positive trace-preserving (CPTP) map $\mathcal{E}$ and any two density operators ϱ and ϱ, S_Q satisfies the CPTP-map contraction property

$$S_Q(\mathcal{E}(\varrho)|\mathcal{E}(\varrho')) \leq S_Q(\varrho|\varrho'). \qquad \text{(G2)}$$

Second, the Kullback-Leibler divergence S_C defined in Eq. (16) fulfills the following property.

(i) Given any two joint probability distributions $P_{X,Y}$ and $P_{X,Y}'$ over classical bits x and y, S_C satisfies the inequality

$$\sum_x P_X(x) S_C(P_{Y|X}(\cdot, x)\|P_{Y|X}'(\cdot, x))$$

$$= \sum_x P_X(x) \sum_y P_{Y|X}(x, y) \frac{\log P_{Y|X}(x, y)}{\log P_{Y|X}'(x, y)}$$

$$= \sum_{x,y} P_{X,Y}(x, y)\left(\frac{\log P_{X,Y}(x, y)}{\log P_{X,Y}'(x, y)} - \frac{\log P_X(x)}{\log P_X'(x)}\right)$$

$$= S_C(P_{X,Y}\|P_{X,Y}') - S_C(P_X\|P_X')$$

$$\leq S_C(P_{X,Y}\|P_{X,Y}'). \qquad \text{(G3)}$$

We are now in a good position to prove the lemmas.

1. Proof of Lemma 1

We begin with Lemma 1.

Proof of Lemma 1.— First, using the definition of $\mathcal{S}_R$ in Eq. (18), we write the left-hand side of Eq. (C1) explicitly as

$$\sum_\omega P_\Omega(\omega) S_A\left(\frac{\mathcal{M}_\omega(\hat\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \frac{\mathcal{M}_\omega(\hat\rho'_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]}\right)$$

$$= \sum_\omega P_\Omega(\omega) \max_{P_{X_f|\Gamma},\{} E_\gamma\} \Bigg[S_C(P_{\Gamma|\Omega}(\cdot, \omega)\|P_{\Gamma|\Omega}'(\cdot, \omega))$$

$$+ \sum_{\gamma, x_f} P_{X_f|\Gamma}(x_f, \gamma) P_{\Gamma|\Omega}(\gamma, \omega) S_Q\left(\frac{\mathbb{1}\otimes E_\gamma[\mathcal{M}_\omega(\hat\rho_{A|X})](x_f)\mathbb{1}\otimes E_\gamma^\dagger}{P_{\Gamma,\Omega}(\gamma, \omega)} \middle\| \frac{\mathbb{1}\otimes E_\gamma[\mathcal{M}_\omega(\hat\rho'_{A|X})](x_f)\mathbb{1}\otimes E_\gamma^\dagger}{P_{\Gamma,\Omega}'(\gamma, \omega)}\right)\Bigg], \quad \text{(G4)}$$

where we have used $P_\Omega(\omega) = \text{Tr}[\mathcal{M}_\omega(\hat\rho_{A|X})]$ and $P_\Omega'(\omega) = \text{Tr}[\mathcal{M}_\omega(\hat\rho'_{A|X})]$, and $P_{\Gamma,\Omega}(\gamma, \omega) = P_{\Gamma|\Omega}(\gamma, \omega) P_\Omega(\omega)$ and $P_{\Gamma,\Omega}'(\gamma, \omega) = P_{\Gamma|\Omega}'(\gamma, \omega) P_\Omega'(\omega)$, with

$$P_{\Gamma|\Omega}(\gamma, \omega) := \text{Tr}\left[\mathbb{1}\otimes E_\gamma \frac{[\mathcal{M}_\omega(\hat\rho_{A|X})](x_f)}{P_\Omega(\omega)} \mathbb{1}\otimes E_\gamma^\dagger\right] = \text{Tr}_B\left[\frac{E_\gamma \mathcal{E}_\omega(\rho_B) E_\gamma^\dagger}{P_\Omega(\omega)}\right] \qquad \text{(G5a)}$$

and

$$P_{\Gamma|\Omega}'(\gamma, \omega) := \text{Tr}\left[\mathbb{1}\otimes E_\gamma \frac{[\mathcal{M}_\omega(\hat\rho'_{A|X})](x_f)}{P_\Omega'(\omega)} \mathbb{1}\otimes E_\gamma^\dagger\right] = \text{Tr}_B\left[\frac{E_\gamma \mathcal{E}_\omega(\rho_B') E_\gamma^\dagger}{P_\Omega'(\omega)}\right], \qquad \text{(G5b)}$$

