Open access • Journal Article • DOI:10.1103/PHYSREVA.76.052116

Entanglement, einstein-podolsky-rosen correlations, bell nonlocality, and steering

Steve James Jones, Howard M. Wiseman, Andrew C. Doherty

Institutions: Griffith University, University of Queensland

Published on: 28 Nov 2007 - Physical Review A (American Physical Society)

Topics: Quantum nonlocality, No-communication theorem, Bell's theorem, EPR paradox and Quantum entanglement

Entanglement, Einstein-Podolsky-Rosen correlations, Bell nonlocality, and steering

S. J. Jones,[1] H. M. Wiseman,[1] and A. C. Doherty[2]
[1]*Centre for Quantum Computer Technology, Centre for Quantum Dynamics, Griffith University, Brisbane, Qld. 4111, Australia*
[2]*School of Physical Sciences, University of Queensland, Brisbane Qld. 4072, Australia*
(Received 4 September 2007; published 28 November 2007)

In a recent work [Phys. Rev. Lett. **98**, 140402 (2007)] we defined "steering," a type of quantum nonlocality that is logically distinct from both nonseparability and Bell nonlocality. In the bipartite setting, it hinges on the question of whether Alice can affect Bob's state at a distance through her choice of measurement. More precisely and operationally, it hinges on the question of whether Alice, with classical communication, can convince Bob that they share an entangled state under the circumstances that Bob trusts nothing that Alice says. We argue that if she can, then this demonstrates the nonlocal effect first identified in the famous Einstein-Podolsky-Rosen paper [Phys. Rev. **47**, 777 (1935)] as a universal effect for pure entangled states. This ability of Alice to remotely prepare Bob's state was subsequently called steering by Schrödinger, whose terminology we adopt. The phenomenon of steering has been largely overlooked, and prior to our work had not even been given a rigorous definition that is applicable to mixed states as well as pure states. Armed with our rigorous definition, we proved that steerable states are a strict subset of the entangled states, and a strict superset of the states that can exhibit Bell nonlocality. In this work we expand on these results and provide further examples of steerable states. We also elaborate on the connection with the original EPR paradox.

DOI: 10.1103/PhysRevA.76.052116 PACS number(s): 03.65.Ud, 03.67.Mn

I. INTRODUCTION

Entanglement is arguably the central concept in the field of quantum information. However, there is an unresolved tension between different notions of what entanglement is, even in the bipartite setting. On the one hand, entangled states are defined as those that cannot be created from factorizable states using local operations and classical communication (LOCC). On the other hand, entanglement is regarded as a resource that enables the two parties to perform interesting or (in more recent times) useful nonlocal protocols. For pure states, which were the only states considered in this context for many decades, these notions coincide, and the word "entangled" (introduced by Schrödinger [1]) is identical with "factorizable."

The first authors to identify an interesting nonlocal effect associated with unfactorizable states were Einstein, Podolsky, and Rosen (EPR) in 1935 [2]. They considered a general unfactorizable pure state of two systems, held by two distant parties (say Alice and Bob)[1],

$$|\Psi\rangle = \sum_{n=1}^{\infty} c_n |u_n\rangle|\psi_n\rangle = \sum_{n=1}^{\infty} d_n |v_n\rangle|\varphi_n\rangle. \tag{1.1}$$

Here $\{|u_n\rangle\}$ and $\{|v_n\rangle\}$ are two different orthonormal bases for Alice's system. If such states exist, then if Alice chose to measure in the $\{|u_n\rangle\}$ (respectively, $\{|v_n\rangle\}$) basis, then she would instantaneously collapse Bob's system into one of the states $|\psi_n\rangle$ (respectively, $|\varphi_n\rangle$). That is, "as a consequence of two different measurements performed upon the first system, the second system may be left in states with two different wave functions" [2]. Now comes the paradox: "the two systems no longer interact, [so] no real change can take place in [Bob's] system in consequence of anything that may be done to [Alice's] system" [2]. That is, if Bob's quantum state is the real state of his system, then Alice cannot choose to make it collapse into either one of the $|\psi_n\rangle$ or one of the $|\phi_n\rangle$ because that would violate local causality. Note that it is crucial to consider more than one sort of measurement for Alice; if Alice were restricted to measuring in one basis (say the $|u_n\rangle$ basis), then it would be impossible to demonstrate any "real change" in Bob's system, because she might know beforehand which of the $|\psi_n\rangle$ is the real state of his system. That is, the paradox exists only if there is not a local hidden state (LHS) model for Bob's system, in which the real state $|\psi_n\rangle$ is hidden from Bob but may be known to Alice.

As the above quotations show, EPR assumed local causality to be a true feature of the world; indeed, they say that no "reasonable" theory could be expected to permit otherwise. They thus concluded that the wave function cannot describe reality; that is, the quantum mechanical (QM) description must be *incomplete*. Their intuition was thus that local causality could be maintained by completing QM. This intuition was supported by the famous example that they then presented as a special case of Eq. (1.1), involving a bipartite entangled state with perfect correlations in position and momentum. The "EPR paradox" in this case is trivially resolved by considering local hidden variables (LHVs) for position and momentum.

Although the argument of EPR against the completeness of QM was correct, their intuition was not. As proven by Bell [3,4], local causality cannot be maintained even if one allows QM to be completed by hidden variables. That is, assuming as always (and with good justification [5]) that QM is correct, Bell's theorem proves that local causality is not a true

[1]All we have changed from EPR's presentaton is to use Dirac's notation rather than wave functions

feature of the world.[2] Interestingly, any unfactorizable pure state can be used not only to demonstrate the EPR paradox [2], but also to demonstrate Bell nonlocality (that is, the violation of local causality). (This fact was perhaps first stated in 1989 by Werner [6]; the first detailed proof was given in 1991 by Gisin [7]; see also [8].)

With the rise of quantum information *experiment*, the idealization of considering only pure states has become untenable. The question of which *mixed* states were Bell nonlocal (that is, allowed a demonstration of Bell nonlocality) was first addressed by Werner [6], in a foundational paper pregnant with implications for, and applications in, quantum information science. Revealing the first hint of the complexity of mixed-state entanglement, still being uncovered [9], Werner showed that not all mixed entangled states can demonstrate Bell nonlocality. Here, for mixtures, an entangled state is defined as one which cannot be written as a mixture of factorizable pure states. Indeed Werner's paper is often cited as that which introduced this definition. That is, he is credited with introducing the dichotomy of entangled states versus separable (i.e., locally preparable) states. However, it is interesting to note that he used neither the term entangled nor the term separable. For a discussion of the history of terms used in this context, and their relation to the present work, see Appendix A.

In a recent paper, the present authors also considered the issue of mixed states and nonlocality [10]. We rigorously defined the class of states that can be used to demonstrate the nonlocal effect which EPR identified in 1935. We proposed the term "steerable" for this class of states (for reasons given in Appendix A), and proved that the set of Bell-nonlocal states is a strict subset of the set of steerable states, which in turn is a strict subset of the set of nonseparable states. This was our main result.

Like "entangled," "steering" is a term introduced by Schrödinger [1] in the aftermath of the EPR paper. Specifically, he credits EPR with calling attention to "the obvious but very disconcerting fact" that for a pure entangled state like Eq. (1.1), Bob's system can be "steered or piloted into one or the other type of state at [Alice's] mercy in spite of [her] having no access to it." He referred to this as a "paradox" [1,11] because if such states can exist, and if the QM description is complete, then local causality must be violated.

In Ref. [10] we first supplied an operational definition of steering in the style of a quantum informational task involving two parties (in contrast to demonstrating Bell nonlocality, which can be defined as a task involving three parties). Next we turned this operational definition into a mathematical definition. Applying this to the case of 2×2 dimensional Werner states enabled us to establish our main result, quoted above. We then completely characterized steerability for $d\times d$-dimensional Werner states and isotropic states. Finally, we completely characterized the Gaussian states that are steerable by Gaussian measurements, and related this to the Reid criterion [12] for the EPR paradox.

In the present paper we expand and extend the material in Ref. [10]. In Sec. II we present the operational definitions of Bell nonlocality and steering as before, and also that for demonstrating nonseparability. In addition we use these operational definitions to show that they lead to a hierarchy of states: Bell nonlocal within steerable within nonseparable. In Sec. III we turn our operational definitions into mathematical definitions, and in addition we explain how our definition of steering conforms to Schrödinger's use of the term. In Sec. IV we derive conditions for steerability for four families of states. As before, we consider Werner states, isotropic states, and Gaussian states, but here we expand the proofs for the benefit of the reader. In addition, we consider another class of states: the "inept states" of Ref. [13]. We also consider a subclass of Gaussian states in more detail: the symmetric two-mode states produced in parametric down conversion. We conclude with a summary and discussion in Sec. V.

II. OPERATIONAL DEFINITIONS

It is useful to begin with some operational definitions for the different properties of quantum states that we wish to consider. This is useful for a number of reasons. First, it presents the ideas that we wish to discuss in an accessible format for those familiar with concepts in modern quantum information. Second, it allows us to present an elementary proof of the hierarchy of the concepts (we will present a more detailed proof of this hierarchy in subsequent sections).

First, let us define the familiar concept of Bell nonlocality [3] as a task, in this case with three parties; Alice, Bob, and Charlie. Alice and Bob can prepare a shared bipartite state, and repeat this any number of times. Each time, they measure their respective parts. Except for the preparation step, communication between them is forbidden (this prevents them from colluding in an attempt to fool Charlie). Their task is to convince Charlie (with whom they can communicate) that the state they can prepare is entangled. Charlie accepts QM as correct, but trusts neither Alice nor Bob. If the correlations between the results they report *can* be explained by a LHV model, then Charlie will not be convinced that the state is entangled; the results could have been fabricated from shared classical randomness. Conversely, if the correlations *cannot* be so explained then the state must be entangled. Therefore they will succeed in their task iff (if and only if) they can demonstrate Bell nonlocality. This task can be thus considered as an operational definition of violating a Bell inequality.

The analogous definition for steering uses a task with only two parties. Alice can prepare a bipartite quantum state and send one part to Bob, and repeat this any number of times. Each time, they measure their respective parts, and communicate classically. Alice's task is to convince Bob that the state she can prepare is entangled. Bob (like Schrödinger) accepts that QM describes the results of the measurements he

[2]For both Bell and EPR, there is an escape, "by denying independent real situations as such to things which are spatially separated from each other," as stated by Einstein in 1946 [41]. That is, Alice, for example, can refuse to admit the reality of Bob's measurement results until she observes them, by talking only about the outcomes of her own future measurements. However, Einstein stated that in his opinion this antirealism was "equally unacceptable" as violating local causality; see Ref. [42] for a discussion.

makes (which, we assume, allow him to do local state tomography). However, Bob does not trust Alice. In this case Bob must determine whether the correlations between his local state and Alice's reported results are proof of entanglement. How he should determine this is explained in detail in Sec. III, but the basic idea is that he should not accept the correlations as proof of entanglement if they can be explained by a LHS model for Bob. If the correlations between Bob's measurement results and the results Alice reports *can* be so explained then Alice's results could have been fabricated from her knowledge of Bob's LHS in each run. Conversely, if the correlations *cannot* be so explained then the bipartite state must be entangled. Therefore we say that Alice will succeed in her task iff she can steer Bob's state.

Finally, the simplest task is for Alice and Bob to determine iff a bipartite quantum state that they share is nonseparable. In this case they can communicate results to one another, they trust each other, and they can repeat the experiment sufficiently many times to perform state tomography. By analyzing the reconstructed bipartite state, they could determine whether it is nonseparable. That is, whether it can be described by correlated LHSs for Alice and Bob. Because Alice and Bob trust each other and can freely communicate, this is really a *one* party task.

Using these operational definitions we can show that Bell nonlocality is a stronger concept than steerability. That is, that Bell-nonlocal states are a subset of the steerable states. The operational definition of Bell nonlocality is based on three parties and requires a completely distrustful Charlie. If we *weaken* this condition by allowing Charlie to trust Bob completely, we arrive at the following situation. Charlie can now, in principle, do state tomography for Bob's local state (as he believes everything told to him by Bob), and he only distrusts the measurement results reported by Alice. In this case, he will only concede that the state prepared by Alice and Bob is entangled if the state is steerable. Thus it is possible to arrive at the operational definition for steering by weakening the operational definition for Bell nonlocality. Thus, the Bell-nonlocal states are a subset of the steerable states.

Similarly, if we weaken the condition for steerability we arrive at the condition for nonseparability as follows. In this case we weaken the condition by allowing for Bob to trust Alice completely. Since Bob now has access to the measurement information for both subsystems (as he believes everything told to him by Alice) he can, in principle, perform state tomography. Clearly, in this situation Bob will only concede that they share an entangled state if the state that Alice prepares really is entangled. Thus, the steerable states are a subset of the entangled states. We illustrate these relations graphically in Fig. 1.

While these operational definitions give a good insight into the relationships between the three classes of states it is also desirable to have a strict mathematical way to define the classes. We present such definitions in the following section.

