Entanglement and Nonlocality are Inequivalent for Any Number of Parties

R. Augusiak,[1] M. Demianowicz,[1] J. Tura,[1] and A. Acín[1,2]

[1]*ICFO—Institut de Ciencies Fotoniques, 08860 Castelldefels (Barcelona), Spain*
[2]*ICREA—Institucio Catalana de Recerca i Estudis Avançats, Lluis Companys 23, 08010 Barcelona, Spain*

Understanding the relation between nonlocality and entanglement is one of the fundamental problems in quantum physics. In the bipartite case, it is known that these two phenomena are inequivalent, as there exist entangled states of two parties that do not violate any Bell inequality. However, except for a single example of an entangled three-qubit state that has a local model, almost nothing is known about such a relation in multipartite systems. We provide a general construction of genuinely multipartite entangled states that do not display genuinely multipartite nonlocality, thus proving that entanglement and nonlocality are inequivalent for any number of parties.

Introduction.—Nonlocality, the phenomenon of the impossibility of describing by local hidden variable (LHV) theories the correlations arising from measuring quantum states, is a fundamental characteristic of quantum mechanics. It was Bell who first pointed out that there exist quantum states for which underlying classical variables cannot account for the measurement statistics on them [1]. Every such state is called nonlocal and it violates some Bell inequality (see Ref. [2]). Having been confronted with the result of Bell's, one might be tempted to identify nonlocality with entanglement—another essential resource of quantum information theory. This belief is even further strengthened by the fact that all pure entangled states do exhibit nonlocal properties [3–5]. The relationship between entanglement and nonlocality, however, turns out to be more involved: while LHV models trivially exist for separable states, not all entangled states are nonlocal.

The first step in the exploration of this intriguing inequivalence was taken by Werner [6]. He introduced a family of highly symmetric states, nowadays known as the Werner states, and provided an explicit LHV model reproducing the measurement statistics obtained when two parties perform local projective measurements on some states from this family. The existence of the model implies that these entangled states do not violate any Bell inequality when subjected to projective measurements. Building on Werner's model, Barrett proved the inequivalence between entanglement and nonlocality for two parties by constructing a new LHV model that is also valid for general measurements performed on some entangled Werner states [7]. Both of these models were later adapted to mixtures of any state ρ and the white noise [8].

Very little is known about the relation between entanglement and nonlocality in the general multipartite scenario. Here, this question becomes much subtler as the multiparty scenario offers a richer variety of different types of entanglement and nonlocality. For instance, for any number of

parties, it is trivial to construct a nonseparable—and thus entangled—state that has a local model: it suffices to combine an entangled and local Werner state for two parties with a fully product state for the remaining ones. However, this is clearly nothing but a manifestation of the inequivalence between entanglement and nonlocality for two parties trivially embedded in a multipartite scenario. Thus, the natural question here is whether, for an arbitrary number of parties N, there always exists a gap between genuine N-partite entanglement and genuine N-partite nonlocality, i.e., whether there are genuinely multipartite entangled (GME) states that do not violate any Bell inequality for genuinely multipartite nonlocality. Quite remarkably, this question remains unsolved even in the seemingly simple case of pure states: despite the fact that all multipartite pure entangled states violate a Bell inequality [5], this inequality does not detect genuine multipartite nonlocality.

The main goal of this Letter is to prove that entanglement and nonlocality are inequivalent for *any* number of parties N. With this aim we provide, for any N, examples of genuinely N-partite entangled states that have an LHV model for the case in which the parties are divided into two groups. Thus, these states cannot violate any Bell inequality for genuine N-partite nonlocality despite being GME.

It should be noted that, for three parties, Tóth and Acín found a fully local model for arbitrary projective measurements on a genuinely entangled state [9]. This model was later extended to general measurements [10], proving ultimately that there are three-partite GME states not displaying nonlocality. It is unknown whether this model can be extended to more parties and, in fact, the straightforward extension of the model to more than three parties does not work [9].