both of which are independent of x_f and a_f. Now, since X_f and Ω are independent variables, we can replace $P_{X_f|\Gamma}$ with $P_{X_f|\Gamma,\Omega}$ and exchange the order of the maximization over $P_{X_f|\Gamma,\Omega}$ and the summation over ω in Eq. (G4). Furthermore, the optimal measurement operators for which the maximization over $\{E_\gamma\}$ is attained for each ω depend, of course, on ω. Hence, we can also exchange the order of the summation over ω and the maximization over the measurement operators if we make this dependence explicit by replacing, in Eqs. (G4) and (G5), $\{E_\gamma\}$ with $\{E_{\gamma,\omega}\}$. With this, we write Eq. (G4) as

$$\sum_\omega P_\Omega(\omega) S_A\left(\frac{\mathcal{M}_\omega(\hat\rho_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \frac{\mathcal{M}_\omega(\hat\rho'_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]}\right)$$

$$= \max_{P_{X_f|\Gamma,\Omega},\{E_{\gamma,\omega}\}} \Bigg\{\sum_\omega P_\Omega(\omega) \Bigg[S_C(P_{\Gamma|\Omega}(\cdot, \omega)\|P_{\Gamma|\Omega}'(\cdot, \omega))$$

$$+ \sum_{\gamma, x_f} P_{X_f,\Gamma|\Omega}(x_f, \gamma, \omega) S_Q\left(\frac{\mathbb{1}\otimes E_{\gamma,\omega}[\mathcal{M}_\omega(\hat\rho_{A|X})](x_f)\mathbb{1}\otimes E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma, \omega)} \middle\| \frac{\mathbb{1}\otimes E_{\gamma,\omega}[\mathcal{M}_\omega(\hat\rho'_{A|X})](x_f)\mathbb{1}\otimes E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}'(\gamma, \omega)}\right)\Bigg]\Bigg\}. \quad \text{(G6)}$$

Next, using Eqs. (3), (10), and (11), we write

$$
S_Q\left(\frac{\mathbb{1}\otimes E_{\gamma,\omega}[\mathcal{M}_\omega(\hat\rho_{A|X})](x_f)\mathbb{1}\otimes E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma,\omega)}\middle\|\frac{\mathbb{1}\otimes E_{\gamma,\omega}[\mathcal{M}_\omega(\hat\rho'_{A|X})](x_f)\mathbb{1}\otimes E_{\gamma,\omega}^\dagger}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)
$$

$$
= S_Q\left(\frac{\sum_{a_f,a,x}P_{X|X_f,\Omega}(x,x_f,\omega)P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma,\omega)}\right.
$$
$$
\left.\middle\|\frac{\sum_{a_f,a,x}P_{X|X_f,\Omega}(x,x_f,\omega)P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho'_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)
$$

$$
\le \sum_x P_{X|X_f,\Omega}(x,x_f,\omega)S_Q\left(\frac{\sum_{a_f,a}P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma,\omega)}\right.
$$
$$
\left.\middle\|\frac{\sum_{a_f,a}P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho'_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right), \tag{G7}
$$

where the inequality is due to Eq. (G1). On the other hand, we note that there always exists a completely positive trace-preserving map $\mathcal{R}_{x,\omega,x_f}\colon \mathcal{L}(\mathcal{H}_E)\to\mathcal{L}(\mathcal{H}_{Ef})$ such that

$$
\mathcal{R}_{x,\omega,x_f}(|a\rangle\langle a|)=\sum_{a_f}|a_f\rangle\langle a_f|P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f). \tag{G8}
$$

Hence, we can apply Eq. (G2) to the von Neumann relative entropy in the right-hand side of Eq. (G7) with the map (G8), to get

$$
S_Q\left(\frac{\sum_{a_f,a}P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma,\omega)}\right.
$$
$$
\left.\middle\|\frac{\sum_{a_f,a}P_{A_f|A,X,\Omega,X_f}(a_f,a,x,\omega,x_f)|a_f\rangle\langle a_f|\otimes E_{\gamma,\omega}K_\omega\varrho'_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)
$$