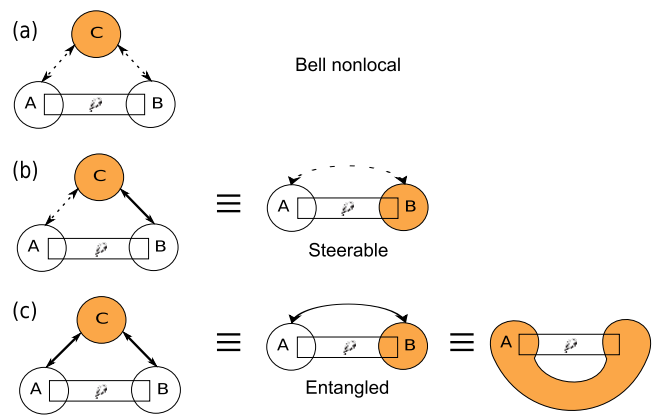

FIG. 1. (Color online) Operational definitions for classes of entangled states. Bell-nonlocal states (a) can be defined via a three-party task (involving Alice [A], Bob [B], and Charlie [C]). Steerable states (b) may be defined using a two-party task. Defining an entangled state (c) essentially requires only one party. In all cases shading indicates the skeptical party, dotted arrows indicate two-way communication, and solid arrows indicate trust and two-way communication.

III. MATHEMATICAL DEFINITIONS

First, we define some terms. Let the set of all observables on the Hilbert space for Alice's system be denoted $\mathfrak{D}_\alpha$. We denote an element of $\mathfrak{D}_\alpha$ by $\hat{A}$, and the set of eigenvalues $\{a\}$ of $\hat{A}$ by $\lambda(\hat{A})$. By $P(a|\hat{A};W)$ we mean the probability that Alice will obtain the result a when she measures $\hat{A}$ on a system with state matrix W. We denote the measurements that Alice is able to perform by the set $\mathfrak{M}_\alpha \subseteq \mathfrak{D}_\alpha$. Note that, following Werner [6], we are restricting to projective measurements. The corresponding notations for Bob, and for Alice and Bob jointly, are obvious. Thus, for example,

$$P(a,b|\hat{A},\hat{B};W) = \mathrm{Tr}[(\hat{\Pi}_a^A \otimes \hat{\Pi}_b^B)W], \tag{3.1}$$

where $\hat{\Pi}_a^A$ is the projector satisfying $\hat{A}\hat{\Pi}_a^A = a\hat{\Pi}_a^A$.

The strongest sort of nonlocality in QM is Bell nonlocality [3]. This is a property of entangled states which violate a Bell inequality. This is exhibited in an experiment on state W if the correlations between a and b cannot be explained by a LHV model. That is, if it is *not* the case that for all $a \in \lambda(\hat{A})$, $b \in \lambda(\hat{B})$, for all $\hat{A} \in \mathfrak{M}_\alpha$, $\hat{B} \in \mathfrak{M}_\beta$, we have

$$P(a,b|\hat{A},\hat{B};W) = \sum_\xi \wp(a|\hat{A},\xi)\wp(b|\hat{B},\xi)\wp_\xi. \tag{3.2}$$

Here, and below, $\wp(a|\hat{A},\xi)$, $\wp(b|\hat{B},\xi)$, and $\wp_\xi$ denote some (positive, normalized) probability distributions, involving the LHV ξ. We say that a *state* is Bell nonlocal if there exists a measurement set $\mathfrak{M}_\alpha \times \mathfrak{M}_\beta$ that allows Bell nonlocality to be demonstrated. If Eq. (3.2) is always satisfied we say W is Bell local.

A strictly weaker [6] concept is that of nonseparability or entanglement. A nonseparable state is one that *cannot* be written as

$$W = \sum_{\xi} \sigma_{\xi} \otimes \rho_{\xi} \wp_{\xi}. \tag{3.3}$$

Here, and below, $\sigma_{\xi} \in \mathfrak{D}_{\alpha}$ and $\rho_{\xi} \in \mathfrak{D}_{\beta}$ are some (positive, normalized) quantum states. We can also give an operational definition, by allowing Alice and Bob the ability to measure a quorum of local observables, so that they can reconstruct the state W by tomography [14]. Since the complete set of observables $\mathfrak{D}$ is obviously a quorum, we can say that a state W is nonseparable if it is *not* the case that for all $a \in \lambda(\hat{A})$, $b \in \lambda(\hat{B})$, for all $\hat{A} \in \mathfrak{D}_{\alpha}$, $\hat{B} \in \mathfrak{D}_{\beta}$, we have

$$P(a,b|\hat{A},\hat{B};W) = \sum_{\xi} P(a|\hat{A};\sigma_{\xi}) P(b|\hat{B};\rho_{\xi}) \wp_{\xi}. \tag{3.4}$$

Bell nonlocality and nonseparability are both concepts that are symmetric between Alice and Bob. However *steering*, Schrödinger's term for the EPR effect [1], is inherently asymmetric. It is about whether Alice, by her choice of measurement $\hat{A}$, can collapse Bob's system into different types of states in the different ensembles $E^A \equiv \{\tilde{\rho}_a^A : a \in \lambda(\hat{A})\}$. Here $\tilde{\rho}_a^A \equiv \mathrm{Tr}_{\alpha}[W(\hat{\Pi}_a^A \otimes \mathbf{I})] \in \mathfrak{D}_{\beta}$ is Bob's state conditioned on Alice measuring $\hat{A}$ with result a. The tilde denotes that this state is unnormalized (its norm is the probability of its realization). Of course Alice cannot affect Bob's unconditioned state $\rho = \mathrm{Tr}_{\alpha}[W] = \Sigma_a \tilde{\rho}_a^A$—that would allow superluminal signaling. Despite this, steering is clearly nonlocal if one believes that the state of a quantum system is a physical property of the system, as did Schrödinger. This is apparent from his statement that "It is rather discomforting that the theory should allow a system to be steered or piloted into one or the other type of state at the experimenter's mercy in spite of his having no access to it."

As this quote also shows, Schrödinger was not wedded to the terminology steering. He also used the term "control" for this phenomenon [11], and the word "driving" in the context of his 1936 result that "…a sophisticated experimenter can…produce a non-vanishing probability of driving the system into any state he chooses" [11]. (By this he means that if a bipartite system is in a pure entangled state, then one party (Alice) can, by making a suitable measurement on her subsystem, create any pure quantum state $|\psi\rangle$ for Bob's subsystem with probability $\langle\psi|\rho^{-1}|\psi\rangle^{-1}$, whenever this is well defined [11].) He regarded steering or driving as a "necessary and indispensable feature" of quantum mechanics [11], but found it "repugnant," and doubted whether it was really true. That is, he was "not satisfied about there being enough experimental evidence for [its existence in Nature]" [11].

What experimental evidence would have convinced Schrödinger? The pure entangled states he discussed are an idealization, so we cannot expect ever to observe precisely the phenomenon he introduced. On the other hand, Schrödinger was quite explicit that a separable but correlated state, which allows "determining the state of the first system by *suitable* [his emphasis] measurement of the second or *vice versa*" could never exhibit steering. Of this situation, he says that "it would utterly eliminate the experimenter's influence on the state of that system which he does not touch." Thus it is apparent that by steering Schrödinger meant something that could not be explained by Alice simply finding out which state Bob's system is in, out of some predefined ensemble of states. In other words, the "experimental evidence" Schrödinger sought is precisely the evidence that would convince Bob that Alice has prepared an entangled state under the conditions described in our first (operational) definition of steering.

To reiterate, we assume that the experiment can be repeated at will, and that Bob can do state tomography. Prior to all experiments, Bob demands that Alice announce the possible ensembles $\{E^A : \hat{A} \in \mathfrak{M}_{\alpha}\}$ she can steer Bob's state into. In any given run (after he has received his state), Bob should randomly pick an ensemble E^A, and ask Alice to prepare it.[3] Alice should then do so, by measuring $\hat{A}$ on her system, and announce to Bob the particular member ρ_a^A she has prepared. Over many runs, Bob can verify that each state announced is indeed produced, and is announced with the correct frequency $\mathrm{Tr}[\tilde{\rho}_a^A]$.

If Bob's system did have a preexisting LHS ρ_{ξ} (as Schrödinger thought), then Alice could attempt to fool Bob, using her knowledge of ξ. This state would be drawn at random from some prior ensemble of LHSs $F = \{\wp_{\xi}\rho_{\xi}\}$ with $\rho = \Sigma_{\xi}\wp_{\xi}\rho_{\xi}$. Alice would then have to announce a LHS $\tilde{\rho}_a^A$ based on her knowledge of ξ, according to some stochastic map from ξ to a. Alice will have failed to convince Bob that she can steer his system if, for all $\hat{A} \in \mathfrak{M}_{\alpha}$, and for all $a \in \lambda(\hat{A})$, there exists an ensemble F and a stochastic map $\wp(a|\hat{A},\xi)$ from ξ to a such that

$$\tilde{\rho}_a^A = \sum_{\xi} \wp(a|\hat{A},\xi) \rho_{\xi} \wp_{\xi}. \tag{3.5}$$

That is, if there exists a *coarse-graining* of ensemble F to ensemble E^A then Alice may simply know Bob's preexisting state ρ_{ξ}. Conversely, if Bob *cannot* find any ensemble F and map $\wp(a|\hat{A},\xi)$ satisfying Eq. (3.5) then Bob must admit that Alice can steer his system.

We can recast this definition as a "hybrid" of Eqs. (3.2) and (3.4): Alice's measurement strategy $\mathfrak{M}_{\alpha}$ on state W exhibits steering if it is *not* the case that for all $a \in \lambda(\hat{A}), b \in \lambda(\hat{B})$, for all $\hat{A} \in \mathfrak{M}_{\alpha}$, $\hat{B} \in \mathfrak{D}_{\beta}$, we can write

$$P(a,b|\hat{A},\hat{B};W) = \sum_{\xi} \wp(a|\hat{A},\xi) P(b|\hat{B};\rho_{\xi}) \wp_{\xi}. \tag{3.6}$$

That is, if the joint probabilities for Alice and Bob's measurements can be explained using a LHS model for Bob and a LHV model for Alice correlated with this state, then we have failed to demonstrate steering. Iff there exists a measurement strategy $\mathfrak{M}_{\alpha}$ that exhibits steering, we say that the state W is *steerable* (by Alice).

[3]This ensures that Bob need not trust Alice that they share the same state W in each run, because Alice gains nothing by preparing different states in different runs, because she never knows what ensemble Bob is going to ask for.

It is straightforward to see that the condition for no steering implies the condition for Bell locality, since if there is a model with $P(b|\hat{B},\rho_\xi)$ satisfying Eq. (3.6), then there is a model with $\wp(b|\hat{B},\xi)$ that satisfies Eq. (3.2); simply make $\wp(b|\hat{B},\xi)=P(b|\hat{B};\rho_\xi)$ for all $\hat{B},\xi$. Since no steering implies no Bell nonlocality, we see that if a state is Bell nonlocal, then it implies that it is also steerable. Hence Bell nonlocality is a stronger concept than steerability.

Similarly, the condition for separability implies the condition for no steering. If there is a model with $P(a|\hat{A};\sigma_\xi)$ satisfying Eq. (3.4), then there is a model with $\wp(a|\hat{A},\xi)$ that satisfies Eq. (3.6); simply make $\wp(a|\hat{A},\xi)=P(a|\hat{A};\sigma_\xi)$ for all $\hat{A},\xi$. Thus, steerability is also a stronger concept than nonseparability. At least one of these relations must be "strictly stronger than," because Bell nonlocality is strictly stronger than nonseparability [6]. In the following sections we prove that in fact steerability is strictly stronger than nonseparability, *and* strictly weaker than Bell nonlocality.

IV. CONDITIONS FOR STEERABILITY

Below we derive conditions for steerability for four families of states W. In each example we parametrize the family of states in terms of a mixing parameter $\eta\in\mathbb{R}$, and a second parameter that may be discrete. In each case, the upper bound for W to be a state is $\eta=1$, and W is a product state if $\eta=0$, and (except in the last case) W is linear in η. For the first two examples (Werner and isotropic states) the conditions derived are both necessary and sufficient for steerability. For the other examples (inept states and Gaussian states) the conditions derived are merely sufficient for steerability.

In terms of the parameter η we can define boundaries between different classes of states. For example, we will make use of η_{Bell}, defined by W^η being Bell nonlocal iff $\eta>\eta_{\text{Bell}}$. Similarly a state W^η is entangled iff $\eta>\eta_{\text{ent}}$. Our goal is then to determine (or at least bound) the steerability boundaries for the above classes of states, defined by W^η being steerable iff $\eta>\eta_{\text{steer}}$.

Crucial to the derivations of the conditions for steerability of these states is the concept of an *optimal ensemble* $F^\star=\{\rho_\xi^\star\wp_\xi^\star\}$; that is, an ensemble such that iff it cannot satisfy Eq. (3.5) then no ensemble can satisfy it. In finding an optimal ensemble $F^\star$ we use the symmetries of W and $\mathfrak{M}_\alpha$:

Lemma 1. Consider a group G with a unitary representation $\hat{U}_{\alpha\beta}(g)=\hat{U}_\alpha(g)\otimes\hat{U}_\beta(g)$ on the Hilbert space for Alice and Bob. Say that $\forall\hat{A}\in\mathfrak{M}_\alpha,\forall a\in\lambda(\hat{A}),\forall g\in G$, we have $\hat{U}_\alpha^\dagger(g)\hat{A}\hat{U}_\alpha(g)\in\mathfrak{M}_\alpha$ and

$$\tilde{\rho}_a^{\hat{U}_\alpha^\dagger(g)\hat{A}\hat{U}_\alpha(g)}=\hat{U}_\beta(g)\tilde{\rho}_a^A\hat{U}_\beta^\dagger(g). \tag{4.1}$$

Then there exists a G-covariant optimal ensemble: $\forall g\in G,\{\rho_\varsigma^\star\wp_\varsigma^\star\}=\{\hat{U}_\beta(g)\rho_\varsigma^\star\hat{U}_\beta^\dagger(g)\wp_\varsigma^\star\}$.