Let us finally remark that several operationally meaningful nonlocality scenarios besides the standard one considered here, consisting of single-copy measurements by each party, have been introduced [11–16]. A notable

example is one that exploits the possibility of performing sequences of measurements on a single copy of the state, which led to the phenomenon of hidden nonlocality [11–13]. Although it is certainly an interesting question whether our results also hold in these scenarios, it should be stressed that, from a fundamental point of view, our scenario is the one at the very root of the notion of nonlocality, as originally put forward by Bell. Moreover, it has clear practical advantages, as, with respect to the other scenarios, it does not require nondestructive measurements, or measurements acting on more than one system.

Preliminaries.—Let us introduce some notation and terminology. Consider N parties $\mathbf{A} := A_1, \ldots, A_N$ (for a small N also denoted A, B, etc.) sharing an N-partite quantum state $\rho_\mathbf{A} \in \mathcal{B}(\mathcal{H}_{N,d})$, where $\mathcal{H}_{N,d} = (\mathbb{C}^d)^{\otimes N}$ and $\mathcal{B}(\mathcal{H})$ stands for the set of bounded linear operators acting on the Hilbert space $\mathcal{H}$. Moreover, by $\mathcal{S}_{M,d}$ and P_{sym}^X we denote, respectively, the symmetric subspace of $\mathcal{H}_{M,d}$ and the projector onto the symmetric subspace of the subsystem X.

Let us then partition the parties into K pairwise disjoint groups S_i, such that by adding them one recovers $\mathbf{A}$, and call it a K-partition; for $K = 2$ we call it a bipartition and denote $S|\bar{S}$, with $\bar{S}$ being the complement of S in $\mathbf{A}$. Denoting by $\mathcal{S}_K$ the set of all K-partitions, we say that $\rho_\mathbf{A}$ is K-separable (biseparable for $K = 2$) if it is a probabilistic mixture of N-partite states separable with respect to some K-partitions, i.e.,

$$\rho_\mathbf{A} = \sum_{S \in \mathcal{S}_K} p_S \sum_i q_S^i \bigotimes_{k=1}^K |\psi_{S_k}^i\rangle\langle\psi_{S_k}^i|. \qquad (1)$$

Here, p_S and q_S^i are probability distributions and $|\psi_{S_k}^i\rangle$ are pure states defined on the S_k subsystem. One then calls $\rho_\mathbf{A}$ fully separable if it is N-separable, and GME if it does not admit any of the above forms of separability; specifically, it is not biseparable.

Analogously to the notions of separability, one introduces those of locality. Imagine that on their share of the state $\rho_\mathbf{A}$, each party A_i performs a measurement $M_i = \{M_{a_i}^{(i)}\}$, where a_i enumerate the outcomes of M_i, while $M_{a_i}^{(i)}$ are the measurement operators; i.e., $M_{a_i}^{(i)} \geq 0$ and $\sum_{a_i} M_{a_i}^{(i)} = \mathbb{1}_d$. If, additionally, $M_{a_i}^{(i)}$ are supported on orthogonal subspaces, we call the corresponding measurement projective (PM), and generalized (GM; also called POVM) otherwise. Now, adopting the definition from Ref. [17], one says that the state $\rho_\mathbf{A}$ is K-local for GMs, or shortly K-local if, for any choice of measurements $\mathbf{M} := M_1, \ldots, M_N$, the probability of obtaining the outcomes $\mathbf{a} := a_1, \ldots, a_N$ decomposes as

$$p(\mathbf{a}|\mathbf{M}) = \sum_{S \in \mathcal{S}_K} p_S \int d\lambda \, \omega_S(\lambda) \prod_{k=1}^K p_k(\mathbf{a}_{S_k}|\mathbf{M}_{S_k}, \lambda). \qquad (2)$$

Here, the sum goes over all possible K-partitions of $\mathbf{A}$, p_S and ω_S are probability distributions, and $p_k(\cdot|\cdot)$ is the probability (also called the response function) that the parties belonging to S_k obtain $\mathbf{a}_{S_k}$ upon measuring $\mathbf{M}_{S_k}$, while having the classical information λ. Accordingly, a state $\rho_\mathbf{A}$ admitting Eq. (2) is said to have a K-local model. In particular, if $K = N$ we say that the state is fully local, while it is bilocal if $K = 2$. Notice that there are also local models reproducing projective measurements only, and mixed ones reproducing GMs for some parties and PMs for the rest. By comparing Eqs. (1) and (2), one can clearly see that every K-separable state is K-local. Multipartite states which do not admit any form of bilocality are called genuinely multipartite nonlocal.