$$
\le S_Q\left(\frac{\sum_a|a\rangle\langle a|\otimes E_{\gamma,\omega}K_\omega\varrho_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P_{\Gamma,\Omega}(\gamma,\omega)}\middle\|\frac{\sum_a|a\rangle\langle a|\otimes E_{\gamma,\omega}K_\omega\varrho'_{A|X}(a,x)K_\omega^\dagger E_{\gamma,\omega}^\dagger}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)
$$
$$
= S_Q\left(\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P_{\Gamma,\Omega}(\gamma,\omega)}\middle\|\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho'_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right), \tag{G9}
$$

where the quantum representation (3) has been invoked again. Then, using Eqs. (G6), (G7), and (G9), we obtain

$$
\sum_\omega P_\Omega(\omega)S_A\left(\frac{\mathcal{M}_\omega(\hat\rho_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho_{A|X})]}\middle\|\frac{\mathcal{M}_\omega(\hat\rho'_{A|X})}{\mathrm{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]}\right)\le\max_{P_{X_f|\Gamma,\Omega},\{E_{\gamma,\omega}\}}\left\{\sum_\omega P_\Omega(\omega)\left[S_C(P_{\Gamma|\Omega}(\cdot,\omega)\|P'_{\Gamma|\Omega}(\cdot,\omega))\right.\right.
$$
$$
+\sum_{\gamma,x_f,x}P_{X_f,\Gamma|\Omega}(x_f,\gamma,\omega)P_{X|X_f,\Omega}(x,x_f,\omega)
$$
$$
\left.\left.\times S_Q\left(\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P_{\Gamma,\Omega}(\gamma,\omega)}\middle\|\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho'_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)\right]\right\}
$$

$$
\le\max_{P_{X_f|\Gamma,\Omega},\{E_{\gamma,\omega}\}}\left\{\left[S_C(P_{\Gamma,\Omega}\|P'_{\Gamma,\Omega})\right.\right.
$$
$$
+\sum_{\gamma,x_f,x,\omega}P_\Omega(\omega)P_{X_f,\Gamma|\Omega}(x_f,\gamma,\omega)P_{X|X_f,\Omega,\Gamma}(x,x_f,\omega,\gamma)
$$
$$
\left.\left.\times S_Q\left(\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P_{\Gamma,\Omega}(\gamma,\omega)}\middle\|\frac{\mathbb{1}\otimes E_{\gamma,\omega}K_\omega\hat\rho'_{A|X}(x)K_\omega^\dagger E_{\gamma,\omega}^\dagger\otimes\mathbb{1}}{P'_{\Gamma,\Omega}(\gamma,\omega)}\right)\right]\right\}, \tag{G10}
$$

where the inequality [(G10)] follows from Eq. [(G3)] and from replacing $P_{X|X_f,\Omega}(x, x_f, \omega)$ with $P_{X|X_f,\Omega,\Gamma}(x, x_f, \omega, \gamma)$, which cannot decrease the value of the resulting maximum.

Finally, using the fact that, because of Bayes' theorem, it holds that

$$\sum_{x_f} P_\Omega(\omega) P_{X_f,\Gamma|\Omega}(x_f, \gamma, \omega) P_{X|X_f,\Omega,\Gamma}(x, x_f, \omega, \gamma) = P_{X,\Gamma,\Omega}(x, \gamma, \omega), \tag{G11}$$

and introducing the joint variable $\Xi := (\Gamma, \Omega)$, with values $\xi := (\gamma, \omega)$, and the joint Kraus operators $T_\xi := E_{\gamma,\omega} K_\omega$, which satisfy the normalization condition $\sum_\xi T_\xi^\dagger T_\xi = \sum_{\gamma,\omega} E_{\gamma,\omega}^\dagger K_\omega^\dagger E_{\gamma,\omega} K_\omega = \mathbb{1}$, we write the inequality [(G10)] as

$$\sum_\omega P_\Omega(\omega) S_A \left(\frac{\mathcal{M}_\omega(\hat{\rho}_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho_{A|X})]} \middle\| \frac{\mathcal{M}_\omega(\hat{\rho}'_{A|X})}{\text{Tr}[\mathcal{M}_\omega(\rho'_{A|X})]} \right)$$

$$\leq \max_{P_{X_f|\Xi},\{T_\xi\}} \left\{ \left[S_C(P_\Xi \| P'_\Xi) + \sum_{x,\xi} P_{X,\Xi}(x, \xi) S_Q \left(\frac{\mathbb{1} \otimes T_\xi \hat{\rho}_{A|X}(x) T_\xi^\dagger \otimes \mathbb{1}}{P_\Xi(\xi)} \middle\| \frac{\mathbb{1} \otimes T_\xi \hat{\rho}'_{A|X}(x) T_\xi^\dagger \otimes \mathbb{1}}{P'_\Xi(\xi)} \right) \right] \right\}. \tag{G12}$$