Proof. For specificity, consider a discrete group with order $|G|$. Say there exists an ensemble $F=\{\rho_\xi\wp_\xi\}$ satisfying Eq. (3.5) for some map $\wp(a|\hat{A},\xi)$. Then under the conditions of Lemma 1, $\tilde{\rho}_a^A$ can be rewritten as

$$|G|^{-1}\sum_{g\in G}\sum_\xi\hat{U}_\beta(g)\rho_\xi\hat{U}_\beta^\dagger(g)\wp(a|\hat{U}_\alpha^\dagger(g)\hat{A}\hat{U}_\alpha(g),\xi)\wp_\xi.$$

Thus the G-covariant ensemble $F^\star=\{\rho_{(g,\xi)}^\star\wp(g,\xi)\}$, with $\rho_{(g,\xi)}^\star=\hat{U}_\beta(g)\rho_\xi\hat{U}_\beta^\dagger(g)$ and $\wp(g,\xi)=\wp_\xi/|G|$, satisfies Eq. (3.5) with the choice

$$\wp^\star(a|\hat{A},(g,\xi))=\wp(a|\hat{U}_\alpha^\dagger(g)\hat{A}\hat{U}_\alpha(g),\xi). \tag{4.2}$$

The analogous formulas for the case of continuous groups are elementary. ■

Once we have determined the optimal ensemble for a given class of states (and a given measurement strategy) it remains to determine if there exists a stochastic map $\wp(a|\hat{A},\xi)$ such that Eq. (3.5) is true. In each steering experiment we assume that Alice really does send Bob an entangled state. To determine if the state is steerable, we take the perspective of a skeptical Bob and imagine that in each case Alice is attempting to cheat; that is, that she sends Bob a random state from the optimal ensemble $F^\star$ and does not perform her measurements. She simply announces her alleged measurement results based on $\wp(a|\hat{A},\xi)$ which defines her cheating strategy. We compare the states that Bob would obtain if Alice really did send half of an entangled state and perform a measurement with those that could be prepared using an optimal ensemble and cheating strategy.

There are two possible reasons why Bob could find that his measurement results are consistent with results reported by Alice. First, Alice could really be sending Bob half of an entangled state and steering his system via her measurements. Or, as the skeptical Bob believes, Alice could really just be sending him different pure states in each run and announcing her results based on her knowledge of this state.

Now if the optimal ensemble (which we are assuming Bob is clever enough to determine) can explain the correlations between Alice's announced results and Bob's results then the state sent by Alice is *not* steerable. However, if the best cheating strategy that Alice could possibly use is insufficient to explain the correlations then Bob must admit that Alice has sent him part of an entangled state. Furthermore, if he makes this admission, the state must be steerable.

A. Werner states

This family of states in $\mathbb{C}_d\otimes\mathbb{C}_d$ was introduced by Werner in Ref. [6]. As mentioned above, we parametrize it by $\eta\in\mathbb{R}$ such that W_d^η is linear in η, it is a product state for $\eta=0$, and is a state at all only for $\eta\leqslant 1$.

$$W_d^\eta=\left(\frac{d-1+\eta}{d-1}\right)\frac{\mathbf{I}}{d^2}-\left(\frac{\eta}{d-1}\right)\frac{\mathbf{V}}{d}. \tag{4.3}$$

Here $\mathbf{I}$ is the identity and $\mathbf{V}$ is the "flip" operator defined by $\mathbf{V}|\varphi\rangle\otimes|\psi\rangle\equiv|\psi\rangle\otimes|\varphi\rangle$. Defining $\Phi=[1-(d+1)\eta]/d$ allows one to reproduce Werner's notation [6] for these states. Werner states are nonseparable iff $\eta>\eta_{\text{ent}}=1/(d+1)$ [6]. For $d=2$, the Werner states violate the Clauser-Horne-Shimony-Holt (CHSH) inequality iff $\eta>1/\sqrt{2}$ [15]. This places an

upper bound on η_{Bell}. For $d>2$ only the trivial upper bound[4] of 1 is known. However, Werner found a lower bound on η_{Bell} of $1-1/d$ [6], which is strictly greater than η_{ent}.

Now let us consider the possibility of steering Werner states. We allow Alice all possible measurement strategies: $\mathfrak{M}_\alpha=\mathfrak{D}_\alpha$, and without loss of generality take the projectors to be rank one: $\hat{\Pi}_a^A=|a\rangle\langle a|$. For Werner states, the conditions of Lemma 1 are then satisfied for the d-dimensional unitary group $\mathfrak{U}(d)$. Specifically, $g\to\hat{U}$, and $\hat{U}_{\alpha\beta}(g)\to\hat{U}\otimes\hat{U}$ [6]. Again without loss of generality we can take the optimal ensemble to consist of pure states, in which case there is a unique covariant optimal ensemble, $F^\star=\{|\psi\rangle\langle\psi|d\mu_{\text{Haar}}(\psi)\}$, where $d\mu_{\text{Haar}}(\psi)$ is the Haar measure over $\mathfrak{U}(d)$.

If Alice were to make any projective measurement of her half of a Werner state and obtain the result a, Bob's unnormalized conditioned state would be given by

$$\begin{aligned}\tilde{\rho}_a^A &= \text{Tr}_A[(\hat{\Pi}_a^A\otimes\mathbf{I})W_d^\eta]\\ &= \langle a|W_d^\eta|a\rangle\\ &= \left(\frac{d-1+\eta}{d(d-1)}\right)\frac{\mathbf{I}}{d}-\left(\frac{\eta}{d(d-1)}\right)|a\rangle\langle a|.\end{aligned}\tag{4.4}$$

This is a state proportional to the completely mixed state *minus* a term proportional to the state Alice's system is projected into by her measurement.

We now determine if it is possible for Alice to simulate this conditioned state using the optimal ensemble $F^\star$ and an optimal cheating strategy defined by $\wp^\star(a|\hat{A},\psi)$. That is, we imagine that in each run of the experiment Alice simply sends Bob a state $\rho_\psi=|\psi\rangle\langle\psi|$ drawn at random from $F^\star=\{|\psi\rangle\langle\psi|d\mu_{\text{Haar}}(\psi)\}$. When asked to perform a measurement $\hat{A}$ and announce her result, she uses $\wp^\star(a|\hat{A},\psi)$ (which is based on her knowledge of $\rho_\psi=|\psi\rangle\langle\psi|$) to determine her answer. In testing whether this is actually what Alice could be doing, we only need to consider the quantity

$$\langle a|\tilde{\rho}_a^A|a\rangle=\frac{(1-\eta)}{d^2}.\tag{4.5}$$

This is due to the form of $\tilde{\rho}_a^A$ noted above in Eq. (4.4).

If (on average) the strategy used by Alice with the ensemble $F^\star$ produces the correct overlap with the state $|a\rangle\langle a|$ then Eq. (3.5) will hold and steering is not possible. Thus Alice makes use of the overlap with $|a\rangle\langle a|$ of the random states $|\psi\rangle$ in determining the optimal $\wp(a|\hat{A},\psi)$.

Since Alice's goal is to simulate $\tilde{\rho}_a^A$, as defined in Eq. (4.4), she will determine which of the eigenstates of $\hat{A}$ has the *least* overlap with $|\psi\rangle\langle\psi|$ in each run of the experiment and announce the eigenvalue associated with that eigenstate as her result. On average Bob would then find that his conditioned state has the least possible overlap with $|a\rangle\langle a|$. Writing this explicitly, the optimal distribution is given by

$$\wp^\star(a|\hat{A},\psi)=\begin{cases}1 & \text{if } \langle\psi|\hat{\Pi}_a^A|\psi\rangle<\langle\psi|\hat{\Pi}_{a'}^A|\psi\rangle \quad \forall\, a'\neq a\\ 0 & \text{otherwise.}\end{cases}\tag{4.6}$$

It is straightforward to see that this ensemble is normalized, that is, $\forall\ \hat{A},\psi$,

$$\sum_a \wp^\star(a|\hat{A},\psi)=1.\tag{4.7}$$

Clearly the optimal distribution $\wp^\star(a|A,\psi)$ is the distribution that will predict the same overlap with $|a\rangle\langle a|$ as that given by Eq. (4.5). This occurs at precisely the steering boundary η_{steer}. When $\eta<\eta_{\text{steer}}$ steering cannot be demonstrated, as it is possible that Alice is using a cheating strategy to simulate Bob's conditioned state. This means that Alice's optimal cheating strategy could actually make Bob believe that his conditioned state has a smaller overlap with $|a\rangle\langle a|$ than would be expected from Eq. (4.5). In this case Alice could correctly simulate $\tilde{\rho}_a^A$ simply by introducing the appropriate amount of randomness to her responses [i.e., increase the overlap to the correct size by choosing a different $\wp(a|\hat{A},\psi)$]. To reiterate, when $\eta<\eta_{\text{steer}}$ it is possible that Alice is performing a classical strategy which is consistent with Bob's results, so he will not believe that the state is genuinely steerable.

To find the form of η_{steer} we compare with Werner's result [6] for the lower bound on η_{ent}. We find that he actually used the construction outlined above. His LHVs for Bob's system were in fact the LHSs used in the optimal ensemble $F^\star$. Werner shows that for any positive normalized distribution $\wp(a|\hat{A},\psi)$,

$$\langle a|\int d\mu_{\text{Haar}}(\psi)|\psi\rangle\langle\psi|\wp(a|\hat{A},\psi)|a\rangle\geqslant 1/d^3.\tag{4.8}$$

The equality is attained for the optimal $\wp^\star(a|\hat{A},\psi)$ specified by Eq. (4.6) (this produces the smallest possible predicted overlap with $|a\rangle\langle a|$).

Now to determine when Eq. (3.5) is satisfied by $F^\star$ (and thus to determine η_{steer}) we simply compare Eq. (4.8) with Eq. (4.5). We find that Alice cannot simulate the correct overlap with $|a\rangle\langle a|$ iff

$$(1-\eta)/d^2<1/d^3.\tag{4.9}$$

Hence we see that for Werner states

$$\eta_{\text{steer}}=1-\frac{1}{d}.\tag{4.10}$$

Recently a new lower bound for η_{Bell} was found for $d=2$ by Acìn *et al.* [16], greater than η_{steer}, as shown in Fig. 2. Reference [16] makes use of a connection with Grothendieck's constant (a mathematical constant from Banach space theory) to develop a local hidden variable model for projective measurements when $d=2$. Acìn *et al.* show that for two-qubit Werner states

[4]This is because no Bell inequality has been found that the Werner states violate for $d>2$. It is only an upper bound because this is not a test of all possible Bell inequalities.

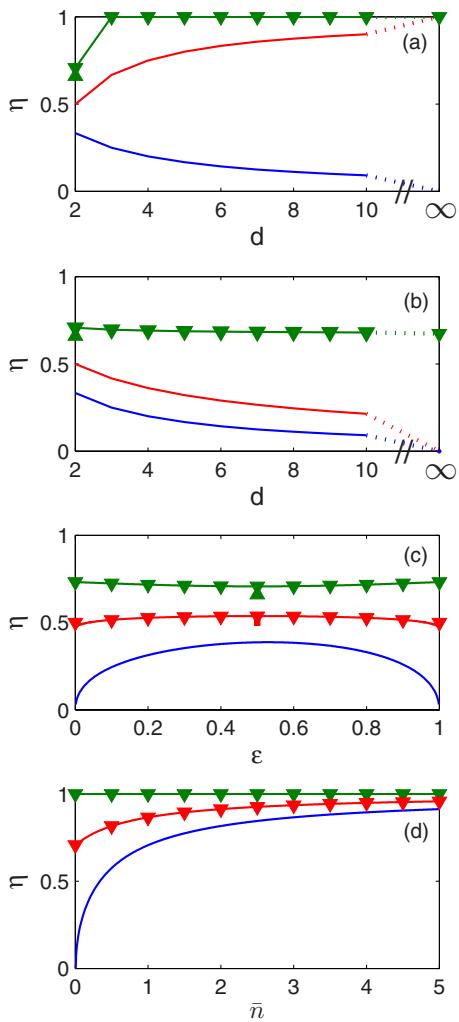

FIG. 2. (Color online) Boundaries between classes of entangled states for Werner (a) and isotropic (b) states W_d^η, inept states W_ϵ^η (c), and two-mode symmetric Gaussian states $W_{\bar{n}}^\eta$ (d). The bottom (blue) line is η_{ent}, above which states are entangled. The next (red) line is η_{steer}, above which states are steerable. In cases (c) and (d) the down arrows indicate that we have only an upper bound on η_{steer}. The top (green) line with down arrows is an upper bound on η_{Bell}, above which states are Bell nonlocal. The up arrows in cases (a) and (b) are lower bounds on η_{Bell} for $d=2$. This lower bound establishes that the classes are strictly distinct. In cases (a) and (b), dots join values at finite d with those at $d=\infty$. The separate point in (c) is explained at the end of Sec. IV C.

$$0.7071 \approx 1/\sqrt{2} \geqslant \eta_{\text{Bell}} \geqslant 1/K_g(3) \approx 0.6595, \quad (4.11)$$

where $K_g(3) \approx 1.5163$ is Grothendieck's constant of order 3. Bounds on $K_g(3)$ ensure that for $d=2$ Werner states $\eta_{\text{Bell}} \geqslant 0.6595$. Using Eq. (4.10), we see that when $d=2$, $\eta_{\text{steer}} = 1/2$. This proves that steerability is strictly weaker than Bell nonlocality as $\eta_{\text{steer}} < 0.6595 \leqslant \eta_{\text{Bell}}$. It is also well known that for $d=2$, $\eta_{\text{ent}} = 1/3$, which is strictly less than η_{steer}. Thus using the $d=2$ Werner states as an example we also see that steerability is strictly stronger than nonseparability. This clear distinction between the three classes can be seen on the left-hand axis of Fig. 2(a).