It has been shown that the above definition of K-locality is inconsistent with an operational interpretation of non-locality given in Refs. [18,19]. To recover consistency it is enough to demand that all $p_k(\mathbf{a}_{S_k}|\mathbf{M}_{S_k}, \lambda)$'s appearing in Eq. (2) fulfill the no-signaling principle. In what follows we will mainly use the first definition because it allows us to state our results in a more general way. Nevertheless, as we will show below, the inequivalence between entanglement and nonlocality also holds for the operational definitions.

Inequivalence of entanglement and nonlocality.—We are now in position to state and prove our main result.

Main result.—Entanglement and nonlocality are inequivalent for any number of parties N, as for any N there exist genuinely entangled N-partite states with bilocal models.

To prove the result we proceed in two steps. First, we show that any bipartite local state can be converted into a multipartite state with a bilocal model. Then, we argue that such a construction may lead to GME states for any N.

As for the first step, we generalize the observation made by Barrett [7]. Let $\varrho_{AB} \in \mathcal{B}(\mathcal{H}_{2,d})$ be arbitrary and let

$$\Lambda_{A \to S} \colon \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}((\mathbb{C}^{d'})^{\otimes L}) \qquad (3)$$

and

$$\Lambda_{B \to \bar{S}} \colon \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}((\mathbb{C}^{d'})^{\otimes N-L}) \qquad (4)$$

be a pair of quantum channels sending operators acting on a single-party Hilbert space $\mathbb{C}^d$ to operators acting on L-partite and $(N - L)$-partite Hilbert spaces of local dimension d', respectively, with $S = A_1 \ldots A_L$ and $\bar{S} = A_{L+1} \ldots A_N$. One proves the following.

Lemma 1.—If ρ_{AB} is local for generalized measurements, then, for any pair of quantum channels $\Lambda_{A \to S}$ and $\Lambda_{B \to \bar{S}}$ defined above, the N-partite state

$$\sigma_\mathbf{A} = (\Lambda_{A \to S} \otimes \Lambda_{B \to \bar{S}})(\rho_{AB}) \qquad (5)$$

is bilocal for any measurements.

Proof.—The reasoning is analogous to the one by Barrett in Ref. [7], but for completeness we present it here.

The fact that ρ_{AB} is local for generalized measurements means that the probabilities of obtaining results a, b upon performing measurements $M_A = \{M_a^A\}$ and $M_B = \{M_b^B\}$, respectively, by the parties A and B, assume the local form (2), which for $N = 2$ simplifies to

$$p(a, b|M_A, M_B) = \int d\lambda\omega(\lambda) p_\rho(a|M_A, \lambda) p_\rho(b|M_B, \lambda).$$
(6)

We have used the subscript ρ to emphasize that the probabilities correspond to ρ_{AB}. Exploiting this model, we will now demonstrate that $\sigma_{\mathbf{A}}$ is bilocal with respect to the bipartition $S|\bar{S} = A_1...A_L|A_{L+1}...A_N$. To this end, let us assume that the parties perform measurements $M_i = \{M_{a_i}^{(i)}\}$, $i = 1, ..., N$ on their shares of $\sigma_{\mathbf{A}}$. Then, denoting by $\Lambda_{S\to A}^\dagger$ and $\Lambda_{\bar{S}\to B}^\dagger$ the dual maps of $\Lambda_{A\to S}$ and $\Lambda_{B\to\bar{S}}$ [20], respectively, we define the following operators,