By Definition 3, the right-hand side of Eq. [(G12)] coincides with the right-hand side of Eq. [(C1)]. ∎

2. Proof of Lemma 2

For the proof of this lemma, it is useful to reexpress Eq. [(18)] in terms of abstract flag states representing the outcomes of Bob's generalized quantum measurements. Introducing an auxiliary extension Hilbert space $\mathcal{H}_{E_B}$ and an orthonormal basis of it, $\{|\gamma\rangle\}$, where each basis member encodes the value γ of the measurement outcomes, and using $\sum_x P_{X|\Gamma}(x, \gamma) = 1$ for all γ, we write

$$S_A(\rho_{A|X} \| \rho'_{A|X}) = \max_{P_{X|\Gamma},\{E_\gamma\},} \left\{ \sum_x P_{X|\Gamma}(x, \gamma) \left[S_C(P_\Gamma \| P'_\Gamma) \right. \right.$$

$$\left. \left. + \sum_\gamma P_\Gamma(\gamma) S_Q \left(\frac{\mathbb{1} \otimes E_\gamma \hat{\rho}_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger}{P_\Gamma(\gamma)} \middle\| \frac{\mathbb{1} \otimes E_\gamma \hat{\rho}'_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger}{P'_\Gamma(\gamma)} \right) \right] \right\}$$

$$= \max_{P_{X|\Gamma},\{E_\gamma\}} \left[\sum_x P_{X|\Gamma}(x, \gamma) S_Q \left(\sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger \middle\| \sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}'_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger \right) \right]. \tag{G13}$$

We can now prove the lemma.

Proof of Lemma 2.— Using Eq. [(G13)], we write the left-hand side of Eq. [(C2)] as

$$S_A \left(\sum_j \mu^{(j)} \rho_{A|X}^{(j)} \middle\| \sum_j \mu^{(j)} \rho'^{(j)}_{A|X} \right)$$

$$= \max_{P_{X|\Gamma},\{E_\gamma\}} \left[\sum_x P_{X|\Gamma}(x, \gamma) S_Q \left(\sum_j \mu^{(j)} \sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}_{A|X}^{(j)}(x) \mathbb{1} \otimes E_\gamma^\dagger \middle\| \sum_j \mu^{(j)} \sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}'^{(j)}_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger \right) \right]$$

$$\leq \max_{P_{X|\Gamma},\{E_\gamma\}} \left[\sum_{x,j} \mu^{(j)} P_{X|\Gamma}(x, \gamma) S_Q \left(\sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}_{A|X}^{(j)}(x) \mathbb{1} \otimes E_\gamma^\dagger \middle\| \sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_\gamma \hat{\rho}'^{(j)}_{A|X}(x) \mathbb{1} \otimes E_\gamma^\dagger \right) \right] \tag{G14}$$

$$\leq \sum_j \mu^{(j)} \max_{\{E_{\gamma,j}\},P_{X|\Gamma,J}} \left[\sum_x P_{X|\Gamma,J}(x, \gamma, j) S_Q \left(\sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_{\gamma,j} \hat{\rho}_{A|X}^{(j)}(x) \mathbb{1} \otimes E_{\gamma,j}^\dagger \middle\| \sum_\gamma |\gamma\rangle\langle\gamma| \otimes E_{\gamma,j} \hat{\rho}'^{(j)}_{A|X}(x) \mathbb{1} \otimes E_{\gamma,j}^\dagger \right) \right], \tag{G15}$$

where Eq. [(G14)] follows from Eq. [(G1)] and, in Eq. [(G15)], we exchanged the order of the maximization and the summation over j by respectively replacing $\{E_\gamma\}$ and $P_{X|\Gamma}$ with $\{E_{\gamma,j}\}$ and $P_{X|\Gamma,J}$, of elements $P_{X|\Gamma,J}(x, \gamma, j)$. Again using Eq. [(G13)], one sees that, by Definition 3, the right-hand side of Eq. [(G15)] coincides with the right-hand side of Eq. [(C2)]. ∎

[1] E. Schrödinger, *Discussion of Probability Relations between Separated Systems*, Proc. Cambridge Philos. Soc. **31,** 555 (1935).