B. Isotropic states

The isotropic states, which were introduced in [17], can be parametrized identically to the Werner states; that is, in terms of their dimension d and a mixing parameter η,

$$W_d^\eta = (1-\eta)\mathbf{I}/d^2 + \eta \mathbf{P}_+. \quad (4.12)$$

Here $\mathbf{P}_+ = |\psi_+\rangle\langle\psi_+|$, where $|\psi_+\rangle = \Sigma_{i=1}^d |i\rangle|i\rangle/\sqrt{d}$ is a maximally entangled state. In fact, for $d=2$ it is straightforward to verify that the isotropic states are identical to Werner states up to local unitaries. Isotropic states are nonseparable iff $\eta > \eta_{\text{ent}} = 1/(d+1)$ [17].

A nontrivial upper bound on η_{Bell} for all d is known; in Ref. [18] it is shown that a Bell inequality is certainly violated by a d-dimensional isotropic state if

$$\eta > \frac{2}{I_d(\text{QM})} \geqslant \eta_{\text{Bell}}, \quad (4.13)$$

where $I_d(\text{QM})$ is defined as

$$I_d(\text{QM}) = 4d \sum_{k=0}^{[d/2]-1} \left(1 - \frac{2k}{d-1}\right)(q_k - q_{-(k+1)}), \quad (4.14)$$

and $q_k = 1/\{2d^3 \sin^2[\pi(k+1/4)/d]\}$. Collins *et al.* [18] go on to show that in the limit as $d \to \infty$ the limiting value this upper bound on η_{Bell} approaches $\pi^2/(16 \times \text{Catalan}) \approx 0.6734$, where Catalan ≈ 0.9159 is Catalan's constant.

In determining steerability we again allow Alice all possible measurement strategies: $\mathfrak{M}_\alpha = \mathfrak{D}_\alpha$, and take the projectors to be rank one: $\hat{\Pi}_a^A = |a\rangle\langle a|$. The isotropic states have the symmetry property that they are invariant under transformations of the form $\hat{U}^* \otimes \hat{U}$, hence the conditions of Lemma 1 are again satisfied for the d-dimensional unitary group $\mathfrak{U}(d)$. In this case, $g \to \hat{U}$ and $\hat{U}_{\alpha\beta}(g) \to \hat{U}^* \otimes \hat{U}$. Thus we can again take the optimal ensemble to be $F^\star = \{|\psi\rangle\langle\psi| d\mu_{\text{Haar}}(\psi)\}$.

Now consider the conditioned state that Bob would obtain if Alice were to make a measurement $\hat{A}$ on her half of W_d^η,

$$\tilde{\rho}_a^A = \text{Tr}_A[(\hat{\Pi}_a^A \otimes \mathbf{I}) W_d^\eta] = \left(\frac{1-\eta}{d}\right)\frac{\mathbf{I}}{d} + \frac{\eta}{d}|a\rangle\langle a|. \quad (4.15)$$

This is a state proportional to the completely mixed state *plus* a term proportional to the state Alice's system would be projected into by her measurement. Note the similarity with the Werner state example, where the conditioned state was proportional to the completely mixed state *minus* a term proportional to $|a\rangle\langle a|$. This difference arises because the isotropic states are symmetrically correlated rather than antisymmetrically correlated as in the Werner state example.

Again we wish to determine if it is possible for Alice to simulate the conditioned state $\tilde{\rho}_a^A$ using the optimal ensemble $F^\star$ and a cheating strategy defined by an optimal distribution $\wp^\star(a|\hat{A},\psi)$.

Imagine that in each run of a steering experiment Alice simply sends Bob a state $|\psi\rangle\langle\psi|$ drawn at random from $F^\star = \{|\psi\rangle\langle\psi| d\mu_G(\psi,m)\}$. When asked to perform a measurement $\hat{A}$ and announce her result, she uses $\wp^\star(a|\hat{A},\psi)$ to determine her answer. In testing whether this is actually what Alice

could be doing, we again only need to consider the quantity

$$\langle a|\tilde{\rho}_a^A|a\rangle = \frac{\eta}{d} + \frac{(1-\eta)}{d^2}. \tag{4.16}$$

In this case Alice's strategy is similar to the Werner state example, except now she wants to simulate the *maximum* possible overlap with $|a\rangle\langle a|$ (due to the form of $\tilde{\rho}_a^A$). Therefore, Bob will only concede that W_d^η is steerable if the maximum overlap with $|a\rangle\langle a|$ predicted using the ensemble $F^\star$ and the optimal cheating strategy $\wp^\star(a|A,\psi)$ is less than that predicted by Eq. (4.16). In this case there would be no possible classical strategy that Alice could possibly be using to simulate the correlations with Bob's results. Identical predictions for the overlap with $|a\rangle\langle a|$ will again occur precisely at the steering boundary η_{steer}, which occurs when $\wp^\star(a|\hat{A},\psi)$ is used.

The optimal $\wp(a|\hat{A},\psi)$ is defined in a similar manner to the Werner state example. However, in each run of the experiment Alice now determines which of the eigenstates of $\hat{A}$ is *closest* to $|\psi\rangle\langle\psi|$ and announces the eigenvalue associated with that eigenstate as her result. That is,

$$\wp^\star(a|\hat{A},\psi) = \begin{cases} 1 & \text{if } \langle\psi|\hat{\Pi}_a^A|\psi\rangle > \langle\psi|\hat{\Pi}_{a'}^A|\psi\rangle \quad \forall\, a' \neq a \\ 0 & \text{otherwise.} \end{cases} \tag{4.17}$$

To test if Eq. (3.5) holds, Alice and Bob would need to run the experiment many times and compare $\langle a|\tilde{\rho}_a^A|a\rangle$ with the quantity

$$\langle a|\int d\mu_{\text{Haar}}|\psi\rangle\langle\psi|\wp^\star(a|A,\psi)|a\rangle. \tag{4.18}$$

This can be written as

$$\langle a|\int_a d\mu_{\text{Haar}}(\psi)|\psi\rangle\langle\psi|a\rangle = \int_a d\mu_{\text{Haar}}(\psi)|\langle a|\psi\rangle|^2, \tag{4.19}$$

where the subscript a on the integral means that in the integral only those states with $|\langle a|\psi\rangle|$ greater than all others will contribute. As shown in Appendix B 1, a random state $|\psi\rangle$ from the ensemble $F^\star$ can be described by the unnormalized state

$$|\tilde{\psi}\rangle = m|\psi\rangle = \frac{1}{\sqrt{d}}\sum_{j=1}^{d} z_j|\phi_j\rangle, \tag{4.20}$$

where the z_j are mutually independent complex Gaussian random variables with zero mean and zero second moments except for $\langle z_j^* z_k\rangle = \delta_{jk}$. That is, we can replace the Haar measure $d\mu_{\text{Haar}}(\psi)$ by $d\mu_G(\psi,m) = d\mu_G(m)d\mu_{\text{Haar}}(\psi)$. In terms of the variables $\{z_j\}$, this can be expressed as

$$d\mu_G(\psi,m) \to \pi^{-d}\exp\left(-\sum_{i=1}^{d} z_i^2\right)d^2z_1\cdots d^2z_d. \tag{4.21}$$

Now using the Gaussian measure $d\mu_G(\psi,m)$ to describe the ensemble $F^\star$, we can rewrite Eq. (4.19) as

$$\int_a d\mu_{\text{Haar}}(\psi)|\langle a|\psi\rangle|^2 = \int_a d\mu_{\text{Haar}}(\psi)|\langle a|\psi\rangle|^2 \frac{\int d\mu_G(m)m^2}{\int d\mu_G(m)m^2}$$

$$= \frac{\int_a d\mu_G(\psi,m)|\langle a|\tilde{\psi}\rangle|^2}{\int d\mu_G(m)m^2}. \tag{4.22}$$

It is straightforward to show that the denominator equals one (see Appendix B 2), and hence we can evaluate the numerator (left to Appendix B 3) to find that

$$\int_a d\mu_G(\psi,m)|\langle a|\tilde{\psi}\rangle|^2 = \frac{H_d}{d^2}, \tag{4.23}$$

where $H_d = 1+1/2+1/3+\cdots+1/d$ is the harmonic series.

Thus we find that for any positive normalized distribution $\wp(a|\hat{A},\psi)$ we must have

$$\langle a|\int d\mu_{\text{Haar}}(\psi)|\psi\rangle\langle\psi|\wp(a|\hat{A},\psi)|a\rangle \leqslant \frac{H_d}{d^2}, \tag{4.24}$$

with the equality obtained for the optimal $\wp^\star(a|\hat{A},\psi)$ as defined in Eq. (4.17). Comparing this with Eq. (4.16) we see that steering can be demonstrated iff

$$\frac{\eta}{d} + \frac{(1-\eta)}{d^2} > \frac{H_d}{d^2}. \tag{4.25}$$

Thus for isotropic states

$$\eta_{\text{steer}} = \frac{H_d - 1}{d-1} \underset{\text{large } d}{\sim} \frac{\ln(d)}{d}. \tag{4.26}$$

For $d=2$ the isotropic states are equivalent (up to local unitaries) to the Werner states, and we again find that $\eta_{\text{steer}} = 1/2$, which is strictly less than η_{Bell} and strictly greater than η_{ent}. For $d>2$, η_{steer} is greater than η_{ent} and significantly less than an upper bound on η_{Bell}. This is shown in Fig. 2(b). For large d we see that both η_{steer} and η_{ent} tend to zero, however, η_{steer} approaches zero more slowly; it is larger than η_{ent} by a factor of $\ln(d)$ [19].

C. Inept states

We now consider a family of states with less symmetry than the previous examples. This makes the analysis more difficult, meaning that we cannot find η_{steer} exactly. However, making use of the symmetry properties of the states allows us to find an upper bound on η_{steer}. We define a family of two-qubit states by

$$W_\epsilon^\eta = \eta|\psi\rangle\langle\psi| + (1-\eta)\rho_\alpha \otimes \rho_\beta, \tag{4.27}$$

where

$$|\psi\rangle = \sqrt{1-\epsilon}|0_\alpha 0_\beta\rangle + \sqrt{\epsilon}|1_\alpha 1_\beta\rangle, \tag{4.28}$$

and the reduced states $\rho_{\alpha(\beta)}$ are found by partial tracing with respect to Bob (Alice). That is,

$$\rho_{\alpha(\beta)} = \mathrm{Tr}_{\beta(\alpha)}[|\psi\rangle\langle\psi|]. \tag{4.29}$$

As in the previous examples, this is a two-parameter family of states; the parameter η is again a mixing parameter, and the parameter ϵ determines how much entanglement is present in the state $|\psi\rangle$. Note that when $\epsilon=1/2$ these states are equivalent to the two-dimensional Werner and isotropic states.

This family of states was studied in Ref. [13] in the context of distributing entanglement. The authors considered an inept company attempting to distribute pure entangled states $|\psi_\epsilon\rangle$ to many pairs of parties. However, they mixed up the addresses some fraction $1-\eta$ of the time, meaning that on average the company would actually distribute mixed entangled states of the form of Eq. (4.27). Hence we will refer to this family of states as "inept" states.

As noted above, the inept states are a family of two-qubit states, which means that it is possible to evaluate η_{ent} analytically. This was done in Ref. [13] leading to the following condition for nonseparability of inept states:

$$\eta > \eta_{\mathrm{ent}} = \frac{\epsilon(1-\epsilon)}{\epsilon(1-\epsilon)+\sqrt{\epsilon(1-\epsilon)}}. \tag{4.30}$$

Reference [13] also considers Bell nonlocality of the state matrix W_ϵ^η by testing if a violation of the CHSH inequality [20] occurs. This was done using the method of [15] for determining the optimal violation of the CHSH inequality for two-qubit states. One finds that the state W_ϵ^η violates the CHSH inequality if and only if

$$\eta > \frac{4\epsilon^2-4\epsilon+1-\sqrt{4\epsilon^2-4\epsilon+3}}{4\epsilon^2-4\epsilon-1} \geqslant \eta_{\mathrm{Bell}}. \tag{4.31}$$

Now, in order to demonstrate steering we must specify a measurement strategy. In the two previous examples we have used the complete set of projective measurements, $\mathfrak{M}_\alpha=\mathfrak{D}$. This would be a suitable measurement strategy to allow us to define an optimal ensemble, however, in order to make our task simpler we will consider a more restricted set of measurements. We note that states defined by Eq. (4.27) have the symmetry property that they are invariant under simultaneous contrary rotations about the z axes. This immediately suggests a restricted measurement scheme; we allow all measurements in the xy plane but only allow a single measurement along the z axis. That is, Alice's measurement scheme is given by $\mathfrak{M}_\alpha=\{\hat{\sigma}_z\}\cup\{\hat{\sigma}_\theta: \theta\in[0,2\pi)\}$, where

$$\hat{\sigma}_\theta = \hat{\sigma}_x\cos(\theta)+\hat{\sigma}_y\sin(\theta). \tag{4.32}$$

In this case the conditions for Lemma 1 are satisfied for the Lie group G generated by $(1/2)\hat{\sigma}_z\otimes\mathbf{I}-(1/2)\mathbf{I}\otimes\hat{\sigma}_z$ (see Appendix C 1).

This is a more restricted scheme than we have considered so far, but will be sufficient to demonstrate steerability if Eq. (3.5) does not hold (since it must hold for *all* measurements to preclude steering). Thus we are only considering an upper bound on η_{steer} (the boundary between steerable and nonsteerable states using all projective measurements).