$$\bar{M}_{\mathbf{a}_S}^A = \Lambda_{S\to A}^\dagger \left(\bigotimes_{i=1}^L M_{a_i}^{(i)}\right), \quad \bar{M}_{\mathbf{a}_{\bar{S}}}^B = \Lambda_{\bar{S}\to B}^\dagger \left(\bigotimes_{i=L+1}^N M_{a_i}^{(i)}\right), \quad (7)$$

acting on $\mathbb{C}^d$ and indexed by the outcomes $\mathbf{a}_S := a_1, ..., a_L$ and $\mathbf{a}_{\bar{S}} := a_{L+1}, ..., a_N$. Since the dual map of a quantum channel is positive and unital (it preserves the identity operator), it is clear to see that the operators (7) form GMs, denoted $\bar{M}_A$ and $\bar{M}_B$. With their aid, let us now define the response functions for the state $\sigma_{\mathbf{A}}$ corresponding to the parties $A_1, ..., A_L$ and $A_{L+1}, ..., A_N$, respectively, as $p_\sigma(\mathbf{a}_S|\mathbf{M}_S, \lambda) = p_\rho(\mathbf{a}_S|\bar{M}_A, \lambda)$ and $p_\sigma(\mathbf{a}_{\bar{S}}|\mathbf{M}_{\bar{S}}, \lambda) = p_\rho(\mathbf{a}_{\bar{S}}|\bar{M}_B, \lambda)$. Then,

$$
\begin{aligned}
p(\mathbf{a}|\mathbf{M}) &= \mathrm{Tr}[(M_{a_1}^{(1)} \otimes ... \otimes M_{a_N}^{(N)})(\Lambda_{A\to S} \otimes \Lambda_{B\to\bar{S}})(\rho_{AB})] \\
&= \mathrm{Tr}[\bar{M}_{a_1...a_L}^A \otimes \bar{M}_{a_{L+1}...a_N}^B \rho_{AB}] \\
&= \int d\lambda\omega(\lambda) p_\rho(\mathbf{a}_S|\bar{M}_A, \lambda) p_\rho(\mathbf{a}_{\bar{S}}|\bar{M}_B, \lambda) \\
&= \int d\lambda\omega(\lambda) p_\sigma(\mathbf{a}_S|\mathbf{M}_S, \lambda) p_\sigma(\mathbf{a}_{\bar{S}}|\mathbf{M}_{\bar{S}}, \lambda), \quad (8)
\end{aligned}
$$

where we have utilized Eq. (7) and the definition of $\sigma_{\mathbf{A}}$. It thus follows that $\sigma_{\mathbf{A}}$ is bilocal for GMs with respect to $A_1...A_L|A_{L+1}...A_N$. ∎

The critical point of our approach will be to observe that the above mapping of local bipartite states to bilocal multipartite ones may lead to GME states. To argue this, we need a technical result concerning genuine multipartite entanglement.

Lemma 2.—Consider an N-partite state $\sigma_{\mathbf{A}} \in \mathcal{B}(\mathcal{H}_{N,d})$ with symmetric subsystems S and $\bar{S}$ ($S \cup \bar{S} = \mathbf{A}$); that is, $P_{\mathrm{sym}}^S \otimes P_{\mathrm{sym}}^{\bar{S}} \sigma_{\mathbf{A}} P_{\mathrm{sym}}^S \otimes P_{\mathrm{sym}}^{\bar{S}} = \sigma_{\mathbf{A}}$ holds. If $\sigma_{\mathbf{A}}$ is not GME, then it is biseparable with respect to the bipartition $S|\bar{S}$, i.e.,

$$\sigma_{\mathbf{A}} = \sum_i p_i \sigma_S^i \otimes \sigma_{\bar{S}}^i, \quad (9)$$

with σ_S^i and $\sigma_{\bar{S}}^i$ being states defined on subsystems S and $\bar{S}$.

Proof.—As the proof is rather technical and lengthy, we present its sketch here, moving the details to the Supplemental Material [21].