[2] H. M. Wiseman, S. J. Jones, and A. C. Doherty, *Steering, Entanglement, Nonlocality, and the Einstein-Podolsky-Rosen Paradox*, Phys. Rev. Lett. **98,** 140402 (2007); S. J. Jones, H. M. Wiseman, and A. C. Doherty, Phys. Rev. A **76,** 052116 (2007).

[3] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen, and G. Leuchs, *Colloquium: The Einstein-Podolsky-Rosen Paradox: From Concepts to Applications*, Rev. Mod. Phys. **81,** 1727 (2009).

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

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

[6] E. G. Cavalcanti, S. J. Jones, H. M. Wiseman, and M. D. Reid, *Experimental Criteria for Steering and the Einstein-Podolsky-Rosen Paradox*, Phys. Rev. A **80,** 032112 (2009).

[7] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng, *Realization of the Einstein-Podolsky-Rosen Paradox for Continuous Variables*, Phys. Rev. Lett. **68,** 3663 (1992); W. P. Bowen, R. Schnabel, and P. K. Lam, *Experimental Investigation of Criteria for Continuous Variable Entanglement*, Phys. Rev. Lett. **90,** 043601 (2003); D.-H. Smith *et al.*, *Conclusive Quantum Steering with Superconducting Transition-Edge Sensors*, Nat. Commun. **3,** 625 (2012); A. J. Bennet, D. A. Evans, D. J. Saunders, C. Branciard, E. G. Cavalcanti, H. M. Wiseman, and G. J. Pryde, *Arbitrarily Loss-Tolerant Einstein-Podolsky-Rosen Steering Allowing a Demonstration over 1 km of Optical Fiber with No Detection Loophole*, Phys. Rev. X **2,** 031003 (2012); V. Händchen, T. Eberle, S. Steinlechner, A. Samblowski, T. Franz, R. F. Werner, and Roman Schnabel, *Observation of One-Way Einstein-Podolsky-Rosen Steering*, Nat. Photonics **6,** 598 (2012); S. Steinlechner, J. Bauchrowitz, T. Eberle, and R. Schnabel, *Strong Einstein-Podolsky-Rosen Steering with Unconditional Entangled States*, Phys. Rev. A **87,** 022104 (2013).

[8] D. J. Saunders, S. J. Jones, H. M. Wiseman, and G. J. Pryde, *Experimental EPR-Steering Using Bell-Local States*, Nat. Phys. **6,** 845 (2010).

[9] B. Wittmann, S. Ramelow, F. Steinlechner, N. K. Langford, N. Brunner, H. Wiseman, R. Ursin, and A. Zeilinger, *Loophole-Free Einstein-Podolsky-Rosen Experiment via Quantum Steering*, New J. Phys. **14,** 053030 (2012).

[10] C. Branciard, E. G. Cavalcanti, S. P. Walborn, V. Scarani, and H. M. Wiseman, *One-Sided Device-Independent Quantum Key Distribution: Security, Feasibility, and the Connection with Steering*, Phys. Rev. A **85,** 010301(R) (2012).

[11] Q. Y. He and M. D. Reid, *Genuine Multipartite Einstein-Podolsky-Rosen Steering*, Phys. Rev. Lett. **111,** 250403 (2013).

[12] J. Barrett, L. Hardy, and A. Kent, *No Signaling and Quantum Key Distribution*, Phys. Rev. Lett. **95,** 010503 (2005); A. Acín, N. Gisin, and L. Masanes, *From Bells Theorem to Secure Quantum Key Distribution*, Phys. Rev. Lett. **97,** 120405 (2006); A. Acín, N. Brunner, N. Gisin, S. Massar, S. Pironio, and V. Scarani, *Device-Independent Security of Quantum Cryptography against Collective Attacks*, Phys. Rev. Lett. **98,** 230501 (2007).

[13] V. Vedral and M. B. Plenio, *Entanglement Measures and Purification Procedures*, Phys. Rev. A **57,** 1619 (1998); M. B. Plenio and S. Virmani, *An Introduction to Entanglement Measures*, Quantum Inf. Comput. **7,** 1 (2007).

[14] F. G. S. L. Brandão and M. B. Plenio, *Entanglement Theory and the Second Law of Thermodynamics*, Nat. Phys. **4,** 873 (2008); *A Reversible Theory of Entanglement and Its Relation to the Second Law*, Commun. Math. Phys. **295,** 829 (2010).