We now consider the optimal ensemble for this restricted set of measurements. We use an ensemble of pure states $F=\{|\psi\rangle\langle\psi|d\mu(\psi)\}$, where

$$|\psi\rangle\langle\psi| = \frac{1}{2}[\mathbf{I}+\sqrt{1-z^2}\cos(\phi)\hat{\sigma}_x+\sqrt{1-z^2}\sin(\phi)\hat{\sigma}_y - z\hat{\sigma}_z], \tag{4.33}$$

and $d\mu(\psi)=(d\phi/2\pi)\wp(z)dz$. It is straightforward to show that this ensemble is of the form of the optimal ensemble since the conditions for Lemma 1 hold (see Appendix C 1). While this ensemble has the form of the optimal, it is not completely specified as $\wp(z)$ is still general. Thus to find the optimal ensemble we need to determine the optimal probability distribution $\wp(z)$.

First consider the reduced states that Bob would obtain if Alice really were to measure $\hat{\sigma}_z$ on her half of W_ϵ^η. If she did so, and obtained the +1 result then Bob's state would be given by

$$\tilde{\rho}_{+\hat{1}}^{\sigma_z} = (\epsilon)\frac{1}{2}[\mathbf{I}-z_+\hat{\sigma}_z]. \tag{4.34}$$

Similarly, for the −1 result, Bob would obtain

$$\tilde{\rho}_{-\hat{1}}^{\sigma_z} = (1-\epsilon)\frac{1}{2}[\mathbf{I}-z_-\hat{\sigma}_z], \tag{4.35}$$

where the constants z_+ and z_- are defined as

$$z_+ = 1-2\eta-2\epsilon(1-\eta),$$

$$z_- = 1-2\epsilon(1-\eta). \tag{4.36}$$

Now we wish to determine if Alice could simulate these conditioned states using the ensemble F and a suitable strategy $\wp(\pm 1|\hat{\sigma}_z,(z,\phi))$. Due to the form of $\tilde{\rho}_{\pm\hat{1}}^{\sigma_z}$ the best strategy for Alice is to split the ensemble F into two subensembles, one to simulate $\tilde{\rho}_{+\hat{1}}^{\sigma_z}$ and the other to simulate $\tilde{\rho}_{-\hat{1}}^{\sigma_z}$. Thus we can separate $\wp(z)$ into two positive distributions

$$\wp(z) = \wp_+(z)+\wp_-(z). \tag{4.37}$$

We imagine that Alice will attempt to simulate measuring $\hat{\sigma}_z$ by randomly generating states $|\psi\rangle\langle\psi|$ using the distribution $\wp(z)$ and sending them to Bob. If in a particular run of the experiment the state she sent Bob was from the subensemble determined by $\wp_+(z)$ then she will announce the result +1. Similarly, if she sent Bob a state from $\wp_-(z)$ then she will announce −1.

Now if Alice uses this strategy, Bob will find on average that

$$\tilde{\rho}_{\pm\hat{1}}^{\sigma_z} = \frac{1}{2}\left[\mathbf{I}\int_{-1}^{+1}dz\wp_\pm(z)-\hat{\sigma}_z\int_{-1}^{+1}dz\wp_\pm(z)z\right]. \tag{4.38}$$

Comparing with Eqs. (4.34) and (4.35) we find that in order for the ensemble F to be able to simulate Alice measuring $\hat{\sigma}_z$ we have the following constraints on $\wp(z)$:

$$\int_{-1}^{1} dz\wp_+(z) = \epsilon, \tag{4.39}$$

$$\int_{-1}^{1} dz\wp_-(z) = 1-\epsilon, \tag{4.40}$$

$$\int_{-1}^{1} dz\wp_+(z)z = \epsilon z_+, \tag{4.41}$$

$$\int_{-1}^{1} dz\wp_-(z)z = (1-\epsilon)z_-. \tag{4.42}$$

Now consider the following conditioned states that Bob would obtain if Alice were to measure $\hat{\sigma}_\theta$:

$$\tilde{\rho}_{\pm 1}^{\sigma_\theta} = \frac{1}{2}[\mathbf{I} \pm \eta\sqrt{\epsilon(1-\epsilon)}\cos(\theta)\hat{\sigma}_x \pm \eta\sqrt{\epsilon(1-\epsilon)}\sin(\theta)\hat{\sigma}_y - (1-2\epsilon)\hat{\sigma}_z]. \tag{4.43}$$

How well could Alice simulate the above state using the ensemble $F^\star$ and a cheating strategy defined by $\wp(\pm 1|\hat{\sigma}_\theta,(z,\phi))$? We know that the ensemble $F^\star$ is symmetric under rotations about the z axis. So in this case Alice would use her knowledge of ϕ to determine the outcome to announce when asked to measure $\hat{\sigma}_\theta$. That is, if the state $|\psi\rangle\langle\psi|$ that she sent Bob is closer to the positive axis defined by $\hat{\sigma}_\theta$ then she will announce the +1 result. Similarly, if $|\psi\rangle\langle\psi|$ is closer to the negative measurement axis then she announces −1. This corresponds to

$$\wp(\pm 1|\hat{\sigma}_\theta,(z,\phi)) = \begin{cases} 1 & \text{if } \phi \in \left[\theta \mp \frac{\pi}{2}, \theta \pm \frac{\pi}{2}\right) \\ 0 & \text{if } \phi \in \left[\theta \pm \frac{\pi}{2}, \theta \mp \frac{\pi}{2}\right). \end{cases} \tag{4.44}$$

From the symmetry under rotations about the z axis we can see that Alice will be able to do equally well using this strategy to simulate states prepared by any measurement $\hat{\sigma}_\theta$ in the xy plane. Thus without loss of generality we set $\theta=0$ and consider the specific case where Alice allegedly measures $\hat{\sigma}_x$. Under these conditions Eq. (4.43) reduces to

$$\tilde{\rho}_{\pm 1}^{\sigma_x} = \frac{1}{2}[\mathbf{I} \pm \eta\sqrt{\epsilon(1-\epsilon)}\hat{\sigma}_x - (1-2\epsilon)\hat{\sigma}_z]. \tag{4.45}$$

If Alice randomly sends Bob states from $F^\star$ and uses Eq. (4.44) to determine her responses, Bob will find on average the state

$$\frac{1}{2}\left[\mathbf{I} \pm \frac{1}{\pi}\int_{-1}^{+1} dz\sqrt{1-z^2}\wp(z)\hat{\sigma}_x - \int_{-1}^{+1} dz\wp(z)z\hat{\sigma}_z\right]. \tag{4.46}$$

We know that when F is optimal, Eq. (4.46) will exactly simulate Eq. (4.45). In determining the optimal F we must find the optimal $\wp(z)$, however, we are constrained in determining $\wp(z)$ by the fact that the ensemble must also simulate the states that Bob would obtain if Alice were to measure $\hat{\sigma}_z$. These constraints are enforced by Eqs. (4.39)–(4.42).

Note that Eqs. (4.39) and (4.40) ensure that the $\hat{\sigma}_z$ term in Eq. (4.46) and Eq. (4.45) will be the same. Therefore, to determine how well Alice's strategy can simulate Eq. (4.45) we only need to consider the coefficient of the $\hat{\sigma}_x$ term. If the coefficient of this term predicted by Eq. (4.46) is as large as in Eq. (4.45) then Alice's strategy simulates Bob's conditioned state perfectly. Thus Bob would not believe that the state W_ϵ^η is genuinely steerable. Hence we need to find the distribution $\wp^\star(z)$, which maximizes the $\hat{\sigma}_x$ coefficient in Eq. (4.46) to determine if steering is possible. That is, we wish to find the $\wp(z)$ that gives the maximum value of $(1/\pi)\int_{-1}^{+1}dz\sqrt{1-z^2}\wp(z)$. This is equivalent to maximizing

$$\frac{1}{\pi}\int_{-1}^{+1} dz\sqrt{1-z^2}[\wp_+(z)+\wp_-(z)], \tag{4.47}$$

subject to the constraints given by Eqs. (4.39)–(4.42).

Writing $\wp_\pm(z)=f_\pm^2(z)$ for real functions $f_\pm(z)$ we can use Lagrange multiplier techniques to perform the optimization. We find that the optimal $\wp(z)$ has the unsurprising form

$$\wp^\star(z) = \epsilon\delta(z-z_+) + (1-\epsilon)\delta(z-z_-), \tag{4.48}$$

where the constants $z_\pm$ are defined in Eq. (4.36) and $\delta(z-z')$ is the Dirac delta function. We see now why the choice of splitting the ensemble into two distributions was the best choice for Alice. The optimal ensemble $F^\star$ is composed of pure states in two rings around the z axis of the Bloch sphere; one in the $+z$ hemisphere defined by z_+, which on average may be used to simulate $\tilde{\rho}_{+1}^{\sigma_{\tilde{z}}}$, and the other in the $-z$ hemisphere defined by z_-, which may simulate $\tilde{\rho}_{+1}^{\sigma_{\tilde{z}}}$. (These comments apply to the case $0\leqslant\epsilon\leqslant 1/2$.)

Using $\wp^\star(z)$ to evaluate Eq. (4.46) we find that

$$\tilde{\rho}_{\pm 1}^{\sigma_x} = \frac{1}{2}\left[\mathbf{I} \pm \frac{1}{\pi}\{\epsilon\sqrt{1-z_+^2} + (1-\epsilon)\sqrt{1-z_-^2}\}\hat{\sigma}_x - (1-2\epsilon)\hat{\sigma}_z\right]. \tag{4.49}$$

Finally, comparing this with $\tilde{\rho}_{\pm 1}^{\sigma_x}$ given by Eq. (4.45) we find that Alice's optimal cheating strategy fails to simulate measurements of $\hat{\sigma}_x$ when

$$\pi\eta\sqrt{\epsilon(1-\epsilon)} - (1-\epsilon)\sqrt{1-[1-2\epsilon(1-\eta)]^2} - \epsilon\sqrt{1-[1-2\eta-2\epsilon(1-\eta)]^2} > 0. \tag{4.50}$$

Thus under these conditions we know that steering is possible using the measurement scheme $\mathfrak{M}_\alpha$. Note that we have not determined η_{steer} as we have not considered all possible projective measurements. However, we can make Eq. (4.50) an equality to provide an equation for η, which is an *upper bound* on η_{steer}. This boundary is plotted in Fig. 2(c).

For $\epsilon=1/2$ we know explicitly that

$$\eta_{\text{Bell}} > \eta_{\text{steer}} > \eta_{\text{ent}} \tag{4.51}$$

[since these states are equivalent to the $d=2$ Werner states (see Appendix C 2)]. This special case yields the isolated points at $\epsilon=1/2$ in Fig. 2(c). For the remaining range of ϵ we

find that our upper bound on η_{steer} is significantly lower than the upper bound on η_{Bell} and significantly higher than η_{ent}. This fact, taken with the known boundary values for $\epsilon = 1/2$ gives us good reason to conjecture that the three boundaries are strictly distinct for all $\epsilon \in [0,1]$.

D. Gaussian states

Finally, we investigate a general (multimode) bipartite Gaussian state W [21]. The mode operators are defined as $\hat{q}_i = \hat{a}_i + \hat{a}_i^\dagger$ and $\hat{p}_i = -i(\hat{a}_i - \hat{a}_i^\dagger)$ for the position and momentum, respectively. Here $\hat{a}_i$ and $\hat{a}_i^\dagger$ are the annihilation and creation operators for the ith mode. For an n-mode state one may define a vector $\hat{R} = (\hat{q}_1, \hat{p}_1, \ldots, \hat{q}_n, \hat{p}_n)$, which allows the commutation relations for the mode operators to be compactly expressed as

$$[R_i, R_j] = 2i\Sigma_{\alpha\beta}^{ij}. \tag{4.52}$$

Here $\Sigma_{\alpha\beta}^{ij}$ are matrix elements of the symplectic matrix $\Sigma_{\alpha\beta} = \oplus_{i=1}^n J_i$, where

$$J_i = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \tag{4.53}$$

A Gaussian state is defined by the mean of the vector of phase-space variables $\hat{R}$, as well as the covariance matrix (CM) $V_{\alpha\beta}$ for these variables. The mean vector can be arbitrarily altered by local unitary operations and hence cannot determine the entanglement properties of W. Thus for our purposes a Gaussian state is characterized by the CM. In (Alice, Bob) block form it appears as

$$\text{CM}[W] = V_{\alpha\beta} = \begin{pmatrix} V_\alpha & C \\ C^\mathsf{T} & V_\beta \end{pmatrix}. \tag{4.54}$$

This represents a valid state if the linear matrix inequality (LMI)

$$V_{\alpha\beta} + i\Sigma_{\alpha\beta} \geqslant 0 \tag{4.55}$$

is satisfied [21].

Rather than addressing steerability in general, we consider the case where Alice can only make Gaussian measurements [21,22], the set of which will be denoted by $\mathfrak{G}_\alpha$. Thus, as for the previous section, since we are considering a restricted class of measurements, if we demonstrate steerability with this measurement scheme it will provide an *upper bound* on η_{steer}.

A measurement $A \in \mathfrak{G}_\alpha$ is described by a Gaussian positive operator with a CM T^A satisfying $T^A + i\Sigma_\alpha \geqslant 0$ [21,22]. When Alice makes such a measurement, Bob's conditioned state ρ_a^A is Gaussian with a CM [23],

$$\text{CM}[\rho_a^A] = V_\beta^A = V_\beta - C^\mathsf{T}(V_\alpha + T^A)^{-1}C, \tag{4.56}$$

which is actually independent of Alice's outcome a.