The assumption that $\sigma_{\mathbf{A}}$ is not GME means that it admits the decomposition (1) with $K = 2$, i.e., $\sigma_{\mathbf{A}} = \sum_{T|\bar{T}\in\mathcal{S}_2} p_{T|\bar{T}} \rho_{T|\bar{T}}$. The sum goes over all bipartitions $T|\bar{T}$ of $\mathbf{A}$ and $\rho_{T|\bar{T}}$ is some state separable with respect to $T|\bar{T}$, i.e., $\rho_{T|\bar{T}} = \sum_i q_{T|\bar{T}}^i |e_T^i\rangle\langle e_T^i| \otimes |f_{\bar{T}}^i\rangle\langle f_{\bar{T}}^i|$. Now, exploiting the assumption that the subspaces S and $\bar{S}$ of $\sigma_{\mathbf{A}}$ are symmetric, one can prove that each $\rho_{T|\bar{T}}$ with $T \neq S$ in the decomposition of $\sigma_{\mathbf{A}}$ is of the form (9) ($\sigma_{S|\bar{S}}$ is already of this form). To this aim, it is enough to realize that every pure state $|e_T^i\rangle|f_{\bar{T}}^i\rangle$ must obey $P_{\mathrm{sym}}^{S(\bar{S})}|e_T^i\rangle|f_{\bar{T}}^i\rangle = |e_T^i\rangle|f_{\bar{T}}^i\rangle$. This, after some algebra, implies that it must also be a product with respect to $S|\bar{S}$. ∎

We are now ready to prove our main result. A straightforward corollary of Lemma 2 is that any N-partite state $\sigma_{\mathbf{A}}$ which does not admit the form (9), i.e., is entangled across some cut $S|\bar{S}$ with symmetric subsystems S and $\bar{S}$, is GME. Take now a bipartite entangled state $\rho_{AB} \in \mathcal{B}(\mathbb{C}^d \otimes \mathbb{C}^d)$ and the quantum channels $\Lambda_{A\to S}: \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathcal{S}_{L,d'})$ and $\Lambda_{B\to\bar{S}}: \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathcal{S}_{N-L,d'})$ that are invertible in the sense that for both of them there exists a channel $\tilde{\Lambda}$ such that $\tilde{\Lambda}\circ\Lambda$ is an identity map on $\mathcal{B}(\mathbb{C}^d)$. Note that these channels now output states acting on the corresponding L and $(N-L)$-partite symmetric subspaces. Clearly, the N-partite state $\sigma_{\mathbf{A}}$ resulting from the application of $\Lambda_{A\to S}$ and $\Lambda_{B\to\bar{S}}$ to ρ_{AB} is symmetric on the subspaces S and $\bar{S}$ and, as ρ_{AB} is entangled, must be GME; if $\sigma_{\mathbf{A}}$ is not GME, then, as the two channels are invertible, ρ_{AB} must be separable. If we further assume that ρ_{AB} is local, the resulting state $\sigma_{\mathbf{A}}$ will have, according to Lemma 1, a bilocal model, proving the desired result.

As a result we have a general method for constructing bilocal genuinely entangled N-partite states for arbitrary N's. This general construction is summarized in Fig. 1.

Applications.—To illustrate our method we consider two paradigmatic classes of states: the isotropic and Werner states [6,22]. The quantum channels are chosen to be $\Lambda_{A\to S}(\cdot) = V_L(\cdot)V_L^\dagger$ and $\Lambda_{B\to\bar{S}}(\cdot) = V_{N-L}(\cdot)V_{N-L}^\dagger$, with $V_M: \mathbb{C}^d \to \mathcal{S}_{M,d}$ being an isometry defined through $V_M|i\rangle = |i\rangle^{\otimes M}$ for any element of the standard basis in $\mathbb{C}^d$.