[15] C. H. Bennett, D. P. DiVincenzo, J. Smolin, and W. K. Wootters, *Mixed-State Entanglement and Quantum Error Correction*, Phys. Rev. A **54,** 3824 (1996).

[16] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight, *Quantifying Entanglement*, Phys. Rev. Lett. **78,** 2275 (1997).

[17] D. Jonathan and M. B. Plenio, *Entanglement-Assisted Local Manipulation of Pure Quantum States*, Phys. Rev. Lett. **83,** 3566 (1999).

[18] 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).

[19] M. Ahmadi, D. Jennings, and T. Rudolph, *The Wigner-Araki-Yanase Theorem and the Quantum Resource Theory of Asymmetry*, New J. Phys. **15,** 013057 (2013).

[20] G. Gour and R. W. Spekkens, *The Resource Theory of Quantum Reference Frames: Manipulations and Monotones*, New J. Phys. **10,** 033023 (2008).

[21] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, *Operational Framework for Nonlocality*, Phys. Rev. Lett. **109,** 070401 (2012).

[22] J. I. de Vicente, *On Nonlocality as a Resource Theory and Nonlocality Measures*, J. Phys. A **47,** 424017 (2014).

[23] M. F. Pusey, *Negativity and Steering: A Stronger Peres Conjecture*, Phys. Rev. A **88,** 032313 (2013).

[24] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, *Quantifying Einstein-Podolsky-Rosen Steering*, Phys. Rev. Lett. **112,** 180404 (2014).

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

[26] J. Bowles, T. Vértesi, M. T. Quintino, and N. Brunner, *One-Way Einstein-Podolsky-Rosen Steering*, Phys. Rev. Lett. **112,** 200402 (2014).

[27] M. T. Quintino, T. Vértesi, and N. Brunner, *Joint Measurability, Einstein-Podolsky-Rosen Steering, and Bell Nonlocality*, Phys. Rev. Lett. **113,** 160402 (2014).

[28] A. B. Sainz, N. Brunner, D. Cavalcanti, P. Skrzypczyk, and T. Vértesi, *Post-Quantum Steering*, arXiv:1505.01430.

[29] O. Oreshkov and J. Calsamiglia, *Distinguishability Measures between Ensembles of Quantum States*, Phys. Rev. A **79,** 032336 (2009).

[30] The terminology, "deterministic maps" refers throughout to probability (trace) preserving classical (quantum) maps, i.e., those that never cause an abortion. For classical maps, for instance, this should not be confused with maps where the output bit is a Kronecker delta function of the input bit. In turn, the term "stochastic" is used throughout to refer to

non-probability-preserving classical maps or non-trace-preserving quantum transformations, which do not occur with certainty.

[31] A. K. Ekert, *Quantum Cryptography Based on Bells Theorem*, Phys. Rev. Lett. **67**, 661 (1991).

[32] A natural question (which we leave open) is whether there exists a definition of relative entropy between assemblages that is nonincreasing under generic assemblage transformations instead of just 1W-LOCCs, so that it can be understood as a measure of distinguishability under fully general strategies. For quantum states, the relative entropy is given by S_Q which is nonincreasing under not only LOCCs but also under any completely positive map. However, note that 1W-LOCC monotonicity of S_A suffices to introduce a steering monotone.

[33] I. Kogias, A. R. Lee, S. Ragy, and G. Adesso, *Quantification of Gaussian Quantum Steering*, Phys. Rev. Lett. **114**, 060403 (2015).

[34] G. Vidal, *Entanglement of Pure States for a Single Copy*, Phys. Rev. Lett. **83**, 1046 (1999).

[35] W. Dür, G. Vidal, and J. I. Cirac, *Three Qubits Can Be Entangled in Two Inequivalent Ways*, Phys. Rev. A **62**, 062314 (2000).

[36] M. Owari, K. Matsumoto, and M. Murao, *Entanglement Convertibility for Infinite-Dimensional Pure Bipartite States*, Phys. Rev. A **70**, 050301(R) (2004).

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

[38] B. Lang, T. Vértesi, and M. Navascués, *Closed Sets of Correlations: Answers from the Zoo*, arXiv:1402.2850.

[39] V. Vedral and M. B. Plenio, *Entanglement Measures and Purification Procedures*, Phys. Rev. A **57**, 1619 (1998).

[40] E. H. Lieb, *Convex Trace Functions and the Wigner-Yanase-Dyson Conjecture*, Adv. Math. **11**, 267 (1973).