Our goal is to determine a sufficient condition for steerability of Gaussian states. We do this by determining the necessary and sufficient condition for steerability with Gaussian measurements. In the previous examples after specifying a measurement scheme we considered Bob's conditioned state if Alice were to perform a measurement and determined when it was possible for this state to be simulated by a cheating strategy. In the following we are working toward the same goal. If Alice were to perform a Gaussian measurement on half of the state W and send the other part to Bob, then Bob's conditioned state would have a covariance matrix defined by Eq. (4.56); however, this is independent of Alice's result a. Thus we do not need to consider a strategy for Alice to announce correctly correlated results to Bob. We simply need to determine when Alice could simulate Bob's conditioned state by sending Bob states from a pure state ensemble (rather than actually sending part of W). We will show that there exists an optimal ensemble of Gaussian states distinguished by their mean vectors (but sharing the same covariance matrix, which we will label U) which Alice could use for this task. If there exists a valid ensemble of Gaussian states defined by U which can simulate V_β^A then Bob will not believe that W is entangled, and hence the state is not steerable.

Before moving to the presentation of our main result consider the following result from linear algebra theory relating to Schur complements of block matrices. The Schur complements of P and Q in a general block matrix

$$B = \begin{pmatrix} P & R \\ R^\mathsf{T} & Q \end{pmatrix} \tag{4.57}$$

are defined as $\Delta_P = Q - R^\mathsf{T}P^{-1}R$ and $\Delta_Q = P - RQ^{-1}R^\mathsf{T}$, respectively. The matrix B is positive semidefinite (PSD), iff both P and its Schur complement are PSD (and likewise for Q and its Schur complement).

The proof of our main theorem is based on the following inequality:

$$V_{\alpha\beta} + \mathbf{0}_\alpha \oplus i\Sigma_\beta \geqslant 0, \tag{4.58}$$

and relies on the following facts:

Lemma 2. If Eq. (4.58) is true then there exists an ensemble defined by covariance matrix U such that

$$U + i\Sigma_\beta \geqslant 0, \tag{4.59}$$

$$V_\beta^A - U \geqslant 0, \tag{4.60}$$

which implies that the state W is not steerable.

Proof. See Appendix D 1 for proof of this lemma.

Lemma 3. If the Gaussian state W defined in Eq. (4.54) is *not* steerable by Alice's Gaussian measurements then there exists a Gaussian ensemble defined by covariance matrix U such that Eqs. (4.59) and (4.60) hold.

Proof. See Appendix D 2 for proof of this lemma.

Lemma 4. If $\forall A \in \mathfrak{G}$ there exists U such that Eqs. (4.59) and (4.60) hold, then

$$V_\alpha + T^A - C(V_\beta + i\Sigma_\beta)^{-1}C^\mathsf{T} \geqslant 0, \tag{4.61}$$

must also hold.

Proof. See Appendix D 3 for proof of this lemma.

We are now in a position to present our main theorem.

Theorem 5. The Gaussian state W defined in Eq. (4.54) is *not* steerable by Alice's Gaussian measurements iff Eq. (4.58) is true.

Proof. By Lemma 2 we know that if Eq. (4.58) is true then W is not steerable.

Now suppose that Eq. (4.58) does not hold, so that

$$V_{\alpha\beta} + 0_\alpha \oplus i\Sigma_\beta \ngeq 0. \tag{4.62}$$

Since we know that $V_\beta + i\Sigma_\beta \geqslant 0$ (that is, V_β is a valid covariance matrix), the Schur complement in this term cannot be PSD (if it were it would imply that Eq. (4.58) were true when we have assumed the opposite). Thus we have

$$G = \Delta_{V_\beta + i\Sigma_\beta} = V_\alpha - C(V_\beta + i\Sigma_\beta)^{-1}C^{\mathsf{T}} \ngeq 0. \tag{4.63}$$

Consider an eigenvector of G, $\vec{\nu}$, associated with a negative eigenvalue, that is, $G\vec{\nu} = -g\vec{\nu}$ where $g > 0$. We can choose a measurement A along an axis such that T^A shares the eigenvector $\vec{\nu}$ so that $T^A\vec{\nu} = t\vec{\nu}$. Now it is possible to arrange the measurement such that $t < g$. This is because it is always possible to make one eigenvalue suitably small (the eigenvalue for the conjugate variable will become large). Now we have chosen the measurement such that $G + T^A$ must have a negative eigenvalue in the $\vec{\nu}$ direction. Hence, for this choice of measurement we have

$$V_\alpha + T^A - C(V_\beta + i\Sigma_\beta)^{-1}C^{\mathsf{T}} \ngeq 0. \tag{4.64}$$

This shows that if Eq. (4.58) does not hold then there exists a measurement A such that Eq. (4.61) does not hold, which by Lemma 4 implies that for this measurement there does not exist an ensemble defined by U such that Eqs. (4.59) and (4.60) hold.

However, from Lemma 3 we know that if W is not steerable then there must exist an ensemble defined by U such that these equations hold. Since they do not hold for all A when Eq. (4.58) is not true, we see that if Eq. (4.58) is not true then we cannot define a suitable ensemble U to prevent steering. Therefore, a Gaussian state W is *not* steerable iff Eq. (4.58) is true. ■

Theorem 5 provides a sufficient condition for demonstrating that the state W is steerable (by any measurements), and hence specifies an *upper* bound on η_{steer}. To illustrate this, it is useful to consider a simple example.

1. Two-mode states and the EPR paradox

We now consider the simplest case where Alice and Bob share a Gaussian state W in which they each have a single mode. It is well known that such a Gaussian state can be brought into standard form using local linear unitary Bogoliubov operations (LLUBOs), so that the CM takes the form [24]

$$V_{\alpha\beta} = \begin{pmatrix} n & 0 & c & 0 \\ 0 & n & 0 & c' \\ c & 0 & m & 0 \\ 0 & c' & 0 & m \end{pmatrix}, \tag{4.65}$$

where $n, m \geqslant 1$.

The Peres-Horodecki criterion for separability can be written as a linear matrix inequality for Gaussian states [25] as

$$\tilde{V}_{\alpha\beta} + i\Sigma_{\alpha\beta} \geqslant 0, \tag{4.66}$$

where $\tilde{V}_{\alpha\beta} = \Lambda V_{\alpha\beta}\Lambda$; $\Lambda = \text{diag}(1,1,1,-1)$. This can be determined by finding when the Schur complement (of the lower block) of $\tilde{V}_{\alpha\beta} + i\Sigma_{\alpha\beta}$ is PSD, which occurs only when

$$\left(m - \frac{c^2 n}{n^2 - 1}\right)\left(m - \frac{c'^2 n}{n^2 - 1}\right) \geqslant \left(1 - \frac{cc'}{n^2 - 1}\right)^2. \tag{4.67}$$

Hence two-mode Gaussian states defined by $V_{\alpha\beta}$ are separable iff Eq. (4.67) is satisfied.

For Gaussian states, which have a positive Wigner function, it is not possible to demonstrate violation of a Bell inequality using Gaussian measurements. This is because the Wigner function gives an explicit hidden variable description, which ensures satisfaction of Bell's inequality.

To determine if the state W is steerable it is a simple matter of testing if $V^{\alpha\beta} + \mathbf{0}_\alpha \oplus i\Sigma_\beta$ is PSD. Again using Schur complements, we find that this is the case iff

$$\left(m - \frac{c^2}{n}\right)\left(m - \frac{c'^2}{n}\right) \geqslant 1. \tag{4.68}$$

Recall that the interest in, and even the name, *steering*, arose in response to the EPR paradox. Therefore, one would expect that any reasonable characterization of steering should include the EPR paradox. This is indeed the case for our formulation of steering. For the class of two-mode Gaussian states that we have been considering, Reid [12] has argued that the EPR "paradox" is demonstrated if the product of the conditional variances $V(q_\beta|q_\alpha)$ and $V(p_\beta|p_\alpha)$ violates the uncertainty principle. This is the case if the conditional variances do *not* satisfy

$$V(q_\beta|q_\alpha)V(p_\beta|p_\alpha) \geqslant 1. \tag{4.69}$$

For a general two-mode Gaussian state W the conditional variances take the form

$$V(q_\beta|q_\alpha) = \min_\mu [(q_\beta - \mu q_\alpha)^2] = m - \frac{c^2}{n}, \tag{4.70}$$

and similarly for $V(p_\alpha|p_\beta)$. Thus Eq. (4.69) is exactly Eq. (4.68). That is, the EPR "paradox" occurs precisely when W is steerable with Gaussian measurements. This example confirms that the EPR "paradox" is merely a particular case of steering. As is well known [26], Reid's EPR condition is strictly stronger than the condition for nonseparability Eq. (4.67). The fact that the EPR "paradox" is an example of steering explains why the EPR condition is stronger than nonseparability; as we have shown in previous examples steering is a strictly stronger concept than nonseparability.

2. Symmetric two-mode states

Finally, we consider the specific case of two-mode Gaussian states prepared by optical parametric amplifiers [26]. When the entanglement is symmetric between the two modes

the covariance matrix describing such states has a particularly simple form. The continuous variable entanglement properties of such a state has recently been characterized experimentally [26]. In this case the covariance matrix of the state W has just two parameters, η and $\bar{n}$.

$$\mathrm{CM}[W_{\bar{n}}^{\eta}] = V_2^{\alpha\beta} = \begin{pmatrix} \gamma & 0 & \delta & 0 \\ 0 & \gamma & 0 & -\delta \\ \delta & 0 & \gamma & 0 \\ 0 & -\delta & 0 & \gamma \end{pmatrix}, \tag{4.71}$$

where $\gamma = 1+2\bar{n}$ and $\delta = 2\eta\sqrt{\bar{n}(1+\bar{n})}$. Here $\bar{n}$ is the mean photon number for each party, and η is a mixing parameter defined analogously with the other examples except that here it is the covariance matrix that is linear in η, not the state matrix.

For such symmetric states the separability condition, Eq. (4.67), becomes

$$\left(\gamma - \frac{\delta^2\gamma}{\gamma^2-1}\right)^2 \geqslant \left(1 + \frac{\delta^2}{\gamma^2-1}\right)^2. \tag{4.72}$$

Substituting for the values of γ and δ we find that the condition for states defined by Eq. (4.71) to be *nonseparable* is simply

$$\eta > \eta_{\mathrm{ent}} = \sqrt{\frac{\bar{n}}{1+\bar{n}}}. \tag{4.73}$$

In determining when symmetric two-mode Gaussian states are steerable, we find that Eq. (4.68) becomes

$$\left(\gamma - \frac{\delta^2}{\gamma}\right)^2 \geqslant 1. \tag{4.74}$$

Hence, as an upper bound on the condition for steerability we have

$$\eta > \sqrt{\frac{1+2\bar{n}}{2(1+\bar{n})}} \geqslant \eta_{\mathrm{steer}}. \tag{4.75}$$

This is an upper bound as we have only considered a restricted class of measurements. These results are plotted for some small values of $\bar{n}$ in Fig. 2(d). Since it is not possible to demonstrate Bell nonlocality for a Gaussian state with Gaussian measurements we have also plotted an upper bound on η_{Bell} in Fig. 2(d).

V. DISCUSSION

We have introduced a rigorous formulation of the concept of steering and given a number of examples to demonstrate where this concept fits in the hierarchy of entangled states. Both our operational and mathematical formulations of steering leads to the notion that steerable states lie between nonseparable states and those entangled states which violate a Bell inequality. In particular, our example for 2×2 Werner states establish that this is a *strict* hierarchy. Our other examples are consistent with this fact.

Recently there has been renewed interest in classifying the resources present in quantum states. For instance, it has been proposed that nonlocality itself is a separate resource from entanglement (see [27], and references therein). This has been motivated by the fact that for suggested measures of nonlocality, the maximally nonlocal states are not necessarily maximally entangled states. Our work provides an interesting addition to the increasingly complex task of characterizing quantum resources. Clearly steerability is another form of nonlocality that a quantum state may possess.

The nonlocality of entangled states has also recently been studied in the context of robustness to noise. Reference [19] determines the maximum amount of noise that an arbitrary bipartite state can accept before its nonlocal correlations (i.e., its ability to violate a Bell inequality) are completely "washed out." They do this by determining when the resulting state's correlations can be explained by a "local model." In fact, the local models defined in Ref. [19] correspond to LHS models for Bob in our terminology. That is, as they recognize [19], the concept of steering is useful for proving new bounds for Bell nonlocality, since the latter is strictly stronger.

The inherent asymmetry in the definition of steerability may suggest applications for asymmetric entangled states. It may appear that a link exists between the recently proposed asymmetric measures of entanglement [28] and steerability. While conceptually appealing, this seems unlikely as states with asymmetric entanglement as defined in Ref. [28] necessarily contain bound entanglement. A connection between steerable states and bound entangled states is unlikely, as we have shown that steerable states exist for $d=2$ (and no bound entangled states exist for $d=2$).

There remain a number of open questions relating to steerability. We have demonstrated the link between the EPR paradox and steerability for two-mode Gaussian states. The EPR paradox has been demonstrated experimentally for Gaussian states, however, it is difficult to prepare an EPR-type experiment for other quantum states. This raises the question: might tests of steerability provide experimental evidence for EPR-type correlations in nonoptical experimental implementations?

From an experimental perspective, the question as to whether it is possible to define steerability witnesses or operators (in analogy with entanglement witnesses and Bell operators) is particularly appealing. This would provide a straightforward experimental test for determining if a given state is steerable. Such a test would simultaneously demonstrate that the given state is entangled. This will be addressed in future work.