Let us begin with the two-qudit isotropic states $\rho_{\mathrm{iso}}(p) = p|\psi_d^+\rangle\langle\psi_d^+| + (1-p)\mathbb{1}_{d^2}/d^2$, where $|\psi_d^+\rangle = \sum_{i=0}^{d-1} |ii\rangle/\sqrt{d}$ is the maximally entangled state. Application of the isometries to $\rho_{\mathrm{iso}}(p)$ leads us to the mixture of the well-known Greenberger-Horne-Zeilinger (GHZ) state of N qudits $|\mathrm{GHZ}_{N,d}\rangle = \sum_{i=0}^{d-1} |i\rangle^{\otimes N}/\sqrt{d}$ and some colored noise:
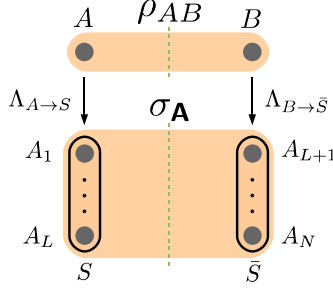
FIG. 1 (color online).　*Schematic depiction of our construction.*
The initial state ρ_{AB} is entangled and local for generalized
measurements. The application of properly chosen local channels
$\Lambda_{A\to S}$ and $\Lambda_{B\to\bar{S}}$, with $S = A_1\dots A_L$ and $\bar{S} = A_{L+1}\dots A_N$, to ρ_{AB}
creates an N-partite state $\sigma_{\mathbf{A}}$ that is genuinely multipartite
entangled and bilocal with respect to the bipartition $S|\bar{S}$. The
dashed green line determines the locality cut.

$$\sigma_{\mathbf{A}}(p) = p|\mathrm{GHZ}_{N,d}\rangle\langle\mathrm{GHZ}_{N,d}| + (1-p)\frac{\mathcal{P}_{L,d}\otimes\mathcal{P}_{N-L,d}}{d^2},$$

$$(10)$$

where $\mathcal{P}_{L,d} = \sum_{i=0}^{d-1}|i\rangle\langle i|^{\otimes L}$ with $1 \le L \le N-1$. Now, as
the isotropic states are local for $p \le (3d-1)(d-1)^{d-1}/d^d(d+1)$ [8], it stems from Lemma 1 that, for the same
range of p and $L = 1,\dots,N-1$, the state $\sigma_{\mathbf{A}}(p)$ is bilocal
with respect to the bipartition $A_1\dots A_L|A_{L+1}\dots A_N$.
Furthermore, isometric channels are always invertible
$(V_M^\dagger V_M = \mathbb{1}_d)$ and thus, as required, preserve entangle-
ment. Hence, the states $\sigma_{\mathbf{A}}(p)$ are GME for the same range
of p as $\rho_{\mathrm{iso}}(p)$ are entangled, i.e., for $p > 1/(d+1)$.
Concluding, the states (10) constitute our first example of
GME states with a bilocal model for any N.

Let us now consider the Werner states $\rho_W(p) = p[2/d(d-1)]P_{\mathrm{asym}} + (1-p)\mathbb{1}_{d^2}/d^2$, where P_{asym} stands
for the projector onto the antisymmetric subspace of
$\mathbb{C}^d\otimes\mathbb{C}^d$. Applying the isometries defined above to
$\rho_W(p)$, one constructs the following N-qudit states:

$$\sigma'_{\mathbf{A}}(p) = p\frac{2\tilde{P}_{L,d}}{d(d-1)} + (1-p)\frac{\mathcal{P}_{L,d}\otimes\mathcal{P}_{N-L,d}}{d^2}, \quad (11)$$

with　$L = 1,\dots,N-1$,　where　$\tilde{P}_{L,d} = \sum_{i<j}|\psi_{ij}\rangle\langle\psi_{ij}|$
with　$|\psi_{ij}\rangle = (1/\sqrt{2})(|i\rangle^{\otimes L}|j\rangle^{\otimes(N-L)} + |j\rangle^{\otimes L}|i\rangle^{\otimes(N-L)})$.
The Werner states have a local model for GMs for $p \le (3d-1)(d-1)^{d-1}/d^d(d+1)$ [7] and so do the states
$\sigma'_{\mathbf{A}}(p)$ for any L. Moreover, $\rho_W(p)$ are entangled for
$p > 1/(d+1)$; thus, $\sigma'_{\mathbf{A}}(p)$ are GME for the same range
of p.

Let us finally notice that our construction can be
generalized to the case when the initial bipartite state is
replaced by a GME state with a local model for GMs. More
concretely, any genuinely entangled K-partite state admit-
ting a fully local model gives rise to a genuinely entangled

N-partite state, with any $N > K$, having a K-local
model [21,23].

Operational definitions of nonlocality.—Let us now
show that the inequivalence between entanglement and
nonlocality also holds true if the operational definitions
of nonlocality from Refs. [18,19] are used. The strongest
one—in the sense that if satisfied, it implies that the other
ones hold—requires that all $p_k(\mathbf{a}_{S_k}|\mathbf{M}_{S_k},\lambda)$ in Eq. (2) obey
the no-signaling principle. Crucially, this condition can be
easily met in our construction.

Let us consider a class of local models in which the
classical variable λ [cf. Eq. (6)] is the classical description of
a quantum state $|\lambda\rangle$ and one of the response functions, say the
one corresponding to B, is also quantum and given by the
formula $p(b|M_B,\lambda) = \mathrm{Tr}[M_b^B|\lambda\rangle\langle\lambda|]$. These models are also
known as local hidden state models in the context of
steering. Consider now an entangled state that admits one
of these models and apply our general construction with
$L = 1$. As proven in Lemma 1, the state (5) has a bilocal
model (8) with respect to the bipartition $A_1|A_2\dots A_N$ in which
the response function corresponding to the group $\bar{S}$ takes the
form　$p_\sigma(\mathbf{a}_{\bar{S}}|\mathbf{M}_{\bar{S}},\lambda) = \mathrm{Tr}[(M_{a_2}^{(2)}\otimes\dots\otimes M_{a_N}^{(N)})\Lambda_{B\to\bar{S}}(|\lambda\rangle\langle\lambda|)]$.
Clearly, $\Lambda_{B\to\bar{S}}(|\lambda\rangle\langle\lambda|)$ is a quantum state, so the latter
probability distribution is quantum and, as such, obeys
the no-signaling principle [21].

Now, the isotropic and Werner states do have local
models with one of the response functions being quantum
(cf. Ref. [24] for a recent review on local models), so the
states $\sigma_{\mathbf{A}}$ [Eq. (10)] and $\sigma'_{\mathbf{A}}$ [Eq. (11)] with $L = 1$ have
bilocal models in which the response functions correspond-
ing to the $N-1$ parties are nonsignaling. Consequently,
the inequivalence between entanglement and nonlocality
holds even when using the operational definitions of K-
locality [18,19].

Conclusions and discussion.—We have proven that
entanglement and nonlocality are inequivalent for any
number of parties N: for any N, we derived states that
do not violate any Bell inequality for genuine
N-partite nonlocality despite being genuinely N-partite
entangled. This has been achieved through an explicit
construction of bilocal genuinely multiparty entangled
states from bipartite local entangled states.

The most interesting open problem now is to understand
the extent to which the inequivalence holds. With the
current state of knowledge, our results show that there exist
GME N-partite states that have a two- or even three-local
model. Now, what is the maximum value of K such that
there exist N-partite states with a K-local model for
any N? Specifically, are there genuinely entangled
N-partite states with a fully local model? This happens
to be the case for $N = 2, 3$, but no results are known beyond
these two cases. A related question is whether there exists
some threshold value of N above which the GME states are
too entangled to allow for a fully local model.

Our result has been established for the standard notion of nonlocality associated with a Bell experiment. It is interesting to study whether the inequivalence between entanglement and nonlocality persists in other scenarios beyond the standard one introduced by Bell and considered here. A number of such operationally meaningful nonlocality scenarios were introduced after Bell's seminal work: the network approach [14], Bell scenarios defined on copies of a state [15] (see also Ref. [16]) or sequential measurements [11–13]. In these approaches, states that are local in the standard setup may display nonlocal properties. It is, in principle, possible that entanglement and nonlocality become equivalent in one of these extended scenarios. However, the study of this topic is much more difficult and the answer is unknown even for two parties. Moreover, the practical relevance of these scenarios is, at the moment, limited, as they all require entangling measurements and/or nondestructive quantum measurements.