Finally, our operational definition of steering in terms of a task involving exchanges of quantum systems with an untrusted partner is reminiscent of the scenarios common in quantum complexity theory such as interactive proof systems and other kinds of quantum games (see, for example, [29]). We do not know of a direct way of mapping steering as we have defined it here onto these problems but it is interesting to ask if steering may play a role in some way comparable to Bell-inequality violation in [30], for example. Secondly, is it possible to define some useful quantum protocol for which the class of steerable states is useful; that is, is there a task for which nonseparable states are an insufficient resource, but steerable states allow the protocol to be implemented?

We conclude by commenting that we expect the answers to these questions (and others) to prove steering a useful concept in the context of quantum information science.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council and the State of Queensland. We would like to thank Rob Spekkens, Volkher Scholz, Antonio Acìn, and Michael Hall for useful discussions.

APPENDIX A: EPR CORRELATIONS, ENTANGLEMENT, AND STEERING: A HISTORY OF TERMS

As stated in the Introduction, although Werner's 1989 paper [6] is often cited as introducing the dichotomy of entangled versus separable states, it is important (for the discussion in this Appendix) to note that he used neither the term entangled nor the term separable. These seem to have not been used in their presently accepted sense until 1996 by Bennett *et al.* [31], and Peres [32], respectively. Rather, he used the terms "EPR-correlated states" versus "classically correlated states." His main result, restated in these terms, was that some "EPR-correlated states" conform with Bell's concept of "locality."

Our recent work [10] also considered the issue of mixed states and EPR correlations. Specifically, we rigorously defined the class of states that can be used to demonstrate the nonlocal effect which EPR identified in 1935. Contrary to Werner's *terminology*, we established that the set of such EPR-correlated states is *not* by definition complementary to the set of locally preparable states. Using this concept of EPR-correlated states, the main results of our paper can, ironically, be expressed entirely in statements contradicting Werner's *natural-language descriptions* of his results. First, it is true (as Werner states) that some EPR-correlated states respect Bell locality, but, contrary to his (natural-language) claims, Werner did *not* prove this. Our proof [10] of this fact makes use of Werner's result, but also requires the much more recent result of Acìn *et al.* [16]. Second, what Werner's result *actually* proves, contrary to his stated dichotomy, is that some states that are not separable (classically correlated) are nevertheless *not* EPR correlated (from which one can conclude also that they are Bell local). To summarize, we used Werner's result to help prove that the set of Bell-nonlocal states is a *strict* subset of the set of EPR-correlated states, which in turn is a *strict* subset of the set of nonseparable states.

We emphasize that we are not disputing at all the mathematical validity of Werner's result, nor its importance, nor his understanding of it. We dispute only his use of the term "EPR-correlated states" to refer to nonseparable states, which he says "is to emphasize the crucial role of such states in the Einstein-Podolsky-Rosen paradox and for the violations of Bell's inequalities." This explanation for the name could equally be used to justify calling nonseparable states "Bell-correlated states," but that would be nonsensical since the point of Werner's paper is that, in the mixed-state case, not all nonseparable states can exhibit correlations that violate a Bell's inequality. Similarly, we maintain that if the term "EPR-correlated states" were to be applied to mixed states, then it should be reserved for those states for which the correlations can actually be used to demonstrate the EPR paradox. Prior to our paper, no rigorous and general definition of this paradox had been given, and so no good definition of "EPR-correlated states" existed. Giving such a definition is no mere semantic exercise; as stated in the preceding paragraph, our work identifies this as a new class of quantum states, distinct both from the Bell-nonlocal ones and the nonseparable ones.

During the past decade Schrödinger's term "entangled states" has replaced Werner's term "EPR-correlated states" (which he credited to Primas [33]) as a synonym for nonseparable states. Nevertheless, there is still potential for confusion if we were to promote the term "EPR-correlated" for the new class of states we defined and categorized in Ref. [10]. For that reason we proposed instead the term "steerable" for this class of states, a term that has been used increasingly in recent years [34–38].

APPENDIX B: ISOTROPIC STATE STEERING

1. Optimal ensemble

First choose an orthonormal basis $|1\rangle, |2\rangle, \ldots, |d\rangle$ to describe the uniform ensemble $F^\star$. Then consider randomly generated unnormalized states

$$|\tilde{\psi}\rangle = \frac{1}{\sqrt{d}}\sum_{j=1}^{d} z_j|\phi_j\rangle, \tag{B1}$$

where z_j are zero-mean Gaussian random variables with the properties $\langle z_j^* z_k\rangle = \delta_{j,k}$ and $\langle z_j z_k\rangle = 0$. Writing $|\tilde{\psi}\rangle = m|\psi\rangle$, we denote the measure for this ensemble as $d\mu_G(\psi,m)$. From Eq. (B1) it is straightforward to see

$$\hat{U}|\tilde{\psi}\rangle = \frac{1}{\sqrt{d}}\sum_{j,j'=1}^{d} z_j U_{jj'}|\phi_{j'}\rangle = \frac{1}{\sqrt{d}}\sum_{j'=1}^{d} w_{j'}|\phi_{j'}\rangle, \tag{B2}$$

where $w_{j'} = \Sigma U_{jj'} z_j$. Due to unitarity, $\langle w_{j'}^* w_{k'}\rangle = \delta_{j'k'}$ and $\langle w_{j'} w_{k'}\rangle = 0$ (that is, the ws satisfy the same statistical relations as the zs). Hence

$$d\mu_G(\psi,m) = d\mu_G(\hat{U}\psi,m), \quad \forall\ \hat{U}, \tag{B3}$$

which means that the measure factorizes into a constant measure over ψ (the Haar measure) and a measure over the weightings m, and can be written as

$$d\mu_G(\psi,m) = d\mu_{\text{Haar}}(\psi)d\mu_G(m). \tag{B4}$$

Hence instead of simply using the Haar measure $d\mu_{\text{Haar}}$ to describe the distribution of the ensemble $F^\star$ we may use the Gaussian measure $d\mu_G(\psi,m)$.

For simplicity we go on to define $z = \sqrt{u}e^{i\theta}$ so that $d\mu_G(\psi,m)$ becomes

$$\wp(u_1,\ldots,u_d,\theta_1,\ldots,\theta_d)du_1\cdots du_d d\theta_1\cdots d\theta_d = \frac{1}{(2\pi)^d}\exp\left(-\sum_{i=1}^{d}u_i\right)du_1\cdots du_d d\theta_1\cdots d\theta_d, \quad \text{(B5)}$$

which is normalized as follows:

$$\frac{1}{(2\pi)^d}\int_0^\infty du_1\cdots du_d \exp\left(-\sum_{i=1}^{d}u_i\right)\int_0^{2\pi}d\theta_1\cdots d\theta_d = 1. \quad \text{(B6)}$$

2. Normalization term

The denominator of Eq. (4.22) evaluates to

$$\begin{aligned}\int d\mu(m)m^2 &= \int\int d\mu(m)m^2 d\mu_{\text{Haar}}(\psi) = \int d\mu_G(\psi,m)\langle\tilde{\psi}|\tilde{\psi}\rangle \\ &= \int d\mu_G(\psi,m)\sum_{i,j=1}^{d}\frac{z_i^* z_j}{d}\langle\phi_i|\phi_j\rangle \\ &= \int d\mu_G(\psi,m)\sum_{i=1}^{d}\frac{|z_i|^2}{d} = E_G\left[\sum_{i=1}^{d}\frac{|z_i|^2}{d}\right] = 1.\end{aligned} \quad \text{(B7)}$$

We have used the facts that $\int d\mu_{\text{Haar}}(\psi)=1$ and $m^2=\langle\tilde{\psi}|\tilde{\psi}\rangle$, while $E_G[x]$ denotes the expected value of x with respect to the Gaussian measure $d\mu_G(\psi,m)$.

3. Evaluating Eq. (4.23)

To calculate the integral in Eq. (4.23) we need the limits of integration. These are determined by considering the states in Hilbert space which are closer to $|a\rangle\langle a|$ than any other basis state as outlined in Eq. (4.17). We can choose the orthornormal basis in Eq. (B1) such that $|a\rangle$ is one of the basis states. Since the states $|\tilde{\psi}\rangle$ are unnormalized we must perform the integral relating to $|a\rangle$ over the complete range from 0 to ∞. That is, the coefficient u_a associated with $|a\rangle$ ranges from 0 to ∞. However, since this must be the largest parameter, the other u_i must only range from 0 to u_a. Using these limits the integral $\int_a d\mu_G(\psi,m)|\langle a|\tilde{\psi}\rangle|^2$ becomes

$$\begin{aligned}\int_a d\mu_G(\psi,m)\left|\langle a|\frac{1}{\sqrt{d}}\sum_{j=1}^{d}z_j|\phi_j\rangle\right|^2 &= \frac{1}{d}\int_a d\mu_G(\psi,m)|z_a|^2 \\ &= \frac{1}{d}\int_a d\mu_G(\psi,m)u_a \\ &= \frac{1}{d(2\pi)^d}\int_0^\infty du_a u_a\int_0^{u_a}du_2\cdots\int_0^{u_a}du_d \\ &\quad\times\int_0^{2\pi}d\theta_1\cdots\int_0^{2\pi}d\theta_d\exp\left(-\sum_{i=1}^{d}u_i\right) \\ &= \frac{(2\pi)^d}{d(2\pi)^d}\int_0^\infty du_a u_a e^{-u_a}\left(\int_0^{u_a}due^{-u}\right)^{d-1} \\ &= \frac{1}{d}\int_0^\infty du_a u_a e^{-u_a}(1-e^{-u_a})^{d-1} \\ &= \frac{1}{d}\int_0^\infty du_a u_a e^{-u_a}\sum_{k=0}^{d-1}\binom{d-1}{k}(-e^{-u_a})^k \\ &= \frac{1}{d}\sum_{k=0}^{d-1}(-1)^k\binom{d-1}{k}\left[\frac{1}{(k+1)^2}\right] \\ &= \frac{1}{d}\sum_{k=1}^{d}(-1)^{k-1}\binom{d-1}{k-1}\frac{1}{k^2} \\ &= \frac{1}{d^2}\sum_{k=1}^{d}\frac{(-1)^{k-1}}{k}\binom{d}{k} \equiv \epsilon_d.\end{aligned} \quad \text{(B8)}$$

It is possible to further simplify ϵ_d. This can be done by considering the following expression:

$$\begin{aligned}\frac{1}{d^2}\int_0^1 dx\frac{1-(1-x)^d}{x} &= \frac{1}{d^2}\int_0^1 dx\frac{1-\sum_{k=0}^{d}\binom{d}{k}(-x)^k}{x} \\ &= \frac{1}{d^2}\sum_{k=1}^{d}\frac{(-1)^{k+1}}{k}\binom{d}{k}.\end{aligned} \quad \text{(B9)}$$

It is not immediately obvious that the above integral is a simpler expression for ϵ_d. However, this expression can be evaluated alternatively using the subsitution $y=1-x$, which gives

$$\begin{aligned}\epsilon_d &= \frac{1}{d^2}\int_0^1 dy\frac{1-y^d}{1-y} = \frac{1}{d^2}\int_0^1 dy(1-y^d)\sum_{k=0}^{\infty}y^k \\ &= \frac{1}{d^2}\left(\sum_{k=0}^{d-1}\frac{1}{k+1}+\sum_{k=d}^{\infty}\frac{1}{k+1}-\sum_{k=0}^{\infty}\frac{1}{k+d+1}\right) \\ &= \frac{1}{d^2}\sum_{k=1}^{d}\frac{1}{k},\end{aligned} \quad \text{(B10)}$$

which is the result in Eq. (4.23).

APPENDIX C: INEPT STATE STEERING

1. Optimal ensemble

We need to show that the conditions for Lemma 1 hold for the ensemble defined by Eq. (4.33). That is, we need to show that Eq. (4.33) defines an optimal ensemble. In this instance G is the group generated by $(1/2)\hat{\sigma}_z\otimes\mathbf{I}-(1/2)\mathbf{I}\otimes\hat{\sigma}_z$, so $g\to\phi\in[0,2\pi)$ and

$$\hat{U}_{\alpha\beta}(\phi)=\exp[-i\phi\hat{\sigma}_z/2]\otimes\exp[i\phi\hat{\sigma}_z/2]. \quad \text{(C1)}$$

For the particular measurement strategy chosen we need to consider only two types of measurement $\hat{\sigma}_z$ and $\hat{\sigma}_\theta$. The condition $\hat{U}_\alpha^\dagger(\phi)\hat{A}\hat{U}_\alpha(\phi)\in\mathfrak{M}_\alpha$ clearly holds for $\hat{A}=\hat{\sigma}_z$ since

$$\exp(i\phi\hat{\sigma}_z)\hat{\sigma}_z\exp(-i\phi\hat{\sigma}_z) = \hat{\sigma}_z. \tag{C2}$$

Therefore Eq. (4.1) holds trivially.

Now it simply remains to test these conditions for measurements of the form of $\hat{A}=\hat{\sigma}_\theta$. In this case we have

$$\begin{aligned}\hat{U}_\alpha^\dagger(\phi)\hat{\sigma}_\theta\hat{U}_\alpha(\phi) &= \exp(i\phi\hat{\sigma}_z)\hat{\sigma}_\theta\exp(-i\phi\hat{\sigma}_z)\\ &= \cos(\theta-\phi)\hat{\sigma}_x + \sin(\theta-\phi)\hat{\sigma}_y\\ &= \hat{\sigma}_{\theta-\phi},\end{aligned} \tag{C3}$$

which is in $\mathfrak{M}_\alpha$. Thus to test Eq. (4.1) we need to evaluate $\tilde{\rho}_a^{\sigma_{\theta-\phi}}$. Using Eq. (4.43) we can see that this simply evaluates to

$$\begin{aligned}\tilde{\rho}_a^{\sigma_{\theta-\phi}} = \frac{1}{2}[&\mathbf{I} + a\eta\sqrt{\epsilon(1-\epsilon)}\cos(\theta-\phi)\hat{\sigma}_x\\ &+ a\eta\sqrt{\epsilon(1-\epsilon)}\sin(\theta-\phi)\hat{\sigma}_y - (1-2\epsilon)\hat{\sigma}_z].\end{aligned} \tag{C4}$$

Finally, we evaluate

$$\begin{aligned}\hat{U}_\beta(\phi)\tilde{\rho}_a^{\sigma_\theta}\hat{U}_\beta^\dagger(\phi) &= \exp(i\phi\hat{\sigma}_z)\tilde{\rho}_a^{\sigma_\theta}\exp(-i\phi\hat{\sigma}_z)\\ &= \frac{1}{2}\begin{pmatrix} e^{-i(\theta-\phi)}2\epsilon & \kappa e^{-i(\theta-\phi)}\\ \kappa e^{i(\theta-\phi)} & 2(1-\epsilon)\end{pmatrix} = \tilde{\rho}_a^{\sigma_{\theta-\phi}},\end{aligned} \tag{C5}$$

where $\kappa=a\eta\sqrt{\epsilon(1-\epsilon)}$. Thus Eq. (4.1) also holds for measurements of $\hat{\sigma}_\theta$. Hence, the conditions of Lemma 1 hold and the ensemble defined by Eq. (4.33) is of the form of the optimal ensemble.

2. Steering bound for $\epsilon=1/2$

We know that for $d=2$ the Werner and isotropic states are equivalent. Now consider the inept states when $\epsilon=1/2$. In this case Eq. (4.27) becomes

$$W_{1/2}^\eta = \eta|\psi\rangle\langle\psi| + (1-\eta)\frac{\mathbf{I}}{4}, \tag{C8}$$

where $|\psi\rangle=\frac{1}{\sqrt{2}}(|0_\alpha 0_\beta\rangle+|1_\alpha 1_\beta\rangle)$. Comparing this with Eq. (4.12) for isotropic states (when $d=2$), one immediately sees that the expressions are identical. Hence, for $\epsilon=1/2$ the inept states are equivalent to the $d=2$ isotropic (and Werner) states.

Setting $\epsilon=1/2$ in Eq. (4.50) we find an upper bound of 0.5468 on η_{steer}. However, we know that the steering boundary for $d=2$ isotropic states occurs at $\eta=1/2$. Thus for $\epsilon=1/2$ we can do better than an upper bound on η_{steer} for inept states; due to the equivalence with isotropic states we know that the true η_{steer} occurs at $\eta=1/2$. We plot this as a separate point at $\epsilon=1/2$ in Fig. 2(c).

APPENDIX D: GAUSSIAN STATE STEERING

1. Proof of Lemma 2

First, suppose Eq. (4.58) is true. Thus the matrix $V_{\alpha\beta}+\mathbf{0}_\alpha\oplus i\Sigma_\beta$ is PSD. Now since Eq. (4.58) is assumed true, and we know that $V_\alpha\geqslant 0$, taking the Schur complement of V_α in Eq. (4.58) we see

$$V_\beta + i\Sigma_\beta - C^\mathsf{T}V_\alpha^{-1}C \geqslant 0, \tag{D1}$$

which implies Eq. (4.59), where $U=V_\beta-C^\mathsf{T}V_\alpha^{-1}C$. This LMI allows us to define an ensemble $F^U=\{\rho_\xi^U\wp_\xi^U\}$ of Gaussian states with $\text{CM}[\rho_\xi^U]=U$, distinguished by their mean vectors (ξ).

Now we wish to see if the ensemble U defined above could be used to simulate Bob's conditioned state V_β^A. This will be the case if V_β^A-U is PSD as explained below. Evaluating this matrix we see that

$$\begin{aligned}V_\beta^A - U &= V_\beta - C^\mathsf{T}(V_\alpha+T^A)^{-1}C - V_\beta + C^\mathsf{T}V_\alpha^{-1}C\\ &= C^\mathsf{T}[V_\alpha^{-1}-(V_\alpha+T^A)^{-1}]C.\end{aligned} \tag{D2}$$

Both C and C^T are positive matrices, so the above expression is PSD if and only if the bracketed term is PSD. To prove this is so, we make use of the Woodbury formula, which can be expressed as

$$X^{-1}-(X+YZ^\mathsf{T})^{-1} = X^{-1}Y(I+Z^\mathsf{T}X^{-1}Y)^{-1}Z^\mathsf{T}X^{-1}. \tag{D3}$$

Thus to check the positivity of $V_\alpha^{-1}-(V_\alpha+T^A)^{-1}$ we set $X=V_\alpha$, $Y=\sqrt{T^A}$, and $Z^\mathsf{T}=\sqrt{T^A}$, and thus

$$V_\alpha^{-1}-(T^A+V_\alpha)^{-1} = V_\alpha^{-1}\sqrt{T^A}(I+\sqrt{T^A}V_\alpha^{-1}\sqrt{T^A})^{-1}\sqrt{T^A}V_\alpha^{-1}. \tag{D4}$$

Now the covariance matrices V_α and T^A are positive by definition, so their inverse and square root, respectively, must also be positive matrices. Since the product and the sum of two positive matrices is PSD, the above expression is PSD if and only if $\sqrt{T^A}V_\alpha^{-1}\sqrt{T^A}$ is PSD, which holds since any matrix ABA^T is PSD whenever B is PSD. Thus the ensemble defined by $U=V_\beta-C^\mathsf{T}V_\alpha^{-1}C$ satisfies Eq. (4.60), which implies that $\forall A\in\mathfrak{G}, \rho_a^A$ is a Gaussian mixture (over ξ) of Gaussian states ρ_ξ^U, all with the same covariance matrix U, but with different mean vectors ξ. Specifically, $\wp(a|\xi,A)$ is a Gaussian distribution in ξ with a mean vector equal to a (which is determined by Alice's measurement A and the bipartite Gaussian state W), and a covariance matrix equal to V_β^A-U. As long as $V_\beta^A-U\geqslant 0$, this distribution is well defined, so that Bob's state ρ_a^A is consistent with Alice merely sending Bob Gaussian states drawn from an ensemble $F^U=\{\rho_\xi^U\wp_\xi^U\}$ in which all states have a CM equal to U, and with mean vectors ξ having a Gaussian distribution $\wp_\xi^U$, which has a covariance matrix $V_\beta-U=C^\mathsf{T}V_\alpha^{-1}C\geqslant 0$. Therefore W is not steerable by Alice for all measurements $A\in\mathfrak{G}$ if Eq. (4.58) is true. ■

2. Proof of Lemma 3

Suppose that W is not steerable. This means that there is some ensemble $F=\{\wp_\xi\rho_\xi\}$, which satisfies Eq. (3.5). Therefore we know that Bob's conditioned state can be written as

$$\rho_a^A = \frac{\sum_\xi \wp_\xi \wp(a|A,\xi)\rho_\xi}{\sum_\xi \wp_\xi \wp(a|A,\xi)}. \tag{D5}$$

This means that the covariance matrix satisfies

$$\mathrm{CM}[\rho_a^A] = V_\beta^A \geqslant \frac{\sum_\xi \wp_\xi \wp(a|A,\xi)\mathrm{CM}[\rho_\xi]}{\sum_\xi \wp_\xi \wp(a|A,\xi)}, \tag{D6}$$

since the CM of a state equal to a weighted sum of states must be at least as great as the weighted sum of the individual CMs. The equality occurs if all the means are the same. Rearranging and taking a sum over a on both sides gives

$$\sum_{\xi,a} \wp_\xi \wp(a|A,\xi) V_\beta^A \geqslant \sum_{\xi,a} \wp_\xi \wp(a|A,\xi)\mathrm{CM}[\rho_\xi]. \tag{D7}$$

From the fact that V_β^A is independent of a, and using the facts that $\Sigma_a \wp(a|A,\xi)=1$ and $\Sigma_\xi \wp_\xi = 1$, one sees that Eq. (D7) simplifies to

$$V_\beta^A \geqslant \sum_\xi \wp_\xi \mathrm{CM}[\rho_\xi]. \tag{D8}$$

Defining $U=\Sigma_\xi \wp_\xi \mathrm{CM}[\rho_\xi]$ satisfies Eq. (4.59) by definition and Eq. (D8) implies Eq. (4.60). Therefore if W is not steerable then there exists an ensemble U such that Eqs. (4.59) and (4.60) are true. ■

3. Proof of Lemma 4

Equation (4.60) defines the Schur complement of $V_\alpha + T^A$ in the following matrix:

$$M = \begin{pmatrix} V_\alpha + T^A & C \\ C^\mathsf{T} & V_\beta - U \end{pmatrix}. \tag{D9}$$

Therefore, since $V_\alpha + T^A \geqslant 0$ (recall that we are considering Gaussian measurements), Eq. (4.60) is equivalent to the condition that the matrix M be PSD. Now we know that the sum of two PSD matrices is PSD, so if $M \geqslant 0$ and using Eq. (4.59) we arrive at

$$M + 0_\alpha \oplus (U + i\Sigma_\beta) = \begin{pmatrix} V_\alpha + T^A & C \\ C^\mathsf{T} & V_\beta + i\Sigma \end{pmatrix} \geqslant 0, \tag{D10}$$

as an equivalent condition to Eqs. (4.59) and (4.60). Finally, we know that $V_\beta + i\Sigma_\beta \geqslant 0$, so the Schur complement of this term in the above matrix must be PSD. ■

[1] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 553 (1935).
[2] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[3] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964). Reproduced in Ref. [39].
[4] J. S. Bell, *Proceedings of the International School of Physics "Enrico Fermi,"* Course IL (Academic, New York, 1971). Reproduced in Ref. [39].
[5] L. Hardy, Contemp. Phys. **39**, 419 (1998).
[6] R. F. Werner, Phys. Rev. A showed **40**, 4277 (1989).
[7] N. Gisin, Phys. Lett. A **154**, 201 (1991).
[8] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
[9] L. Masanes, Y.-C. Liang, and A. C. Doherty, e-print arXiv:quant-ph/0703268.
[10] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).
[11] E. Schrödinger, Proc. Cambridge Philos. Soc. **32**, 446 (1936).
[12] M. D. Reid, Phys. Rev. A **40**, 913 (1989).
[13] S. J. Jones, H. M. Wiseman, and D. T. Pope, Phys. Rev. A **72**, 022330 (2005).
[14] G. M. D'Ariano, L. Maccone, and M. G. A. Paris, J. Phys. A **34**, 93 (2001).
[15] R. Horodecki, P. Horodecki, and M. Horodecki, Phys. Lett. A **200**, 340 (1995).
[16] A. Acìn, N. Gisin, and B. Toner, Phys. Rev. A **73**, 062105 (2006).
[17] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
[18] D. Collins, N. Gisin, N. Linden, S. Massar, and S. Popescu, Phys. Rev. Lett. **88**, 040404 (2002).
[19] M. L. Almeida, S. Pironio, J. Barrett, G. Tòth, and A. Acìn, Phys. Rev. Lett. **99**, 040403 (2007).
[20] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[21] G. Giedke and J. I. Cirac, Phys. Rev. A **66**, 032316 (2002).
[22] G. Giedke, J. Eisert, J. I. Cirac, and M. B. Plenio, Quantum Inf. Comput. **3**, 211 (2003).
[23] K. Zhou, J. C. Doyle, and K. Glover, *Optimal and Robust Control* (Prentice-Hall, Englewood Cliffs, NJ, 1996).
[24] L. M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **84**, 2722 (2000).
[25] R. Simon, Phys. Rev. Lett. **84**, 2726 (2000).
[26] W. P. Bowen, R. Schnabel, P. K. Lam, and T. C. Ralph, Phys. Rev. A **69**, 012304 (2004).
[27] N. Brunner, N. Gisin, and V. Scarani, New J. Phys. **7**, 88 (2005).
[28] K. Horodecki, M. Horodecki, and P. Horodecki, e-print arXiv:quant-ph/0512224.
[29] G. Gutoski and J. Watrous, in *Proceedings of the Thirty-Ninth Annual Symposium on Theory of Computing*, edited by D. S. Johnson and U. Feige (ACM Press, New York 2007), pp. 565–574.
[30] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, in *Proceedings of the 19th IEEE Conference on Computational Complexity*

(IEE Computer Society, Washington D.C., 2004).
[31] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[32] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[33] H. Primas, in *Moderne Naturphilosophie* (Königshausen und Neumann, Würzburg, 1984).
[34] M. Vujićič and F. Herbut, J. Phys. A **21**, 2931 (1998).
[35] F. Verstraete, Ph.D. thesis, Kath. Univ. Leuven, 2002 (unpublished).
[36] R. Clifton, J. Bub, and H. Halvorson, Found. Phys. **33**, 1561 (2003).
[37] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).
[38] K. A. Kirkpatrick, Found. Phys. Lett. **19**, 95 (2006).
[39] J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, England, 1987).
[40] *Albert Einstein: Philosopher-Scientist*, edited by P. A. Schilpp (Library of the Living Philosophers, Evanston, 1949).
[41] A. Einstein, Autobiographical Notes, in *Albert Einstein: Philosopher-Scientist* (Ref. [40]), pp. 1–94, with translation (the odd-numbered pages) by P. A. Schilpp.
[42] H. M. Wiseman, Contemp. Phys. **47**, 79 (2006).