Let us conclude by noting that our construction also implies that genuine multipartite entanglement is inequivalent to steering—another intriguing phenomenon of quantum information theory [25]. That is, by applying it to a bipartite state that has a local model with a quantum response function, the construction produces a GME state which is unsteerable (in at least one direction) across the same bipartition with respect to which it is bilocal. This will be worked out in detail in a further publication [23].

[1] J. S. Bell, Physics (Long Island City, N.Y.) **1**, 195 (1964).
[2] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[3] N. Gisin, Phys. Lett. A **154**, 201 (1991).
[4] N. Gisin and A. Peres, Phys. Lett. A **162**, 15 (1992).
[5] S. Popescu and D. Rohrlich, Phys. Lett. A **166**, 293 (1992).
[6] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[7] J. Barrett, Phys. Rev. A **65**, 042302 (2002).
[8] M. L. Almeida, S. Pironio, J. Barrett, G. Tóth, and A. Acín, Phys. Rev. Lett. **99**, 040403 (2007).
[9] G. Tóth and A. Acín, Phys. Rev. A **74**, 030306(R) (2006).
[10] M. T. Quintino (private communication).
[11] S. Popescu, Phys. Rev. Lett. **74**, 2619 (1995).
[12] N. Gisin, Phys. Lett. A **210**, 151 (1996).
[13] F. Hirsch, M. T. Quintino, J. Bowles, and N. Brunner, Phys. Rev. Lett. **111**, 160402 (2013).
[14] D. Cavalcanti, M. L. Almeida, V. Scarani, and A. Acín, Nat. Commun. **2**, 184 (2011).
[15] C. Palazuelos, Phys. Rev. Lett. **109**, 190401 (2012).
[16] Y.-C. Liang, Ll. Masanes, and D. Rosset, Phys. Rev. A **86**, 052115 (2012).
[17] G. Svetlichny, Phys. Rev. D **35**, 3066 (1987).
[18] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Phys. Rev. Lett. **109**, 070401 (2012).
[19] J.-D. Bancal, J. Barrett, N. Gisin, and S. Pironio, Phys. Rev. A **88**, 014102 (2013).
[20] For a linear map $\Lambda\colon \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})$, we define the dual map as a linear map $\Lambda^\dagger\colon \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})$ that satisfies $\mathrm{Tr}[X\Lambda(Y)] = \mathrm{Tr}[\Lambda^\dagger(X)Y]$ for any $X \in \mathcal{B}(\mathcal{K})$ and $Y \in \mathcal{B}(\mathcal{H})$. Recall that if Λ is positive (specifically, completely positive), its dual $\Lambda^\dagger$ is also positive. Moreover, if Λ is trace preserving, i.e., $\mathrm{Tr}[\Lambda(X)] = \mathrm{Tr}X$ for any X, the dual map $\Lambda^\dagger$ is unital, i.e., $\Lambda^\dagger(\mathbb{1}_\mathcal{K}) = \mathbb{1}_\mathcal{H}$, with $\mathbb{1}_\mathcal{X}$ being the identity operator acting on $\mathcal{X}$.
[21] See the Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.115.030404 for detailed proofs of Lemma 1 and 2, for a generalization of our method to the case of K-partite initial states, and for a detailed proof of the inequivalence under the operational definitions of nonlocality.
[22] M. Horodecki and P. Horodecki, Phys. Rev. A **59**, 4206 (1999).
[23] R. Augusiak et al. (to be published).
[24] R. Augusiak, M. Demianowicz, and A. Acín, J. Phys. A **47**, 424002 (2014).
[25] E. Schrödinger, Proc. Cambridge Philos. Soc. **31**, 555 (1935); H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).